# Feedback Between Credit and Liquidity Risk in the US Corporate Bond Market[☆]

Rob C. Sperna Weiland[a,*], Roger J.A. Laeven[b], Frank de Jong[c]

[a]*Department of Finance, Amsterdam Business School, University of Amsterdam, Plantage Muidergracht 12, 1018 TV Amsterdam, The Netherlands*

[b]*Department of Quantitative Economics, Amsterdam School of Economics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands*

[c]*Department of Finance, Tilburg School of Economics and Management, Tilburg University, Warandelaan 2, 5037 AB Tilburg, The Netherlands*

This Version: September 11, 2017

**Abstract**

We analyze the dynamic interactions between credit and liquidity risk and their impact on bond prices and risk. We propose a novel way of modeling credit-liquidity interactions through mutually exciting processes and develop a corresponding Bayesian estimation procedure. Using US corporate bond transaction data, we show that there is evidence of feedback between credit and liquidity risk and that this feedback is stronger for lower-rated bonds. We find that, on average, the credit-induced liquidity component accounts for 8% (AAA/AA) to 17% (B and lower) of total yield spreads, but in the most distressed periods it can account for over 40%.

*Keywords:* Credit risk, Liquidity risk, Corporate bonds, Mutually exciting processes,

Jumps, MCMC

*JEL:* C11, C58, G12

[☆]We are very grateful to Cynthia Wu and participants at the 10th Annual Society for Financial Econometrics (SoFiE) Conference in New York for their comments and suggestions. This research was funded in part by the Netherlands Organization for Scientific Research (NWO) under grant NWO Vidi 2009 (Laeven).

[*]Corresponding author.

*Email addresses:* R.C.SpernaWeiland@uva.nl (Rob C. Sperna Weiland), R.J.A.Laeven@uva.nl (Roger J.A. Laeven), F.DeJong@tilburguniversity.edu (Frank de Jong)

## 1. Introduction

It is well documented that credit and liquidity risk in the corporate bond market are interrelated and that liquidity conditions tend to worsen when credit risk manifests itself. Dick-Nielsen et al. (2012) and Friewald et al. (2012), for example, show that corporate bond market liquidity dried up substantially during the recent subprime crisis, and especially so for speculative grade bonds. In these studies, as well as in the majority of the empirical literature, the interaction between credit and liquidity risk is mainly addressed in a "static" way, by comparing liquidity measures across rating classes or by comparing tranquil periods with periods of distress. As argued by theoretical papers within the structural credit risk literature, however, it is plausible that credit-liquidity interactions are "dynamic" and that they contribute for a non-trivial part to bond yield spreads (see, e.g., He and Milbradt, 2014, and Chen et al., 2017).[1]

In this paper, we aim to provide more insight into the dynamic interaction between bond market credit and liquidity risk and its effects on yield spreads and risk. In order to do so, we introduce a new model, in which dynamic feedback loops between credit and liquidity risk are possible. The idea behind the feedback mechanism that we employ, is that adverse credit shocks increase the probability of occurrences of liquidity shocks, and vice versa. In this

[1]There are several reasons to expect dynamic interactions between credit and liquidity risk in the corporate bond market. As documented by Dick-Nielsen et al. (2012) and Bongaerts et al. (2017), credit spreads increase while corporate bond turnover decreases in distress times. The decrease in corporate bond turnover can partly be explained by a flight-to-liquidity type of behavior, where investors refrain from trading relatively illiquid corporate bonds and instead use more liquid assets for rebalancing and risk-shifting purposes. Due to the over-the-counter nature of the corporate bond market, the lower turnover increases dealers' search costs of finding counterparties and therefore induces them to quote higher bid-ask spreads, i.e., transaction costs increase when credit quality deteriorates. This supports a dynamic interaction from credit to liquidity. Another reason why higher credit risk can imply lower liquidity is the cost of market making. If higher credit risk is associated with higher bond price volatility, bond market makers will charge wider bid-ask spreads to cover their inventory costs (Ho and Stoll, 1981). Higher credit risk and higher volatility may also lead to increased information asymmetry among corporate bond traders and for this reason have a detrimental effect on liquidity, see, e.g., the standard microstructure model of Glosten and Milgrom (1985). He and Milbradt (2014) argue that there is also a channel from corporate bond market liquidity to credit risk. Starting from the observation that bond transaction costs increase in times of distress, they show that a decrease in bond market liquidity results in rollover losses, which in turn increases default risk.

way, our model is able to capture, in a dynamic way, episodes of distress in which credit risk builds up and market liquidity dries up. In contrast to the structural credit risk literature, however, we take an intensity-based approach, which has as main advantages that it is more parsimonious and, more importantly, much better able to describe bond price data.

The novelty of our modeling approach is that we model the default and liquidity intensities as bivariate self- and mutually exciting processes. The self-excitation in the credit and liquidity intensities captures the empirical observations that credit events tend to cluster in time (see Aït-Sahalia et al., 2014), and that liquidity risk is persistent (see, e.g., Acharya and Pedersen, 2005 and references therein). The mutual excitation between credit and liquidity gives rise to an explicit dynamic feedback mechanism over time that generates positive default-liquidity loops in which credit and liquidity shocks cluster in a flexible potentially asymmetric fashion.

We derive (semi-) closed-form bond price formulae, which depend on both intensity processes. More specifically, the liquidity intensity appears through a liquidity discount factor, reflecting the observation that higher illiquidity drives down prices (see, e.g., Duffie and Singleton, 1999, and Longstaff et al., 2005). As a consequence, credit (liquidity) shocks affect bond prices in two ways, namely by a self-exciting (cross-exciting) jump in the credit intensity, reflecting the increased likelihood of going into default, and by a cross-exciting (self-exciting) jump in the liquidity intensity, which increases the illiquidity discount. Compared to other intensity-based models, and in line with the theoretical models of He and Milbradt (2014) and Chen et al. (2017), our model thus facilitates a more refined decomposition of yield spreads, allowing us to explicitly quantify the price impact of and risk associated with pure credit risk, pure liquidity risk, and the two credit-liquidity interaction terms.

For our empirical analysis we use corporate bond transaction data from the Trade Reporting

and Compliance Engine (TRACE) database, which contains price information on practically all secondary market corporate bond transactions. For every transaction we obtain the corresponding bond contract specifications and credit ratings from the Mergent Fixed Income Database. Since most structural feedback mechanisms between credit and liquidity risk identified by the literature are of a firm-specific nature, we first illustrate our methodology in a firm-specific analysis. We use data on Ford Motor Company, because it is among the companies with the most, and most actively traded, bonds outstanding. Next to this case study, we consider bond portfolios sorted by credit rating. We conjecture that credit-liquidity interactions are particularly prevailing in distressed periods, and therefore we focus on the crisis period starting from July 2007 up to July 2009.

The main challenges in estimating our model stem from the fact that both the credit and liquidity intensity processes, as well as the credit and liquidity shocks, are inherently latent. Because of this latency, and the fact that we work in a bond context where data sampling is irregular, we cannot rely on existing estimation procedures dealing with self- and mutual excitation (see, e.g., Aït-Sahalia et al., 2015, and Boswijk et al., 2016). Instead, building on Eraker (2004), we propose a Bayesian Markov Chain Monte Carlo (MCMC) procedure based on a discretization of the intensity processes. We add an additional measurement equation to our model that explicitly incorporates liquidity information embedded in bid-ask spreads. We show in a Monte Carlo study that our estimation methodology yields good performance.

Our results show that the model captures the bond price data well and is able to identify multiple clusters of credit and liquidity shocks. We find strong evidence of both self- and mutually exciting behavior of credit and liquidity risk. We show that adverse credit (liquidity) shocks significantly affect bond market liquidity (credit risk) and that these cross-excitation effects tend to be stronger for bonds with a lower credit rating. In the Ford case study, we find a very pronounced feedback effect from credit to liquidity risk but only a small effect of

liquidity shocks on credit risk, despite the relatively low credit rating of Ford in our sample period.[2]

In line with the empirical literature, we find that the widening of yield spreads during the financial crisis can for a large part be explained by worsening liquidity conditions. However, our decomposition reveals that this illiquidity is in fact for a substantial part driven by deteriorating credit conditions. We find that the credit-induced liquidity component is especially prevalent for bonds with a lower credit rating and that, on average, it accounts for approximately 8% (AAA/AA) to 17% (B and lower) of total 10-year yield spreads. The decomposition fluctuates substantially over time, however, and we show that the credit-induced liquidity interaction term is substantially more important in the most turbulent times, where it can account for over 40% of total yield spreads. The liquidity-induced credit component is typically somewhat smaller and accounts for up to 6% of total yield spreads.

In the Ford case study, we also investigate the bond yield decomposition along the term structure dimension and show that the credit-liquidity interaction terms influence bonds of different maturities in different ways. We find a clear upwards sloping term structure effect of the credit-induced liquidity interaction term in which bonds with longer maturities are (relatively) more affected than bonds with shorter maturities. For example, shortly after the default of Lehman Brothers, the credit-driven liquidity component accounted for almost 60% of the 20-year yield spreads, whereas this component accounted for 35% of the 3-year yield spreads.

Finally, we show that our results may have important implications for standard risk management procedures such as Value-at-Risk computations. The self- and mutual excitation mechanisms induce a fat tail in the loss domain of the bond return distribution, and, there-

[2]The bonds of Ford were all rated B or lower throughout our sample. Since Ford did not roll over any debt during our sample period, this finding provides suggestive evidence that the debt rollover channel proposed by the theoretical credit risk literature (see, e.g., He and Xiong, 2012, and He and Milbradt, 2014) may indeed be an important channel to explain feedback effects of bond market liquidity on credit risk.

fore, ignoring credit-liquidity interactions may result in a considerable underestimation of capital requirements. We find, for example, that the 10-day 99% VaR of a 10-year Ford bond at the start of our sample decreases from 19% to 6% if one ignores credit-liquidity interactions.[3] These results highlight that a substantial part of the capital requirements is due to credit-liquidity interactions, and we show that this effect typically becomes more pronounced for bonds with lower credit ratings.

Our research is part of the literature on the interaction of credit and liquidity risk in the (corporate) bond market. Our paper belongs to the intensity-based credit risk literature. Within this literature, credit and liquidity risk are typically modeled by stochastic diffusive processes for the credit and liquidity intensities (see, e.g., Janosi et al., 2002, Duffie et al., 2003, Driessen, 2005, Liu et al., 2006, and Buhler and Trapp, 2010). By using (possibly correlated) diffusive components, these models typically only consider contemporaneous and symmetric interactions,[4] and are not able to address the (potentially asymmetric) lead-and-lag relationship of the interaction effects. One notable exception to this modeling approach is Monfort and Renne (2014), who model default and liquidity interactions through credit- and liquidity-related regimes which influence the process parameters of the default and liquidity intensities. Our model has a dynamic feedback mechanism through mutual excitation between credit and liquidity risk that does not assume symmetry between the prevalence of credit-liquidity interactions. Furthermore, our model enables us to study the causal impact of credit shocks on liquidity, and vice versa, over time, and to explicitly quantify the impact of the interaction terms—credit-induced liquidity and liquidity-induced credit—on bond yield spreads over time and across maturity.

[3]That is, for every dollar invested in this 10-year bond, the total VaR-induced capital requirements are 19 cents, of which 13 cents can be attributed to credit-liquidity interactions.

[4]Buhler and Trapp (2010) assume that default and liquidity intensities are linear combinations of two underlying independent factors, which they label the pure default and pure liquidity factors. In this way, they allow for asymmetry between the contemporaneous effects of credit on liquidity, and vice versa.

A few papers in the structural credit risk literature also address credit-liquidity interactions. Ericsson and Renault (2006) introduce a structural model in which the illiquidity of the market for distressed debt influences renegotiations in financial distress, and show that their model implies a decomposition of yield spreads into a pure credit component, a pure liquidity component, and an interaction term. He and Milbradt (2014) develop a structural model in which a default-liquidity loop arises through a debt rollover channel (see also He and Xiong, 2012). Assuming a relatively illiquid post-default secondary bond market, they show that higher default risk worsens a bond's pre-default secondary market liquidity. This, in turn, amplifies equity holders' rollover losses and increases the endogenous probability of default. In calibrating their model, they find that credit-liquidity interaction terms account for approximately 25% of the total credit spread. Nagler (2017) provides some empirical evidence that the rollover channel from bond liquidity to credit risk exists and has a substantial price impact around liquidity shocks. Chen et al. (2017) study the default-liquidity spiral across the business cycle. They find that default-liquidity interactions can account for 16% to 46% of the changes in spreads over time.

Our research is also related to the literature on MCMC methods in a financial context (see, e.g., Johannes et al., 1999, Eraker et al., 2003, Eraker, 2004, and Johannes and Polson, 2009). More specifically, we develop an estimation methodology that builds upon and significantly extends the procedure of Eraker (2004) for estimating different classes of jump diffusion models of stock price dynamics using joint option and stock price data. The main differences between our setting and his are that we consider irregularly sampled bond price data instead of option and stock data, that our state processes are latent instead of (partially) observable, and that our model set-up is more general, since we allow for self- and mutual excitation between the state variables instead of perfectly correlated jumps as in Eraker (2004). The result of these differences is that identification and estimation are more challenging in our model, and therefore we propose a different order of sampling in the MCMC algorithm.

The remainder of this paper is structured as follows: Section 2 describes the model set-up. Sections 3 and 4 contain the data description and estimation methodology, respectively. Section 5 presents the estimation results of our empirical analysis. Section 6 illustrates some of the implications of our findings, and Section 7 concludes. Proofs and some auxiliary details of the estimation methodology are relegated to two Appendices [A, B] and four Internet Appendices [IA.1, IA.2, IA.3, IA.4].

## 2. The Model

Throughout we consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\geq 0}, \mathbb{Q})$ satisfying the usual conditions of right-continuity and completeness. To model credit and liquidity interactions, we employ mutually exciting jump processes defined on this probability space. That is, we define two point processes $N_t^c$ and $N_t^l$, denoting the credit and liquidity shock arrival processes, respectively, with corresponding stochastic intensity processes $\lambda_t^c$ and $\lambda_t^l$, respectively. The jump intensities describe the $\mathcal{F}_t$-conditional probabilities of credit and liquidity shocks occurring in an interval of time $\Delta$ as follows:

$$\begin{cases} \Pr\left[N_{t+\Delta}^c - N_t^c = 0 \middle| \mathcal{F}_t\right] = 1 - \lambda_t^c \Delta + o(\Delta), \\ \Pr\left[N_{t+\Delta}^c - N_t^c = 1 \middle| \mathcal{F}_t\right] = \lambda_t^c \Delta + o(\Delta), \\ \Pr\left[N_{t+\Delta}^l - N_t^l = 0 \middle| \mathcal{F}_t\right] = 1 - \lambda_t^l \Delta + o(\Delta), \\ \Pr\left[N_{t+\Delta}^l - N_t^l = 1 \middle| \mathcal{F}_t\right] = \lambda_t^l \Delta + o(\Delta), \end{cases} \tag{1}$$

and $\Pr\left[N_{t+\Delta}^k - N_t^k > 1 | \mathcal{F}_t\right] = o(\Delta)$, for $k = \{c, l\}$. The feedback between the two point processes comes from the fact that both intensities respond to jumps in both processes,

supposed to occur in the following mean-reverting way:

$$\begin{aligned} \mathrm{d}\lambda_t^c &= \alpha_c(\lambda_\infty^c - \lambda_t^c)\mathrm{d}t + \sigma_c\sqrt{\lambda_t^c}\mathrm{d}W_t^c + \beta_{1,1}\mathrm{d}N_t^c + \beta_{1,2}\mathrm{d}N_t^l, \\ \mathrm{d}\lambda_t^l &= \alpha_l(\lambda_\infty^l - \lambda_t^l)\mathrm{d}t + \sigma_l\sqrt{\lambda_t^l}\mathrm{d}W_t^l + \beta_{2,1}\mathrm{d}N_t^c + \beta_{2,2}\mathrm{d}N_t^l, \end{aligned} \tag{2}$$

where $W^c$ and $W^l$ are independent Brownian motions capturing small fluctuations in credit and liquidity risk. In the case of a large credit shock (i.e., $\mathrm{d}N_t^c = 1$), the credit intensity jumps up by $\beta_{1,1}$ and the liquidity intensity by $\beta_{2,1}$. Similarly, in the case of a liquidity shock (i.e., $\mathrm{d}N_t^l = 1$), the liquidity intensity jumps by $\beta_{2,2}$ and the credit intensity by $\beta_{1,2}$. Hence, occurrences of large credit and liquidity shocks directly increase the probabilities of future credit and liquidity shocks via their direct impacts on the intensity processes. The self-excitation of the credit intensity captures the empirical observation that credit events tend to cluster. Similarly, the self-excitation in liquidity captures the persistence in liquidity risk. The mutual excitation between credit and liquidity provides a parsimonious way of modeling feedback between the two. Unlike contemporaneously correlated diffusive intensities, which are typically employed in the literature, this feedback mechanism is dynamic in nature. Furthermore, we do not impose any symmetry (i.e., we do not impose $\beta_{1,2} = \beta_{2,1}$) between the effects of credit risk on liquidity and vice versa.

It can be shown that the quadruple $\left(N^c, N^l, \lambda^c, \lambda^l\right)$ is jointly Markov (see, e.g., Aït-Sahalia et al., 2015), and, under suitable parameter restrictions, stationary and nonnegative. More specifically, to assure nonnegativity of the processes, we assume that $\lambda_\infty^k \geq 0$, $\alpha_k \geq 0$, $\sigma_k \geq 0$ $(k = c, l)$, $\beta_{i,j} \geq 0$ $(i, j = 1, 2)$, and that the Feller-conditions in the credit and liquidity intensity processes are satisfied (i.e., $2\alpha_c\lambda_\infty^c \geq \sigma_c^2$ and $2\alpha_l\lambda_\infty^l \geq \sigma_l^2$). In order to assure stationarity, we assume sufficiently strong mean reversion and we impose the sufficient conditions $\alpha_c > \beta_{1,1} + \beta_{1,2}$ and $\alpha_l > \beta_{2,1} + \beta_{2,2}$.

It would be feasible to let the diffusive parts be correlated, but we choose to only consider

interactions between liquidity and credit risk through the mutual excitation parameters $\beta_{1,2}$ and $\beta_{2,1}$. In our opinion, it is most interesting to focus on the feedback mechanisms that arise in episodes of distress (i.e., large credit or liquidity shocks), rather than the contemporaneous correlation between the two in tranquil periods. Furthermore, we find that the benefits of adding a correlation parameter do not outweigh the econometric costs of having to estimate this additional parameter.

The nature of our jump model, which aims to capture the subtle features of the credit and liquidity intensity dynamics, requires a sample with a sufficient number of jumps to properly estimate the jump intensity dynamics and the jump process parameters. This follows as a result of the Peso problem. In our empirical application, we therefore focus on the crisis period from July 2007 until July 2009. It would be tenuous, however, to accurately estimate risk premia based on this relatively short sample period. Since the main aim of this paper is to quantify credit-liquidity interactions, and not the evaluation of risk premia, we therefore choose to abstract away from risk premia. That is, we assume risk premia to be equal to zero.

### *2.1. Bond Pricing*

We derive our bond price formula by extending the results of Duffie and Singleton (1999). In that paper, the authors derive a pricing formula for a defaultable coupon bond using the doubly stochastic Poisson framework to model default risk. The authors find that the time $t$ value of an $n$-year defaultable bond with semi-annual coupon $c$ and face value 1, under the assumption of recovery of face value with recovery rate $(1 - L)$, is given by

$$\begin{aligned} P_t = c\sum_{j=1}^{2n} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{t+0.5j}(r_s+\lambda_s)\mathrm{d}s}\middle|\mathcal{F}_t\right] + \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{t+n}(r_s+\lambda_s)\mathrm{d}s}\middle|\mathcal{F}_t\right] \\ + \int_t^{t+n}(1-L)\mathbb{E}^{\mathbb{Q}}\left[\lambda_s e^{-\int_t^{t+s}(r_u+\lambda_u)\mathrm{d}u}\middle|\mathcal{F}_t\right]\mathrm{d}s, \end{aligned} \tag{3}$$

where $r$ is the risk-free interest rate, $\lambda$ the default intensity, and $\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T \lambda_s \mathrm{d}s}\middle|\,\mathcal{F}_t\right]$ the survival probability up to time $T > t$ under the assumption that a default occurs at the first jump of the underlying jump process. The first part of equation (3) is the current value of all future coupon payments, under the condition that there is survival until at least the coupon payment dates. The second part of (3) is the current value of the principal repayment. The third part of (3) is the value of the recovery payment in case of default. Typically a recovery rate of 40% is assumed (i.e., $L = 0.6$).

Equation (3) is derived under the standard assumption that the company defaults at the first jump of the underlying jump process. We assume, however, that upon the occurrence of a credit event, a default occurs with probability $0 < \gamma \leq 1$. Letting $\tau$ denote the time-to-default, the $\mathcal{F}_t$-probability of survival to time $T > t$ is then given by

$$p(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{Q}\left(\tau > T\middle|\,\mathcal{F}_T\right)\middle|\,\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[(1-\gamma)^{N_T^c - N_t^c}\middle|\,\mathcal{F}_t\right], \tag{4}$$

where we will use the convention $0^0 = 1$ and $0^c = 0$ for $c > 0$ as this makes formula (4) also valid for $\gamma = 1$.

Furthermore, (3) does not yet take liquidity effects into account and, therefore, in line with the intensity-based credit risk modeling literature, we introduce an additional liquidity discount factor which can be seen as a convenience yield that investors require for holding an illiquid (defaultable) asset (see, e.g., Duffie and Singleton, 1999). Finally, we assume that the risk-free rate is independent of the credit and liquidity intensities, and we denote the time $t$ price of a risk-free zero-coupon bond with maturity $T$ by $D(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s \mathrm{d}s\right)\middle|\,\mathcal{F}_t\right]$. All-in-all, we obtain the following bond pricing formula, which

takes into account both credit and liquidity risk:

$$\begin{aligned} V_t = c\sum_{j=1}^{2n} D(t,t+0.5j)\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{t+0.5j}\rho\lambda_s^l \mathrm{d}s}(1-\gamma)^{N^c_{t+0.5j}-N^c_t}\middle|\mathcal{F}_t\right] \\ + D(t,t+n)\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{t+n}\rho\lambda_s^l \mathrm{d}s}(1-\gamma)^{N^c_{t+n}-N^c_t}\middle|\mathcal{F}_t\right] \\ + \int_t^{t+n}(1-L)D(t,s)\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{s}\rho\lambda_u^l \mathrm{d}u}\lambda_s^c\gamma(1-\gamma)^{N^c_s-N^c_t}\middle|\mathcal{F}_t\right]\mathrm{d}s. \end{aligned} \tag{5}$$

The parameter $\rho$ in the liquidity discount factor is a pre-fixed auxiliary scaling parameter which we add for identification purposes. The expectations appearing in (5) can be calculated explicitly (up to a solution of the system of ODEs) by using the framework of Duffie et al. (2000). The proofs and computations are given in Internet Appendix IA.1.

