The Pricing of Correlated Default Risk: *Evidence from the Credit Derivatives Market*

NIKOLA TARASHEV AND HAIBIN ZHU

NIKOLA TARASHEV
is a senior economist, Research and Policy Analysis at Bank for International Settlements in Basel, Switzerland.
nikola.tarashev@bis.org

HAIBIN ZHU
is a senior economist, Research and Policy Analysis at Bank for International Settlements in Basel, Switzerland.
haibin.zhu@bis.org

Portfolio credit risk has three key components: probability of default (PD), loss given default (LGD), and the probability distribution of joint defaults.[1] The last component, which has received the least attention owing to the traditional focus of academic researchers and market practitioners on single-name credit events, has recently gained in importance as a result of the rapid development of innovative products in structured finance. Such products, which allow investors to trade portfolio credit risk, include collateralized debt obligations (CDO), CDOs of CDOs (or CDO^2), nth-to-default credit default swaps (CDS) and CDS indices.[2] The prices of these financial instruments rely heavily on estimated probabilities of joint defaults (see Hull and White [2004]; Gibson [2004]) but there is no consensus on how market participants construct such estimates.

There are two popular approaches to deriving the probability distribution of joint defaults on the basis of firm-level data. One approach estimates this distribution directly from historical data on defaults (Daniels et al. [2008]; Demey et al. [2004]; Jarrow and van Deventer [2005]; Das et al. [2007]). Since defaults are rare events, however, this approach could lead to large estimation errors, especially for portfolios comprising investment-grade entities. The second approach derives the probability of joint defaults indirectly. This approach builds on the Merton [1974] framework and exploits 1) the notion that a default occurs when a borrower's assets fall below some threshold and 2) the insight that, since equity is a call option on the value of the firm, equity prices reflect asset values. Under the indirect approach, the probability distribution of joint defaults is driven by individual PDs and asset-return correlations, which are estimated from balance-sheet and stock market data.[3] Although it is inherently model dependent, the second approach possesses an important advantage because it incorporates large data sets.[4]

In this article, we follow the spirit of the second approach but propose a method for estimating the probability distribution of joint defaults on the basis of data from the rapidly growing CDS market. Our estimation method allows us to investigate to what extent the pricing of the credit risk of *individual* companies could help one understand market valuation of *portfolio* credit risk. As a concrete instance of such valuation, we consider five-year spreads on different tranches of a popular CDS index: the Dow Jones CDX North America investment-grade index. In order to dissect these tranche spreads, we use data on single-name CDS spreads and market expectations of LGDs, which imply two sets of credit risk parameters that vary over time and across firms: *risk-neutral* PDs (i.e., PDs incorporating a premium for the risk of individual defaults) and *physical* (i.e., actual) asset return correlations.[5] Together with the data on LGDs, these credit risk parameters deliver "predicted" tranche spreads that can be compared directly to their empirical counterparts from the CDS index market.

The two sets of estimated credit risk parameters relate to different aspects of the probability distribution of joint defaults and, as a result, have different impacts on predicted tranche spreads. On the one hand, an increase in PDs entails higher expected default rates, which raises the spreads of all tranches. On the other hand, a change in asset return correlations does not alter expected default rates but affects the probability of default clustering. In particular, higher correlations raise the probability of many defaults, which boosts the spreads of senior tranches, and also raise the probability of no defaults, which depresses the spreads of the first-loss (or equity) tranche.

We find that the index and single-name CDS markets employ similar risk-neutral PDs. There is a close match between the cross-sectional average of our PD estimates and the average PD implied by spreads on the overall (or single-tranche) CDS index. Furthermore, this match is stable over time and underpins the finding that PDs implied by the single-name CDS market explain to a large extent the intertemporal evolution of empirical tranche spreads of the CDS index.

That said, predicted spreads fall short of accounting for their empirical counterparts tranche by tranche. For instance, observed spreads of senior tranches imply prices of protection against catastrophe credit events that are much higher than those implied by predicted senior-tranche spreads. At the same time, observed equity tranche spreads reveal lower than predicted prices of protection against a small number of defaults. This points to a failure of the single-name CDS market to fully incorporate market perceptions of or attitude toward the risk of clustered credit losses. Given the constraints of the available data, we explain this failure by arguing that the *risk-neutral* asset return correlations, which are used for pricing portfolio credit risk, are higher than the physical correlations underpinning predicted spreads.

We interpret the wedge between risk-neutral and physical asset return correlations as reflecting a correlation risk premium. We measure this premium by subtracting a homogenized version of the physical asset return correlations, which stays constant across pairs of firms and fits predicted tranche spreads as closely as possible on each day, from the corresponding risk-neutral correlation, which is fitted to observed tranche spreads. Over our sample period, from 13 November 2003 to 18 March 2005, we estimate homogenized physical correlations that average 0.15 and risk-neutral correlations that average 0.25. The two sets of estimates translate into a sizeable and time-varying correlation risk premium.

We also study the joint intertemporal behavior of the correlation risk premium and the homogenized physical correlation. We find that the premium tends to decrease when the contemporaneous realization of physical correlations increases. Thus, when asset return correlations are higher, the market tends to anticipate lower upward correlation risk or have greater appetite for such risk.

By analyzing the pricing of portfolio credit risk on the basis of information from the single-name credit market, this article stands apart from the related literature, which has mainly focused on fitting flexible models to observed spreads of CDS index tranches. Longstaff and Rajan [2008], for example, use a multi-factor model of loan losses and derive that three credit risk factors—two of which occur with a low probability but have industry- and economy-wide impacts—are needed in order to explain fully tranche spreads of the CDS index considered here. Pursuing a similar objective, Kalemanova et al. [2007] and Moosbrucker [2006] demonstrate that tranche spreads of CDS indices are consistent with Lévy processes driving default trigger variables. In a different exercise, Hull et al. [2006] construct "implied correlations"—which are similar to the risk-neutral correlations derived here—under various modeling assumptions and explore how such assumptions can help account for empirical tranche spreads. However, in contrast to the analysis here, all these studies do not draw parallels between different credit markets and estimate neither physical asset return correlations nor a correlation risk premium. To the best of our knowledge, such a premium has been studied, but in the context of the equity market, only by Driessen et al. [2006]. In accordance with our results, this study also finds a large and time-varying correlation risk premium.

THE CDS INDEX MARKET

The market for a CDS index, which allows traders to buy and sell protection against portfolio credit risk, delivers two sets of prices. The first set is a time series of *index spreads*, which are effectively the prices of protecting the *entire* notional amount of the index against losses caused by defaults of the entities in this index. Under risk neutrality, index spreads reveal the market's *expectation* of such losses but are insensitive to the market's perception of and attitude towards the probability of default clustering.

These perceptions and attitude, do affect, however, the second set of prices, which comprises several time series of *tranche spreads*. Each time series consists of the effective prices of protection against a particular range (or "tranche") of credit losses on the notional amount of the index. For example, the tranche relating to the first losses—and, thus, carrying the highest level of credit risk—is known as the equity tranche. If none of the entities in the index defaults, the investor in this tranche (i.e., the protection seller) receives quarterly a fixed premium payment (or "spread") on the tranche's principal, which is typically 3% of the total notional amount of the index. If defaults occur, this investor stands ready to compensate its counterparty (i.e., the protection buyer) for any credit losses that do not exceed the outstanding principal of the equity tranche. At the same time, this principal and the associated premium payments are reduced for the remainder of the contract's life in order to reflect ongoing credit losses.[6] Similarly, an investor in the so-called mezzanine tranche is typically responsible only for losses between 3% and 7% of the total notional amount, while investors in the two senior and two super-senior tranches are responsible only for losses between 7% and 10%, 10% and 15%, 15% and 30%, and 30% and 100% of the total notional amount, respectively. Thus, the higher the seniority of the tranche, the less likely it is that the corresponding investor will need to make payments to the protection buyer.

A tranche spread is of great use to the analysis in this article because it pertains to a particular segment of the probability distribution of defaults and, thus, reveals the market's perception of and attitude towards default clustering. To see why, observe that, in a CDS index consisting of 100 equally-weighted entities with LGDs of 50%, the spread of the equity tranche is effectively the price of protection against the first six defaults in the underlying portfolio. For a given expectation of default losses, weaker interdependence of defaults across entities raises the probability of there being a few (i.e., up to six) defaults and, as a result, raises the spread of the equity tranche. Conversely, stronger interdependence of defaults increases the probability of default clustering—e.g., of there being zero or a lot of defaults—which lowers the equity tranche spread. At the same time, greater default clustering raises the spreads of the senior tranches, because (referring back to the stylized example) these spreads are the prices of protection against the 14th to the 20th and the 20th to the 30th defaults, respectively.

PREDICTED TRANCHE SPREADS

In order to obtain *predicted* tranche spreads for a CDS index, it suffices to input an estimate of the probability distribution of joint defaults and data on LGDs and risk-free rates in the numerical methodology developed in Gibson [2004]. For each particular tranche, this methodology delivers the expected present value of the principal, EP, and the expected present value of contingent payments, EC, made by the protection seller. Denoting the tranche spread by θ, the present value of the expected fee revenue of the protection seller, $\theta \cdot EP$, has to equal EC. The tranche spread is then calculated as:

$$\theta = \frac{EC}{EP}$$

Since this article relies on the Gibson [2004] methodology, the heart of the empirical exercise is the construction of the probability distribution of joint defaults. Under typical assumptions on the stochastic distribution of borrowers' asset returns (i.e., Gaussian or Student-t), the probability distribution of joint defaults has two key components: the PDs of constituent entities and the corresponding asset return correlations. The rest of this section describes how we estimate these two components on the basis of single-name CDS spreads.

Estimating Risk-Neutral PDs

In order to extract risk-neutral PDs from single-name CDS spreads, we follow the framework of Duffie [1999]. In a typical single-name CDS contract—written on firm i at date t—the protection buyer agrees to make constant periodic premium payments, determined by the CDS spread $s_{i,t}$, to the protection seller until the contract matures at time $t + T$ or a default occurs, whichever happens first. If a default occurs before $t + T$, the protection seller compensates the protection buyer for the realized credit loss.