From (5), we see that credit (liquidity) shocks affect bond prices in two ways, namely by a self-exciting (cross-exciting) jump in the credit intensity, reflecting the increased likelihood of going into default, and by a cross-exciting (self-exciting) jump in the liquidity intensity, which increases the illiquidity discount. Bond prices thus contain information on the dynamics of both credit and liquidity shocks, and, therefore, provide an appropriate basis for our empirical analysis.

## 3. Data

In our empirical analysis, we consider both US bond portfolios sorted by rating and a case study on Ford Motor Company, over the period from July 2007 until July 2009. We focus on Ford, since it is among the companies with the most bonds outstanding, and many of these bonds are traded relatively frequently. Furthermore, as we will show below, bid-ask spreads of Ford's bonds rose to very high levels, and many of these bonds were subjected to multiple rating changes during our sample period, making it an interesting case to study credit-liquidity interactions.

Our empirical study is based on bond transaction data, which we obtain from the TRACE (Trade Reporting and Compliance Engine) database. The TRACE system was launched on July 1, 2002, and contains the transaction reports of secondary over-the-counter corporate bond transactions done by members of the Financial Industry Regulatory Authority. Initially, not all transaction reports were disseminated. From October 2004 onwards the standard TRACE database covers practically all secondary market corporate bond transactions. We use the Enhanced TRACE data, which differs from the standard TRACE data in a few respects. First, the Enhanced TRACE data contains all transactions (including the previously non-disseminated transactions) back to July 1, 2002. Second, the Enhanced TRACE data does not cap the transaction volumes (in standard TRACE data, transactions are capped at \$5,000,000 for investment-grade bonds and \$1,000,000 for speculative-grade bonds). Third, the Enhanced TRACE data contains historical buy and sell information.[5] We filter out erroneous trades using the methods proposed by Dick-Nielsen (2014). Furthermore, we use a price sequence filter that filters out price outliers, transactions with prices that are more than 20 percent away from the median transaction price that day, and transactions with prices that are more than 20 percent away from the previous trading price.

In our sample, we focus on fixed coupon bonds, and we exclude bonds that are putable, convertible, denominated in other currencies than US dollars, or have a maturity below 1 or beyond 30 years. We obtain this bond specific information from the Mergent Fixed Income Securities Database (FISD). This database also contains bond credit rating histories and bond information relevant for pricing such as the coupon rates, coupon frequencies, principal values and day count conventions. We normalize all transaction data to have a principal value of \$1,000. Finally, we discard retail-sized transactions with volumes below \$10,000 for the Ford case study, and below \$100,000 for the rating portfolios.

[5]These data improvements, compared to the standard TRACE data, come at the cost of a delay in the availability of the data.

The full Ford sample consists of 72 bonds, traded on 506 days, with a total of 40,174 transactions. We use this full set of transactions to estimate daily bid-ask spreads (see Section 3.1). In estimating our model, however, we employ only the closing prices of every bond on every day it is traded, since the use of the full set of transactions would result in excessive computing time. Therefore, we have an unbalanced panel dataset consisting of 10,903 end-of-day transactions. The corresponding summary statistics are reported in Table 1. Figure 1 illustrates the price path of a typical bond over the period 2007 - 2010 and highlights several major credit events that occurred during that period.[6] This figure already suggests that credit events tend to cluster in time and that bond prices typically drop around the time of a credit event.

The rating portfolios are subdivided into AAA/AA, A, BBB, BB, and B and lower rated bonds, and are updated every quarter. A bond is included in the rating class portfolio corresponding to its rating prevailing at its first transaction in that quarter. In order to prevent the portfolios from containing many infrequently traded bonds, we include in every quarter only bonds of firms that have two or more bonds that trade on at least 20 different days, and subsequently only consider these relatively frequently traded bonds. Since not all bonds within the same rating class are equally credit risky, we correct the transaction prices for bond-specific fixed effects which measure the differences in baseline credit-worthiness. This procedure is explained in Internet Appendix IA.2. The summary statistics of the final rating portfolios are given in Table 2.

### *3.1. Bid-Ask Spreads and Measurement Equation*

In order to get a better insight in how the bond market liquidity of Ford and the rating portfolios evolves throughout the sample period, we construct estimates of bid-ask spreads

[6] We define a credit event as any major adverse information shock regarding creditworthiness of a company. Concrete examples are credit rating downgrades or negative press releases.

using the buy- and sell indicators provided by TRACE. First, we sort the data by bond and transaction time on every transaction date. Next, we construct pairs of consecutive trades on each date and for every bond $i$, and run the following linear regression (over all transaction pairs on a certain day):

$$P^i_{t_k} - P^i_{t_{k-1}} = \alpha_t + \beta_t \left( d^i_{t_k} - d^i_{t_{k-1}} \right) + \epsilon^i_{t_k}. \tag{6}$$

Here $P^i_{t_k}$ is bond $i$'s log-price at time $t_k$, $d^i_{t_k}$ is a dummy that is equal to $+1$ if the transaction was a buy order and $-1$ if the transaction was a sell order, and $\epsilon^i_{t_k}$ is an i.i.d. normally distributed error term. In this regression equation, the parameter $\beta_t$ represents the average relative bid-ask spread on that day.[7] Table 3 displays the summary statistics of the estimated relative bid-ask spreads of Ford and the rating portfolios. In both the case of Ford as well as for the rating portfolios, we observe strong increases of relative bid-ask spreads during the peak of the crisis. Furthermore, the overall liquidity of Ford bonds was particularly poor compared to other firms (of the same credit rating) during our sample period.

Since we explicitly model, and wish to estimate, liquidity jumps, it is insightful to look at the bid-ask spread data to see when potential liquidity jumps have occurred. In principle, days on which the bid-ask spread is much higher compared to the previous day, are days on which it is likely that a liquidity shock has occurred. However, since we do not have the true bid-ask spread data, but only the (noisy) estimated bid-ask spread data, we have to account for the estimation error before we can use this data to identify potential ("likely") liquidity shocks. More specifically, we assume that $\beta_t$ is the true, but unobserved, bid-ask

[7] If we use actual prices in (6), we obtain estimates of the bid-ask spread in dollars, but since prices drop dramatically during the sample period, relative bid-ask spreads are more informative.

spread and that it follows an $AR(1)$ process:

$$\beta_t = \phi\beta_{t-1} + \eta_t,$$

with $0 \leq \phi \leq 1$, and $\eta_t$ an i.i.d. error term with zero mean and finite variance. The bid-ask spread at time $t$ is, however, estimated with error $u_t$ according to

$$\hat{\beta}_t = \beta_t + u_t.$$

On a daily frequency, we can assume $\phi$ to be close to one, and, therefore, innovations in the bid-ask spread are captured by the error term $\eta_t$. Hence, by filtering the $\eta_t$ series, we can indicate potential liquidity jumps by identifying those dates on which (the filtered) $\eta_t$ is very large. Figure 2 displays the filtered $\eta_t$ series for the Ford bond portfolio, and the days that are identified as potential liquidity jumps by taking a level of 2.25 standard deviations above the mean as jump boundary. In total, we identify 11 potential liquidity shocks. The details of the filtering procedure are described in Appendix A. In a similar way, we identify 5, 4, 6, 8, and 7 potential liquidity shocks for the AAA/AA, A, BBB, BB, and B and lower rated portfolios, respectively. In our estimation procedure detailed in the next section, we will use the liquidity jumps as identified by the procedure above through additional measurement equations. We will assume that the liquidity jumps, as identified by the bid-ask spread data, are noisy indicators of actual liquidity jumps, and exploit this information in identifying liquidity jumps. Note that this procedure does not pre-fix liquidity jump times or imposes them to be necessarily in the set of dates identified by the bid-ask spread data.

## 4. Estimation Methodology

In this section, we develop our estimation approach for estimating the latent jump times, values of the intensity processes, and parameters based on a Markov Chain Monte Carlo (MCMC) method. Our estimation methodology builds upon Eraker (2004) and extends it to accommodate bond price data, which are irregularly sampled, to deal with fully latent state processes, and to explicitly take into account self- and mutual excitation in the latent jump intensities. As we will show below, the key to the latter challenge of dealing with the excitation is to first include both the latent jump times and the latent jump intensities into the state vector, inducing a joint Markovian structure, and next to properly discretize and orderly sample (in a specific way that accommodates their latency) the intensity processes defined in (2).

### *4.1. Joint Posterior Density*

Following Eraker (2004), we will first find an expression for the joint density of the observed data, model parameters, and the state variables. This joint density is important in deriving the MCMC sampler and therefore we will explicitly show how it is derived.

First we write the bond prices $P_t$ as a function $F$ of a parameter vector $\Theta$, the state variables $X_t = \{N_t^c, N_t^l, \lambda_t^c, \lambda_t^l\}$, and other arguments such as time to maturity and coupon frequency collected in a vector $\chi$. That is,

$$P_t = F(X_t, \chi, \Theta),$$

where $F$ is given by the bond price formula (5), and

$$\Theta = \{\alpha_c, \lambda_\infty^c, \sigma_c^2, \alpha_l, \lambda_\infty^l, \sigma_l^2, \beta_{1,1}, \beta_{1,2}, \beta_{2,1}, \beta_{2,2}, \gamma, \rho\}.$$

Since we use only the closing transaction prices, we have an unbalanced panel dataset with

daily bond price observations and we let $S(t)$ denote the set of bonds for which we have an observation at day $t$, $t = 1, \ldots, N$. We assume that the (log) bond prices $Y_{t,j}$, $j \in S(t)$ are observed with normally distributed pricing errors,[8] that is,

$$Y_{t,j} = \log\left(F(X_t, \chi_j, \Theta)\right) + \eta_{t,j}, \qquad \forall j \in S(t), t = 1, \ldots, N,$$

with $\eta_{t,j} \sim \mathcal{N}(0, h^2)$.

Apart from (log-) bond prices, we also use the liquidity jump indicator mentioned in Section 3. We let

$$I_t^l = \mathbb{1}(\Delta \mathrm{BA}_t > c), \tag{7}$$

where $\Delta \mathrm{BA}_t$ is the difference in the (relative) bid-ask spread between days $t-1$ and $t$, which we measure in the way described in Section 3 and Appendix A. As mentioned above, we take $c > 0$ a fixed threshold value, which we set at the mean change in the relative bid-ask spread plus 2.25 times the standard deviation of the change bid-ask spread. That is, on days where the change in the bid-ask spread is more than 2.25 standard deviations higher than the average change of the bid-ask spread over the sample period, the indicator function is 1, and it is 0 otherwise. The indicator function $I_t^l$ defined in (7) is then viewed as a noisy measure of the latent liquidity jump occurrence, and we specify the corresponding measurement equation as follows:

$$I_t^l = \begin{cases} Bernoulli(\phi_{high}) & \text{if } \Delta N_t^l = 1, \\ Bernoulli(\phi_{low}) & \text{if } \Delta N_t^l = 0. \end{cases} \tag{8}$$

The idea behind this measurement equation is that it helps us to identify liquidity jumps, which aids in overall parameter and state identification. Note that we do not impose or

[8]For the rating portfolios we assume that fixed effects corrected log-prices are observed with normally distributed pricing errors.

fix any jump times, nor do we restrict potential jumps to be on those dates for which the indicator is 1. We merely state that, on a day for which the liquidity jump indicator $I_t^l$ is 1, it is more likely to identify a liquidity jump ($\Delta N_t^l = 1$), since $\phi_{high} > \phi_{low}$.

Let $Y = \{Y_{t,j} : t = 1, \ldots, N, j \in S(t)\}$ denote the vector of all bond price observations, $I^l = \{I_t^l | t = 1, \ldots, N\}$ the vector of all liquidity jump indicators, $\bar{Y} = \{Y, I^l\}$ the vector with all measurement observations, $X = \{X_t : t = 1 \ldots, N\}$ the vector with all states, and $\bar{\Theta} = \{\Theta, \phi_{high}, \phi_{low}, h^2\}$ the vector with all parameters. Then the specification of the pricing errors and the liquidity jump indicator functions allow us to write the conditional density of the joint observations as

$$\begin{aligned} p(\bar{Y}|X, \bar{\Theta}) &\propto \prod_{t=1}^{N} p(I_t^l | X_t, \bar{\Theta}) \prod_{j \in S(t)} \phi(Y_{t,j}; F(X_t, \chi_j, \Theta), h^2) \\ &=: \prod_{t=1}^{N} p(I_t^l | X_t, \bar{\Theta}) p(Y_t | X_t, \Theta), \end{aligned}$$

where $\phi(x; m, s^2)$ denotes a normal density with mean $m$ and variance $s^2$ evaluated at $x$, and $p(I_t^l | X_t, \bar{\Theta})$ is a Bernoulli density with the right success probability parameter (depending on the value of $\Delta N_t^l$).

The full joint posterior density $p(\bar{Y}, X, \bar{\Theta})$ is then given by

$$p(\bar{Y}, X, \bar{\Theta}) \propto p(\bar{Y}|X, \bar{\Theta}) p(X|\bar{\Theta}) p(\bar{\Theta}), \tag{9}$$

where $p(\bar{\Theta})$ is the prior for $\bar{\Theta}$. We choose our priors to be proper but in such a way that they impose little information. More details can be found in Internet Appendix IA.3.

Using the Markovian property of the joint intensities and jump processes (see Aït-Sahalia

et al., 2014), we can rewrite this as a product over the observations times. We get

$$p(\bar{Y}, X, \bar{\Theta}) \propto p(\bar{\Theta}) \prod_{t=1}^{N} p(I_t^l | X_t, \bar{\Theta}) p(Y_t | X_t, \bar{\Theta}) p(X_t | X_{t-1}, \bar{\Theta}), \tag{10}$$

where $p(X_t | X_{t-1}, \bar{\Theta})$ is the transition density of the state processes.

Since the transition density of the state processes is unknown, we consider with slight abuse of notation the following discretized version of (2):

$$\begin{aligned}
\lambda_{t+1}^c - \lambda_t^c &= \alpha_c \lambda_\infty^c \Delta_{t+1} - \alpha_c \lambda_t^c \Delta_{t+1} + \sigma_c \sqrt{\lambda_t^c \Delta_{t+1}} \epsilon_{t+1}^c + \beta_{1,1} N_{t+1}^c + \beta_{1,2} N_{t+1}^l, \\
\lambda_{t+1}^l - \lambda_t^l &= \alpha_l \lambda_\infty^l \Delta_{t+1} - \alpha_l \lambda_t^l \Delta_{t+1} + \sigma_l \sqrt{\lambda_t^l \Delta_{t+1}} \epsilon_{t+1}^l + \beta_{2,1} N_{t+1}^c + \beta_{2,2} N_{t+1}^l,
\end{aligned}$$

where $\Delta_{t+1}$ is the time interval between $t$ and $t+1$ (i.e., a business day), $\epsilon_{t+1}^c$ and $\epsilon_{t+1}^l$ are independent standard normal random variables, and $N_{t+1}^k = 1$ $(k = c, l)$ indicates a jump arrival. The jump times $N_{t+1}^k$ are Bernoulli random variables with non-constant intensities $\lambda_t^k \Delta_{t+1}$ $(k = c, l)$. For this discretization to be appropriate, the observation intervals $\Delta_t$ should not be too far apart. Eraker et al. (2003) show in a simulation study that for a daily frequency the discretization bias is negligible. Using this discretization, we have that the transition density can be decomposed as

$$\begin{aligned}
p(X_t | X_{t-1}, \bar{\Theta}) &= p(\lambda_t^c, \lambda_t^l | N_t^c, N_t^l, X_{t-1}, \bar{\Theta}) p(N_t^c, N_t^l | X_{t-1}, \bar{\Theta}) \\
&= p(\lambda_t^c, \lambda_t^l | N_t^c, N_t^l, X_{t-1}, \bar{\Theta}) p(N_t^c | X_{t-1}, \bar{\Theta}) p(N_t^l | X_{t-1}, \bar{\Theta}),
\end{aligned} \tag{11}$$

where $p(\lambda_t^c, \lambda_t^l | N_t^c, N_t^l, X_{t-1}, \Theta)$ is bivariate Gaussian, and $p(N_t^k | X_{t-1}, \Theta)$ $(k = c, l)$ are Bernoulli. That is, the transition density is a mixture of normal densities. Note that the above discretization and the decomposition in (11) allow us to deal with the mutual and self-excitation, since we can sequentially draw $N_t^c$ and $N_t^l$ from the Bernoulli densities $p(N_t^c | X_{t-1}, \Theta)$ and $p(N_t^l | X_{t-1}, \Theta)$, respectively, and then draw $\lambda_t^c$ and $\lambda_t^l$ from

$p(\lambda_t^c, \lambda_t^l | N_t^c, N_t^l, X_{t-1}, \Theta)$ using the newly drawn $N_t^c$ and $N_t^l$ in the conditioning information.

### *4.2. MCMC Sampling*

Ultimately, we are interested in the joint conditional posterior density $p(\bar{\Theta}, X|\bar{Y})$. The MCMC algorithm, which we describe below, will draw from this conditional distribution, and with these draws we can estimate the marginal posterior density of the parameters $p(\bar{\Theta}|\bar{Y})$, since, by Bayes theorem, $p(\bar{\Theta}|\bar{Y}) \propto p(\bar{\Theta}, X|\bar{Y})$. Hence, we can take the sample average of $\bar{\Theta}^{(1)}, \bar{\Theta}^{(2)}, \ldots, \bar{\Theta}^{(G)}$, $\hat{\bar{\Theta}} = \frac{1}{G}\sum_{g=1}^{G} \bar{\Theta}^{(g)}$, as our estimate for $\bar{\Theta}$. In a similar way, we can filter the latent jump intensities by considering the sample averages $\lambda_t^k$ for all $t = 1, \ldots, N$ $(k = c, l)$. Furthermore, to filter the jump times, i.e., to decide whether a jump occurred at time $t$, $t = 1, \ldots, N$, we define a threshold, $\omega > 0$, and say that a jump occurred at time $t$ if $\frac{1}{G}\sum_{g=1}^{G} N_t^{k(g)} > \omega$ $(k = c, l)$ (see Johannes et al., 1999).

Since the joint conditional posterior density is high-dimensional and nonstandard, we cannot sample from this density directly. In order to overcome these problems, we employ a Gibbs sampler, allowing us to sequentially draw all random variables of the joint posterior density. Furthermore, by using appropriate Metropolis-Hastings algorithms, we are able to deal with the problem of non-standard densities. The general idea of our estimation method is outlined next, the details can be found in Internet Appendix IA.3.

### *4.3. MCMC Algorithm*

The Gibbs sampler consists of the following steps, initialized by an appropriate set of starting values for $X$ and $\bar{\Theta}$ when $g = 0$:

For $g = 1, \ldots, G, t = 1, \ldots, N$, simulate

1. $X_t^{(g+1)}$ from $p(X_t | X_{1:N_{\setminus t}}^{(g)}, \bar{\Theta}^{(g)}, \bar{Y})$, and
2. $\bar{\Theta}^{(g+1)}$ from $p(\bar{\Theta}|X^{(g+1)}, \bar{Y})$,

where $X_{1:N_{\setminus t}}$ denotes the collection of state vectors $X_s$ at all $s = 1, \ldots, N$ except at $s = t$. Note that we sample both the jump intensities and jump times at every transaction date before considering the next transaction date. Eraker (2004) first samples all jump times for all transaction dates and then, given these newly drawn jump times, samples the volatility process for all transaction times. This difference in the order of sampling is in principle asymptotically irrelevant, as the Markov chain converges regardless of the order of drawing in each iteration. However, as we will explain below, the order matters in the case of latent states, since one might get stuck in a situation where jumps are never sampled. By sampling all states day-by-day, we can alternate between two different sampling schemes, allowing us to circumvent this problem.

*4.3.1. Drawing Intensities and Jump Times*

Let us first consider simulating $X_t^{(g+1)}$ from $p(X_t|X_{1:N_{\setminus t}}^{(g)}, \bar{\Theta}^{(g)}, \bar{Y})$. We take advantage of the fact that $p(X_t|X_{1:N_{\setminus t}}^{(g)}, \bar{\Theta}^{(g)}, \bar{Y})$ is again characterized by its full conditionals. Therefore, we can split drawing $X_t^{(g+1)}$ into the following three steps:

1. draw $N_t^{c(g+1)}$ from $p(N_t^c|\lambda_t^{c(g)}, \lambda_t^{l)(g)}, N_t^{l(g)}, X_{1:N_{\setminus t}}^{(g)}, \bar{\Theta}^{(g)}, \bar{Y})$;
2. draw $N_t^{l(g+1)}$ from $p(N_t^l|\lambda_t^{c(g)}, \lambda_t^{l)(g)}, N_t^{c(g+1)}, X_{1:N_{\setminus t}}^{(g)}, \bar{\Theta}^{(g)}, \bar{Y})$;
3. draw $\lambda_t^{c(g+1)}$ and $\lambda_t^{l(g+1)}$ from $p(\lambda_t^c, \lambda_t^l|N_t^{c(g+1)}, N_t^{l(g+1)}, X_{1:N_{\setminus t}}^{(g)}, \bar{\Theta}^{(g)}, \bar{Y})$.

In Internet Appendix IA.3 we show that the probability of drawing jumps does not depend on the observed bond prices, but only on the (previously drawn) paths of the intensities and the liquidity jump time indicator. As a consequence, we may never sample jumps even when the bond data would clearly indicate large price drops, since we decoupled the drawing of jumps from the data likelihood. This problem is directly linked to the fact that our intensity processes are latent, unlike in Eraker (2004), where one of the state processes (stock prices) is observable.

In order to overcome this problem, we alternate between the above simulation scheme and one where we draw the whole vector $X_t$ jointly. Under the latter scheme, we do not split drawing $X_t$ into three substeps as above, but draw the whole vector in one go. Since the $\lambda^k$ $(k = c, l)$ influence bond prices, the acceptance of a new draw for $X_t$ then depends on both the transition densities as well as the bond price data and this increases the likelihood of drawing an initial jump. A disadvantage of this simulation scheme, however, is that the liquidity jump indicator measurement equations do not play a direct role in the proposal of new (liquidity) jumps, which effectively diminishes the information contained in these measurement equations. Because of this, we alternate between this simulation scheme and the previously explained simulation scheme.

*4.3.2. Drawing Parameters*

Next we consider drawing $\bar{\Theta}^{(g+1)}$ from $p(\bar{\Theta}|X^{(g+1)}, \bar{Y})$. Where possible, we draw parameters in a Gibbs step using conjugate priors. This can be done for $h^2$, $\phi_{high}$ and $\phi_{low}$, for which we specify an inverse Gamma prior, and Beta priors, respectively. Since the other parameters appear in both the transition densities as well as in the bond price formulas, their conditional densities are not known and we need to rely on Metropolis steps. From a computational perspective, it is better to update these parameters in blocks. Indeed, every time one of these parameters gets updated, the ODEs of the bond price formulas have to be recomputed. On the other hand, we do not want the blocks to be too large, since drawing from a high-dimensional density can result in low acceptance probabilities. We therefore draw the parameters in the following blocks: $\{\alpha_c, \lambda^c_\infty, \sigma^2_c, \gamma\}$, $\{\alpha_l, \lambda^l_\infty, \sigma^2_l, \rho\}$, and $\{\beta_{1,1}, \beta_{1,2}, \beta_{2,1}, \beta_{2,2}\}$. Similar to Eraker (2004), we use for each block a random-walk Metropolis step with a Gaussian proposal density with as mean vector the previous draw and with a diagonal covariance matrix. Details can again be found in Internet Appendix IA.3.

In order to test the performance of our estimation methodology, we conduct a Monte Carlo

study, first for a restricted version of the model in which the liquidity intensity is absent and next for the full model specification. The results, detailed in Appendix B, show that our estimation methodology performs well in a realistic setting and that most parameters are estimated with high precision.

## 5. Estimation Results

### *5.1. Parameter assumptions and practical considerations*

Although they are theoretically identified, it can be difficult in practice to accurately estimate all parameters in our multivariate model. We therefore pre-identify some of the parameters and treat them as fixed and given in our analysis. First, we set $\gamma = 0.1$. This is the value we obtain from estimating a restricted model specification, which ignores liquidity discounting and only contains the credit risk component, on the Ford sample.[9] Since the long-run average instantaneous default probability is given by the product $\lambda^c_\infty \times \gamma$, fixing $\gamma$ greatly aids the identification of $\lambda^c_\infty$. To facilitate comparability, we also take $\gamma = 0.1$ for the rating portfolios. Second, we fix $\lambda^l_\infty$ and choose it such that the model is able to fit the average prices in the relatively tranquil period in the first part of the sample. In this period, there are no large (liquidity) discounts, and a relatively low value of $\lambda^l_\infty$ captures this observation. We take the same value of $\lambda^l_\infty$ for all rating portfolios, since the bid-ask spreads of the rating portfolios are approximately equal during this period. The liquidity of Ford throughout our sample period was, however, particularly poor compared to other firms, and therefore we fix $\lambda^l_\infty$ to a higher value for Ford than for the rating portfolios. We also fix the auxiliary scaling parameter $\rho$, and set $\rho = 0.25$. This value is chosen such that the resulting liquidity jump parameters $\beta_{2,1}$ and $\beta_{2,2}$ are of a similar magnitude as the credit

[9] More specifically, under the restricted model specification we set all the parameters governing the liquidity intensity process to zero.

jump parameters $\beta_{1,1}$ and $\beta_{1,2}$. Finally, we set $\sigma_l^2 = 0$. This effectively implies that there is no Brownian component in the liquidity intensity and thus ignores small fluctuations in liquidity. The improved accuracy in overall parameter and jump identification is, however, substantial, and, because we are mainly interested in the interaction between large credit and liquidity shocks through mutually exciting jumps, we think this restriction is justifiable.

### *5.2. Ford*

Table 4 reports the posterior means, posterior standard deviations, and 99 percent highest posterior density (HPD) intervals[10] of the parameter estimates for Ford.

Table 4 shows that most parameters are estimated with high precision as indicated by the relatively low posterior standard deviations compared to the (positive-valued) posterior means. Furthermore, the 99 percent HPD intervals are far away from zero for all parameters, except $\beta_{1,2}$ and $\phi_{low}$, indicating that nearly all parameters are positive.

Central to our analysis are the jump size parameters $\beta_{ij}, i, j \in \{1, 2\}$. We find relatively large values for $\beta_{1,1}$, $\beta_{2,1}$ and $\beta_{2,2}$, which support the existence of self-excitation in both liquidity and credit risk, and cross-excitation from credit risk to liquidity. The effect of liquidity shocks on credit risk is small ($\beta_{1,2} = 0.0898$). In principle, this is in line with what can be expected from the structural credit risk literature, in which the main channel through which bond market liquidity affects credit risk is a debt rollover mechanism (see, e.g., He and Xiong, 2012, He and Milbradt, 2014, Nagler, 2017, and Chen et al., 2017). This channel was unlikely to be in play for Ford. Indeed, within our sample period, no bonds of Ford matured, and there were only two new bond issuances (one on April 28, 2008, and one on May 28, 2009).

[10]The 99 percent HPD interval is the smallest possible interval that contains 99% of the mass of the posterior distribution. It is analogous (but not equal) to confidence intervals in frequentist statistics.

To address more formally the issue of significance of the jump size parameters, we compare our model with nested model specifications using Bayes factors.[11] In particular, we compare our full model with (1) a model without excitation (i.e., $\beta_{1,1} = \beta_{1,2} = \beta_{2,1} = \beta_{2,2} = 0$), (2) a model without self-excitation (i.e., $\beta_{1,1} = \beta_{2,2} = 0$), and (3) a model without cross-excitation (i.e., $\beta_{1,2} = \beta_{2,1} = 0$). In all cases the Bayes factors are essentially zero, and, hence, favor the full model specification over the nested model specifications.

The upper panels of Figure 3 plot the evolution of the bond prices in absolute (upper left panel) and relative (upper right panel) terms after either a liquidity or a credit shock.[12] These figures reveal that the initial price effect of a credit shock is larger than the initial effect of a liquidity shock. However, the price effect of the credit shock decays faster. These differences enable us to disentangle credit and liquidity jumps and to econometrically identify shocks.

An important role in credit and liquidity jump identification is played by the parameters $\alpha_c$ and $\alpha_l$ dictating the speed of mean reversion. We find that these parameters differ substantially ($\alpha_c = 2.1713$ and $\alpha_l = 4.1628$), which implies that the decay of liquidity shocks in the liquidity intensity is much faster than the decay of shocks in the credit intensity (see the lower panels of Figure 3), facilitating the identification of the type of shock.

Apart from the identification through the decay patterns, we are also able to distinguish jumps in the credit intensity process from typical Brownian innovations. Consider the average daily volatility of the continuous part, which is given by $\sqrt{\sigma_c^2 \times \lambda_\infty^c \times \frac{1}{252}} \approx 0.05$. This average daily volatility is much smaller than the credit self-excitation parameter ($\beta_{1,1} = 1.0136$), and, therefore, it is relatively easy to separate a credit jump from a typical fluctu-

[11] The details of how the Bayes factors are computed are given in Internet Appendix IA.4.

[12] We take the long-term average model-implied price of a 10-year bond and consider the price effect of a single credit or liquidity jump. That is, we compute the 10-year bond price using the estimated parameters and putting the credit and liquidity intensities equal to their long-term averages $\lambda_\infty^c = 0.35$ and $\lambda_\infty^l = 0.015$, respectively, and then impose that either one credit or one liquidity shock occurs.

ation caused by the Brownian component.

The liquidity jump indicator parameters $\phi_{high}$ and $\phi_{low}$ indicate that there is a reasonable degree of correspondence between the liquidity shocks identified by the estimation algorithm and by the bid-ask measurement equation. $\phi_{low} \approx 0.015$, for example, implies that on days where our estimation algorithm did not detect a liquidity jump, the probability that the bid-ask data would have detected a liquidity jump is only approximately 1.5 percent. Furthermore, $\phi_{high} \approx 0.5$ implies that on days where our algorithm detects a liquidity jump, the bid-ask spread data also detects a liquidity jump with probability 0.5. This shows that the measurement equation provides valuable information on the occurrences of liquidity jumps, but that the algorithm still has some freedom to identify liquidity jumps on dates different from the measurement equation. The long-term mean of the credit intensity process, $\lambda^c_\infty$, is 0.35, hence the average 1-year (risk-neutral) default probability is $\lambda^c_\infty \times \gamma = 0.035$.

The estimated credit and liquidity intensity paths are displayed in the upper panels of Figures 4 and 5, respectively. Both the credit and liquidity intensities increase during the peak of the crisis, reflecting an increase in credit risk and a worsening market liquidity. To verify whether the estimated credit and liquidity intensity paths indeed represent adequate measures of credit and liquidity risk, we compare them to other measures of credit and liquidity risk. In the upper panel of Figure 6, we plot both the (normalized) estimated credit intensities and the (normalized) 10-year CDS spreads of Ford.[13] The correlation between the two is 0.79. The overall pattern suggests that the estimated credit intensities are indeed meaningful representations of credit risk. In a similar fashion, the upper left panel of Figure 7 compares the (normalized) estimated liquidity intensities and the (normalized) estimated relative bid-ask spreads of Ford. Again, the high correlation between the two suggests that our liquidity factor is indeed a representative measure of liquidity risk.

[13]This maturity is chosen because it is closest to the average maturity of bonds used in our sample (see Table 1).

The lower panels of Figures 4 and 5 display the estimated credit and liquidity jump probabilities, respectively. Specifically, we compute, for every day in the sample, the fraction of MCMC iterations in which a credit or liquidity jump was drawn. We observe that the estimated credit jump probabilities exhibit a few broad peaks at which the jump probability is high for a number of days in a row, reflecting the clustering of credit shocks. The estimated liquidity jump probability series exhibits somewhat sharper and more delineated peaks, which is not surprising because we have an explicit measurement equation that helps to identify liquidity jumps.

Since the MCMC algorithm sometimes draws jumps randomly, we define actual jumps to be those days on which the estimated jump probability is larger than 0.4. In total we identify 16 credit jumps and 6 liquidity jumps, which are displayed in Figure 8. We find the estimated credit shocks to be consistent with the price path and credit events indicated in Figure 1. For example, we find a clear cluster of credit shocks around the default of Lehman Brothers. Furthermore, we find a high, but not one-to-one, correspondence between the liquidity jumps we estimated and the liquidity jumps indicated by the bid-ask spread data. Concerning the model fit, we find that, even though our model is parsimonious and the data spans a very turbulent time period, bond prices are fitted quite well. The average relative pricing errors in our Ford sample range between 4.10% and 11.15% with an average of 6.91% across all bonds in the sample. Compared to the univariate model we analyzed as part of our Monte Carlo study, this relative price error is a great improvement, since that model produced an average relative pricing error of around 19%. This shows that taking into account liquidity and the interactions between credit and liquidity greatly improves the model's capability of fitting the observed price paths in this turbulent period. Figure 9 displays the true and model-implied price paths of a representative bond in our sample. It shows that a large part of our pricing error comes from the smoothing nature of our estimators (i.e., the posterior means condition on all observations), which induces somewhat more smoothed model-implied price

paths compared to the more noisy true price path.

### *5.3. Rating Portfolios*

Table 5 presents the estimation results for the rating portfolios. As expected, we find that the long-run average level of the credit intensity, $\lambda_\infty^c$, typically increases as the credit rating worsens. Furthermore, the pattern in the $\sigma_c^2$ parameter estimates indicates that there is higher credit risk volatility for bonds with a lower credit rating. We observe that both speed-of-mean-reversion parameters $\alpha_c$ and $\alpha_l$ are each estimated at similar values across all rating classes.

Concerning the jump size parameters, we find a clear monotonically increasing pattern in the cross-excitation parameters $\beta_{1,2}$ and $\beta_{2,1}$ as the rating class gets worse. Hence, credit-liquidity interactions are more pronounced for bonds with a lower credit rating. The liquidity self-excitation parameter $\beta_{2,2}$ also shows an increasing pattern with the exception of the AAA-rated bonds. There is, however, no clear pattern across rating classes in the credit self-excitation parameter $\beta_{1,1}$. As indicated by the posterior standard deviations and HPD intervals, the jump parameters are typically estimated more precisely for the lower rating classes.

Finally, the estimated parameters $\phi_{high}$ and $\phi_{low}$ are all very similar across the different rating portfolios and also compare well to the estimates in the Ford case study. This indicates that the potential liquidity jumps indicated by our bid-ask spread estimates for the rating portfolios are roughly as informative about liquidity jumps as in the Ford case study.

In Figures 6 and 7 we also compare the estimated credit and liquidity intensities of the different rating classes with the corresponding credit default swap indices and estimated relative bid-ask spreads, respectively.[14] Again, the results indicate that the estimated credit

[14]In comparing the credit intensities, we only focus on the BB and B and lower rating classes, since these are the only rating classes for which comparable credit default swap indices are available.

and liquidity intensities seem to represent credit and liquidity risk well.

In estimating the model for the rating portfolios, we are not looking at firm-specific credit and liquidity risk, but at average credit and liquidity risk for bonds in a certain rating class. Of course, not all firms are hit by shocks at the same time, and therefore the effects of these shocks are dampened and jump identification becomes more tenuous. This is why we take slightly different boundaries for estimating credit and liquidity jumps than in the Ford case study, and also allow for some heterogeneity between credit and liquidity shocks. More specifically, we set the credit and liquidity jump boundaries to 0.3 and 0.1[15], respectively.

Comparing the estimation results of Ford with the B and lower rating class reveals that many parameters take similar values. The estimated mean-reversion parameters of the credit and liquidity intensities, $\alpha_c$ and $\alpha_l$, the long-run average level of the credit intensity, $\lambda^c_\infty$, the credit intensity volatility parameter, $\sigma^2_c$, and the excitation parameters $\beta_{1,1}$, $\beta_{2,1}$, and $\beta_{2,2}$ all compare very well. There are, however, some differences in the long-term average liquidity intensity parameters $\lambda^l_\infty$ (by assumption), and the excitation parameter $\beta_{1,2}$, which represents the effect of liquidity shocks on credit risk. The latter might be explained by the fact that Ford did not roll over any debt during our sample period. Hence, the main feedback channel from liquidity to credit risk proposed in the literature was not in play for Ford. This finding could thus serve as indirect support for the importance of the debt rollover channel in explaining possible feedback from liquidity to credit risk, since we do find this effect to be strong for the rating portfolios.

[15]For comparability reasons we take the same jump boundaries across rating classes. These boundaries are chosen such that we identify at least one credit and liquidity jump for all rating classes. In total, we identify 2, 2, 4, 1, and 7 liquidity jumps for the AAA/AA, A, BBB, BB, and B and lower rated portfolios, respectively. Comparing this to the number of potential liquidity shocks identified by the bid-ask spread estimates (we identified 5, 4, 6, 8, and 7 potential liquidity shocks for the AAA/AA, A, BBB, BB, and B and lower rated portfolios, respectively), we again conclude that the liquidity jump measurement equation is informative but not enforcing. Apart from the BB rating portfolio, we also observe a clear increasing pattern in the number of estimated credit and liquidity jumps as credit rating gets lower. Since we use the same jump boundaries across the different rating classes, this clearly indicates that bonds of lower credit ratings are typically hit by more large shocks than bonds of higher credit ratings.

Upon comparing the relative pricing errors of our rating portfolios to the most closely related structural credit risk models we find that our model is much better able to capture bond price data. For example, Chen et al. (2017) mention that their model produces relative pricing errors of over 100% for Baa and Ba-rated bonds during the 2008-2009 crisis, whereas we obtain relative pricing errors of only 5.10% and 4.88% for BBB- and BB-rated bonds, respectively.

## 6. Model Implications

In this section, we illustrate some implications of our model and estimation results in more depth. More specifically, we show how our model gives rise to a natural decomposition of yield spreads into a pure credit, a pure liquidity, a credit-induced liquidity, and a liquidity-induced credit component, and look at risk management implications of incorporating credit-liquidity interactions in Value-at-Risk computations.

### *6.1. Credit-Liquidity Decomposition*

The empirical literature typically decomposes yield spreads into a credit and a liquidity part in an additive way. As argued by He and Milbradt (2014) such a decomposition overlooks the important role played by credit and liquidity interactions. In calibrating their model-implied yield spread decomposition, He and Milbradt (2014) and Chen et al. (2017) find evidence that credit-liquidity interaction terms contribute for a non-trivial part to yield spreads. A formal econometric investigation of the contribution of interaction terms to yield spreads is, however, outside of the scope of their studies and is the topic of this section.

Similar to He and Milbradt (2014), we propose a yield spread decomposition that nests the common additive default-liquidity decomposition by further subdividing the credit and liquidity components into pure credit, liquidity-induced credit, pure liquidity, and credit-

induced liquidity parts. Throughout, we consider yield spreads with respect to the corresponding Treasury yields, which we denote by $ys$. The decomposition scheme is given by

$$ys = \underbrace{ys_{\text{pureCRED}} + ys_{\text{LIQ}\,\to\,\text{CRED}}}_{\text{Credit Component } ys_{\text{CRED}}} + \underbrace{ys_{\text{pureLIQ}} + ys_{\text{CRED}\,\to\,\text{LIQ}}}_{\text{Liquidity Component } ys_{\text{LIQ}}}. \tag{12}$$

We first decompose yield spreads into a credit and liquidity component. The credit component $ys_{\text{CRED}}$ is obtained by computing yield spreads that would occur if there was no liquidity discounting. We can compute these yield spreads by setting the parameters governing the liquidity intensity process equal to zero, and by using the estimated credit intensities and credit intensity process parameters of the full model. In this way, the resulting spreads only price in default risk and there is no liquidity discount. The liquidity component is then given by the difference between the full yield spread and the credit component, i.e., $ys_{\text{LIQ}} = ys - ys_{\text{CRED}}$.

Note that, even though there is no formal liquidity discounting in computing the $ys_{\text{CRED}}$ component, the estimated credit intensities still contain some liquidity influences through the cross-excitation parameter $\beta_{1,2}$. Similarly, in computing $ys_{\text{LIQ}}$, there is no formal default component, but the liquidity intensities still contain the spillover effects from credit risk through the cross-excitation parameter $\beta_{2,1}$. Precisely these features facilitate a further decomposition of the credit component into pure credit and liquidity-induced credit subcomponents and the liquidity component into pure liquidity and credit-induced liquidity subcomponents.

To see this, note that we can decompose the credit intensities into a component attributable to pure credit risk and a component attributable to liquidity risk, i.e., $\lambda^c_t = \lambda^c_{\text{pureCRED},t} + \lambda^c_{\text{LIQ}\,\to\,\text{CRED},t}$. We can compute $\lambda^c_{\text{LIQ}\,\to\,\text{CRED},t}$, because we know that every time there is a liquidity jump, the credit intensity jumps up by $\beta_{1,2}$. Since we have explicitly estimated the liquidity jump times (see, e.g., Figure 8 in the case of Ford), we can compute at any point

in time the total accumulating effect of liquidity jumps on the credit intensity, taking into account that these jump effects decay over time because of the mean-reverting behavior of the credit intensity process. The first panel of Figure 10 illustrates the decomposition of $\lambda_t^c$ into the pure credit part $\lambda^c_{\text{pureCRED},t}$ and the liquidity-induced credit part $\lambda^c{}_{\text{LIQ}\,\rightarrow\,\text{CRED},t}$ in the case of Ford. We observe that, since $\beta_{1,2}$ is relatively small and there are only a limited number of liquidity jumps (6 in total), the liquidity-induced credit part of the credit intensities is small compared to the pure credit part.

Having decomposed the credit intensities into pure credit and liquidity-induced credit intensity components, we define the pure credit yield spread component to be the yield spreads induced by the restricted model without liquidity intensities and liquidity intensity parameters, and that only take into account the pure credit intensity part $\lambda^c_{\text{pureCRED},t}$. The liquidity-driven credit component of the yield spread is defined as the difference between the total credit component and the pure credit component, i.e., $ys_{\text{LIQ}\,\rightarrow\,\text{CRED}} = ys_{\text{CRED}} - ys_{\text{pureCRED}}$, and captures the increase in credit risk due to market illiquidity.

In a similar way, we can find the pure and credit-induced liquidity components of the yield spread by first decomposing the liquidity intensities into a pure liquidity part and a credit-induced liquidity part. The second panel of Figure 10 shows the decomposition of $\lambda_t^l$ into the pure liquidity part $\lambda^l_{\text{pureLIQ},t}$ and the credit-induced liquidity part $\lambda^l{}_{\text{CRED}\,\rightarrow\,\text{LIQ},t}$ in the case of Ford. In order to compute the credit-induced liquidity component of the yield spreads, we take the yield spread of a bond that is not exposed to credit risk, but only to credit-induced liquidity risk. That is, we do not take into account the credit intensities and restrict all credit intensity parameters to zero and use the credit-induced liquidity intensity values in computing the liquidity discount. The pure liquidity yield spread is given by the difference between the total liquidity component and the credit-induced liquidity component, $ys_{\text{pureLIQ}} = ys_{\text{LIQ}} - ys_{\text{CRED}\,\rightarrow\,\text{LIQ}}$.

*6.1.1. Decomposition over the cross-section and over time*

We first explore the yield spread decomposition across different rating classes and consider how the contribution of the different components to the total yield spread evolved over time. This sheds light on what drove yield spread changes during the financial crisis. We consider hypothetical 10-year zero coupon bonds, and compute, for every day in our sample, the yield spread decomposition by applying the decomposition scheme described above.

Table 6 summarizes the cross-sectional decomposition results and displays the average contribution, both in absolute and relative terms, of the different components throughout the whole sample period and during the period September 15, 2008 until March 31, 2009 covering the peak of the crisis. We find that, in absolute terms, all components are increasing as the credit rating gets worse except for the pure credit components of the A- and BBB-rated bonds and the liquidity-induced credit component of the BB-bonds. This reflects that for bonds with lower credit ratings, not only credit and liquidity are typically worse, but that also the interactions between the two are more pronounced. Compared to the average decomposition during the full sample period, we find that the yield spread contribution of all components increase (in absolute terms) during the peak of the crisis. This effect is again particularly pronounced for the lower credit rating classes.

In line with the empirical literature (see, e.g., Bao et al., 2011, Dick-Nielsen et al., 2012 and Friewald et al., 2012), Table 6 also shows that the increase in yield spreads during the peak of the crisis was mainly the result of deteriorating liquidity. Our decomposition reveals, however, that a substantial part of this liquidity effect is actually credit-driven. This credit-induced liquidity component is particularly large for the BB and B and lower rating classes, and becomes even more important during the peak of the crisis. For example, for B and lower rated bonds, the relative (absolute) yield spread contribution of this component is on average 16.6% (1.12 percentage points) during the whole sample period, but 26.3% (2.32

percentage points) during the peak of the crisis. The liquidity-induced credit component also becomes more important in both absolute and relative terms during the peak of the crisis, but exhibits a less clear pattern across rating classes.

The decomposition results in Table 6 also show that Ford is not a typical representative firm of the B and lower rating class. The total yield spreads are much larger for Ford than for the average B and lower rated bonds. In particular, the pure credit and pure liquidity components are much larger (in absolute terms), indicating that Ford was more credit risky than a typical B and lower firm, and that the liquidity of its bonds was worse than that of the average B and lower firm. The latter finding is consistent with our bid-ask spread estimates (see Table 3), whereas the former finding is consistent with the fact that Ford bonds were typically rated CCC or lower. Another major difference between Ford and the B and lower rated bonds is that the liquidity-induced credit risk component is much lower for Ford. This is the result of the low estimated value of the cross-excitation parameter $\beta_{1,2}$, reflecting that liquidity shocks hardly affect credit risk in the case of Ford.

Apart from the average decompositions over a period of time, our model set-up also allows us to consider the actual evolution of the decomposition over time. To illustrate this, we plot this evolution, both in absolute and relative terms, for 3- and 10-year bonds of Ford in Figures 11 and 12, respectively. We see that the yield spreads of the 3-year maturity bond are much higher, and increased by more, during the peak of the crisis than those of the 10-year maturity bond. For both maturities, we find that, at any point in time, the liquidity-driven credit component is negligible. Most interestingly, these figures show that the credit-induced liquidity component became mainly significant during the peak of the crisis, and at some days even accounted for up to 40% of the total yield spreads. This illustrates that the interaction terms may actually be much more important in turbulent times than the average decomposition results would suggest.