Under market clearing, the present value of CDS premium payments (the left-hand side of Equation (1)) has to equal the present value of protection payments (the right-hand side):

$$s_{i,t}\int_{t}^{t+T} e^{-r_\tau \tau}\Gamma_{i,\tau}d\tau = LGD_{i,t}\int_{t}^{t+T} e^{-r_\tau \tau}q_{i,\tau}d\tau \qquad (1)$$

where r_τ stands for the risk-free rate of return, $q_{i,\tau}$ denotes the (annualized) unconditional risk-neutral default intensity of borrower i, $\Gamma_{i,\tau} \equiv 1 - \int_0^\tau q_{i,v}dv$ is the associated risk-neutral survival probability over the following τ years, and $LGD_{i,t} \in [0,1]$ is the date-t expectation of loss given default.[7] Under the standard simplifying assumptions that r_τ and $q_{i,\tau}$ do not change between time t and time $t + T$, Equation (1) implies that the one-year *risk-neutral* PD equals:

$$PD_{i,t}(1) = q_{i,t} = \frac{a_t s_{i,t}}{a_t LGD_{i,t} + b_t s_{i,t}} \qquad (2)$$

where $a_t \equiv \int_{t}^{t+T} e^{-r_t \tau}d\tau$ and $b_t \equiv \int_{t}^{t+T} \tau e^{-r_t \tau}d\tau$.

We use Equation (2) in order to estimate *daily* time series of borrower-specific risk-neutral one-year PDs on the basis of daily time series of CDS spreads, expected LGDs, and risk-free rates of return. Importantly, each of these PD estimates assumes a flat structure of the risk-neutral default intensity over the life of the corresponding CDS contract.

Estimating Physical Asset Return Correlations

We model the cross-sectional interdependence of default events as driven by the correlation of entity-specific random variables that trigger default. Each "default-trigger" variable is a *one-dimensional summary* of credit quality and is extracted from the corresponding CDS spread. As such, this variable comprises all the information that is deemed relevant by the single-name CDS market for the pricing of credit risk at the level of an individual entity. Such information includes: 1) balance sheet and stock market information about the entity, i.e., information that determines credit outlook in traditional structural model à la Merton [1974] and 2) information about systemic "frailty" factors, e.g., of the type examined recently by Das et al. [2007] and Duffie et al. [2007] in their study of default contagion.

In order to be able to draw straightforward parallels with the extant literature on portfolio credit risk (see Hull and White [2004]), we henceforth refer to the default-trigger variable as "the value of the firm's *assets*" and incorporate it in the following model. We start by assuming that, under the risk-neutral measure, the asset value process of entity i is:

$$\frac{dV_{i,t}}{V_{i,t}} = \mu dt + \sigma_i dW_{i,t} \qquad (3)$$

where μ denotes the drift, σ_i the asset volatility, and the shock $W_{i,t}$ is a standard Wiener process. Then, given a default boundary D_i, we define the distance-to-default variable $DD_{i,t} \equiv \frac{\ln V_{i,t} - \ln D_i}{\sigma_i}$. By Ito's Lemma, $dDD_{i,t}$ has a drift $\mu_i^* = \frac{\mu - \sigma_i^2/2}{\sigma_i}$ and a unit variance. Postulating that entity i defaults τ years into the future if $DD_{i,\tau} < 0$, we obtain that the probability of default equals:

$$PD_{i,t}^M\left(\tau; DD_{i,t}, \mu_i^*\right) = \Phi\left(\frac{-DD_{i,t} - \tau\mu_i^*}{\sqrt{\tau}}\right) \qquad (4)$$

where $\Phi(\cdot)$ stands for the standard normal CDF. Given two time series of PD estimates, the corresponding asset return correlation estimate equals:

$$\begin{aligned}
\rho_{ij} &\equiv corr(\Delta \ln V_{i,t}, \Delta \ln V_{j,t}) \\
&= corr(\Delta DD_{i,t}, \Delta DD_{j,t}) \\
&= corr(\Delta\Phi^{-1}(PD_{i,t}^M(1; DD_{i,t}, \mu_i^*)), \Delta\Phi^{-1}(PD_{j,t}^M(1; DD_{j,t}, \mu_j^*))) \\
&\approx corr(\Delta\Phi^{-1}(q_{i,t}), \Delta\Phi^{-1}(q_{j,t}))
\end{aligned} \qquad (5)$$

where the PD horizon τ is set to one year and Δ denotes the first difference in discrete time.

Repeating this exercise for each pair of entities $\{i, j\}$, we obtain an estimate of the correlation matrix of asset returns for the portfolio underpinning the date-t CDS index contract. We then assume that this estimate captures market perceptions of the "representative" asset-return correlations over the life of the corresponding contract and, thus, is used in setting tranche spreads.[8]

A calculation of asset return correlations on the basis of Equation (5) warrants several remarks. First, the model underlying this procedure assumes that the default boundary (D_i), asset volatility (σ_i), and drift (μ_i^*) remain constant over time. In order to relax this assumption, we would need to incorporate additional stochastic processes and estimate their parameters. Given the available data, the errors produced by such an estimation would be large enough to render the exercise useless. Second, the underlying model implies that $DD_{i,t}$ follow unit root processes, which is supported by the data.[9] Third, since $\Delta\ln(V_{i,t})$ stands for an *actual* asset return, Equation (5) delivers a measure of the *physical* asset return correlation. Finally, the fourth line in Equation (5) holds only as an approximation because $q_{i,t}$ is derived in Equation (2) under the assumption of a flat term structure of the default intensity, whereas $PD_{i,t}^{M}(\tau; DD_{i,t}, \mu_i^*)$ in Equation (4) violates this assumption. Robustness checks, outlined in the following subsection, reveal that the approximation in the fourth line of Equation (5) is remarkably good.

Alternative Estimators of Asset Return Correlations

The adopted mapping from PDs to physical asset return correlations can be criticized on the grounds of an inconsistency between two of its underlying Equations (2) and (4). In this subsection, we propose two alternative correlation estimators, which circumvent this inconsistency.

The first alternative is to estimate asset return correlations on the basis of distance-to-default variables, which are extracted directly from single-name CDS spreads and, in accordance with the Merton model, incorporate a non-flat term structure of the default intensity. Specifically, using the fact that $\Gamma_{i,\tau} = 1 - PD_{i,\tau}$ and $q_{i,\tau} = \frac{dPD_{i,t}^{M}}{d\tau}$, we incorporate the PD from Equation (4) into Equation (1). Given a value of the risk-neutral drift μ_i^*, the resulting equation implies that the CDS spread depends only on the distance-to-default variable. This allows us to generate a one-to-one mapping between time series of CDS spreads and time series of $DD_{i,t}$ and, by extension, estimate asset return correlations within a coherent framework.

The other alternative is in the spirit of the first but incorporates the relationship between PD and distance-to-default variables that is predicted by a model in which a default can occur at any point in time before the contract's maturity date. In this so-called first-passage model, the probability that firm i defaults over the next τ years equals:

$$PD_{i,t}^{FP}(\tau; DD_{i,t}, \mu_i^*) = 1 - \Phi\left(\frac{DD_{i,t} + \tau\mu_i^*}{\sqrt{\tau}}\right) + \exp(-2DD_{i,t} \cdot \mu_i^*)\Phi\left(\frac{-DD_{i,t} + \tau\mu_i^*}{\sqrt{\tau}}\right) \quad (6)$$

Using equation (6) to substitute for $\Gamma_{i,\tau}$ and $q_{i,\tau}$ in Equation (1) leads to an expression in which there is again a one-to-one mapping between date-t spreads and $DD_{i,t}$. Thus, for given values of the risk-neutral drifts and a panel of CDS spreads, we obtain a panel of distance-to-default values, which implies an alternative set of asset-return correlation estimates.

In order to carry out either of the two alternative procedures, we need to calibrate the risk-neutral drift μ_i^*. Given that estimates of these parameters are extremely noisy, we experiment with several values suggested by the literature, keeping each of these values constant in the cross section. We find that the correlation estimates are virtually insensitive to the exact value of μ_i^* and below report results only for $\mu_i^* = 0$.

Our overall empirical procedure can be summarized as follows. For each day and pair of entities in our sample, we estimate physical asset return correlations on the basis of the associated time series of PD estimates over the previous six months and a methodology outlined above.[10] Then, for each day in the sample, we rely on our PD and asset return correlation estimates and Monte Carlo simulations (outlined in Appendix A) in order to generate the probability distribution of joint defaults in the portfolio underlying the current CDS index. Used as an input to the Gibson [2004] methodology, this probability distribution leads to *predicted* spreads of the CDS index tranches.

It is worth emphasizing that, being calculated as described above, predicted spreads of CDS index tranches could be expected to be close to their empirical counterparts only under the hypothesis that the levels of asset-return correlations under the physical measure differ little from the corresponding levels under the risk-neutral measure. We maintain this hypothesis for our comparison between predicted and observed tranche spread. Of course, this hypothesis need not be true if the market requires non-negligible compensation for bearing correlation risk. We investigate this point later in this article.

DATA

We work with two large data sets, which are described in this section. In addition to these data, we obtain 5-year Treasury rates from Bloomberg, which we use to proxy for the risk-free rate of return.

The first data set is provided by JP Morgan Chase and pertains to 5-year contracts written on the CDS index

Dow Jones CDX North America investment-grade index (CDX.NA.IG.5Y). These standardized contracts are highly liquid on the secondary market. We use spreads for the entire "on-the-run" CDX.NA.IG.5Y index, as well as spreads for the equity, mezzanine and two senior tranches of the same index.[11] At each point in time the CDS index consists of 125 entities that represent major industrial sectors and are actively traded in the single-name CDS market as well. All entities have equal shares in the total notional principal of the index. The composition of the index is updated semi-annually—in a new release—in order to reflect events such as defaults, rating changes, and mergers and acquisitions. We consider three releases of the CDX.NA.IG.5Y index, launched respectively on November 13, 2003, March 23, 2004, and September 21, 2004. The total number of entities that appear in at least one of these releases is 136.

The second data set pertains to the single-name CDS market and is provided by Markit, which has constructed a network of leading market participants who contribute pricing information across several thousand credits on a daily basis. Markit aggregates the information it receives and releases daily a "consensus" CDS spread and expected LGD for each credit in its database.[12] In line with the contractual terms of the CDX.NA.IG.5Y index, we use time series of 5-year senior unsecured single-name CDS spreads associated with the no-restructuring clause (see ISDA [2003]) and denominated in U.S. dollars. For the period between April 24, 2003, and September 27, 2005, we download CDS spreads and the associated LGD estimates for all 136 entities that belonged to at least one of the CDS index releases we consider.

The data on single-name CDS spreads warrant two remarks. First, the spreads in our sample do appear highly responsive to changes in credit conditions. This property surfaces in that market staleness, defined as the realization of the same spread on two consecutive days, characterizes only 13% of our sample. By contrast, it is typical for spreads on the corporate bond market to be stale for weeks. The high quality of the CDS spreads prompts us to use daily data in computing asset return correlations. Second, the LGDs provided by Markit reflect market participants' consensus view on *expected* losses, and therefore need not match *realized* losses. The reported LGDs exhibit little cross-sectional difference and time variation (see Exhibit 1). In the cross section of 136 time averages of LGDs, the 1st and 99th percentiles equal 60% and 63% respectively. In addition, the time series of cross-sectional averages of LGDs fluctuates within a similarly narrow band. In the light of this and in order to eliminate potential noise in the LGD data, we set LGDs to be the same across entities and smooth the resulting time series via an HP filter with a parameter $\lambda = 64000$.

EMPIRICAL FINDINGS

In this section, we analyze predicted tranche spreads. These spreads are underpinned by two sets of credit risk parameter estimates—risk-neutral PDs and physical asset return correlations. We maintain the hypothesis that asset return correlations under the physical measure do not differ from those under the risk-neutral measure (i.e., that there is no correlation risk premium), which allows for a direct comparison between predicted and observed tranche spreads of the CDS index. This comparison reveals substantial differences between the two sets of spreads, suggesting a rejection of the maintained hypothesis. Then, we argue that these differences are robust to a number of alternative estimates of physical correlations and several pricing frameworks that might have been adopted by the market.

PD and Correlation Estimates

Using Equations (2) and (5), we estimate time series of *risk-neutral* PDs and *physical* asset return correlations, respectively. Exhibit 1 reports summary statistics for the two sets of credit risk parameters. The statistics relate to all 136 entities that belong to at least one of the first three releases of the CDS index CDX.NA.IG.5Y.

In general, both PDs and correlations change substantially over time. For example, the cross-sectional average of PDs peaks in August 2004 at about 1.2% and reaches its lowest levels, roughly 0.75%, towards the end of the sample period. As for asset return correlations, they are on a general upward trend during the sample period, increasing from around 7% in late 2003 to 17.5% in March 2005.

Despite the fact that all sample firms are investment-grade entities, the cross-sectional dispersion in the two sets of credit risk parameters is quite pronounced. Across the 136 entities, the time average of risk-neutral PD estimates has a mean of 94 basis points and a standard deviation of 78 bps, with the maximum level (425 bps) being 18 times higher than the minimum level (23 bps). Similarly, there is marked heterogeneity across pairwise correlation estimates. Correlations can be as high as 80–90% for firms in the same business area, as is the case for Ford

EXHIBIT 1
Risk Parameter Estimates

	mean	std. dev.	min	5%	25%	50%	75%	95%	max
LGDs (%)									
Daily averages	61.6	0.9	60.3	60.3	60.5	61.7	62.3	62.7	63.6
Averages over time	61.6	0.7	59.0	60.5	61.1	61.5	61.9	62.7	63.7
PDs (basis points)									
Daily averages	93.9	11.8	75.0	77.4	83.0	93.5	105.0	111.4	116.6
Averages over time	93.9	78.2	23.3	34.3	53.0	68.5	92.2	268.9	425.0
Pairwise correlations (%)									
Daily averages	12.0	2.9	6.6	7.0	10.5	12.6	13.7	16.7	17.5
Averages over time	12.0	7.5	-14.8	0.9	7.0	11.5	16.4	24.5	89.1

Note: The summary statistics refer to the sample period between November 13, 2003, and March 18, 2005, and reflect all 136 entities that belong to any of the first three releases of the CDS index CDX.NA.IG.5Y. In each panel, the first row is based on a time series of daily cross-sectional averages, whereas the second row is based on a cross section of time averages. The second and third variables are risk-neutral PDs and physical asset-return correlations, respectively, estimated on the basis of single-name CDS spreads. For each day in the sample, the correlation estimates rely on data covering the previous six months.

Motor Credit Company and General Motors Acceptance Corporation. At the other extreme, there are negative correlations, such as the one between Intel and Amerada Hess Corporation. In principle, as argued by Hull and White [2004], heterogeneity across PDs and pairwise correlations can have important implications for the probability distribution of joint defaults and, by extension, for the pricing of portfolio credit risk instruments.

Comparing Predicted with Observed Tranche Spreads

In order to construct predicted spreads for the four (i.e., equity, mezzanine, and two senior) tranches of the CDS index CDX.NA.IG.5Y, we use our LGD data and the PD and correlation estimates obtained from the single-name CDS market. In applying the Gibson [2004] methodology, we assume that all random variables are normal and there is no correlation risk premium. We refer to the resulting predicted tranche spreads as "baseline spreads" in order to distinguish them from alternatives reported in the next section. Exhibit 2 plots the time series of these spreads alongside their counterparts observed in the data and Exhibit 3 reports the corresponding summary statistics (baseline predicted spreads: first row in each panel; observed spreads: last row in the top panel).

Baseline predicted spreads differ substantially from their counterparts in the data. For the equity tranche, predicted spreads are too high over most of the sample and deviate from observed spreads by 10% on average. The differences are more pronounced in the two senior tranches where predicted spreads are too low over the entire sample period and underpredict observed spreads by 38% for the first and 63% for the second senior tranche. As for the mezzanine tranche, the predicted spreads match well the observed spreads on average (306.3 versus 303.9 bps), but this result masks large pricing differences on individual days. In fact, the mean absolute percentage error between the predicted and observed spreads for the mezzanine tranche is substantial, averaging 16% of observed spreads.

At the same time, predicted tranche spreads exhibit statistically significant, albeit not quite large, explanatory power for the evolution of observed spreads over time. Regression results, reported in Exhibit 4 (Panel A), reveal that changes in predicted tranche spreads account for 45% or more of the variability of changes in observed spreads for the equity, mezzanine and first senior tranches.[13] The goodness-of-fit measure drops to 16% for the second senior tranche.[14]

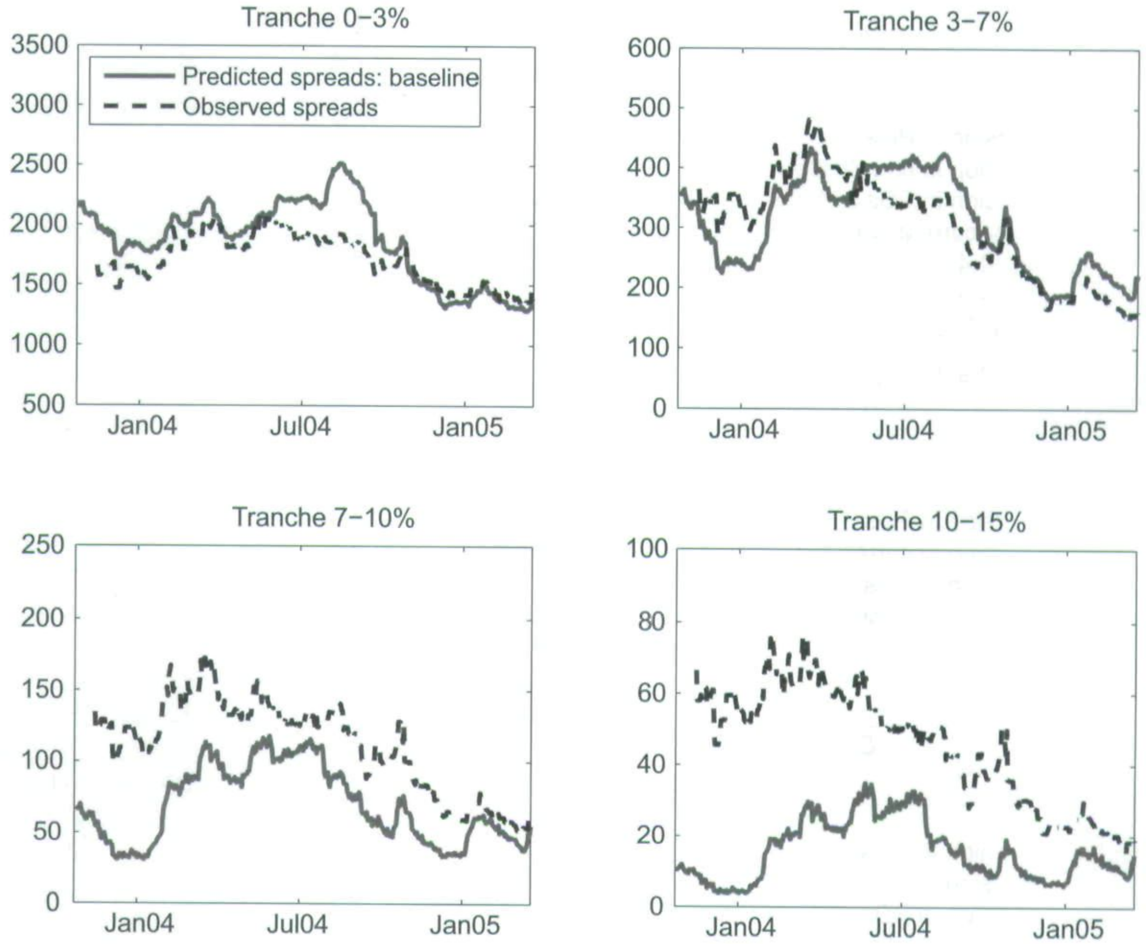
Robustness Checks

It is possible that the wedge between observed and baseline predicted spreads is due to the fact that market participants rely on estimates of physical correlations or pricing frameworks that are different from the ones we have considered so far. A series of robustness checks reveals that this conjecture is *not* borne out.

Robustness of predicted spreads to alternative estimates of physical correlations. Following the procedures outlined previously, we examine two alternative estimates

EXHIBIT 2

Predicted Versus Observed Spreads of CDS Index Tranches



Note: The observed tranche spreads in the CDS index market are provided by JP Morgan Chase. Predicted tranche spreads are based on firm-specific risk-neutral PDs and physical asset return correlations implied by the single-name CDS market as well as on a particular pricing algorithm. This algorithm incorporates Monte Carlo simulations and assumes normally distributed asset returns (see Appendix A).

of physical asset return correlations that circumvent the approximation in the fourth line of Equation (5). It turns out that the sample average of asset return correlations under the first (second) alternative procedure is 12.39% (12.41%), while the corresponding average of the original estimates—which underpin the baseline predicted spreads—equals 12.42%. A comparison across the time series of cross-sectional averages of original and alternative correlations reveals a maximum difference of 0.3 percentage points. In turn, considering the cross section of the time averages of correlation estimates, we find that the maximum difference is roughly 0.8 percentage points. Such differences have a negligible impact on predicted tranche spreads.