As mentioned above, the general idea behind our decomposition is based on He and Milbradt (2014) and Chen et al. (2017) and it is interesting to compare our decomposition results with theirs.[16] Some caution is, however, required in a direct comparison, because, among other aspects, Chen et al. (2017) seem to have difficulties capturing the bond price data well during crisis periods, which may significantly affect their decomposition results. For example, they mention that they obtain relative pricing errors of over 100% for Baa- and Ba-rated bonds in the 2008-2009 crisis period, whereas the pricing errors obtained by our model are much lower (see Table 5). Nevertheless, we obtain results broadly similar to Chen et al. (2017) on the (relative) size of the liquidity-induced credit component. We typically find slightly lower (relative) contributions of the credit-induced liquidity component for all rating classes. This is likely to be a consequence of the fact that they keep the pure liquidity component fixed, whereas we allow it to fluctuate over time and across ratings. As can be seen from Table 6, we indeed find large differences in the pure liquidity component across ratings and over time. The fact that Chen et al. (2017) fix the pure liquidity component to the same value across ratings also drives their finding of an increasing relative contribution of the pure credit risk component and a decreasing relative contribution of the pure liquidity component as the rating worsens. By contrast, we find the exact opposite results.

### *6.1.2. Decomposition across maturities*

Next to the evolution of the decompositions over time for given maturities, we also investigate the term structures of the different components that make up the decompositions. We show in particular that the term structures of the relative decompositions can differ substantially from tranquil to turbulent periods. Focusing again on Ford, Figures 13 and 14 plot the term structures of the absolute and relative decompositions on different dates in the sample.

[16] We compare our results with Chen et al. (2017), since they extend the analysis of He and Milbradt (2014) and claim to have more reliable results. We focus on their results in "bad states" of the economy, since they label our complete sample period as a period in which the economy is in a bad state.

More specifically, we display the term structures of the different components on July 2, 2007, representing the start of our sample period, on September 15, 2008, which is the default date of Lehman Brothers, and on November 17, 2008, two months after the default of Lehman Brothers.

Figure 13 shows that at the start of our sample period the term structure of yield spreads was relatively flat. At the default date of Lehman Brothers, however, all components, except the liquidity-induced credit component, have clear downwards sloping term structures, and the yield spread components are much larger. The negative slope of the term structures is caused by the mean-reverting behavior of the intensity processes, which dampens the effects of shocks in the long run. Therefore, the yield impact (in absolute terms) of shocks on long maturity bonds is smaller than for short maturity bonds. The downwards sloping pattern becomes even more pronounced two months after the default of Lehman Brothers. We thus clearly see that credit and liquidity conditions worsen as the crisis develops, and that especially the credit-induced liquidity component becomes more prevalent. The liquidity-induced credit component is negligible for Ford.

Figure 14 plots the term structures of the components in relative terms. We observe that in relative terms, both the pure credit and pure liquidity components become less important as the crisis develops. At the same time there is a large increase in the relative importance of the credit-induced liquidity component, which, furthermore, has a clear upwards sloping term structure. Credit-induced liquidity thus affects yield spreads of bonds with a longer maturity more than bonds with a shorter maturity, and makes up almost 60% of the 20-year yield spread two months after the default of Lehman Brothers. This thus illustrates that, in turbulent times, the relative decomposition is substantially different for different maturities and deviates from the long-run average decomposition reported in Table 6.

*6.2. Implications for Value-at-Risk*

Our results also have important implications for standard risk management procedures such as Value-at-Risk (VaR) computations. The self- and mutual excitation mechanisms induce a return distribution that has a fatter tail in the loss domain compared to distributions implied by typical diffusion models or jump-diffusion models with i.i.d. shocks. This is the case, since shocks have an adverse price impact and since occurrences of credit or liquidity shocks make it more likely that additional shocks will follow. This excitation mechanism will especially matter in distressed periods, since then overall probabilities of shock occurrences are high.

To illustrate the effects of the self- and mutual excitation components on VaR capital requirements, we compute the 10-day 99% VaR for a hypothetical 10-year semi-annual coupon bond with coupon rate of 6.64% (i.e., the Ford sample average) under different model specifications. More specifically, we consider the full model specification, the model without cross-excitation ($\beta_{1,2} = \beta_{2,1} = 0$), and the model without both self- and cross-excitation (i.e., $\beta_{1,1} = \beta_{1,2} = \beta_{2,1} = \beta_{2,2} = 0$). For every model specification, we simulate 100,000 10-day sample paths of the intensity and jump processes using the estimated values of the model parameters and taking the intensity values on September 15, 2008, as starting points, and compute the implied 10-day return distribution of the bond. To compute the 99% VaR, we then take the first percentile of the simulated return distribution.

Figure 15 displays the resulting return distributions and 99% VaR levels for BBB-rated bonds. We clearly see that the full model has a much fatter tail in the loss domain than the restricted models. The model without mutual excitation in turn has a fatter tail than the model without any excitation. This thus shows that both self- and cross-excitation significantly affect the tails of the return distribution, and, hence, the VaR computations.

Table 7 reports the resulting VaR capital requirements under the different model specifica-

tions for both Ford and the rating portfolios. That is, we compute the amount of required capital in cents per dollar invested in the respective bonds. For example, the VaR of BBB-rated bonds implied by the full model is $-8.29\%$. The VaR of the model without mutual excitation is $-4.56\%$. Therefore, if one ignores the interactions between credit and liquidity risk, the resulting VaR capital requirement is $8.29 - 4.56 = 3.73$ cents per dollar too low. In a similar fashion, the capital requirements resulting from the model without excitation is 0.70 cents per dollar invested. This implies that ignoring the self-exciting components results in an additional underestimation of capital of $4.56 - 0.70 = 3.86$ cents per dollar invested.

As can be seen from Table 7, the cross-excitation channel affects the VaR capital requirements for lower rated bonds the most, as the differences between the full model VaR and the VaR resulting from the model without cross-excitation become larger when the rating deteriorates. For example, in the B and lower case, omitting to take into account the credit-liquidity interactions amounts in an underestimation of capital requirements of 15.27 cents per dollar invested. In relative terms, this implies that 76.5% of the total VaR capital requirements can be attributed to the cross-excitation channel. The self-excitation channel, on the other hand, is mainly important for higher credit ratings. For example, the VaR capital requirements attributable to self-excitation for AAA/AA-rated bonds amount to 6.82 cents per dollar invested, making up 88% of the total capital requirements.

## 7. Conclusions

In this paper, we study dynamic interactions between credit and liquidity risk in the corporate bond market and propose a novel way of modeling these interactions through self- and mutually exciting processes. This interaction mechanism gives rise to dynamic feedback loops between credit and liquidity risk, and captures the empirical observation that liquidity

dries up when credit conditions deteriorate. We construct bond price formulas and develop a corresponding Bayesian estimation procedure to estimate the model parameters using bond transaction data.

In our empirical analysis, we consider a case study on Ford Motor Company, and also analyze US bond portfolios sorted by rating. We find strong evidence of asymmetric feedback between credit and liquidity risk and show that these effects typically tend to be stronger for bonds with a lower credit rating. Our model enables a decomposition of bond yield spreads into a pure credit, a pure liquidity, a credit-induced liquidity, and a liquidity-induced credit component and allows us to explicitly study the price effects of the interaction terms. We find that the widening of yield spreads during the financial crisis is mainly caused by worsening liquidity conditions, which, in turn, are for a large part the consequence of deteriorating credit conditions. The credit-induced liquidity component is most significant for bonds with lower credit ratings. The effect of liquidity on credit risk is a bit smaller, but can still make up a non-trivial part of the yield spreads.

In the Ford case study, we show that the yield spread decomposition can change considerably over time. For example, the credit-induced liquidity component accounts on average for 14.4% of the 10-year yield spread, but we show that in the most turbulent period of our sample this fraction goes up to 40%. We also investigate the interaction components along the term structure and find that, in relative terms, the effect of the credit-induced liquidity component is particularly pronounced for long maturity bonds.

Finally, we highlight that our results have important implications for risk management, and show that ignoring credit-liquidity interactions results in a significant underestimation of risk and leads to too low capital buffers, especially for bonds with lower credit ratings.

## References


Acharya, V. V. and Pedersen, L. H. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics*, 77:375 – 410.

Aït-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. J. A. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117:585 – 606.

Aït-Sahalia, Y., Laeven, R. J. A., and Pelizzon, L. (2014). Mutual excitation in Eurozone sovereign CDS. *Journal of Econometrics*, 183:151 – 167.

Bao, J., Pan, J., and Wang, J. (2011). The illiquidity of corporate bonds. *Journal of Finance*, 66:911 – 946.

Bongaerts, D., De Jong, F., and Driessen, J. (2017). An asset pricing approach to liquidity effects in corporate bond markets. *Review of Financial Studies*, 30:1229 – 1269.

Boswijk, H., Laeven, R., and Lalu, A. (2016). Asset returns with self-exciting jumps: Option pricing and estimation with a continuum of moments. Working paper.

Buhler, W. and Trapp, M. (2010). Time-varying credit risk and liquidity premia in bond and CDS markets. Working paper.

Chen, H., Cui, R., He, Z., and Milbradt, K. (2017). Quantifying liquidity and default risks of corporate bonds over the business cycle. *Review of Financial Studies*, forthcoming.

Chib, S. and Jeliazkov, I. (2001). Marginal likelihood from the Metropolis-Hastings output. *Journal of the American Statistical Association*, 96:270 – 281.

De Jong, F. and Schotman, P. (2010). Price discovery in fragmented markets. *Journal of Financial Econometrics*, 8:1 – 28.

Dick-Nielsen, J. (2014). How to clean enhanced trace data. Supplement to "Liquidity biases in TRACE" (2009).

Dick-Nielsen, J., Feldhütter, P., and Lando, D. (2012). Corporate bond liquidity before and after the onset of the subprime crisis. *Journal of Financial Economics*, 103:471 – 492.

Driessen, J. (2005). Is default event risk priced in corporate bonds? *Review of Financial Studies*, 18:165 – 195.

Duffie, D., Pan, J., and Singleton, K. J. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68:1343 – 1376.

Duffie, D., Pedersen, L. H., and Singleton, K. J. (2003). Modeling sovereign yield spreads: A case study of Russian debt. *Journal of Finance*, 58:119 – 160.

Duffie, D. and Singleton, K. J. (1999). Modeling term structures of defaultable bonds. *Review of Financial*

*Studies*, 12:687 – 720.

Eraker, B. (2004). Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *Journal of Finance*, 59:1367 – 1403.

Eraker, B., Johannes, M., and Polson, N. (2003). The impact of jumps in volatility and returns. *Journal of Finance*, 58:1269 – 1300.

Ericsson, J. and Renault, O. (2006). Liquidity and credit risk. *Journal of Finance*, 61:2219 – 2250.

Errais, E., Giesecke, K., and Goldberg, L. R. (2010). Affine point processes and portfolio credit risk. *SIAM Journal of Financial Mathematics*, 1:642 – 665.

Friewald, N., Jankowitsch, R., and Subrahmanyam, M. G. (2012). Illiquidity or credit deterioration: A study of liquidity in the US corporate bond markets during financial crises. *Journal of Financial Economics*, 105:18 – 36.

Glosten, L. and Milgrom, J. (1985). Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14:71 – 100.

He, Z. and Milbradt, K. (2014). Endogenous liquidity and defaultable bonds. *Econometrica*, 82:1443 – 1508.

He, Z. and Xiong, W. (2012). Rollover risk and credit risk. *Journal of Finance*, 67:391 – 430.

Ho, T. and Stoll, H. (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics*, 9:47 – 73.

Janosi, T., Jarrow, R. A., and Yildirim, Y. (2002). Estimating expected losses and liquidity discounts implicit in debt prices. *Journal of Risk*, 5:1 – 38.

Johannes, M., Kumar, R., and Polson, N. (1999). State dependent jump models: How do US equity indices jump? Working paper.

Johannes, M. and Polson, N. (2009). MCMC methods for financial econometrics. In Aït-Sahalia, Y. and Hansen, L., editors, *Handbook of Financial Econometrics*, pages 1 – 72.

Liu, J., Longstaff, F. A., and Mandell, R. (2006). The market price of risk in interest rate swaps: The roles of default and liquidity risks. *Journal of Business*, 79:2337 – 2360.

Longstaff, F. A., Mithal, S., and Neis, E. (2005). Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *Journal of Finance*, 60:2213 – 2253.

Monfort, A. and Renne, J.-P. (2014). Decomposing Euro-area sovereign spreads: Credit and liquidity risks. *Review of Finance*, 18:2103 – 2181.

Nagler, F. (2017). Yield spreads and the corporate bond rollover channel. Working paper.

# Appendices

## Appendix A Constructing Liquidity Jump Time Indicator

Let $\beta_t$ be the true, but unobserved, bid-ask spread which follows an $AR(1)$ process

$$\beta_t = \varphi\beta_{t-1} + \eta_t,$$

with $0 \leq \varphi \leq 1$. In the following, we assume that $\varphi$ is known. In typical higher frequency applications, for example with daily data, $\varphi$ will be close to one. The bid-ask spread can be estimated with error by

$$\hat{\beta}_t = \beta_t + u_t.$$

Now let $y_t = \hat{\beta}_t - \varphi\hat{\beta}_{t-1}$ so that

$$y_t = \eta_t + u_t - \varphi u_{t-1}.$$

The variance and first order auto-covariance of $y_t$ are

$$Var(y_t) = \sigma_\eta^2 + (1+\varphi^2)\sigma_u^2, \quad Cov(y_t, y_{t-1}) = -\varphi\sigma_u^2,$$

and denote $\rho = -Cov(y_t, y_{t-1})/Var(y_t)$. We want to make a best estimate of $\eta_t$ given observations of the series $\hat{x}_t$ or, equivalently, $y_t$. Leads and lags of $y_t$ also contain information about $\eta_t$ as they are correlated with $y_t$. We can use a regression setup to estimate $\eta_t$, using two leads and lags of $y_t$, as follows. Let $Y_t = (y_{t-2}, y_{t-1}, y_t, y_{t+1}, y_{t+2})$. Then the best estimate of $\eta_t$ given $Y_t$ is

$$\mathbb{E}[\eta_t|Y_t] = Cov(\eta_t, Y_t)Var(Y_t)^{-1}Y_t'.$$

Notice that

$$Var(Y_t) = \begin{pmatrix} 1 & -\rho & 0 & 0 & 0 \\ -\rho & 1 & -\rho & 0 & 0 \\ 0 & -\rho & 1 & -\rho & 0 \\ 0 & 0 & -\rho & 1 & -\rho \\ 0 & 0 & 0 & -\rho & 1 \end{pmatrix} Var(y_t)$$

and

$$Cov(\eta_t, Y_t) = (0, 0, \sigma_\eta^2, 0, 0).$$

The inverse of the variance matrix is

$$Var(Y_t)^{-1} = \frac{1}{Var(y_t)(1-2\rho^2)} \begin{pmatrix} 1-\rho^2 & \rho & \rho^2 & \rho^3 & \rho^4 \\ \rho & 1 & \rho & \rho^2 & \rho^3 \\ \rho^2 & \rho & 1 & \rho & \rho^2 \\ \rho^3 & \rho^2 & \rho & 1 & \rho \\ \rho^4 & \rho^3 & \rho^2 & \rho & 1-\rho^2 \end{pmatrix}.$$

Hence, we find the best estimate of $\eta_t$ as

$$\hat{\eta}_t = \frac{\sigma_\eta^2}{Var(y_t)(1-2\rho^2)}(\rho^2 y_{t-2} + \rho y_{t-1} + y_t + \rho y_{t+1} + \rho^2 y_{t+2}).$$

Finally, we need an estimator of $\sigma_\eta^2$. Following De Jong and Schotman (2010), we use the non-parametric estimator

$$\hat{\sigma}_\eta^2 = \sum_{k=-2}^{2} Cov(y_t, y_{t-k}).$$

Putting this together and denoting $\rho_k = Cov(y_t, y_{t-k})/Var(y_t)$, we find

$$\hat{\eta}_t = \frac{\sum_{k=-2}^{2} \rho_i}{1-2\rho^2}(\rho^2 y_{t-2} + \rho y_{t-1} + y_t + \rho y_{t+1} + \rho^2 y_{t+2}),$$

with $\rho = -\rho_1$.

## Appendix B Simulation Study

In order to assess the performance of our estimation methodology, we conduct a Monte Carlo study, which we detail in this Appendix.

In every simulation, we construct a dataset of bond prices by first simulating the intensity and jump time processes, and then applying bond formula (5) to these series and adding small normally distributed measurement errors to the resulting (log) prices. In simulating the state vector series, we use realistic parameter values that mimic the values obtained in the empirical analysis, and we approximate the continuous-time dynamics with an Euler discretization scheme using a very fine time grid. We choose the bond characteristics such that every simulated sample closely reflects the actual data (see Table 1). More specifically, every sample consists of daily observations of 23 semi-annual coupon bonds with coupon rates equal to the sample average of 6.64% and maturities $3, 4, 5, \ldots, 25$ years.

Apart from bond prices, the estimation procedure uses liquidity jump indicators constructed from bid-ask spread data. Since it is hard to simulate such data, it is difficult to present a Monte Carlo study of our complete estimation methodology. Therefore, we focus on two nested model specifications, which we believe to give good insight in the performance of our estimation procedure.

The first model specification we consider, is one in which we abstract away from liquidity issues and assume there is only credit risk. In this univariate model specification, we do not have a liquidity discount factor, and, hence, no liquidity jumps. In line with our full model, we still assume that the credit intensity process contains self-exciting jumps. We first take this univariate model specification to the empirical data and use the obtained parameter values in the Monte Carlo experiment. Since the univariate model is only relevant for

this Monte Carlo study, we do not discuss the detailed results in this paper. We do want to mention that the model fit is poor compared to the multivariate model (relative pricing errors are around 20%), which supports the use of a multivariate model that includes liquidity risk.

In the second model specification, we use the full model, but we assume that the hypothetical liquidity jump indicator is fully informative about liquidity jumps. That is, we assume that liquidity jump times are known, and, hence, we do not need to simulate the bid-ask spread data explicitly. We use the model parameters as found in the empirical application detailed in Section 5.

For both model specifications, we run 100 simulations consisting of 252 daily (i.e., one business year) observations of our bond panel and apply our estimation procedure to the constructed datasets. Admittedly, the number of simulation samples is small compared to the typical number used in Monte Carlo studies, but, given the high computational costs involved in our estimation procedure, it provides a feasible implementation and gives a reasonable indication of the performance of the estimation procedure. The results for the univariate model can be found in Table 8 and the results for the multivariate model in Table 9.

## Tables

**Table 1. Descriptive statistics transaction data Ford.** This table reports descriptive statistics of the transaction prices, maturities, number of transactions per day, and coupon rates for the Ford sample. The sample period is July 2007 until July 2009.

| | Price ($) | Maturity (years) | #Obs. per day | Coupon rate (%) |
|---|---|---|---|---|
| Mean | 600.85 | 10.99 | 21.55 | 6.64 |
| Std. Dev. | 213.23 | 6.57 | 6.61 | 0.83 |
| Min | 93.3 | 3.13 | 3 | 5.65 |
| Max | 952.15 | 25.1 | 64 | 9.375 |

**Table 2. Descriptive statistics transaction data rating portfolios.** This table reports descriptive statistics of the (adapted) transaction prices, maturities, number of transactions per day, and coupon rates for the rating portfolios. The sample period is July 2007 until July 2009.

| | | AAA/AA | A | BBB | BB | B and lower |
|---|---|---|---|---|---|---|
| #Firms | | 219 | 377 | 353 | 236 | 298 |
| #Transactions | | 12,215 | 22,149 | 26,982 | 26,240 | 32,807 |
| Price ($) | Mean | 1001.5 | 982.3 | 958.2 | 896 | 799.4 |
| | Std. Dev. | 43.8 | 82.2 | 84.7 | 91.4 | 151 |
| | Min | 490.2 | 17.0 | 326.4 | 488.2 | 136 |
| | Max | 1261.8 | 1434.7 | 1307.3 | 1218.8 | 1387.2 |
| Maturity (years) | Mean | 8.71 | 14.9 | 14 | 7.5 | 6.26 |
| | Std. Dev. | 9.1 | 10.94 | 10.25 | 4.54 | 3.11 |
| | Min | 1 | 1.02 | 1 | 1.01 | 1.01 |
| | Max | 30 | 30 | 30 | 29.75 | 26.62 |
| #Obs. per day | Mean | 24.6 | 44.8 | 54.5 | 52.8 | 66 |
| | Std. Dev. | 10.4 | 21.3 | 21.5 | 16.2 | 22.3 |
| | Min | 2 | 5 | 2 | 1 | 1 |
| | Max | 90 | 131 | 160 | 130 | 196 |
| Coupon rate (%) | Mean | 5.03 | 5.93 | 6.45 | 7.10 | 8.26 |
| | Std. Dev. | 1.33 | 0.94 | 1.03 | 1.02 | 1.60 |
| | Min | 1.625 | 3.625 | 3.6 | 4.8 | 5 |
| | Max | 9.375 | 10.35 | 10.5 | 12 | 13.5 |

**Table 3. Descriptive statistics relative bid-ask spreads.** This table reports descriptive statistics of the estimated relative (to transaction prices) bid-ask spreads of Ford and the rating portfolios over the whole sample period and the period September 15, 2008 until March 31, 2009, reflecting the peak of the crisis. We report the mean, median, standard deviation, minimum, and maximum of the relative bid-ask spread over the corresponding period.

| Rating | Mean | Median | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| | Full sample period | | | | |
| AAA/AA | 0.71% | 0.62% | 0.32% | 0.00% | 2.32% |
| A | 0.78% | 0.72% | 0.30% | 0.00% | 2.33% |
| BBB | 0.83% | 0.76% | 0.32% | 0.00% | 2.08% |
| BB | 0.69% | 0.63% | 0.29% | 0.31% | 3.37% |
| B and lower | 0.88% | 0.69% | 0.53% | 0.31% | 4.66% |
| Ford | 2.44% | 1.95% | 1.65% | 0.00% | 10.16% |
| | Peak of crisis | | | | |
| AAA/AA | 1.09% | 1.00% | 0.33% | 0.00% | 2.32% |
| A | 1.13% | 1.03% | 0.30% | 0.00% | 2.33% |
| BBB | 1.19% | 1.15% | 0.29% | 0.00% | 2.08% |
| BB | 0.94% | 0.84% | 0.37% | 0.34% | 3.37% |
| B and lower | 1.41% | 1.18% | 0.71% | 0.55% | 4.66% |
| Ford | 4.00% | 3.91% | 2.02% | 0.00% | 10.16% |