In the rest of the article, we keep on working with the original estimates of PDs and asset return correlations—given by Equation (2) and the fourth line of Equation (5), respectively—for the following reasons. First, as stated in the previous paragraph, the approximation in the fourth line of Equation (5) generates negligible errors in our correlation estimates. Second, in adopting this approximation, we work with PD estimates that do not depend on an estimate of the drift parameters $\{\mu_i^*\}$. To see this, compare the adopted Equation (2) with the alternative specifications in Equations (4) and (6). This is important because, unlike the correlation estimates, the PD estimates are sensitive to the values of $\{\mu_i^*\}$ but the available data do not allow us to pin down this value with satisfactory precision. Finally, it is reportedly popular market practice to adopt the

EXHIBIT 3
Predicted versus Observed Tranche Spreads

A. Predicted tranche spreads (averages), in basis points

	0-3%	3-7%	7-10%	10-15%
Baseline	1849.7	306.3	68.5	15.9
No dispersion in PDs	1849.3	315.3	74.5	18.7
No dispersion in correlations	1960.7	290.3	52.0	10.0
No dispersion in PDs and correlations	1929.5	307.4	62.0	13.4
One-factor correlation structure	1915.3	289.7	62.2	14.5
t-copula: (4,4)	2087.0	210.5	46.2	20.2
t-copula: (4,∞)	1899.7	244.8	60.1	24.5
Adjusted PD levels	1887.8	316.3	71.9	16.9
Observed tranche spreads	*1705.4*	*303.9*	*111.1*	*45.5*

B. Mean absolute error, in basis points

	0-3%	3-7%	7-10%	10-15%
Baseline	178.4	45.7	42.7	29.6
No dispersion in PDs	177.3	54.7	36.9	26.8
No dispersion in correlations	256.2	52.1	59.1	35.5
No dispersion in PDs and correlations	231.0	56.2	49.1	32.1
One-factor correlation structure	222.1	46.2	48.9	31.1
t-copula: (4,4)	371.8	95.2	65.3	25.5
t-copula: (4,∞)	222.5	62.3	51.4	21.6
Adjusted PD levels	171.1	53.1	39.1	28.1

C. Mean absolute percentage error, in percent

	0-3%	3-7%	7-10%	10-15%
Baseline	10.2	15.5	37.5	62.8
No dispersion in PDs	10.2	18.9	31.7	55.2
No dispersion in correlations	14.5	16.5	52.9	76.5
No dispersion in PDs and correlations	13.1	18.7	43.1	67.7
One-factor correlation structure	12.6	14.8	43.2	66.1
t-copula: (4,4)	21.0	30.8	57.8	51.4
t-copula: (4,∞)	12.7	18.9	44.0	41.9
Adjusted PD levels	9.8	20.1	32.8	58.9

Note: The reported averages are based on daily tranche spreads calculated (or observed) for the period between November 13, 2003, and March 18, 2005 (369 business days). Column headings specify the particular tranche of the CDS index. Row headings indicate alternative assumptions behind predicted spreads. The "baseline" numbers are based on firm-specific PDs and pairwise asset return correlations estimated from single-name CDS spreads, assuming that asset returns are Gaussian and Monte Carlo simulations (see Appendix A). The other results reflect variations on the baseline scenario and are obtained by: 1) removing the dispersion in PDs on each day; 2) removing the dispersion in correlation coefficients on each day; 3) removing the dispersion in both PDs and correlation coefficients on each day; 4) adopting a single-factor correlation structure that best fits the original correlation matrix (see Appendix B) and relying on a Gaussian copula (Appendix A); 5) adopting the same single-factor correlation structure as in (4) but relying on a "t-copula" (the two numbers in parentheses refer to the degrees of freedom of the common and idiosyncratic factors, respectively); and 6) adjusting the level of individual PDs so that the cross-sectional average PD equals the PD implied by the index spread on each day.

constant-default-intensity assumption of Equation (2) in pricing credit derivatives products. If this is true, then choosing a non-flat term structure of default intensities, as implied by Equations (4) and (6), would be supported by structural credit risk models but would introduce errors in predicted tranche spreads.

Robustness of predicted spreads to alternative pricing frameworks. In this subsection, we report the results of three robustness checks, which explore the implications of different pricing frameworks. Each of these frameworks incorporates a particular feature of reported market practice.

EXHIBIT 4
Explaining the Time Variation in Tranche Spreads

	in levels				in first differences			
	0-3%	*3-7%*	*7-10%*	*10-15%*	*0-3%*	*3-7%*	*7-10%*	*10-15%*
A. Dependent variable: observed spreads, in basis points (bps)								
Predicted spreads (bps)	**0.54**	**0.88**	**0.87**	**0.85**	**0.55**	**0.78**	**0.81**	**0.64**
	(15.6)	(11.3)	(8.7)	(4.2)	(7.7)	(9.8)	(7.6)	(3.8)
Adjusted R^2	0.78	0.64	0.52	0.19	0.46	0.58	0.45	0.16
B. Dependent variable: observed spreads (bps)								
PD (bps)	**15.34**	**5.39**	**2.16**	**0.89**	**15.74**	**4.99**	**2.26**	**0.86**
	(23.9)	(10.9)	(13.8)	(8.1)	(11.3)	(9.0)	(9.3)	(6.6)
Adjusted R^2	0.89	0.63	0.73	0.48	0.65	0.53	0.55	0.37
C. Dependent variable: observed spreads (bps)								
PD (bps)	**15.40**	**5.05**	**2.04**	**0.79**	**15.51**	**4.99**	**2.25**	**0.86**
	(23.5)	(11.3)	(15.2)	(8.8)	(10.9)	(8.7)	(9.0)	(6.4)
Correlation (%)	1.66	**-9.26**	**-3.28**	**-2.52**	4.63	0.05	0.25	0.03
	(0.5)	(4.5)	(5.3)	(6.0)	(0.7)	(0.0)	(0.2)	(0.1)
Adjusted R^2	0.89	0.71	0.81	0.66	0.65	0.53	0.55	0.37
D. Dependent variable: baseline predicted spreads (bps)								
PD (bps)	**24.49**	**6.26**	**2.01**	**0.58**	**24.23**	**6.22**	**1.68**	**0.39**
	(64.8)	(77.9)	(36.3)	(23.7)	(48.7)	(44.3)	(21.2)	(10.2)
Correlation (%)	**-27.94**	**3.00**	**4.15**	**1.77**	**-45.55**	**4.85**	**6.56**	**3.03**
	(16.8)	(8.5)	(17.0)	(16.4)	(20.0)	(7.6)	(18.0)	(17.5)
Adjusted R^2	0.99	0.99	0.95	0.91	0.97	0.97	0.93	0.88

Note: All variables are at the weekly frequency, equal weekly averages of daily data, and cover the period between November 13, 2003, and March 18, 2005. t-statistics are reported in parentheses and significant coefficients (at the 95% confidence level) are in bold. The dependent variable and the regressors are expressed in levels (left panel) or first differences (right panel). Column headings specify the particular tranche of the CDS index, while row headings refer to cross-sectional averages of risk-neutral PDs and physical asset-return correlations estimated from single-name CDS spreads. All regressions include a constant term, which is omitted from the exhibit.

For the first check, we calculate predicted spreads after removing the heterogeneity in PDs and/or pairwise correlations. This proxies for a scenario in which investors in the CDS index market lack information on firm-specific PDs or are unable to derive asset return correlation for each pair of firms. The results of this exercise, summarized in Exhibit 3 (rows 2–4 in each panel), reveal limited revisions to predicted tranche spreads. In fact, shutting off the cross-sectional difference in correlations raises (lowers) the predicted spreads for the equity (two senior) tranches, leading to an even worse match with observed spreads.

Our second exercise is motivated by market commentary, which refers regularly to common factor models of asset returns.[15] We investigate whether the use of such a model for pricing purposes could have a material impact on predicted tranche spreads. We start by estimating a one-factor structure of the estimated correlation matrix (see Appendix B) and then employ this structure in a Gaussian copula to derive tranche spreads (see Appendix A). As reported in Exhibit 3 (row 5 in each panel), adopting a one-factor model has a limited impact on our results and, if anything, renders the difference between the two sets of spreads even larger.[16]

The last robustness check examines the pricing implications of alternative assumptions regarding the distribution of asset returns. Researchers (see Hull and White [2004]) and market practitioners have argued for the use of Student-*t* distributions, which perform better than a Gaussian distribution in accounting for the fat tails of asset returns observed in the data. In the light of this, we use a one-factor model of asset returns and assume that the common factor and/or idiosyncratic factors follow a Student-*t* distribution with four degrees of freedom.[17] To derive predicted tranche spreads under this assumption, we employ the so-called *t*-copula. The results are reported in Exhibit 3 (rows 6–7 in each panel). The fat tails implied by a Student-*t* distribution raise the probability of a large number of defaults, which reduces somewhat the gap between predicted and observed spreads for the second senior tranche. However, this is at the cost of larger gaps for the other three tranches.

THE PRICING OF PORTFOLIO CREDIT RISK

This section delves further into the differences between baseline predicted and observed tranche spreads. Specifically, we investigate whether the risk-neutral PDs and physical asset return correlations we extract from the single-name CDS market, albeit correctly estimated, differ from the credit risk parameters employed by CDS index market.

Market Estimates of PDs

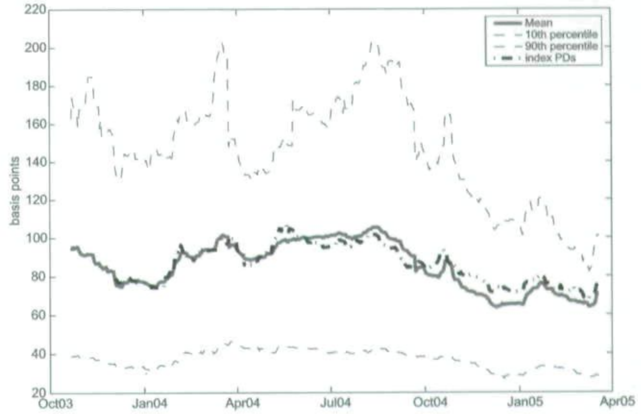
For the calculation of predicted tranche spreads of the CDS index, we have used individual risk-neutral PDs estimated on the basis of data from the single-name CDS market. However, structural differences between the index and single-name markets—driven, for example by differences in the respective settlement clauses—may induce investors to effectively adopt different sets of risk-neutral PDs. Hence, it is natural to ask if PDs are behind the tranche-specific gaps between predicted and observed spreads.