**Table 4. Parameter estimates Ford.** This table reports the posterior means, standard deviations (in parenthesis), and 99 percent HPD intervals (in square brackets) for the parameters of the mutually exciting credit-liquidity model based on Ford bond prices in the period July 2007 until July 2009. Throughout, we kept $\lambda_\infty^l = 0.015$, $\gamma = 0.1$, $\sigma_l^2 = 0$, and $\rho = 0.25$ fixed.

| | Post. Mean | Post. Std. Dev. | 99% HPD interval |
|---|---|---|---|
| $\alpha_c$ | 2.1713 | (0.0920) | [2.0434, 2.3956] |
| $\lambda_\infty^c$ | 0.3500 | (0.0595) | [0.2933, 0.5148] |
| $\sigma_c^2$ | 1.5226 | (0.3282) | [1.2056, 2.3366] |
| $\alpha_l$ | 4.1628 | (0.3385) | [3.5747, 4.6861] |
| $\beta_{1,1}$ | 1.0136 | (0.0761) | [0.7859, 1.1365] |
| $\beta_{1,2}$ | 0.0898 | (0.2262) | [0.0000, 0.8382] |
| $\beta_{2,1}$ | 1.0344 | (0.2613) | [0.6618, 1.5214] |
| $\beta_{2,2}$ | 0.9659 | (0.0439) | [0.8923, 1.0637] |
| $\phi_{high}$ | 0.5003 | (0.1439) | [0.1702, 0.8447] |
| $\phi_{low}$ | 0.0149 | (0.0066) | [0.0000, 0.0332] |
| $h$ | 0.1197 | (0.0221) | [0.1106, 0.1994] |

**Table 5. Parameter estimates rating portfolios.** This table reports the posterior means, standard deviations (in parenthesis), and 99 percent HPD intervals (in square brackets) for the parameters of the mutually exciting credit-liquidity model for the rating portfolios in the period July 2007 until July 2009. The last row reports the mean relative pricing error (MRPE). Throughout, we kept $\lambda_\infty^l = 0.005$, $\gamma = 0.1$, $\sigma_l^2 = 0$, and $\rho = 0.25$ fixed.

| | AAA/AA | A | BBB | BB | B and lower |
|---|---|---|---|---|---|
| $\alpha_c$ | 2.2308 | 2.4516 | 2.6513 | 2.5555 | 2.3128 |
| | (0.2899) | (0.2660) | (0.1281) | (0.1950) | (0.0572) |
| | [1.8840, 2.8037] | [2.2106, 3.0395] | [2.4995, 2.8760] | [2.2330, 2.8720] | [2.2375, 2.4427] |
| $\lambda_\infty^c$ | 0.1294 | 0.2127 | 0.1768 | 0.2550 | 0.3366 |
| | (0.0098) | (0.0125) | (0.0047) | (0.0030) | (0.0020) |
| | [0.1169, 0.1472] | [0.1930, 0.2400] | [0.1715, 0.1849] | [0.2501, 0.2604] | [0.3334, 0.3430] |
| $\sigma_c^2$ | 0.5772 | 0.9984 | 0.9309 | 1.2931 | 1.5549 |
| | (0.1214) | (0.0924) | (0.0267) | (0.1128) | (0.0374) |
| | [0.4422, 0.8173] | [0.8907, 1.2500] | [0.8898, 0.9932] | [1.1215, 1.4750] | [1.5223, 1.6514] |
| $\alpha_l$ | 4.0533 | 4.2179 | 4.2087 | 5.2540 | 4.2632 |
| | (0.3256) | (0.1267) | (0.0648) | (0.3263) | (0.1165) |
| | [3.2914, 4.6007] | [3.9718, 4.5577] | [4.0867, 4.3503] | [4.7145, 5.8495] | [4.0997, 4.4654] |
| $\beta_{1,1}$ | 1.4392 | 1.0849 | 1.1548 | 0.7876 | 0.8506 |
| | (0.2200) | (0.1118) | (0.2419) | (0.0213) | (0.0326) |
| | [0.8013, 1.6217] | [0.7733, 1.2149] | [0.5917, 1.4017] | [0.7406, 0.8218] | [0.7812, 0.9213] |
| $\beta_{1,2}$ | 0.4708 | 1.0790 | 1.1064 | 1.1898 | 1.3926 |
| | (0.0543) | (0.0462) | (0.1090) | (0.1592) | (0.0968) |
| | [0.3323, 0.5634] | [0.9916, 1.1667] | [0.7525, 1.2292] | [0.9158, 1.4527] | [1.0613, 1.5007] |
| $\beta_{2,1}$ | 0.0833 | 0.1041 | 0.1557 | 0.8914 | 0.9694 |
| | (0.1217) | (0.1454) | (0.1680) | (0.0365) | (0.0834) |
| | [0.0000, 0.3937] | [0.0000, 0.4716] | [0.0000, 0.5018] | [0.8101, 0.9462] | [0.8782, 1.1541] |
| $\beta_{2,2}$ | 0.9675 | 0.4744 | 0.7152 | 0.7611 | 0.9730 |
| | (0.0408) | (0.0370) | (0.0584) | (0.0257) | (0.0169) |
| | [0.8902, 1.0279] | [0.3934, 0.5492] | [0.6368, 0.8101] | [0.7181, 0.8126] | [0.9348, 1.0160] |
| $\phi_{high}$ | 0.5261 | 0.5178 | 0.4382 | 0.4940 | 0.4036 |
| | (0.1459) | (0.1458) | (0.1557) | (0.1516) | (0.1385) |
| | [0.1919, 0.8813] | [0.1725, 0.8579] | [0.1089, 0.8232] | [0.1631, 0.8545] | [0.1018, 0.7516] |
| $\phi_{low}$ | 0.0104 | 0.0082 | 0.0119 | 0.0138 | 0.0132 |
| | (0.0046) | (0.0040) | (0.0047) | (0.0051) | (0.0053) |
| | [0.0018, 0.0243] | [0.0008, 0.0203] | [0.0025, 0.0256] | [0.0038, 0.0294] | [0.0028, 0.0290] |
| $h$ | 0.0476 | 0.1036 | 0.0643 | 0.0623 | 0.1385 |
| | (0.0009) | (0.0007) | (0.0005) | (0.0003) | (0.0021) |
| | [0.0461, 0.0495] | [0.1020, 0.1055] | [0.0633, 0.0654] | [0.0615, 0.0631] | [0.1359, 0.1447] |
| ARPE | 3.19% | 5.49% | 5.10% | 4.88% | 10.25% |
| # Credit jumps | 3 | 10 | 11 | 8 | 16 |
| # Liquidity jumps | 2 | 2 | 4 | 1 | 7 |

**Table 6. Average yield spread decomposition.** This table reports the average contribution of the different components to the 10-year bond yield spreads of Ford and the rating portfolios over the whole sample period and the period September 15, 2008 until March 31, 2009, reflecting the peak of the crisis. We report both the average absolute contribution in percentage points as well as the relative contribution to the total credit spread (in parentheses).

| Rating | $ys_{\text{pureCRED}}$ | $ys_{\text{LIQ} \to \text{CRED}}$ | $ys_{\text{pureLIQ}}$ | $ys_{\text{CRED} \to \text{LIQ}}$ | $ys$ |
|---|---|---|---|---|---|
| | Full sample period | | | | |
| AAA/AA | 1.31 (73.6%) | 0.05 (2.8%) | 0.28 (15.7%) | 0.14 (7.9%) | 1.78 (100%) |
| A | 1.65 (70.8%) | 0.08 (3.4%) | 0.41 (17.6%) | 0.19 (8.2%) | 2.33 (100%) |
| BBB | 1.47 (55.1%) | 0.15 (5.6%) | 0.82 (30.7%) | 0.23 (8.6%) | 2.67 (100%) |
| BB | 1.75 (41.7%) | 0.02 (0.5%) | 2.03 (48.3%) | 0.40 (9.5%) | 4.20 (100%) |
| B and lower | 2.25 (33.4%) | 0.21 (3.1%) | 3.15 (46.8%) | 1.12 (16.6%) | 6.73 (100%) |
| Ford | 3.99 (39.5%) | 0.01 (0.1%) | 4.65 (46.0%) | 1.46 (14.4%) | 10.11 (100%) |
| | Peak of crisis | | | | |
| AAA/AA | 1.41 (61.6%) | 0.16 (7.0%) | 0.54 (23.6%) | 0.18 (7.9%) | 2.29 (100%) |
| A | 1.86 (55.9%) | 0.25 (7.5%) | 0.93 (27.9%) | 0.29 (8.7%) | 3.33 (100%) |
| BBB | 1.56 (40.2%) | 0.36 (9.3%) | 1.64 (42.3%) | 0.32 (8.2%) | 3.88 (100%) |
| BB | 2.01 (36.5%) | 0.03 (0.5%) | 2.78 (50.5%) | 0.69 (12.5%) | 5.51 (100%) |
| B and lower | 2.43 (27.6%) | 0.51 (5.8%) | 3.56 (40.4%) | 2.32 (26.3%) | 8.82 (100%) |
| Ford | 4.65 (31.0%) | 0.03 (0.2%) | 6.94 (46.2%) | 3.41 (22.7%) | 15.01 (100%) |

**Table 7. VaR capital requirements per dollar invested.** This table reports the VaR capital requirement in cents per dollar invested in a 10-year semi-annual coupon bond based on different model specifications. For every rating class/Ford, we simulate 100,000 10-day paths of the intensity and jump processes using the parameter estimates of the corresponding rating class/Ford and taking the intensity values on September 15, 2008, as a starting point.

| | Full model | No Cross-Excitation | No Excitation |
|---|---|---|---|
| AAA/AA | 8.83 | 7.72 | 0.09 |
| A | 7.48 | 5.03 | 0.08 |
| BBB | 8.29 | 4.56 | 0.07 |
| BB | 10.50 | 3.21 | 1.37 |
| B and lower | 19.97 | 4.70 | 1.79 |
| Ford | 19.00 | 6.38 | 1.16 |

**Table 8. Monte Carlo Results Univariate Model.** This table reports the results of the Monte Carlo study using the univariate model specification with self-exciting credit intensity process following the dynamics $\mathrm{d}\lambda_t^c = \alpha_c(\lambda_\infty^c - \lambda_t^c)\mathrm{d}t + \sigma_c\sqrt{\lambda_t}\mathrm{d}W_t^c + \beta_{1,1}\mathrm{d}N_t^c$. The study consists of 100 simulated bond panels with maturities 3,4,...,25 years, each containing 252 observations (i.e., 1 year of simulated data at daily frequency assuming 252 business days per year). The true parameter values, Monte Carlo sample means, standard deviations, and minimum and maximum values are reported in separate columns.

| | True | MC Mean | MC Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| $\alpha_c$ | 1.75 | 1.8068 | 0.0830 | 1.6004 | 1.9387 |
| $\lambda_\infty^c$ | 0.55 | 0.5581 | 0.0328 | 0.4602 | 0.6191 |
| $\sigma_c^2$ | 1.85 | 1.8073 | 0.0983 | 1.6664 | 2.0196 |
| $\beta_{1,1}$ | 1.7 | 1.7252 | 0.0829 | 1.5455 | 1.9021 |
| $\gamma$ | 0.1 | 0.1001 | 0.0124 | 0.0819 | 0.1457 |
| $h$ | 0.01 | 0.0116 | 0.0003 | 0.0113 | 0.0129 |

**Table 9. Monte Carlo Results Multivariate Model.** This table reports the results of the Monte Carlo study using the bivariate model specification with mutually exciting credit and liquidity intensity processes with the following dynamics: $\mathrm{d}\lambda_t^c = \alpha_c(\lambda_\infty^c - \lambda_t^c)\mathrm{d}t + \sigma_c\sqrt{\lambda_t}\mathrm{d}W_t^c + \beta_{1,1}\mathrm{d}N_t^c + \beta_{1,2}\mathrm{d}N_t^l$ and $\mathrm{d}\lambda_t^l = \alpha_l(\lambda_\infty^l - \lambda_t^l)\mathrm{d}t + \beta_{2,1}\mathrm{d}N_t^c + \beta_{2,2}\mathrm{d}N_t^l$. The study consists of 100 simulated bond panels with maturities 3,4,...,25 years, each containing 252 observations (i.e., 1 year of simulated data at daily frequency assuming 252 business days per year). Throughout, we assume that liquidity jump times are known, and we keep $\rho = 0.25, \lambda_\infty^l = 0.015$, and $\gamma = 0.1$ fixed. The true parameter values, Monte Carlo sample means, standard deviations, and minimum and maximum values are reported in separate columns.

| | True | MC Mean | MC Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| $\alpha_c$ | 2.17 | 2.2091 | 0.1884 | 1.7920 | 2.6498 |
| $\lambda_\infty^c$ | 0.35 | 0.3596 | 0.0189 | 0.3175 | 0.4084 |
| $\sigma_c^2$ | 1.52 | 1.4573 | 0.1216 | 1.1368 | 1.7260 |
| $\alpha_l$ | 4.16 | 4.2857 | 0.2732 | 3.7130 | 4.9681 |
| $\beta_{1,1}$ | 1.01 | 1.0260 | 0.1062 | 0.7539 | 1.3111 |
| $\beta_{1,2}$ | 0.09 | 0.0905 | 0.0624 | 0.0157 | 0.3170 |
| $\beta_{2,1}$ | 1.03 | 1.0658 | 0.0672 | 0.9168 | 1.2393 |
| $\beta_{2,2}$ | 0.97 | 0.9449 | 0.0608 | 0.8082 | 1.1068 |
| $h$ | 0.05 | 0.0498 | 0.0006 | 0.0488 | 0.0518 |

## Figures

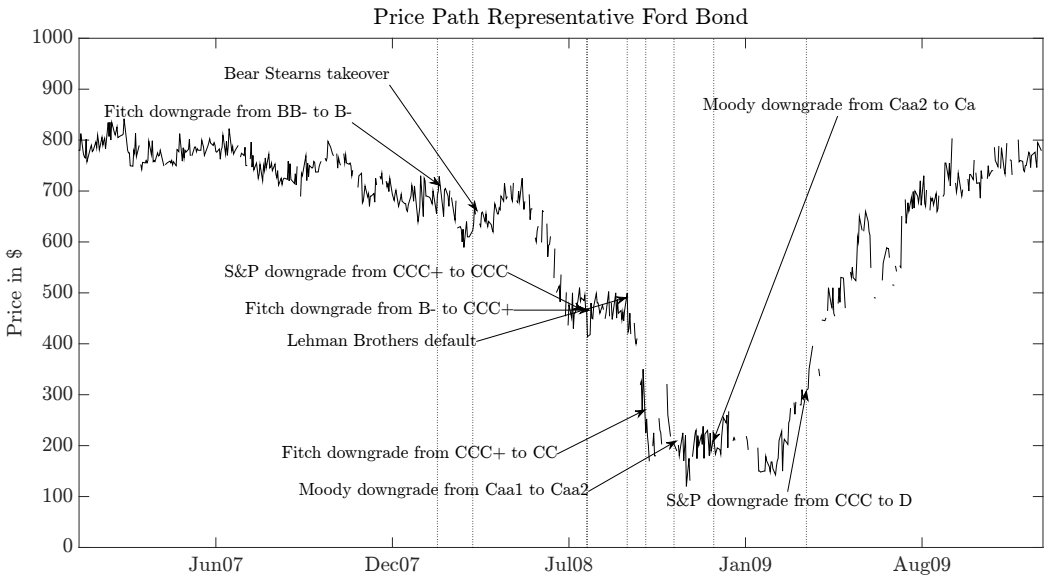


**Figure 1. Representative Ford bond price path.** The figure plots the price path of a representative bond of Ford, and indicates several credit events that Ford was exposed to. The reported prices are the final transaction prices at each date the bond was traded. Source: TRACE and Mergent FISD.

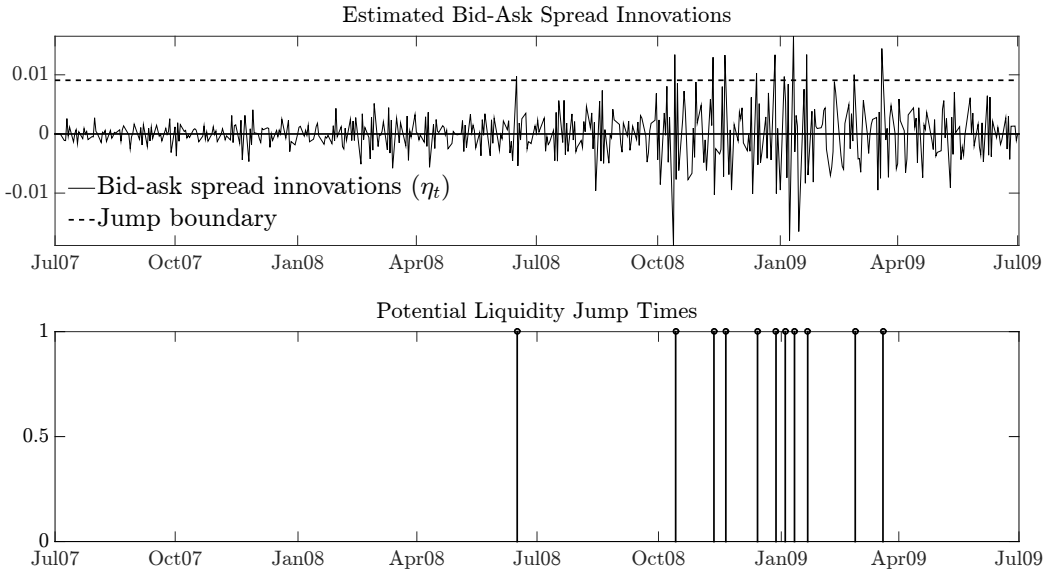


**Figure 2. Liquidity jump indicator Ford.** The upper panel of the figure displays the filtered $\eta_t$ series for the Ford bond portfolio, representing the innovations in the relative bid-ask spreads. The dashed horizontal line represents the jump indicator boundary, which is set as 2.25 standard deviations above the sample average of $\eta_t$. The lower panel of the figure displays the dates on which the jump indicator boundary is exceeded.

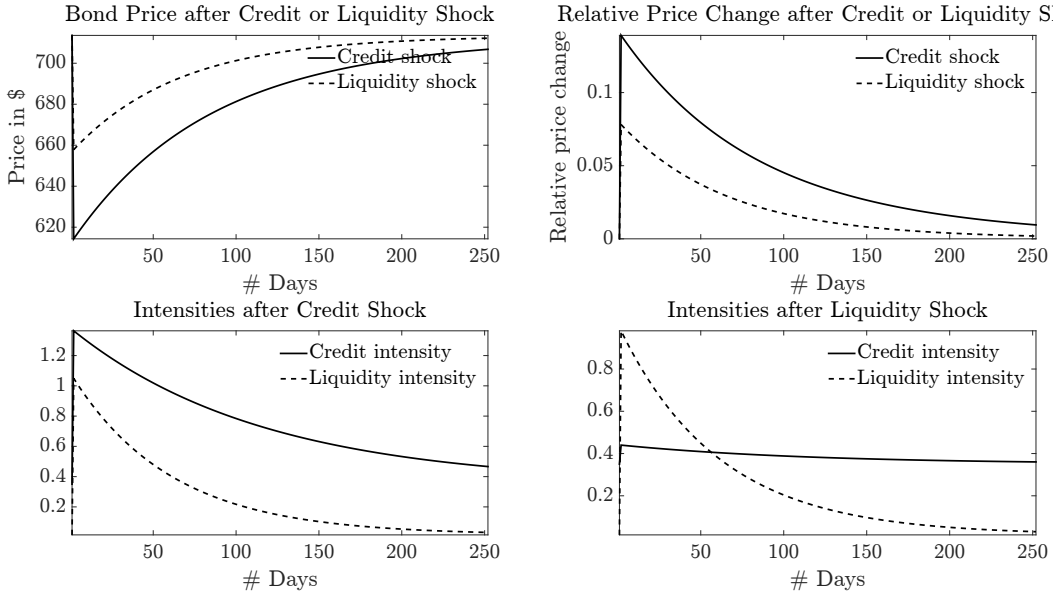


**Figure 3. Price decay credit and liquidity shocks Ford.** The figure plots the effects of credit and liquidity shocks on bond prices and the credit and liquidity intensities. In every panel, the credit and liquidity intensities start at their long-term means $\lambda^c_\infty = 0.3500$ and $\lambda^l_\infty = 0.015$, respectively, and we assume that either a credit shock or a liquidity shock occurs (at Day 0). The upper left panel plots the price patterns after the credit or liquidity shock. The upper right panel plots the evolution of the relative price differences (compared to the initial price) after the credit or liquidity shock. The bottom left panel plots the credit and liquidity intensity paths after the credit jump. The bottom right panel plots the credit and liquidity intensity paths after a liquidity jump.

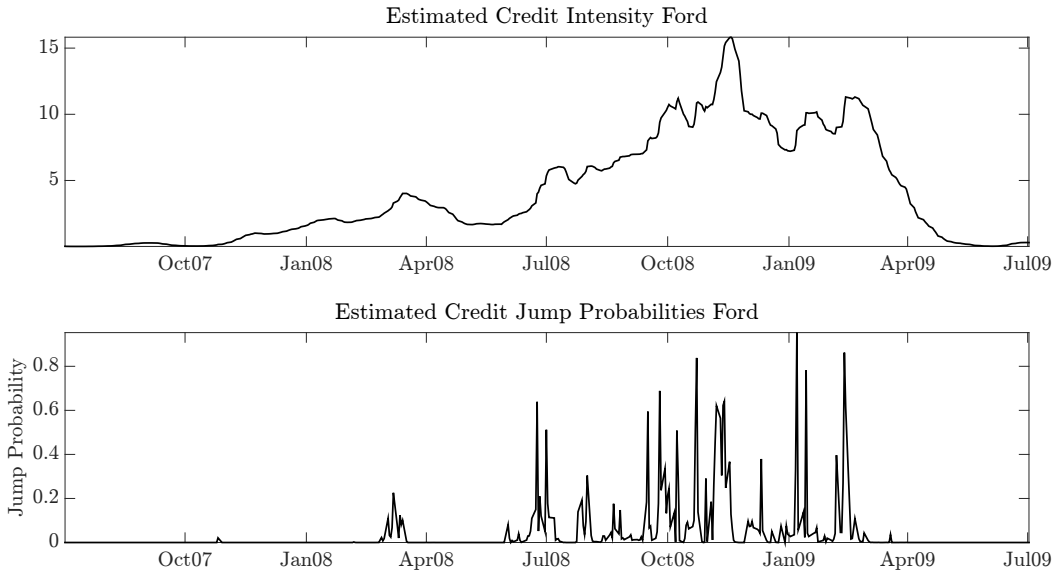


**Figure 4. Estimated credit intensities and credit jump probabilities Ford.** The upper panel plots the estimated credit intensities over the period July 2007 until July 2009. The bottom panel displays the estimated probabilities of credit jump occurrences.