Intuitively, adopting a different set of PD estimates in the index market is not likely to explain these gaps. The reason is rooted in the fact that an overall rise in PDs translates directly into higher portfolio credit risk, which raises the spreads for all index tranches. Given that predicted spreads overshoot observed ones for the equity tranche but undershoot for the senior tranches, PD estimates that bring predicted spreads in line with observed spreads for some tranches are bound to increase the gap between the two sets of spreads for other tranches.

That said, CDS index spreads allow us to extract the cross-sectional averages of risk-neutral PDs used for pricing in the CDS index market. This PD series is plotted in Exhibit 5 alongside average PDs implied by the single-name CDS spreads. The two time series differ on average by only 1.21 bps, which is roughly 1.4% of the average PD implied by single-name CDS spreads. Importantly, these differences do not have a material impact on predicted tranche spreads (see Exhibit 3, row 8 in each panel) and we can conclude that the PD estimates used in the index market are consistent with the PDs embedded in the single-name CDS market.

Moreover, the close match between the two series of average PDs drives the similarity of the time paths of predicted and observed tranche spreads (see Exhibit 2). To substantiate this claim, we conduct regression analysis. As Exhibit 4 shows, PDs possess significant explanatory power for the level and changes in both predicted and observed tranche spreads (panels B to D). Strikingly, the goodness-of-fit measures indicate that changes in observed spreads are explained better by changes in PD estimates than by changes in predicted spreads. This is evidence that not PDs but another estimated component of predicted spreads weakens their co-movement with observed spreads. This other component is the correlation of asset returns, which we analyze in the following subsection.

EXHIBIT 5
Probabilities of Default



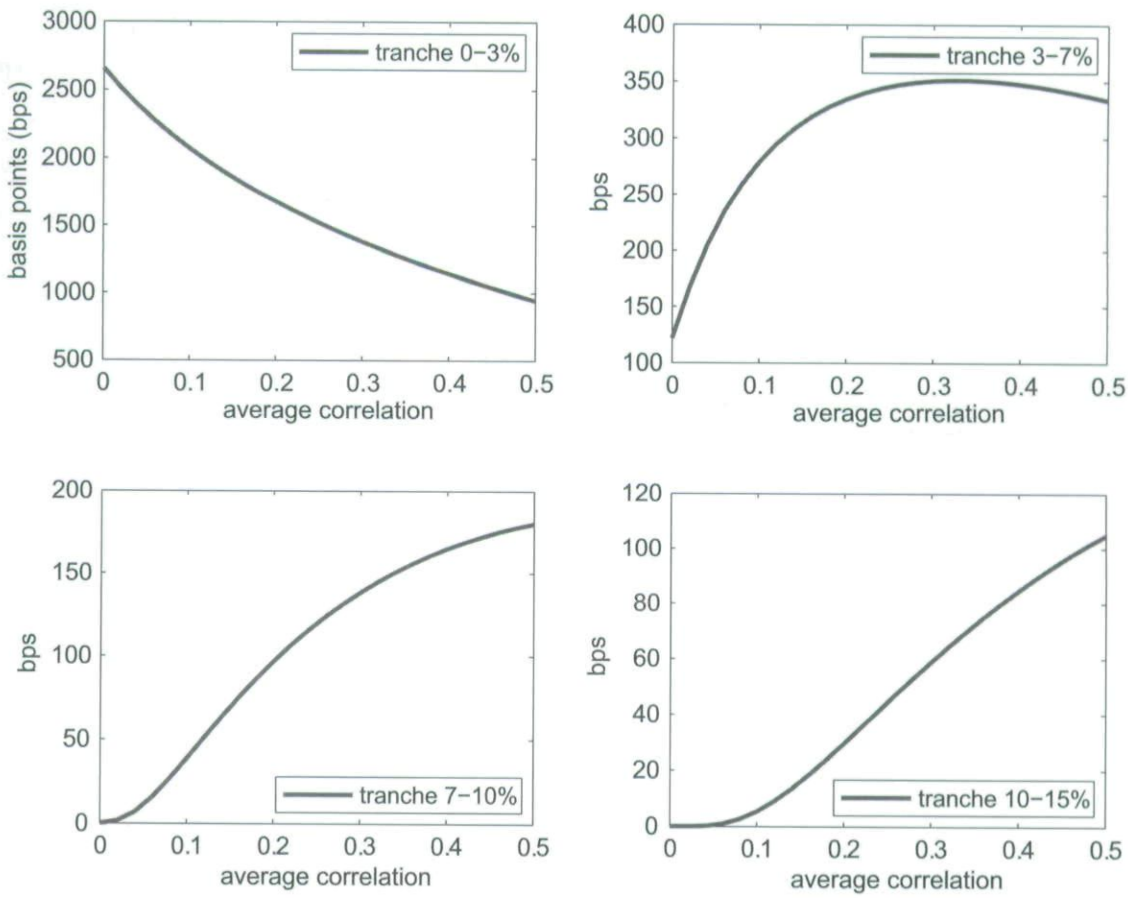
Note: The mean and the two percentiles are based on daily cross sections of risk-neutral PDs, which are estimated from the single-name CDS spreads of the 125 entities in the "on-the-run" release of the CDS index CDX.NA.IG.5Y. The series "index PDs" consists of the average PDs (across the same 125 entities) implied by the CDS index spreads.

Risk-Neutral Versus Physical Asset Return Correlations

In calculating predicted spreads, we have assumed that the *physical* correlation of asset returns implied by the single-name CDS market is used for pricing in the index market. Such an assumption is likely to be violated in practice. In fact, Driessen et al. [2006] find strong evidence that the *risk-neutral* correlations used by investors in equity-index options are substantially higher than the physical correlations implied by the single-name equity market. In order to examine whether a similar phenomenon exists in the credit market, we compare a homogenized version of our physical correlation estimates with the corresponding risk-neutral correlations, which we estimate on the basis of observed tranche spreads.

Correlation and the average level of tranche spreads. Risk-neutral asset return correlations that are larger than their physical counterparts can help explain the gap between predicted and observed spreads. This is suggested by Exhibit 6, which illustrates the sensitivity of spreads to changes in (homogenous) correlation coefficients. When the correlation coefficient increases, the default of a particular firm is more likely to be driven by the deterioration of common risk factors and hence is more likely to be accompanied by defaults of other firms. This raises (lowers) the probability of a large (small) number of defaults. As a result, higher asset return correlations lead to higher (lower) spreads for the senior (equity) tranches of a CDS index but have an indeterminate impact on the mezzanine tranche. In light of Exhibit 2, this implies that replacing the physical correlations underpinning

EXHIBIT 6
The Sensitivity of Tranche Spreads to Average Correlations



Note: This illustrative example uses the cross section of time averages of risk-neutral PD estimates and the average LGD and risk-free rate in our sample. All pairwise correlations are held fixed in the cross section.

predicted spreads with higher risk-neutral correlations would help us account to a greater extent for observed spreads.

Correlation and changes in tranche spreads over time. Differences between the physical correlation of asset returns and its risk-neutral counterpart could also account for the intertemporal evolution of the differences between predicted and observed tranche spreads. To see this, recall Exhibit 6, which implies that if physical correlations match closely risk-neutral ones, then an increase in their level should lower equity tranche spreads and raise senior tranche spreads. However, regression results reported in Exhibit 4 (panel C) show that physical correlations enter either with the wrong sign ("level" regressions) or are statistically insignificant ("first differences" regressions) as explanatory variables of observed spreads.

Exhibit 7 provides further evidence that physical correlations, which underpin predicted spreads, differ from the risk-neutral correlations behind observed spreads. This exhibit contains results from regressions of the gap

EXHIBIT 7
Explaining Errors in Predicted Tranche Spreads

	in levels				in first differences			
	0-3%	*3-7%*	*7-10%*	*10-15%*	*0-3%*	*3-7%*	*7-10%*	*10-15%*
	A. dependent variable: prediction error (bps)							
Dependent variable (-1)	**0.96**	**0.97**	**0.97**	**0.97**	**0.36**	0.20	0.07	-0.03
	(28.3)	(31.9)	(27.4)	(31.2)	(3.2)	(1.7)	(0.6)	(0.2)
Adjusted R^2	0.92	0.94	0.92	0.93	0.12	0.03	0.00	0.00
	B. dependent variable: prediction error (bps)							
Correlation (%)	**-38.30**	**11.39**	**7.40**	**4.41**	**-41.46**	**6.04**	**5.75**	**2.59**
	(5.8)	(6.3)	(18.0)	(12.7)	(4.8)	(2.5)	(5.5)	(4.0)
Adjusted R^2	0.32	0.35	0.82	0.70	0.24	0.07	0.30	0.18
	C. dependent variable: prediction error (bps)							
PD (bps)	**10.31**	0.75	-0.30	**-0.37**	**6.28**	**1.46**	-0.26	**-0.32**
	(8.5)	(1.6)	(1.4)	(2.8)	(3.1)	(2.8)	(0.9)	(2.1)
Adjusted R^2	0.50	0.02	0.01	0.09	0.11	0.08	0.00	0.05
	D. dependent variable: prediction error (bps)							
Dependent variable (-1)	**0.84**	**0.89**	**0.69**	**0.82**	0.10	0.15	0.01	-0.08
	(14.4)	(25.2)	(11.7)	(16.4)	(1.0)	(1.3)	(0.1)	(0.8)
Correlation (%)	-2.41	**2.39**	**2.59**	**0.94**	**-48.17**	**4.68**	**6.28**	**3.05**
	(0.8)	(3.5)	(5.4)	(3.7)	(6.2)	(2.0)	(6.0)	(4.9)
PD (bps)	**2.07**	**0.27**	-0.01	-0.03	**7.63**	1.00	**-0.61**	**-0.51**
	(3.1)	(2.2)	(0.2)	(0.9)	(4.3)	(1.9)	(2.6)	(3.6)
Adjusted R^2	0.93	0.95	0.94	0.94	0.46	0.12	0.34	0.30

Note: All variables are at the weekly frequency, equal averages of daily data and cover the period between November 13, 2003, and March 18, 2005. t-statistics are reported in parentheses and significant coefficients (at the 95% confidence level) are in bold. The dependent variable and the regressors are expressed in levels (left panel) or first differences (right panel). The dependent variable, prediction errors, is defined as basline predicted tranche spreads minus observed spreads in the CDS index market (See Exhibit 3 for a description of baseline predicted spreads). Explanatory variables, specified in row headings, include the first lag of the dependent variable and cross-sectional averages of physical correlations and risk-neutral PDs. All regressions include a constant term, which is omitted from the exhibit. Column headings specify the particular tranche of the CDS index.

between predicted and observed tranche spreads and reveals that physical correlation of asset returns is by far the most important driver of this gap. Specifically, the physical correlation enters with a statistically significant positive coefficient the regressions for the two senior tranches. This reveals that increases in the physical correlation, which raise the predicted spread for the senior tranches, are not associated with a concurrent rise in observed spreads for those tranches. A symmetric reasoning applies to the "equity tranche" regressions.[18]

Correlation risk premium. This subsection describes our measure of a correlation risk premium in the CDS index market. This premium reflects the market-determined compensation for bearing the risk that physical asset return correlations may *increase* above their expected level.

Thus, a change in the correlation risk premium has different impacts on the spreads of different index tranches. An increase in this premium signifies a higher value of protection against a large number of defaults, which inflates the spreads of senior tranches. The flipside of this is that an increase in the correlation risk premium is tantamount to a lower value of protection against just a few defaults, which depresses the spread of the equity tranche.

In order to derive the correlation risk premium, we start by extracting a time series of risk-neutral correlations from observed tranche spreads in the index market. On each day, the risk-neutral correlation has two general properties. First, it is constrained to be the same for all pairs of entities and is used to calculate four "fitted" spreads: for the equity, mezzanine, and two senior tranches. Second, the specific value of the risk-neutral correlation is picked so that it minimizes the mean squared *percentage* difference between the fitted spreads it implies and the corresponding observed spreads.[19] Importantly, our estimate of risk-neutral correlations is not allowed to vary across tranches and incorporates PDs that vary in the cross section. For these two reasons, risk-neutral correlations are different from the popular "compound" or "base" correlations that are tranche-specific and underpinned by homogenous PDs.[20]

Not surprisingly, fitted spreads perform better than predicted ones in matching observed tranche spreads. This can be visualized by comparing Exhibits 8 and 2. The overall improvement in the match, which is particularly visible for the two senior tranches, reveals that the risk-neutral correlations we estimate capture to a large extent the market's valuation of the risk of default clustering. At the same time, the rather rough match in the equity and mezzanine tranches[21] suggests that asset return correlations miss some information that is necessary for a full account of all tranche spreads of the CDS index.[22] However, our goal in this article is not to provide such a full account but to employ data from the single-name credit market for the understanding of the pricing of portfolio credit risk. This is what our estimates of risk-neutral correlations allow us to do.

The second component of the correlation risk premium is a "homogenized" physical correlation. This correlation is defined in the same way as its risk-neutral counterpart but is estimated on the basis of predicted, as opposed to observed, spreads. As a result, on each day, the homogenized physical correlation condenses into a scalar the information contained in the physical correlation matrix estimated from single-name CDS spreads. Importantly, the homogenized physical correlation fits predicted spreads extremely well across all tranches, entailing an average absolute percentage error of less than 2%.[23]

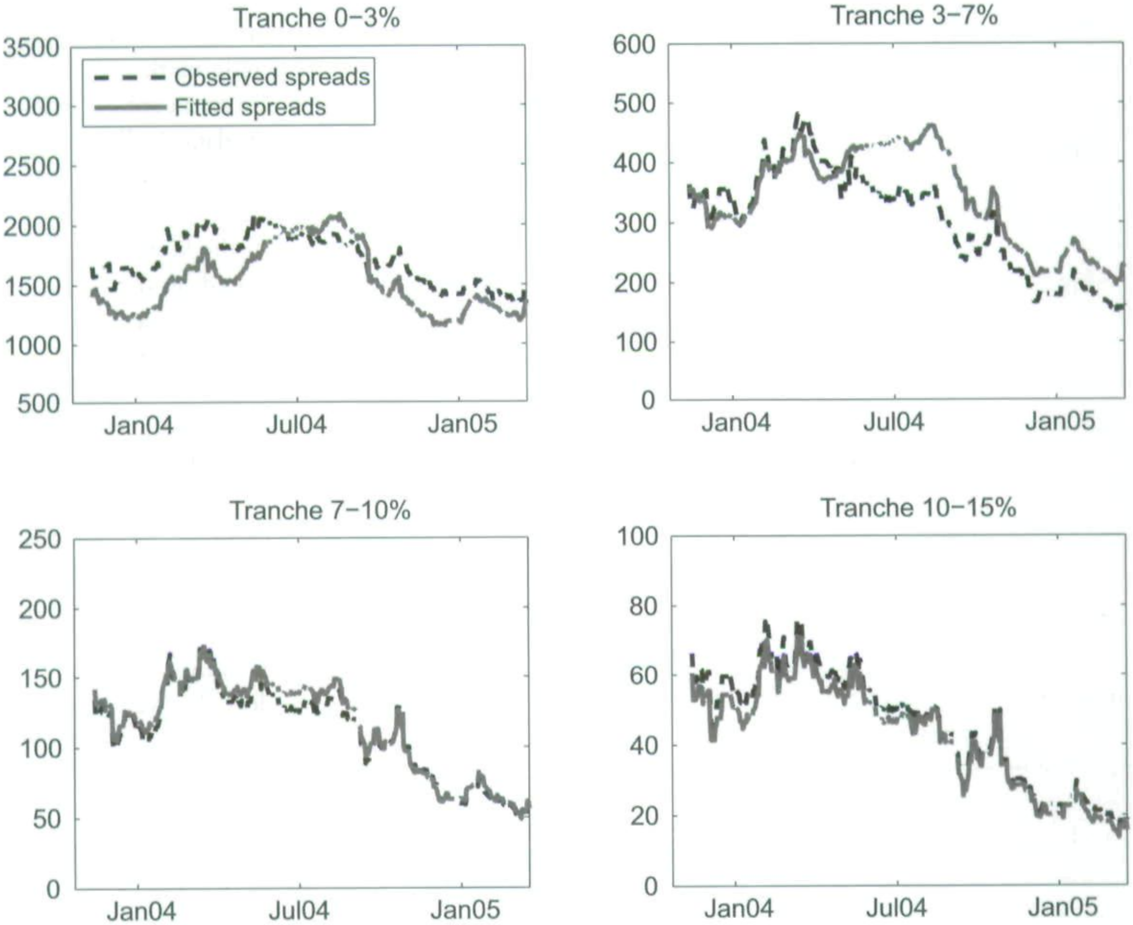
We define the correlation risk premium as the risk-neutral minus the homogenized physical correlation. The time series of this premium is plotted in Exhibit 9, alongside the time series of the two correlation estimates. During the sample period, the average of the correlation risk premium equals 10.1%, i.e., 66% of the homogenized physical correlation, which averages 15.3%. This result is in line with the finding of Driessen et al. [2006] that between 1996 and 2003, the risk-neutral correlation implied by equity-index options is 63% higher than its physical counterpart.

Intertemporal links between the correlation risk premium and physical correlations. Exhibit 9 (bottom panel) highlights the negative relationship between the correlation risk premium and the physical correlation. For instance, the premium peaks at the end of 2003, exactly when the physical correlation attains its lowest levels. More generally, the correlation risk premium is on a downward path over the entire sample period, while the physical correlation is on an upward path. In the rest of this subsection, we examine more closely the joint behavior of the correlation risk premium and the physical correlation.

In interpreting the findings of such an examination, it is important to keep in mind that the correlation risk premium is constructed in this article on the basis of one forward-looking and one backward-looking variable. On the one hand, the risk-neutral correlations are extracted directly from market spreads and, thus, reflect investors' forward-looking perceptions and attitudes

E X H I B I T 8

Matching Observed Tranche Spreads



Note: Fitted spreads are based on daily estimates of the risk-neutral correlations, which do not vary either across pairs of firms or across tranches. On each day, a single risk-neutral correlation minimizes the mean squared percentage error between observed and fitted spreads across the four tranches.

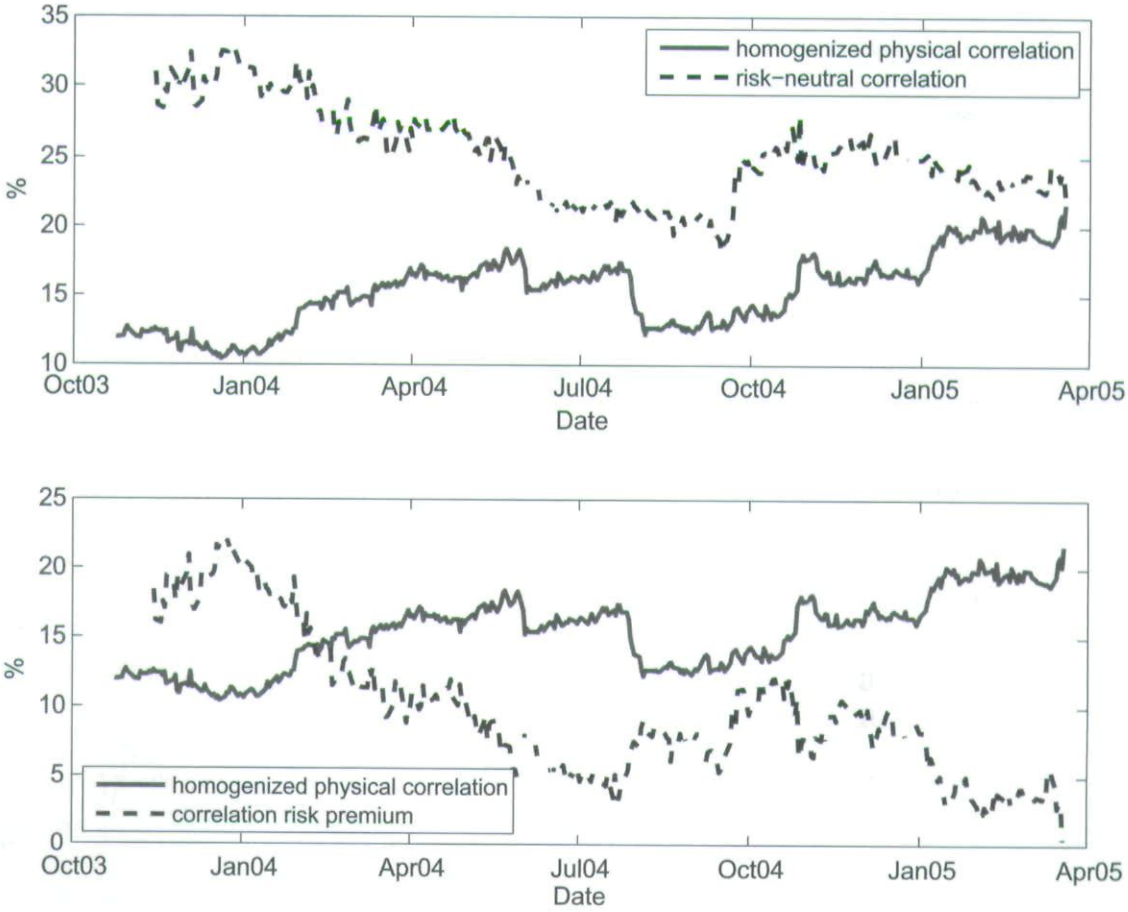
toward correlation risk. On the other hand, realized physical correlations are estimated from historical data, which renders them backward looking. As a result, the difference between these two variables, i.e., our estimate of the correlation risk premium, has three general components. The first component is due to a discrepancy between the currently realized value of physical correlations and the market's expectation of these correlations over the life of CDS contracts. The second component reflects the risk that future realized correlations are higher than their expected value. Finally, the third component is driven by market participants' appetite for correlation risk.