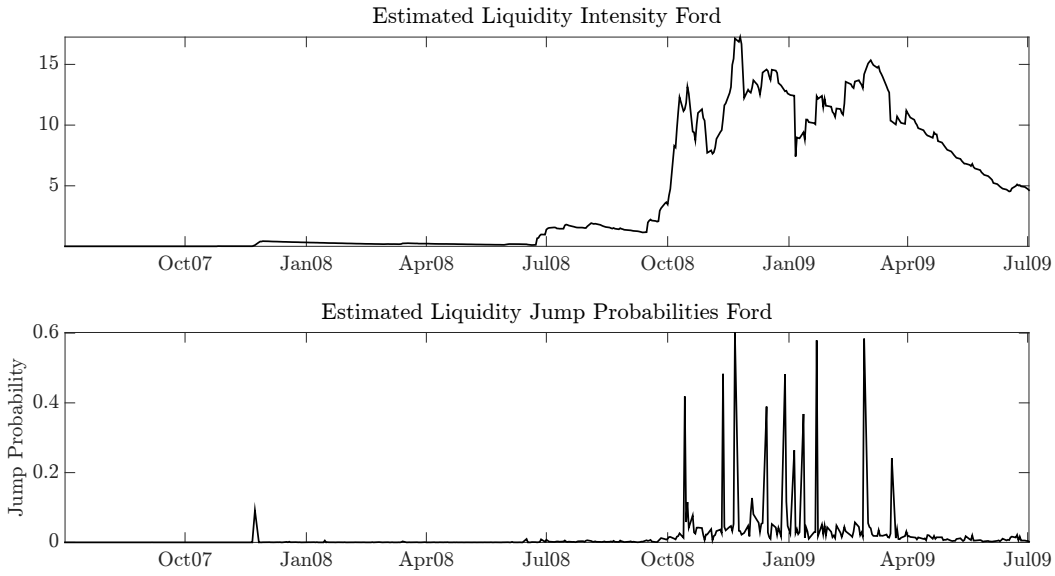


**Figure 5. Estimated liquidity intensities and liquidity jump probabilities Ford.** The upper panel plots the estimated liquidity intensities over the period July 2007 until July 2009. The bottom panel displays the estimated probabilities of liquidity jump occurrences.

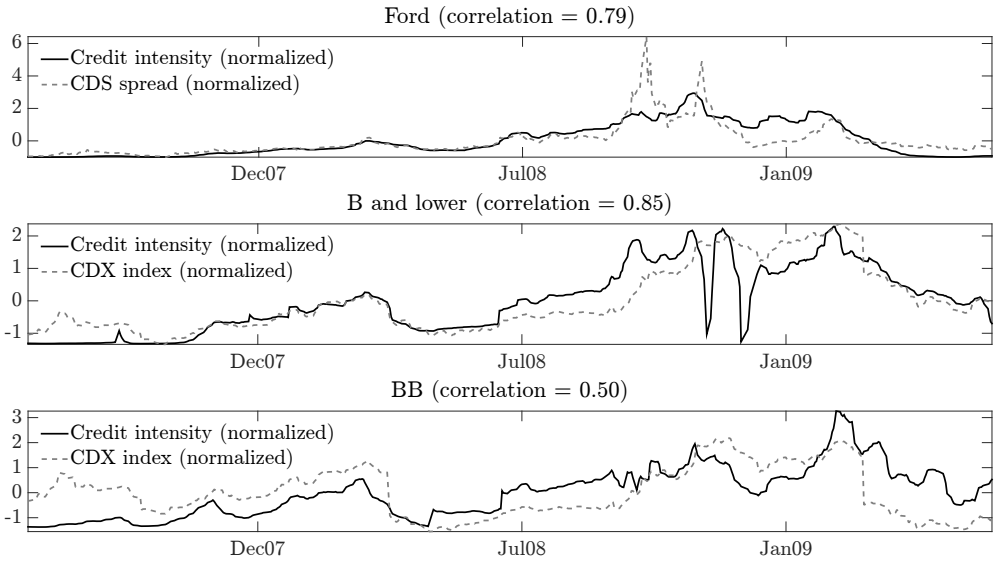


**Figure 6. Credit identification.** The figure compares the (normalized) estimated credit intensities with market-implied measures of credit risk. For Ford (upper panel) we use the (normalized) 10-year CDS spread. For the B and lower (middle panel) and BB (bottom panel) rating classes, we use the (normalized) 5-year maturity Markit CDS indices (CDX) of high yield firms with a B and BB rating, respectively. For each panel, the corresponding correlation between the (normalized) credit intensities and CDS/CDX data is reported.

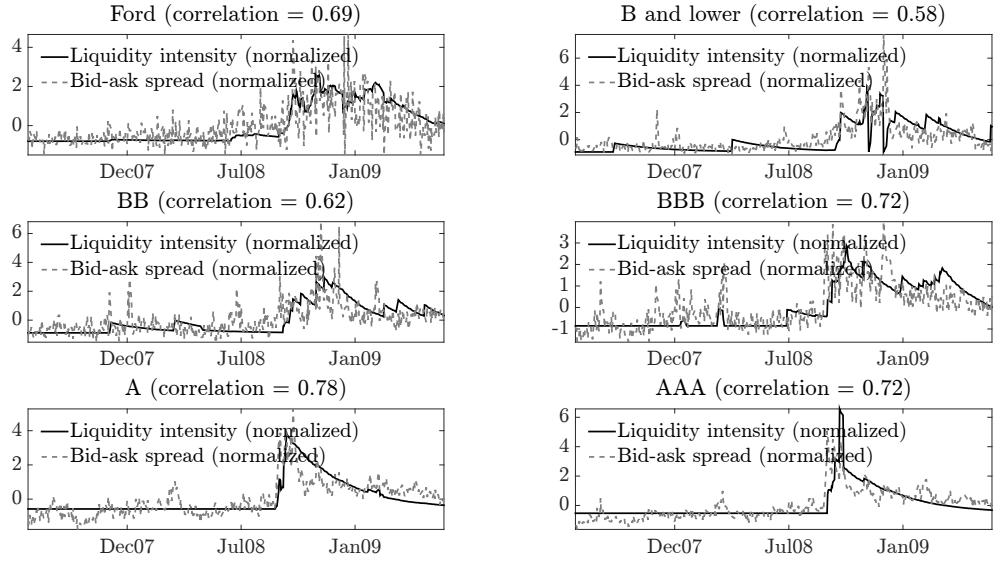


**Figure 7. Liquidity identification.** The figure plots the (normalized) estimated liquidity intensities and corresponding (normalized) estimated relative bid-ask spreads for Ford and all rating classes. For each panel, the corresponding correlation between the (normalized) liquidity intensities and relative bid-ask spreads is reported.

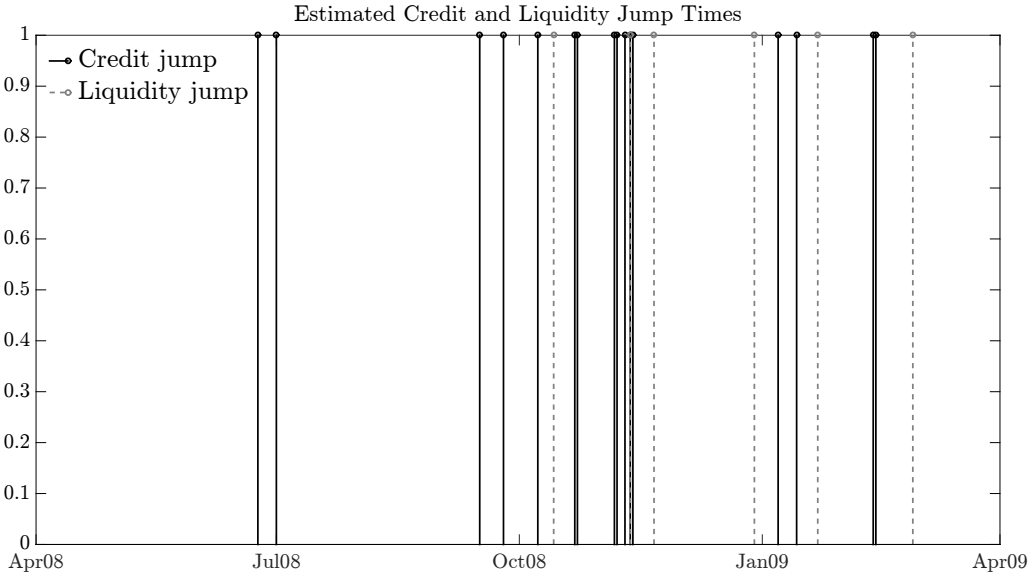


**Figure 8. Estimated liquidity and credit jump times Ford.** The figure plots the estimated credit and liquidity jumps for Ford. A jump is detected if the estimated jump probability is larger than 0.4.

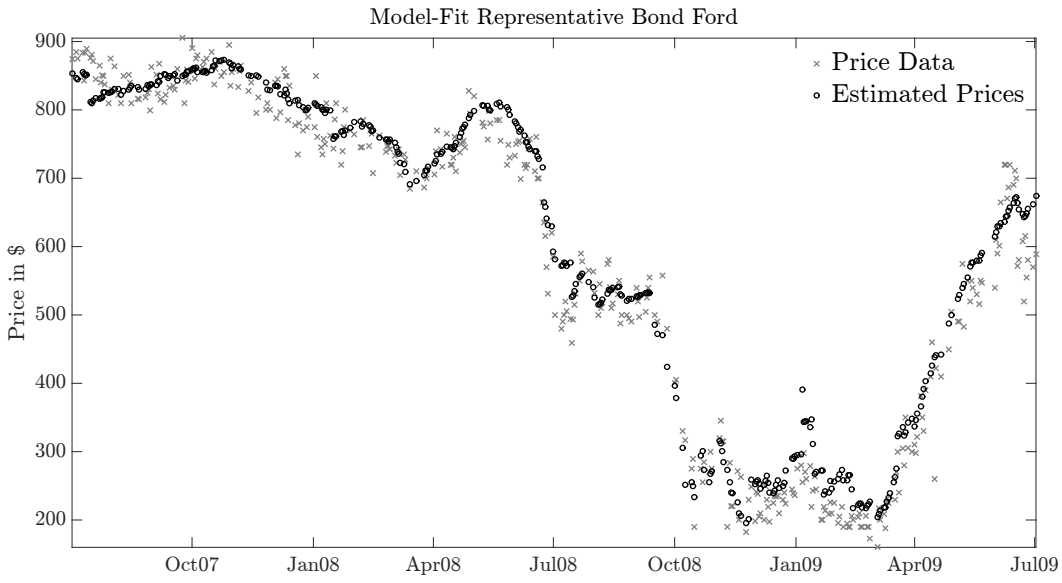


**Figure 9. Estimated prices representative Ford bond.** The figure plots the actual and model-implied price paths of a representative bond of Ford.

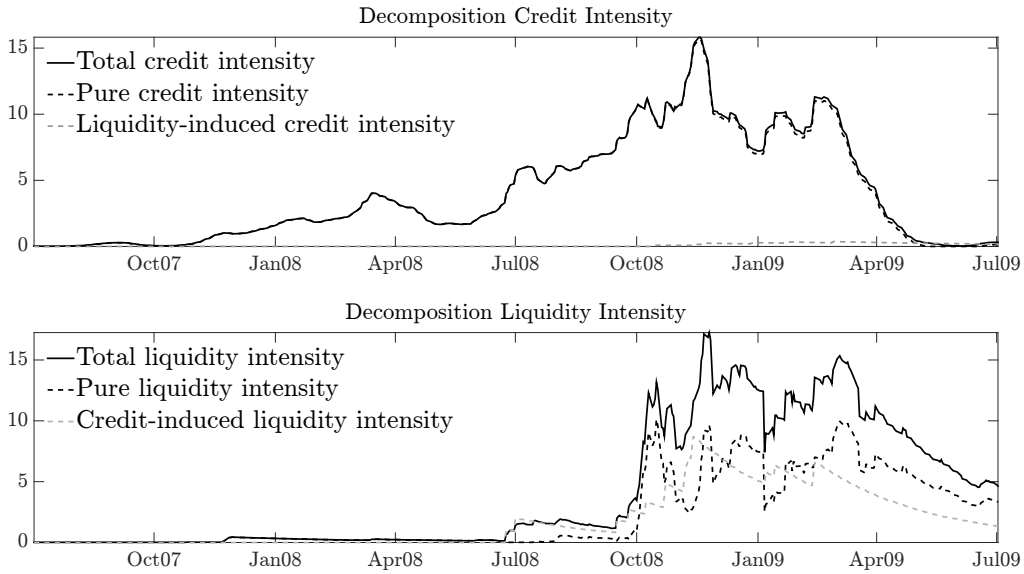


**Figure 10. Decomposition estimated intensities Ford.** The top panel plots the estimated credit intensities, the pure credit intensities, and the liquidity-induced credit intensities for the Ford sample. The bottom panel plots the estimated liquidity intensities, the pure liquidity intensities, and the credit-induced liquidity intensities for the Ford sample.

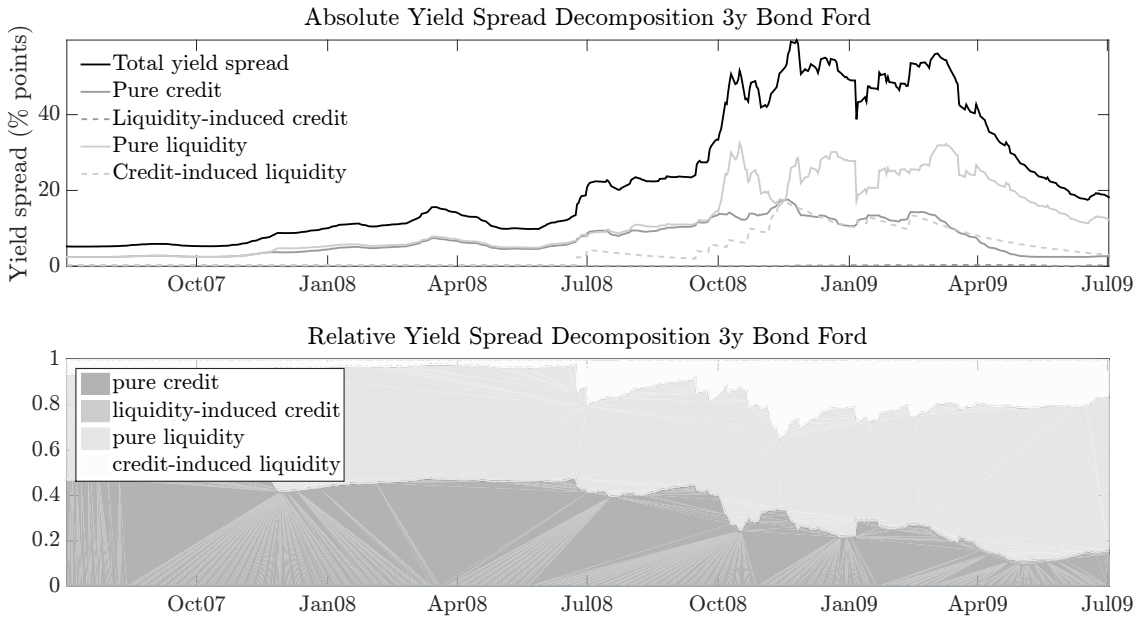


**Figure 11. Decomposition of 3y Ford bond yield spread.** The upper panel of the figure plots the absolute contribution of the pure credit, pure liquidity, credit-driven liquidity, and liquidity-driven credit components to the 3-year Ford bond yield spread over time. The lower panel displays the relative size of the different components in terms of the total yield spread over time. Note that the liquidity-driven credit component is small and therefore hardly visible.

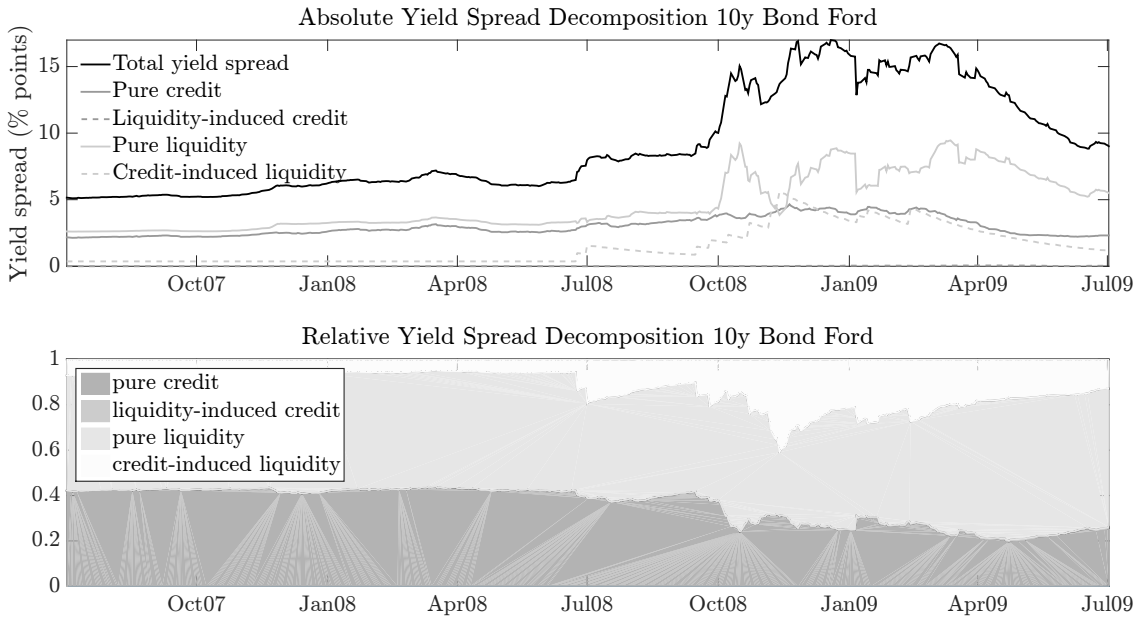


**Figure 12. Decomposition of 10y Ford bond yield spread.** The upper panel of the figure plots the absolute contribution of the pure credit, pure liquidity, credit-driven liquidity, and liquidity-driven credit components to the 10-year Ford bond yield spread over time. The lower panel displays the relative size of the different components in terms of the total yield spread over time. Note that the liquidity-driven credit component is small and therefore hardly visible.

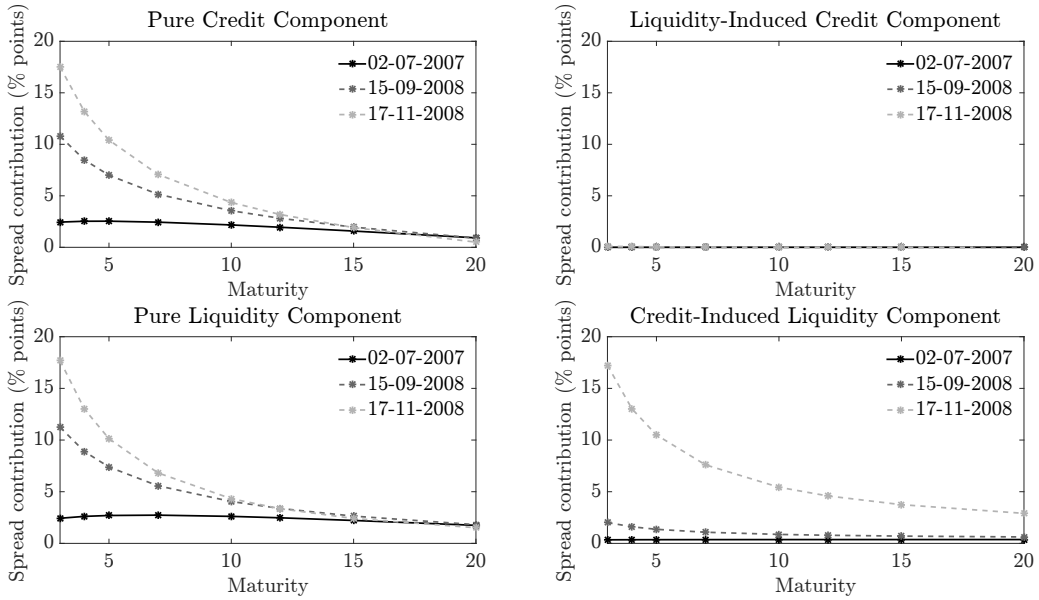


**Figure 13. Absolute term structure per component for Ford.** The figure plots the absolute term structures for the pure credit, pure liquidity, credit-driven liquidity, and liquidity-driven credit components of Ford bond yield spreads on different dates in the sample. The dates represent the start of the sample (July 2, 2007), the default date of Lehman Brothers (September 15, 2008), and two months after the default of Lehman Brothers (November 17, 2008).

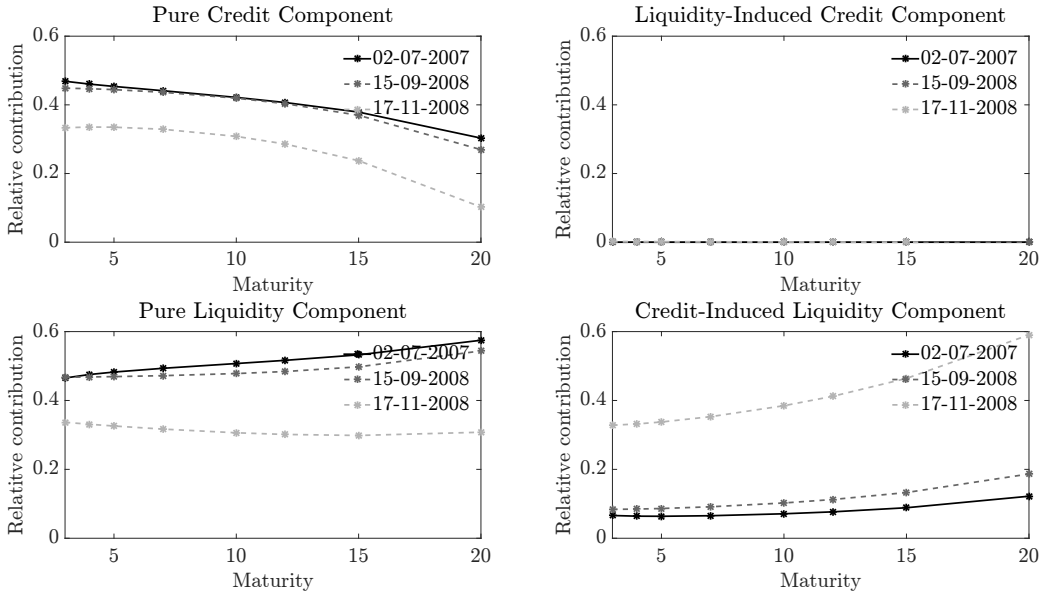


**Figure 14. Relative term structure per component for Ford.** The figure plots the relative term structures for the pure credit, pure liquidity, credit-driven liquidity, and liquidity-driven credit components of Ford bond yield spreads on different dates in the sample. The dates represent the start of the sample (July 2, 2007), the default date of Lehman Brothers (September 15, 2008), and two months after the default of Lehman Brothers (November 17, 2008).

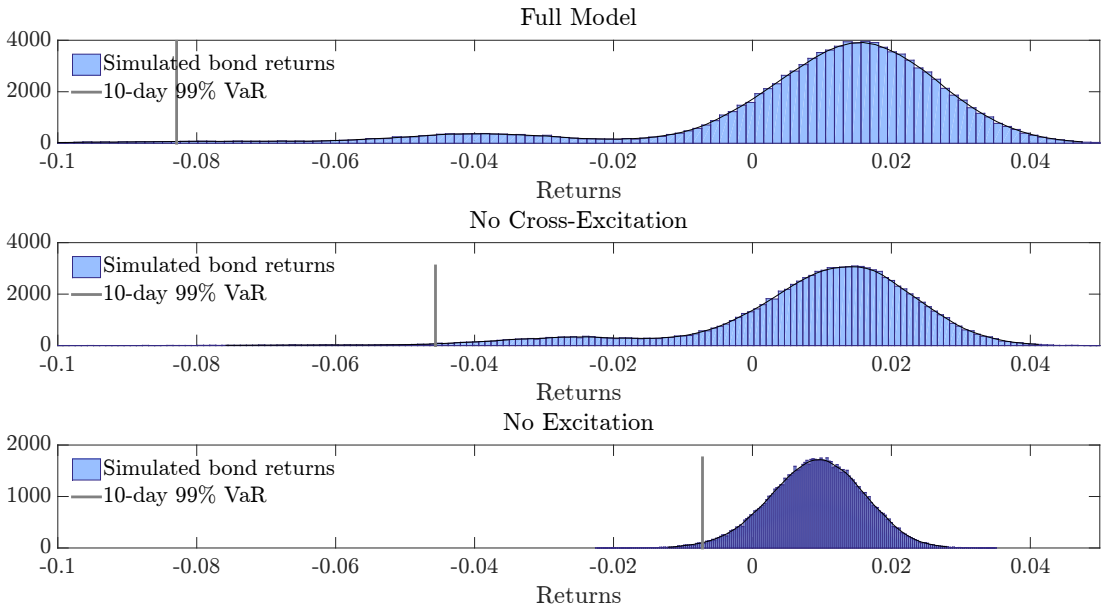


**Figure 15. Simulated 10-day bond return distributions BBB-rated bonds.** This figure plots the simulated 10-day bond returns of BBB-rated bonds using the full model (upper panel), the model without cross-excitation (middle panel), and the model without cross- and self-excitation (bottom panel). In every panel, the first percentile (i.e., the 10-day 99% VaR) of the simulated return distribution is indicated by a vertical line.

# Internet Appendices

### IA.1. Closed-Form Pricing Formula

In this Appendix, we show how to compute the expectations appearing in (5) in closed-form (up to the solution of a system of ODEs) by exploiting the fact that the state vector consisting of the jump and intensity processes (i.e., $X_t = (N_t^c, N_t^l, \lambda_t^c, \lambda_t^l)$) falls into the class of affine jump-diffusions (see also Aït-Sahalia et al., 2014, and Errais et al., 2010). A generalized jump-diffusion process $X_t$ in a state space $D \subset \mathbb{R}^{2\times m}$ is defined to be the solution of the stochastic differential equation

$$\mathrm{d}X_t = \mu^X(X_t)\mathrm{d}t + \sigma^X(X_t)\mathrm{d}W_t^X + \sum_{j=1}^{m} \mathrm{d}J_{j,t}, \tag{IA.1}$$

where $\mu^X : D \to \mathbb{R}^{2\times m}$, $\sigma^X : D \to \mathbb{R}^{(2\times m)\times(2\times m)}$, $W^X$ is a Brownian motion in $\mathbb{R}^{2\times m}$ and $J_i$, $i = 1, \ldots, m$, are pure jump processes with jump intensities $\lambda_{i,t}^X = \lambda_i^X(X_t)$, for some $\lambda_i^X : D \to [0, \infty)$ and with fixed jump size distribution $\nu^i$ with jump transforms $\theta^i(c) = \int_{\mathbb{R}^n} \exp(c \cdot z)\, \mathrm{d}\nu(z)$ for $c \in \mathbb{C}^n$, and where $\mu^X$, $\sigma^X \sigma^{X\prime}$ and $\lambda_i^X$ are affine on $D$:

$$\begin{aligned}
\mu^X &= K_0 + K_1 x, \text{ for } K = (K_0, K_1) \in \mathbb{R}^{2\times m} \times \mathbb{R}^{(2\times m)\times(2\times m)}, \\
\left(\sigma^X \sigma^{X\prime}\right)_{ij} &= (H_0)_{ij} + (H_1)_{ij} \cdot x, \text{ for } H = (H_0, H_1) \in \mathbb{R}^{(2\times m)\times(2\times m)} \times \mathbb{R}^{(2\times m)\times(2\times m)\times(2\times m)}, \\
\lambda_i^X &= l_0 + l_1 \cdot x, \text{ for } l = (l_0, l_1) \in \mathbb{R} \times \mathbb{R}^{(2\times m)}.
\end{aligned}$$

Now let $X_t = (N_t^c, N_t^l, \lambda_t^c, \lambda_t^l)'$. We then have that de dynamics of $X$ are given by

$$\begin{aligned}
\mathrm{d}X_t &= \mathrm{d}\begin{pmatrix} N_t^c \\ N_t^l \\ \lambda_t^c \\ \lambda_t^l \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 0 \\ \alpha_c(\lambda_\infty^c - \lambda_t^c) \\ \alpha_l(\lambda_\infty^l - \lambda_t^l) \end{pmatrix} \mathrm{d}t + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_c\sqrt{\lambda_t^c} & 0 \\ 0 & 0 & 0 & \sigma_l\sqrt{\lambda_t^l} \end{pmatrix} \mathrm{d}\begin{pmatrix} 0 \\ 0 \\ W_t^c \\ W_t^l \end{pmatrix} \\
&\quad + \begin{pmatrix} 1 \\ 0 \\ \beta_{1,1} \\ \beta_{2,1} \end{pmatrix} \mathrm{d}N_t^c + \begin{pmatrix} 0 \\ 1 \\ \beta_{1,2} \\ \beta_{2,2} \end{pmatrix} \mathrm{d}N_t^l. \qquad \text{(IA.2)}
\end{aligned}$$

From this specification it is clear that the joint state vector $X$ is affine, since we can write $n = 2 \times 2 = 4$, $m = 2$, $K_0 = (0, 0, \alpha_c\lambda_\infty^c, \alpha_l\lambda_\infty^l)^\top$,

$$K_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\alpha_c & 0 \\ 0 & 0 & 0 & -\alpha_l \end{pmatrix},$$

$H_0 = \mathbf{0}$,

$$H_1^1 = H_1^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, H_1^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_c^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, H_1^4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_l^2 \end{pmatrix},$$

$l_0^c = l_0^l = 0$, $l_1^c = (0,0,1,0)^\top$, and $l_1^l = (0,0,0,1)^\top$. Furthermore, since $N_t^c$ and $N_t^l$ are counting processes and the coefficients $\beta_{i,j}$ for $i,j = 1,2$ are constants, $\theta^c(c) = \exp(c_1 + \beta_{1,1}c_3 + \beta_{2,1}c_4)$ and $\theta^l(c) = \exp(c_2 + \beta_{1,2}c_3 + \beta_{2,2}c_4)$.

Using results of Duffie et al. (2000), we can now find closed-form expressions for the transform $u \mapsto \mathbb{E}\left[e^{-\int_t^T R(X_s)\mathrm{d}s} e^{u\cdot X_T} \middle| \mathcal{F}_t\right]$ and extended transform $(u,v) \mapsto \mathbb{E}\left[e^{-\int_t^T R(X_s)\mathrm{d}s}(v\cdot X_T) e^{u\cdot X_T} \middle| \mathcal{F}_t\right]$, where $R(x) = \rho_0 + \rho_1 \cdot x$ for $(\rho_0, \rho_1) \in \mathbb{R} \times \mathbb{R}^{2\times m}$.

**Proposition 1.**

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T \rho\lambda_s^l \mathrm{d}s}(1-\gamma)^{N_T^c} \middle| \mathcal{F}_t\right] = e^{\alpha(t)+\beta_1(t)N_t^c+\beta_2(t)N_t^l+\beta_3(t)\lambda_t^c+\beta_4(t)\lambda_t^l},$$

*with*

$$\begin{aligned}
\dot{\alpha}(t) &= -\alpha_c\lambda_\infty^c\beta_3(t) - \alpha_l\lambda_\infty^l\beta_4(t), \\
\dot{\beta}_1(t) &= \dot{\beta}_2(t) = 0, \\
\dot{\beta}_3(t) &= \alpha_c\beta_3(t) - \frac{1}{2}(\sigma_c^2\beta_3^2(t)) - \left(e^{\beta_1(t)+\beta_{1,1}\beta_3(t)+\beta_{2,1}\beta_4(t)} - 1\right), \\
\dot{\beta}_4(t) &= \rho + \alpha_l\beta_4(t) - \frac{1}{2}(\sigma_l^2\beta_4^2(t)) - \left(e^{\beta_2(t)+\beta_{1,2}\beta_3(t)+\beta_{2,2}\beta_4(t)} - 1\right), \\
\alpha(T) &= \beta_2(T) = \beta_2(s) = \beta_3(T) = \beta_4(T) = 0, \\
\beta_1(T) &= \log(1-\gamma) = \beta_1(s), \quad t \le s \le T.
\end{aligned}$$

*Proof.* Consider an affine jump-diffusion process $X$ in some state space $D \subset \mathbb{R}^n$ solving the

stochastic differential equation

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t + \sum_{i=1}^{m} \mathrm{d}Z_t^i,$$

where $Z^i$ are pure jump processes whose jumps have a fixed probability distribution $\nu^i$ on $\mathbb{R}^n$ and arrive with intensity $\lambda^i(X_t)$ for some $\lambda^i : D \to [0, \infty)$. Let us fix an affine process $R : D \to \mathbb{R}$. Then we have that the complete affine structure of the model is captured by:

$$\begin{aligned}
\mu(x) &= K_0 + K_1 x, \text{ for } K = (K_0, K_1) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}. \\
\sigma(x)\sigma(x)^\top &= H_0 + \sum_{k=1}^{n} H_1^{(k)} x_k, \text{ for } H = (H_0, H_1) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n}. \\
\lambda^i(x) &= l_0^i + l_1^i \cdot x, \text{ for } l = (l_0, l_1) \in \mathbb{R} \times \mathbb{R}^n. \\
R(x) &= \rho_0 + \rho_1 \cdot x, \text{ for } \rho = (\rho_0, \rho_1) \in \mathbb{R} \times \mathbb{R}^n.
\end{aligned}$$

Let us furthermore denote the jump-transforms, which determine the jump-size distributions, as $\theta^i(c) = \int_{\mathbb{R}^n} \exp(c \cdot z)\, \mathrm{d}\nu^i(z)$ for $c \in \mathbb{C}^n$.

We want to compute an expression of the form

$$\phi^X(u, X, t, T) = \mathbb{E}^X \left[ \exp\left( - \int_t^T R(X_s)\mathrm{d}s \right) e^{u \cdot X_T} \middle| \mathcal{F}_t \right].$$

According to Proposition 1 of Duffie et al. (2000), we have, under some technical assumptions on the processes being well-behaved, that we can write

$$\phi^X(u, x, t, T) = e^{\alpha(t) + \beta(t) \cdot x},$$

where $\beta$ and $\alpha$ satisfy the following (complex-valued) ODEs:

$$\begin{aligned}\dot{\beta}(t) &= \rho_1 - K_1^\top \beta(t) - \frac{1}{2}\beta(t)^\top H_1 \beta(t) - \sum_{i=1}^{m} l_1^i(\theta^i(\beta(t)) - 1) \\ \dot{\alpha}(t) &= \rho_0 - K_0 \cdot \beta_t - \frac{1}{2}\beta(t)^\top H_0 \beta(t) - \sum_{i=1}^{m} l_0^i(\theta^i(\beta(t)) - 1),\end{aligned}$$

with boundary conditions $\beta(T) = u$ and $\alpha(T) = 0$.

Since we want to compute $\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T \rho\lambda_s^l \mathrm{d}s}(1-\gamma)^{N_T^c}\middle|\mathcal{F}_t\right]$, we have that $\rho_0 = 0$, $\rho_1 = (0,0,0,\rho)^\top$ and $u = (\log(1-\gamma), 0, 0, 0)^\top$. Together with the system matrices, $K_0$, $K_1$, $H_0$, $H_1$, $l_0^c$, $l_0^l$, $l_1^c$, $l_1^l$, $\theta^c(c)$, and $\theta^l(c)$ as specified above, the result thus follows by applying Proposition 1 of Duffie et al. (2000). □

In a similar way we can find a system of ODEs for $\mathbb{E}^{\mathbb{Q}}\left[\lambda_T^c \gamma e^{-\int_t^T \lambda_u^l \mathrm{d}u}(1-\gamma)^{N_T^c}\middle|\mathcal{F}_t\right]$:

**Proposition 2.**

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[\lambda_T^c \gamma e^{-\int_t^T \lambda_u^l \mathrm{d}u}(1-\gamma)^{N_T^c}\middle|\mathcal{F}_t\right] &= e^{\alpha(t)+\log(1-\gamma)N_t^c+\beta_3(t)\lambda_t^c+\beta_4(t)\lambda_t^l} \\ &\quad \times \left(A(t) + B_3(t)\lambda_t^c + B_4(t)\lambda_t^l\right),\end{aligned}$$

*with*

$$
\begin{aligned}
-\dot{A}(t) &= \alpha_c \lambda_\infty^c B_3(t) + \alpha_l \lambda_\infty^l B_4(t), \\
-\dot{B}_3(t) &= -\alpha_c B_3(t) + \beta_3(t)\sigma_c^2 B_3(t) \\
&\quad + \beta_{1,1} e^{\beta_1(t)+\beta_{1,1}\beta_3(t)+\beta_{2,1}\beta_4(t)} B_3(t) + \beta_{2,1} e^{\beta_1(t)+\beta_{1,1}\beta_3(t)+\beta_{2,1}\beta_4(t)} B_4(t), \\
-\dot{B}_4(t) &= -\alpha_l B_4(t) + \beta_4(t)\sigma_l^2 B_4(t) \\
&\quad + \beta_{1,2} e^{\beta_2(t)+\beta_{1,2}\beta_3(t)+\beta_{2,2}\beta_4(t)} B_3(t) + \beta_{2,2} e^{\beta_2(t)+\beta_{1,2}\beta_3(t)+\beta_{2,2}\beta_4(t)} B_4(t), \\
\alpha(T) &= \beta_2(T) = \beta_2(s) = \beta_3(T) = \beta_4(T) = A(T) = B_1(s) \\
&= B_2(s) = B_4(T) = 0, \quad t \le s \le T, \\
\beta_1(T) &= \log(1-\gamma) = \beta_1(s), \quad t \le s \le T, \\
B_3(T) &= \gamma,
\end{aligned}
$$

*where $\alpha(t)$ and $\beta(t)$ satisfy the ODEs presented in Proposition 1.*

*Proof.* Proposition 3 of Duffie et al. (2000) with
$u = (\log(1-\gamma), 0, 0, 0)^\top$ and $v = (0, 0, \gamma, 0)^\top$ yields the result. □

### IA.2. Price Correction Rating Portfolios

Since not all bonds within the same rating class are equally credit risky, we correct for bond-specific fixed effects which measure the differences in baseline credit-worthiness. A simple fixed effects correction would, however, also correct for coupon and maturity differences between bonds, because these differences are also constant over time. Our pricing formulas use maturity and coupon rate as input, and are thus able to deal with these differences without correcting for them. Therefore, we propose to correct prices for bond-specific fixed effects within groups obtained by double-sorting on maturity and coupon rates. That is,

we put bonds with similar maturity and coupon rates in the same bucket and apply a fixed effects correction per bucket. Since we update the rating portfolios every quarter, we apply the price corrections per rating class, per quarter. For every rating class and quarter, we create nine different buckets, consisting of different combinations of maturity and coupon rates (low, middle, high, based on lower, middle, and upper 33th percentiles), and apply the following fixed effects price correction per bucket:

Let the log-price of bond $i$ be given by

$$\log P_{i,t} = \alpha_i + f(\chi_i, X_t, \Theta) + \epsilon_{i,t}, \tag{IA.3}$$

where $\alpha_i$ is the bond-specific fixed effect, $f$ our price function depending on bond characteristics $\chi_i$ (maturity, coupon rate, etc.), the states $X_t$ (intensity values) at time $t$ and the parameters $\Theta$, and $\epsilon_{i,t}$ a normally i.i.d. measurement error. We have that the time series average of the log-price of bond $i$ is given by

$$\frac{1}{T}\sum_{t=1}^{T} \log P_{i,t} = \alpha_i + \frac{1}{T}\sum_{t=1}^{T} f(\chi_i, X_t, \Theta). \tag{IA.4}$$

Denoting $\frac{1}{T}\sum_{t=1}^{T} \log P_{i,t} = \overline{\log P_{i,t}}$ and $\frac{1}{T}\sum_{t=1}^{T} f(\chi_i, X_t, \Theta) = \overline{f(\chi_i, X_t, \Theta)}$, we then get

$$\frac{1}{N}\sum_{i=1}^{N} \overline{\log P_{i,t}} = \frac{1}{N}\sum_{i=1}^{N} \alpha_i + \overline{f(\chi_i, X_t, \Theta)} = \overline{f(\chi_i, X_t, \Theta)}, \tag{IA.5}$$

where $N$ is the number of bonds in the same maturity-coupon bucket. The last equality follows from $\sum_{i=1}^{N} \alpha_i = 0$ and the fact that we assume $\chi_i = \chi_j \; \forall i, j$ in this bucket. This gives the following estimator for $\alpha_i$:

$$\hat{\alpha_i} = \overline{\log P_{i,t}} - \frac{1}{N}\sum_{i=1}^{N} \overline{\log P_{i,t}}.$$

All-in-all, we get $\log P_{i,t} - \hat{\alpha}_i = \widetilde{\log P_{i,t}} = f(\chi_i, X_t, \Theta) + \epsilon_{i,t}$, and use these prices in estimating our model.

The above procedure sometimes produces an overly strong price correction when the price development of a bond in a certain quarter is not comparable to the other bonds in its same maturity-coupon bucket. For example, if a bond defaults or gets downgraded, its prices for the remainder of the quarter will typically be much lower than those of the average bond in the same bucket. Therefore, the fixed effects correction produces adapted prices that are much higher than its original price, since we subtract the relatively low bond's time series average and add the relatively high cross-sectional bucket average to the original transaction prices. In these cases, the price correction may be so large that we do not consider these transactions representative for that rating class and maturity-coupon bucket. Therefore, when the price correction is more than 50% away from the original price, we throw away these observations. This results in discarding 0, 140, 340, 630, and 2459 transactions for the AAA/AA, A, BBB, BB, and B and lower rated portfolios, respectively.

## IA.3. Details Metropolis-Hastings Steps

In this Appendix, we provide more details of the Metropolis-Hastings steps in our estimation methodology. We first consider the drawing of new states and next consider the drawing of new parameters.

*Metropolis Step for Simulating the States*

As explained in the main text, we alternate between two simulation schemes for the intensities and jump times. In the first scheme, we draw at every time $t$ the credit jump, the liquidity jump, and the intensities in consecutive steps. In the second scheme, we draw at

every time $t$ the whole state vector $X_t$ in one step. Below we will provide the details of the Metropolis steps used in both schemes.

*State Simulation Scheme 1*

In the first scheme, we take advantage of the fact that $p(X_t|X^{(g)}_{1:N_{\setminus t}}, \bar{\Theta}^{(g)}, \bar{Y})$ is again characterized by its full conditionals. Therefore, we can split drawing $X_t^{(g+1)}$ into the following three steps:

1. draw $N_t^{c(g+1)}$ from $p(N_t^c|\lambda_t^{c(g)}, \lambda_t^{l)(g)}, N_t^{l(g)}, X^{(g)}_{1:N_{\setminus t}}, \bar{\Theta}^{(g)}, \bar{Y})$;
2. draw $N_t^{l(g+1)}$ from $p(N_t^l|\lambda_t^{c(g)}, \lambda_t^{l)(g)}, N_t^{c(g+1)}, X^{(g)}_{1:N_{\setminus t}}, \bar{\Theta}^{(g)}, \bar{Y})$;
3. draw $\lambda_t^{c(g+1)}$ and $\lambda_t^{l(g+1)}$ from $p(\lambda_t^c, \lambda_t^l|N_t^{c(g+1)}, N_t^{l(g+1)}, X^{(g)}_{1:N_{\setminus t}}, \bar{\Theta}^{(g)}, \bar{Y})$.

We have that the full posterior of $N_t^c$ is a Bernoulli density with success probability

$$p(N_t^c = 1|\lambda_t^{c(g)}, \lambda_t^{l(g)}, N_t^{l(g)}, X^{(g)}_{1:N_{\setminus t}}, \bar{\Theta}^{(g)}, \bar{Y})$$
$$= \frac{p(\lambda_t^{c(g)}, \lambda_t^{l(g)}|N_t^c = 1, N_t^{l(g)}, X_{t-1}^{(g+1)}, \bar{\Theta}^{(g)})p(N_t^c = 1|X_{t-1}^{(g+1)})}{\sum_{s=0,1} p(\lambda_t^{c(g)}, \lambda_t^{l(g)}|N_t^c = s, N_t^{l(g)}, X_{t-1}^{(g+1)}, \bar{\Theta}^{(g)})p(N_t^c = s|X_{t-1}^{(g+1)})}, \quad \text{(IA.6)}$$

which is easy to compute, since $p(\lambda_t^{c(g)}, \lambda_t^{l(g)}|N_t^c = s, N_t^{l(g)}, X_{t-1}^{(g+1)}, \bar{\Theta}^{(g)})$ is a bivariate normal density and $p(N_t^c = s|X_{t-1}^{(g+1)})$ is Bernoulli with success probability $\lambda_{t-1}^{c(g+1)}\Delta_t$. The credit jump probability does not depend on $\bar{Y}$, since the bond prices only depend on the intensities and parameter vector, and the liquidity jump time indicator $I_t^l$ only depends on $N_t^l$ and the parameters and not on $N_t^c$.

Similarly, the full posterior of $N_{t_i}^l$ is Bernoulli with success probability given by

$$p(N_t^l = 1|\lambda_t^{c(g)}, \lambda_t^{l)(g)}, N_t^{c(g+1)}, X_{1:N_{\setminus t}}^{(g)}, \bar{\Theta}^{(g)}, \bar{Y})$$
$$= \frac{p(I_t^l|N_t^l = 1, \bar{\Theta})p(\lambda_t^{c(g)}, \lambda_t^{l(g)}|N_t^l = 1, N_t^{c(g+1)}, X_{t-1}^{(g+1)}, \bar{\Theta}^{(g)})p(N_t^l = 1|X_{t-1}^{(g+1)})}{\sum_{s=0,1} p(I_t^l|N_t^l = s, \bar{\Theta})p(\lambda_t^{c(g)}, \lambda_t^{l(g)}|N_t^l = s, N_t^{c(g+1)}, X_{t-1}^{(g+1)}, \bar{\Theta}^{(g)})p(N_t^l = s|X_{t-1}^{(g+1)})}, \tag{IA.7}$$

which is again easy to compute, since $p(\lambda_t^{c(g)}, \lambda_t^{l(g)}|N_t^l = s, N_t^{c(g+1)}, X_{t-1}^{(g+1)}, \Theta^{(g)})$ is bivariate normal, $p(I_t^l|N_t^l = s, \bar{\Theta})$ is Bernoulli with success probability $\phi_{high}$ if $N_t^l = 1$ and $\phi_{low}$ if $N_t^l = 0$, and $p(N_t^l = s|X_{t-1}^{(g+1)})$ is Bernoulli with success probability $\lambda_{t-1}^{l(g+1)}\Delta_t$. Note that the posterior liquidity jump probability depends on $\bar{Y}$, since $N_t^l$ appears in the density of $I_t^l$.