To the extent that the first component of the correlation risk premium estimate is small, the joint behavior of this estimate and realized physical correlations would reveal important insights about the CDS index market. In particular, there would be contemporaneous co-movement between the correlation risk premium and physical correlations if these correlations affect 1) market perceptions of the future volatility of asset return correlations or 2) the prevailing attitudes toward correlation risk. We investigate this co-movement via a regression analysis, in which we incorporate two additional (control) variables: lagged changes in the correlation risk premium and changes in the cross-sectional averages of PDs. As reported in Exhibit 10, change in the physical correlation is the only significant regressor and accounts for roughly 40% of the time variation in changes of the correlation risk premium. By contrast, changes in PDs—a standard proxy for the credit cycle—exhibit no statistical significance.

Exhibit 10 also reveals that the correlation risk premium tends to decline when the concurrent realization

Exhibit 9

Correlation Risk Premium



Note: On each day, both the risk-neutral and the homogenized physical correlations are constant across pairs of firms and the four (i.e., equity, mezzanine, and two senior) tranches. The risk-neutral correlation is fitted to baseline predicted observed tranche spreads, while the physical correlation is fitted to baseline predicted tranche spreads. The correlation risk premium is defined as the risk-neutral correlation minus the physical correlation.

of the physical correlation rises. This finding sheds light on market perceptions of and attitudes towards correlation risk. First, it is possible that investors perceive a negative relationship between the level and the volatility of correlation over time.[24] Thus, when our backward-looking estimate of physical correlations is high, this perception induces investors to demand a lower premium, as they attribute a lower probability to future large increases in the correlation level. In addition, a second possibility is that investors with higher tolerance for correlation risk tend to dominate the market when correlations are high. The concurrent drop of the correlation risk premium would then be a result of the lower price of risk demanded by such investors.

CONCLUSION

This article has analyzed the pricing of portfolio credit risk in the CDS index market on the basis of information obtained from the single-name CDS market. This analysis has revealed the existence of a large correlation risk premium, defined as the difference between the *risk-neutral* correlation of asset returns, used for pricing in the index market, and the corresponding *physical* correlation. This premium changes over time and co-varies negatively with current estimates of the physical correlation. The intertemporal behavior of the correlation risk premium, which would be usefully revisited by future research on the basis of longer data series, reveals information about market perceptions of and attitudes towards correlation risk.

Exhibit 10

Explaining the Correlation Risk Premium

	dependent variable: Δ correlation risk premium (%)			
Lagged dependent variable			0.09	0.05
			(0.8)	(0.5)
Δ homogenized physical correlation (%)		**-1.04**		**-1.01**
		(6.8)		(6.4)
Δ PD (bps)	-0.07			-0.04
	(1.4)			(0.8)
Adjusted R^2	0.02	0.40	0.01	0.41

Note: All variables are at the weekly frequency, equal first differences of weekly averages and cover the period between November 13, 2003, and March 18, 2005. t-statistics are reported in parentheses and significant coefficients (at the 95% confidence level) are in bold. The dependent variable, correlation risk premium, equals risk-neutral correlations minus homogenized physical correlations. On each day, the risk-neutral and the homogenized physical correlations are constrained to be the same across all pairs of firms and across the four index tranches. Given this constraint, the risk-neutral and physical correlations are fitted to observed and predicted tranche spreads, respectively, and minimize mean squared percentage errors. Explanatory variables include the first lag of the dependent variable, homogenized physical correlations and cross-sectional averages of risk-neutral PDs. All regressions include a constant term, which is omitted from the exhibit.

Appendix A

ESTIMATING THE PROBABILITY DISTRIBUTION OF DEFAULTS

This appendix outlines two methods for estimating the probability distribution of joint defaults in a given portfolio when the PDs of individual entities and the asset return correlations across entities are known. The first method relies on Monte Carlo simulations and does not impose any restriction on the structure of the correlation matrix. The second, copula, method requires the correlation structure to be driven by a finite number of common factors.

Monte-Carlo Simulation

Under this method, the probability distribution of defaults in a portfolio of N entities is derived as follows.

1. Generate N random draws x_0 from independent standard normal distributions.
2. Calculate $x = R'x_0$, where R denotes the Cholesky factor of the asset return correlation matrix for the N entities.
3. Denoting the i-th member of x by x_i ($i = 1, \ldots, N$) and the associated default probability by $PD_{i,t}$, entity i is said to default if and only if $x_i < \Phi^{-1}(PD_{i,t})$.
4. Repeat steps 2 to 4 a large number of times to estimate the probability of $n \in \{0, \ldots, N\}$ defaults.

Copula

The copula method, developed by Li [2000], Laurent and Gregory [2005], and Andersen and Sidenius [2005], relies on a common factor structure of asset returns. For simplicity, we describe the copula method assuming there is only one common factor but the method can be generalized to multiple common factors. Without loss of generality, all factors are assumed to have zero mean and unit standard deviation.

Denoting the common and idiosyncratic factors, the loading coefficient on the common factor and the unconditional PD by M_t, $Z_{i,t}$, α_i, and $PD_{i,t}$, we calculate the probability distribution of joint defaults in three steps. In the first step, we calculate the conditional default probability for entity i on date t: $PD(i \mid M_t)$. Postulating that a default is triggered when the asset value $V_{i,t} = \alpha_i M_t + \sqrt{1-\alpha_i^2} \cdot Z_{i,t}$ falls below some threshold, we obtain:

$$PD(i \mid M_t) = G\left(\frac{F^{-1}(PD_{i,t}) - \alpha_i M_t}{\sqrt{1-\alpha_i^2}}\right) \tag{7}$$

where G and F are the cumulative distribution functions of $Z_{i,t}$ and $V_{i,t}$, respectively. In general, there need not be analytical expressions for these distributions.

The second step is to calculate the conditional probability of an arbitrary number of defaults. Suppose we know the probability of $k \in \{0,1, \ldots, K\}$ defaults in a portfolio of K entities: $p^K(k \mid M_t)$. Then, adding one more entity leads to the following update of the distribution of defaults:

$$\begin{aligned}
p^{K+1}(0 \mid M_t) &= p^K(0 \mid M_t)(1 - PD(K+1 \mid M_t)) \\
p^{K+1}(k \mid M_t) &= p^K(k \mid M_t)(1 - PD(K+1 \mid M_t)) \\
&\quad + p^K(k-1 \mid M_t)PD(K+1 \mid M_t) \text{ for } k = 1, \cdots, K \\
p^{K+1}(K+1 \mid M_t) &= p^K(K \mid M_t)PD(K+1 \mid M_t)
\end{aligned}$$

This recursion starts with the initial condition $p^0(0 \mid M_t) = 1$.

The final step is to calculate the unconditional probability of k defaults in a portfolio of N entities:

$$p^N(k,t) = \int_{-\infty}^{\infty} p^N(k \mid M_t)\varphi(M_t)dM_t$$

where φ is the probability density function of M_t.

Appendix B

ESTIMATING A COMMON-FACTOR STRUCTURE OF ASSET RETURN CORRELATIONS

This appendix describes how we fit a common-factor structure to a given correlation matrix with entries ρ_{ij}, where i and $j \in \{1, \ldots, N\}$ and N is the size of the cross section. This matrix is to be approximated under the assumption that asset returns are underpinned by F common factors $M_t = [M_{1,t}, \ldots, M_{F,t}]'$ and N idiosyncratic factors $Z_{i,t}$:

$$\Delta \ln(V_{i,t}) = A_i M_t + \sqrt{1 - A_i' A_i} \cdot Z_{i,t}$$

where $A_i \equiv [\alpha_{i,1}, \ldots, \alpha_{i,f}, \ldots, \alpha_{i,F}]$, is the vector of common factor loadings, $\alpha_{i,f} \in [-1,1]$ and $\Sigma_{f=1}^{F} \alpha_{i,f}^2 \leq 1$. Without loss of generality, all common and idiosyncratic factors are assumed to be mutually independent and to have zero means and unit standard deviations.

We estimate the loading coefficients $\alpha_{i,f}$ ($i = 1, \ldots, N$, $f = 1, \ldots, F$) by minimizing the mean squared difference between the target correlations and the factor-driven correlations:

$$\min_{A_1 \cdots A_N} \sum_{i=2}^{N} \sum_{j<i} (\rho_{ij} - A_i A_j')^2$$

Andersen et al. [2003] propose an efficient algorithm to solve this optimization problem. Importantly, besides the "zero mean-unit variance" normalization, this estimation method imposes no restriction on the distribution of the common and idiosyncratic factors.

ENDNOTES

We are grateful to Dave Berger, Claudio Borio, Sanjiv Das, Joost Driessen, Song Han, Peter Hoerdahl, Nikunj Kapadia, David Lando, Srichander Ramaswamy, Kostas Tsatsaronis, Hao Zhou and seminar participants at the Bank for International Settlements, the 2006 FDIC Derivatives Conference, the 2006 CREDIT conference in Venice, the 2006 Meeting of the European Econometric Society in Vienna, the 2007 Federal Reserve Board credit risk conference, the Swiss Finance Institute conference on "Portfolio Management and Derivatives, " the 2007 BCBS conference on the "Interaction of Market and Credit Risk" in Berlin and the 2008 MFA conference in San Antonio for helpful comments. We also thank Marcus Jellinghaus for valuable help with the data. The views presented here are our own and do not necessarily represent those of the Bank for International Settlements.

[1]The mainstream of the credit risk literature focuses on PD: see Duffie and Singleton [2003] for an overview. The growing literature on LGD includes Altman and Kishore [1996], Jarrow [2001], Covitz and Han [2004], Pan and Singleton [2008] and Carey and Gordy [2006].

[2]For a general discussion of products used for trading portfolio credit risk, see BCBS [2004].

[3]The proprietary Global Correlation model by Moody's KMV is one such example. See Das and Ishii [2001] and Crosbie [2005] for details.

[4]Zhou [2001] attempts to relate the two approaches to estimating the probability distribution of joint defaults. Specifically, he studies analytically the link between asset return and default correlations in a first-passage credit risk model.

[5]In comparison to bond spreads, CDS spreads are widely considered as embedding less noisy information about market valuation of default risk. This is because CDS spreads respond more quickly to changes in credit conditions (Blanco et al. [2005]; Zhu [2006]) and are affected to a lesser extent by non-credit factors (Longstaff et al. [2005]).

[6]For the CDS index contract we consider below, a default triggers an immediate adjustment to the payments by the protection seller and buyer. In our calculations, however, we impose the simplifying assumption that such adjustments are made quarterly.

[7]Equation (1) assumes that expected LGD stays flat over the life of the CDS contract (i.e., between t and $t + T$). At the same time, it allows for time variation in expected LGDs.

[8]Ideally, one would estimate market perceptions of future asset-return correlations on the basis of a model that takes explicit account of the potential variability of these parameters over time (see Engle [2002] for example). Unfortunately, the available data sample does not allow us to carry out such an estimation.

[9]More precisely, a battery of Phillips-Perron tests fail to reject the unit-root null for 132 of the 136 distance-to-default time series we construct. In addition, a unit root process provides a reasonable approximation to the dynamics in the remaining four series.

[10]We also calculate time series of asset return correlations on the basis of three months of data. These alternative estimates are more volatile, both in the cross section and over time, but do not alter our main conclusions.

[11]We abstract from the two super-senior tranches because the spreads on these tranches are likely to be affected substantially by non-credit factors, such as administrative costs and a liquidity premium. Although the analysis of such factors is important, it is beyond the scope of this article.

[12]Given that it is designed to be consistent with the associated CDS spread, each expected LGD reported by Markit is derived under the risk-neutral measure.

[13]Exhibits 4 and 7 report two versions of each regression: one version uses the variables in levels and the other one in

first differences. While the coefficient estimates in the "levels" regressions are easier to interpret, the associated significance levels and goodness-of-fit measures are questionable because one cannot reject the unit-root hypothesis for both dependent and explanatory variables. This issue is addressed by the "first difference" regressions.

[14]All regressions in this article are based on weekly averages of the dependent and explanatory variables. This restricts the impact of market microstructure noise, which is likely to surface at high frequencies. Running the regressions with daily variables tends to preserve the sign and significance of the coefficients but, not surprisingly, lowers the goodness-of-fit measures (i.e. adjusted R^2).

[15]Also see Collin-Dufresne et al. [2003], Das et al. [2007], and Giesecke [2004] for discussions of why the assets of different firms may be driven by common factors.

[16]The one-factor approximation matches the mean of the original correlation matrix extremely well—producing an average discrepancy of 12 bps—but tends to underestimate the dispersion in correlation coefficients. To address this issue, we also estimate a two- and a three-factor correlation structures. These generalizations improve substantially the fit of the original correlation matrix but entail negligible revisions of predicted trance spreads.

[17]On the basis of an ad hoc value for asset return correlations, Hull and White [2004] find that assuming Student-t distributions with four degrees of freedom for both the common and idiosyncratic factors helps one account well for the tranche spreads of Dow Jones iTraxx EUR 5Y (the European counterpart of the CDX.NA.IG.5Y index).

[18]Exhibit 7 also reveals that PD coefficients are statistically significant and negative in the regressions of the errors in the predicted spreads for senior tranches. Although PDs have a low explanatory power in these regressions, negative coefficients are puzzling because they seem to suggest that a rise in PDs lowers the tranche spread. Background analysis reveals that this result is driven by a non-linear interaction between firm-specific PDs and pairwise asset return correlations.

[19]By minimizing percentage differences, we effectively equalize the units of the pricing errors across different tranches.

[20]For the estimation of risk-neutral correlations, we use the same LGDs and risk-free rates that underpin our estimates of risk-neutral PDs and physical asset return correlations.

[21]Specifically, the absolute percentage errors between fitted and observed spreads average 13%, 18%, 4%, and 7% for the equity, mezzanine, and the two senior tranches, respectively. The differences in the match across tranches is a manifestation of the so-called "correlation smile" phenomenon (see Amato and Gyntelberg [2005], for a review).

[22]To attain a better match, researchers have—explicitly or implicitly—modeled default trigger variables as driven by more general processes than the ones studied in this article (see Longstaff and Rajan [2008], Kalemanova et al. [2007], and Moosbrucker [2006], for example). These general processes effectively provide a greater number of degrees of freedom.

[23]The simple cross-sectional average of physical correlation coefficients is on average 2.9 percentage points lower than the homogenized correlation we work with. That said, switching between the two physical correlations affects immaterially the intertemporal properties of the correlation risk premium.

[24]Unfortunately, owing to the short sample period, we are not able to test for such intertemporal properties of the physical correlation of asset returns.

REFERENCES

Altman, E., and V. Kishore. "Almost Everything You Want to Know about Recoveries on Default Bonds." *Financial Analysts Journal*, 52 (1996), pp. 57–64.

Amato, J., and J. Gyntelberg. "CDS Index Tranches and the Pricing of Credit Risk Correlations." *BIS Quarterly Review*, March 2005, pp. 73–87.

Andersen, L., and J. Sidenius. "Extensions to the Gaussian Colupa: Random Recovery and Random Factor Loadings." *Journal of Credit Risk*, 1 (2005), pp. 29–70.

Andersen, L., J. Sidenius, and S. Basu. "All Your Hedges in One Basket." *Risk*, 16 (2003), pp. 67–72.

Blanco, R., S. Brennan, and I.W. March. "An Empirical Analysis of the Dynamic Relationship between Investment-Grade Bonds and Credit Default Swaps." *Journal of Finance*, 60 (2005), pp. 2255–2281.

Carey, M., and M. Gordy. "The Bank as Grim Reaper: Debt Composition and Recoveries on Defaulted Debt." Working paper, Federal Reserve Board, 2006.

Collin-Dufresne, P., R. Goldstein, and J. Helwege. "Is Credit Event Risk Priced? Modeling Contagion via the Updating of Beliefs." Working paper, Carnegie Mellon University, 2003.

Covitz, D., and S. Han. "An Empirical Analysis of Bond Recovery Rates: Exploring a Structural View of Default." Working paper, Federal Reserve Board, 2004.

"Credit Risk Transfer." Basel Committee on Banking Supervision Joint Forum, 2004.

Crosbie, P. "Global Correlation Factor Structure: Modeling Methodology." Moody's KMV Documents, 2005.

Daniels, R., S.J. Koopman, and A. Lucas. "A Non-Gaussian Panel Time Series Model for Estimating and Decomposing Default Risk." *Journal of Business and Economic Statistics*, 2008 (forthcoming).

Das, A., and S. Ishii. "Methods for Calculating Asset Correlations: A Technical Note." Moody's KMV Documents, 2001.

Das, S., D. Duffie, N. Kapadia, and L. Saita. "Common Failings: How Corporate Defaults are Correlated." *Journal of Finance*, 62 (2007), pp. 93–117.

Demey, P., J. Jouanin, C. Roget, and T. Roncalli. "Maximum Likelihood Estimate of Default Correlations." *Risk*, 17 (2004), pp. 104–108.

Driessen, J., P. Maenhout, and G. Vilkov. "Option-Implied Correlation and the Price of Correlation Risk." Working paper, 2006.

Duffie, D. "Credit Swap Valuation." *Financial Analysts Journal*, 55 (1999), pp. 73–87.

Duffie, D., L. Saita, and K. Wang. "Multi-Period Corporate Default Prediction with Stochastic Covariates." *Journal of Financial Economics*, 83 (2007), pp. 635–665.

Duffie, D., and K. Singleton. "Credit Risk: Pricing, Measurement, and Management." Princeton University Press, 2003.

Engle, R. "Dynamic Conditional Correlation–A Simple Class of Multivariate GARCH Models." *Journal of Buiness and Economic Statistics*, 20 (2002), pp. 339–350.

Gibson, M. "Understanding the Risk of Synthetic CDOs." Working paper, Federal Reserve Board, 2004.

Giesecke, I. "Correlated Default with Incomplete Information." *Journal of Banking and Finance*, 28 (2004), pp. 1521–1545.

Hull, J., M. Predescu, and A. White. "The Valuation of Correlation-Dependent Credit Derivatives Using a Structural Model." Working paper, University of Toronto, 2006.

Hull, J., and A. White. "Valuation of a CDO and an n-th to Default CDS without Monte Carlo Simulation." *Journal of Derivatives*, 12 (2004), pp. 8–23.

"International Swaps and Derivatives Association (ISDA) Credit Derivatives Defintions." ISDA Supplements and Commentaries, 2003.

Jarrow, R. "Default Parameter Estimation Using Market Prices." *Financial Analysts Journal*, 57 (2001), pp. 75–92.

Jarrow, R., and D. van Deventer. "Estimating Default Correlations Using a Reduced-Form Model." *Risk*, 18 (2005), pp. 83–88.

Kalemanova, A., B. Schmid, and R. Werner. "The Normal Inverse Gaussian Distribution for Synthetic CDO Pricing." *Journal of Derivatives*, 14 (2007).

Laurent, J., and J. Gregory. "Basket Default Swaps, CDOs and Factor Copulas." *Journal of Risk*, 7 (2005), pp. 103–122.

Li, D. "On Default Correlation: A Copula Approach." *Journal of Fixed Income*, 9 (2000), pp. 43–54.

Longstaff, F., S. Mithal, and E. Neis. "Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit-Default-Swap Market." *Journal of Finance*, 60 (2005), pp. 2213–2253.

Longstaff, F., and A. Rajan. "An Empirical Analysis of the Pricing of Collateralized Debt Obligations." *Journal of Finance*, 63 (2008), pp. 529–563.

Merton, R. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *Journal of Finance*, 29 (1974), pp. 449–470.

Moosbrucker, T. "Explaining the Correlation Smile Using Variance Gamma Distributions." *Journal of Fixed Income*, 16 (2006), pp. 71–86.

Pan, J., and K. Singleton. "Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads." *Journal of Finance*, 63 (2008, forthcoming).

Zhou, C. "An Analysis of Default Correlation and Multiple Defaults." *Review of Financial Studies*, 14 (2001), pp. 555–576.

Zhu, H. "An Empirical Comparison of Credit Spreads Between the Bond Market and the Credit Default Swap Market." *Journal of Financial Services Research*, 29 (2006), pp. 211–235.