After having drawn $N_t^{c(g+1)}$ and $N_t^{l(g+1)}$, we draw the new intensities $\lambda_t^{c(g+1)}$ and $\lambda_t^{l(g+1)}$ from the density

$$p(\lambda_t^c, \lambda_t^l|N_t^{c(g+1)}, N_t^{l(g+1)}, X_{1:N_{\setminus t}}^{(g)}, \bar{\Theta}^{(g)}, \bar{Y})$$
$$\propto p(Y_t|\lambda_t^c, \lambda_t^l, N_t^{c(g+1)}, N_t^{l(g+1)}, \bar{\Theta}^{(g)})p(\lambda_{t+1}^{c(g)}, \lambda_{ti+1}^{l(g)}|\lambda_t^c, \lambda_t^l, N_{t+1}^{c(g)}, N_{t+1}^{l(g)}, \bar{\Theta}^{(g)})$$
$$\times p(\lambda_t^c, \lambda_t^l|\lambda_{t-1}^{c(g+1)}, \lambda_{t-1}^{l(g+1)}, N_t^{c(g+1)}, N_t^{l(g+1)}, \bar{\Theta}^{(g)}),$$

which is the product of a multivariate normal density and two bivariate normal densities. In this density, $I^l$ and $\phi_{high}$ and $\phi_{low}$ do not play a role, since they do not depend on the intensity values. This density is non-standard and therefore we use a Metropolis step with proposal density $p(\lambda_t^c, \lambda_t^l|\lambda_{t-1}^{c(g+1)}, \lambda_{t-1}^{l(g+1)}, N_t^{c(g+1)}, N_t^{l(g+1)}, \bar{\Theta}^{(g)})$, which is a bivariate normal distribution with mean

$$\begin{pmatrix} \lambda_{t-1}^{c(g+1)} + \alpha_c^{(g)}\lambda_\infty^{c(g)}\Delta_t - \alpha_c^{(g)}\lambda_{t-1}^{c(g+1)}\Delta_t + \beta_{1,1}^{(g)}N_t^{c(g+1)} + \beta_{1,2}^{(g)}N_t^{l(g+1)} \\ \lambda_{t-1}^{l(g+1)} + \alpha_l^{(g)}\lambda_\infty^{l(g)}\Delta_t - \alpha_l^{(g)}\lambda_{t-1}^{l(g+1)}\Delta_t + \beta_{2,1}^{(g)}N_t^{c(g+1)} + \beta_{2,2}^{(g)}N_t^{l(g+1)} \end{pmatrix} \tag{IA.8}$$

and variance matrix

$$\begin{pmatrix} \sigma_c^{2(g)} \lambda_{t-1}^{c(g+1)} \Delta_t & 0 \\ 0 & \sigma_l^{2(g)} \lambda_{t-1}^{l(g+1)} \Delta_t \end{pmatrix}. \tag{IA.9}$$

Using this proposal density, the acceptance criterion becomes

$$\min \left( \frac{p(Y_t|X_t^{(g+1)}, \bar{\Theta}^{(g)}) p(\lambda_{t+1}^{c(g)}, \lambda_{t+1}^{l(g)}|N_{t+1}^{c(g)}, N_{t+1}^{l(g)}, X_t^{(g+1)}, \bar{\Theta}^{(g)})}{p(Y_t|X_t^{(g)}, \bar{\Theta}^{(g)}) p(\lambda_{t+1}^{c(g)}, \lambda_{t+1}^{l(g)}|N_{t+1}^{c(g)}, N_{t+1}^{l(g)}, X_t^{(g)}, \bar{\Theta}^{(g)})}, 1 \right).$$

For the end point $t = N$ the acceptance criterion simplifies, since the terms $p(\lambda_{t+1}^{c(g)}, \lambda_{t+1}^{l(g)}|N_{t+1}^{c(g)}, N_{t+1}^{l(g)}, X_t^{(g+1)}, \bar{\Theta}^{(g)})$ and $p(\lambda_{t+1}^{c(g)}, \lambda_{t+1}^{l(g)}|N_{t+1}^{c(g)}, N_{t+1}^{l(g)}, X_t^{(g)}, \bar{\Theta}^{(g)})$ do not appear anymore in the numerator and denominator, respectively. For the starting point $t = 1$, we use a slightly different proposal density, since we cannot condition on $X_{t-1}$. We therefore draw $\lambda_1^{c(g+1)}$ and $\lambda_1^{l(g+1)}$ from a bivariate normal density with mean vector $\left(\lambda_2^{c(g)}, \lambda_2^{l(g)}\right)^\top$ and variances $\sigma\sqrt{\lambda_2^{c(g)}\Delta_2}$ and $\sigma\sqrt{\lambda_2^{l(g)}\Delta_2}$, respectively. Denoting this proposal density as $q(\lambda_1^{c(g+1)}, \lambda_1^{l(g+1)}|\lambda_2^{c(g)}, \lambda_2^{l(g)}, \Theta)$ gives the following acceptance criterion:

$$\min \left( \frac{p(Y_1|X_1^{(g+1)}, \bar{\Theta}^{(g)}) p(\lambda_2^{c(g)}, \lambda_2^{l(g)}|N_2^{c(g)}, N_2^{l(g)}, X_1^{(g+1)}, \bar{\Theta}^{(g)}) q(\lambda_1^{c(g)}, \lambda_1^{l(g)}|\lambda_2^{c(g)}, \lambda_2^{l(g)}, \bar{\Theta}^{(g)})}{p(Y_1|X_1^{(g)}, \bar{\Theta}^{(g)}) p(\lambda_2^{c(g)}, \lambda_2^{l(g)}|N_2^{c(g)}, N_2^{l(g)} X_1^{(g)}, \bar{\Theta}^{(g)}) q(\lambda_1^{c(g+1)}, \lambda_1^{l(g+1)}|\lambda_2^{c(g)}, \lambda_2^{l(g)}, \bar{\Theta}^{(g)})}, 1 \right).$$

*State Simulation Scheme 2*

In the second simulation scheme we sample the complete vector $X_t$ at once from $p(X_t|X_{1:N_{\setminus t}}, \bar{\Theta}, \bar{Y})$. From (10) we have

$$p(X_t|X_{1:N_{\setminus t}}^{(g-1)}, \bar{\Theta}, \bar{Y}) \propto p(I_t^l|X_t, \bar{\Theta}) p(Y_t|X_t, \bar{\Theta}) p(X_t|X_{t-1}^{(g)}, \bar{\Theta}) p(X_{t+1}^{(g-1)}|X_t, \bar{\Theta}), \tag{IA.10}$$

where we exploit the Markovian property of the state processes, and, hence, only need to consider the terms in (10) where (parts of) $X_t$ enters directly. We observe that in this density, both the likelihoods as well as the transition densities play a role and this allows us

to circumvent the problems encountered in the sampling procedure above.

We showed in Equation (11) that the transition density can be written as

$$\begin{aligned} p(X_t|X_{t-1},\bar{\Theta}) &= p(\lambda_t^c,\lambda_t^l|N_t^c,N_t^l,X_{t-1},\bar{\Theta})p(N_t^c,N_t^l|X_{t-1},\bar{\Theta}) \\ &= p(\lambda_t^c,\lambda_t^l|N_t^c,N_t^l,X_{t-1},\bar{\Theta})p(N_t^c|X_{t-1},\bar{\Theta})p(N_t^l|X_{t-1},\bar{\Theta}). \end{aligned} \tag{IA.11}$$

We have that $p(Y_t|X_t,\bar{\Theta})$ is multivariate normal with dimension equal to the number of observations at time $t$. More specifically, let there be $k$ observations at time $t$. Then we have that

$$p(Y_t|X_t,\bar{\Theta}) \sim \mathcal{N}\left(\begin{pmatrix} F(X_t,\chi_1,\Theta) \\ F(X_t,\chi_2,\Theta) \\ \vdots \\ F(X_t,\chi_k,\Theta) \end{pmatrix}, \begin{pmatrix} h^2 & 0 & 0 & \cdots \\ 0 & h^2 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h^2 \end{pmatrix}\right).$$

Furthermore, $p(I_t^l|N_t^l,\bar{\Theta})$ is Bernoulli with success probability $\phi_{high}$ if $N_t^l = 1$ and $\phi_{low}$ if $N_t^l = 0$.

In order to draw $X_t^{(g+1)}$ from (IA.10), we use the following proposal density

$$\begin{aligned} q(X_t|X_{1:N_{\setminus t}}^{(g)},\bar{\Theta}^{(g)}) &= p(X_t|X_{t-1}^{(g+1)},\bar{\Theta}^{(g)}) \\ &= p(\lambda_t^c,\lambda_t^l|N_t^c,N_t^l,X_{t-1}^{(g+1)},\bar{\Theta}^{(g)})p(N_t^c|X_{t-1}^{(g+1)},\bar{\Theta}^{(g)})p(N_t^l|X_{t-1}^{(g+1)},\bar{\Theta}^{(g)}). \end{aligned}$$

From the discretization of the intensity processes, we get that this is a mixture of bivariate normal distributions. To see this, note that $p(\lambda_t^c,\lambda_t^l|N_t^c,N_t^l,X_{t-1}^{(g+1)},\bar{\Theta}^{(g)})$ is bivariate normal with mean

$$\begin{pmatrix} \lambda_{t-1}^{c(g+1)} + \alpha_c^{(g)}\lambda_\infty^{c(g)}\Delta_t - \alpha_c^{(g)}\lambda_{t-1}^{c(g+1)}\Delta_t + \beta_{1,1}^{(g)}N_t^c + \beta_{1,2}^{(g)}N_t^l \\ \lambda_{t-1}^{l(g+1)} + \alpha_l^{(g)}\lambda_\infty^{l(g)}\Delta_t - \alpha_l^{(g)}\lambda_{t-1}^{l(g+1)}\Delta_t + \beta_{2,1}^{(g)}N_t^c + \beta_{2,2}^{(g)}N_t^l \end{pmatrix}$$

and variance matrix

$$\begin{pmatrix} \sigma_c^{2(g)}\lambda_{t-1}^{c(g+1)}\Delta_t & 0 \\ 0 & \sigma_l^{2(g)}\lambda_{t-1}^{l(g+1)}\Delta_t \end{pmatrix},$$

Now, depending on the four possible combinations of $N_t^c$ and $N_t^l$ (i.e., these are both 1, both 0 or one of them is 1 and the other 0), we have that $p(\lambda_t^c, \lambda_t^l | N_t^c, N_t^l, X_{t_{i-1}}^{(g+1)}, \bar{\Theta}^{(g)}) p(N_t^c | X_{t-1}^{(g+1)}, \bar{\Theta}^{(g)}) p(N_t^l | X_{t-1}^{(g+1)}, \bar{\Theta}^{(g)})$ is a mixture of 4 (bivariate) normal distributions.

When we draw from this distribution, we can first draw $N_t^{c(g+1)}$ and $N_t^{l(g+1)}$ from independent Bernoulli distributions with success probabilities $\lambda_{t-1}^{c(g+1)}\Delta_t$ and $\lambda_{t-1}^{l(g+1)}\Delta_t$, respectively, and then, given their outcomes, draw $p(\lambda_t^c, \lambda_t^l | N_t^{c(g+1)}, N_t^{l(g+1)}, X_{t-1}^{(g+1)}, \Theta^{(g)})$ from a normal distribution with the appropriate mean (depending on outcomes of the draws of $N_t^{c(g+1)}$ and $N_t^{l(g+1)}$).

So let $X_t^{(g+1)} = \left\{ \lambda_t^{c(g+1)}, \lambda_t^{l(g+1)}, N_t^{c(g+1)}, N_t^{l(g+1)} \right\}$ denote the new draw of the intensities and jump processes obtained from $p(\lambda_t^{c(g+1)}, \lambda_t^{l(g+1)} | N_t^{c(g+1)}, N_t^{l(g+1)}, X_{t-1}^{(g+1)}, \Theta^{(g)})$ $\times p(N_t^{c(g+1)} | X_{t-1}^{(g+1)}, \Theta^{(g)}) p(N_t^{l(g+1)} | X_{t-1}^{(g+1)}, \Theta^{(g)})$. Then we have that the acceptance criterion simplifies, since the proposal density is a part of the true density and therefore multiple terms cancel out. More specifically, we will have the following acceptance criterion:

$$\min \left( \frac{p(I_t^l | N_t^{l(g+1)}, \bar{\Theta}) p(Y_t | X_t^{(g+1)}, \Theta^{(g)}) p(\lambda_{t+1}^{c(g)}, \lambda_{t+1}^{l(g)}, N_{t+1}^{c(g)}, N_{t+1}^{l(g)} | X_t^{(g+1)}, \Theta^{(g)})}{p(I_t^l | N_t^{l(g)}, \bar{\Theta}) p(Y_t | X_t^{(g)}, \Theta^{(g)}) p(\lambda_{t+1}^{c(g)}, \lambda_{t+1}^{l(g)}, N_{t+1}^{c(g)}, N_{t+1}^{l(g)} | X_t^{(g)}, \Theta^{(g)})}, 1 \right),$$

where $p(\lambda_{t+1}^{c(g)}, \lambda_{t+1}^{l(g)}, N_{t+1}^{c(g)}, N_{t+1}^{l(g)} | X_t^{(g+1)}, \bar{\Theta}^{(g)}) = p(\lambda_{t+1}^{c(g)}, \lambda_{t+1}^{l(g)} | N_{t+1}^{c(g)}, N_{t+1}^{l(g)}, X_t^{(g+1)}, \bar{\Theta}^{(g)})$ $\times p(N_{t+1}^{c(g)} | X_t^{(g+1)}, \bar{\Theta}^{(g)}) p(N_{t+1}^{l(g)} | X_t^{(g+1)}, \bar{\Theta}^{(g)})$. Since all densities are standard (i.e., multivariate normal, bivariate normal or bernoulli), we can easily evaluate this acceptance criterion. For the end points $t = 1$ and $t = N$ similar comments apply as in simulation scheme 1.

*Metropolis Step for Simulating the Parameters*

We next explain the Metropolis step for estimating the parameters in more detail.

Where possible, we update the parameters using a Gibbs step with conjugate priors. Let us first consider $\phi_{high}$ and $\phi_{low}$, for which we specify the priors to be $Beta(a_{high}, b_{high})$ and $Beta(a_{low}, b_{low})$, respectively. Let $S(N^{l(g+1)} = 1)$ denote the days for which $N_t^{l(g+1)} = 1$. Then we use standard results that a Beta prior on a Bernoulli success probability, where the Bernoulli trial is repeated several times, is again a Beta distribution with updated parameters. We get $\phi_{high}^{(g+1)} \sim Beta(a_{high}^{posterior}, b_{high}^{posterior})$ with $a_{high}^{posterior} = a_{high} + \sum_{t \in S(N^{l(g+1)}=1)} I_t^l$ and $b_{high}^{posterior} = b_{high} + \sum_{t=1}^{N} N_t^{l(g+1)} - \sum_{t \in S(N^{l(g+1)}=1)} I_t^l$, and $\phi_{low}^{(g+1)} \sim Beta(a_{low}^{posterior}, b_{low}^{posterior})$ with $a_{low}^{posterior} = a_{low} + \sum_{t \notin S(N^{l(g+1)}=1)} I_t^l$ and $b_{low}^{posterior} = b_{low} + \left(N - \sum_{t=1}^{N} N_t^{l(g+1)}\right) - \sum_{t \notin S(N^{l(g+1)}=1)} I_t^l$.

For $h^2$ we use an inverse gamma prior specification. Since $h^2$ only appears as variance in the measurement equations for the bonds, we get that the posterior distribution for $h^2$ is again inverse gamma with updated parameters. This follows from $p(h^2|\bar{\Theta}\backslash\{h^2\}, X, \bar{Y}) \propto p(Y|\bar{\Theta}, X)p(h^2)$, where $p(Y|\bar{\Theta}, X)$ is multivariate normal with diagonal variance matrix with $h^2$ as variance and $p(h^2)$ the prior inverse gamma density. Explicit computations are standard and are therefore omitted.

All the other parameters appear in both the transition densities as well as the bond price formulas and therefore it is not possible to specify conjugate prior specifications for them. We use Metropolis steps to update these parameters in multiple blocks, since this is computationally less demanding than updating one parameter at the time. Following Eraker (2004), we propose to use a Gaussian proposal density with as mean vector the current draw of the parameters in the block, and with a diagonal covariance matrix.

We choose the priors on all parameters to be proper but uninformative in the sense that the prior variances should be high compared to the estimated posterior variances. We use the

following priors:

- $\alpha_c$, $\alpha_l$, $\lambda_\infty^c$, $\lambda_\infty^l$, $\sigma_c^2$, and $\sigma_l^2$: Gamma(0.02,10)
- $\beta_{1,1}$, $\beta_{1,2}$, $\beta_{2,1}$, and $\beta_{2,2}$: Normal(0.02,10)
- $\gamma$: Uniform(0,1)
- $h^2$: $IG(100, 1)$
- $\phi_{high}$: $Beta(5, 5)$
- $\phi_{low}$: $Beta(1, 10)$

In line with theoretical parameter restrictions for nonnegativity and stationarity of the processes, we impose some parameters to be nonnegative by using gamma priors. For $\beta_{i,j}$ $(i, j = 1, 2)$ we take, however, normally distributed priors. This choice does not restrict them from taking negative values, although we can easily assure their positivity by rejecting all parameter draws for which one of these parameters becomes negative. The reason for choosing normal priors is that in order to compute the Bayes factors for testing model restrictions, we need that the priors are well-defined for zero values. Since $\gamma$ is the probability of going into default in the case a credit shock arrives, we put a Uniform(0,1) prior on it. In general, the means and variances of all parameters are chosen arbitrarily, but such that the means are small and positive for all parameters and the variances relatively large compared to their means. All-in-all, results are robust against prior specification, since typically the likelihood contribution of the priors is small compared to the likelihood contribution of the data in the acceptance criteria. Furthermore, the posterior standard deviations are also much smaller than the prior standard deviations, indicating that our priors do not impose much information.

### IA.4. Bayes Factors

In order to test whether the model should contain self- and cross-excitation, we use Bayes factors to compare the full model with different models in which we put restrictions on the self- and cross-excitation parameters. We let $\mathcal{M}_1$ denote the full model, $\mathcal{M}_2$ the model with restrictions $\beta_{1,1} = \beta_{1,2} = \beta_{2,1} = \beta_{2,2} = 0$ (no excitation), $\mathcal{M}_3$ the model with restrictions $\beta_{1,2} = \beta_{2,1} = 0$ (no cross-excitation), and $\mathcal{M}_4$ the model with restrictions $\beta_{1,1} = \beta_{2,2} = 0$ (no self-excitation). The Bayes factors are given by the ratios of the two marginal likelihoods of the models under consideration. That is, if we compare model $\mathcal{M}_i$ $(i = 2, 3, 4)$ with model $\mathcal{M}_1$, we get

$$BF_{\mathcal{M}_i|\mathcal{M}_1} = \frac{p_{\mathcal{M}_i}(Y)}{p_{\mathcal{M}_1}(Y)}. \tag{IA.12}$$

Let us denote the parameter vector of the full model by $\Theta = \{\theta_1, \theta_2\}$, where $\theta_2$ are the parameters that are restricted in model $\mathcal{M}_i$. For example, when comparing model $\mathcal{M}_2$ with the full model, we have that $\theta_2 = \{\beta_{1,1}, \beta_{1,2}, \beta_{2,1}, \beta_{2,2}\}$ and the restriction is that $\theta_2 = \theta_2^*$, with $\theta_2^* = \{0, 0, 0, 0\}$. Since models $\mathcal{M}_2$, $\mathcal{M}_3$, and $\mathcal{M}_4$ are all nested in model $\mathcal{M}_1$, and since the priors on the common parameters, $\theta_1$, are the same and independent of the restricted parameters, $\theta_2$, we can use Savage-Dickey density ratios to compute the Bayes factors. We get

$$BF_{\mathcal{M}_i|\mathcal{M}_1} = \frac{p_{\mathcal{M}_1}(\theta_2|Y)|_{\theta_2=\theta_2^*}}{p_{\mathcal{M}_1}(\theta_2)|_{\theta_2=\theta_2^*}}, \tag{IA.13}$$

that is, the Bayes factor can be computed as the ratio of model $\mathcal{M}_1$'s marginal posterior density of $\theta_2$ and marginal prior density of $\theta_2$, both evaluated in $\theta_2^*$. The computation of the denominator is simple, as we just have to compute the prior marginal density of the full model evaluated in $\theta_2^*$.

In order to compute the numerator of (IA.13), we follow Chib and Jeliazkov (2001), and we

denote the subkernel of the Metropolis-Hastings chain for $\theta_2$ conditional on $(\theta_1, X)$ by

$$p(\theta_2, \theta_2^*|Y, \theta_1, X) = \alpha(\theta_2, \theta_2^*|Y, \theta_1, X)q(\theta_2, \theta_2^*|Y, \theta_1, X),$$

where $q(\theta_2, \theta_2'|Y, \theta_1, X)$ is the proposal density for the transition from $\theta_2$ to $\theta_2'$ and

$$\alpha(\theta_2, \theta_2'|Y, \theta_1, X) = \min\left(\frac{p(\theta_2'|Y, \theta_1, X)p(\theta_1, \theta_2')q(\theta_2', \theta_2|Y, \theta_1, X)}{p(\theta_2|Y, \theta_1, X)p(\theta_1, \theta_2)q(\theta_2, \theta_2'|Y, \theta_1, X)}, 1\right)$$

denotes the probability of a move. The following (local) reversibility condition holds (see Chib and Jeliazkov, 2001):

$$p(\theta_2, \theta_2^*|Y, \theta_1, X)p(\theta_2|Y, \theta_1, X) = p(\theta_2^*|Y, \theta_1, X)p(\theta_2^*, \theta_2|Y, \theta_1, X). \tag{IA.14}$$

Multiplying both sides of (IA.14) by $p(\theta_1, X|Y)$ and integrating over $\psi = (\theta_1, \theta_2, X)$ gives

$$\int p(\theta_2, \theta_2^*|Y, \theta_1, X)p(\theta_2|Y, \theta_1, X)p(\theta_1, X|Y)\mathrm{d}\psi \\ = \int p(\theta_2^*|Y, \theta_1, X)p(\theta_2^*, \theta_2|Y, \theta_1, X)p(\theta_1, X|Y)\mathrm{d}\psi,$$

or

$$\int p(\theta_2, \theta_2^*|Y, \theta_1, X)p(\theta_1, \theta_2, X|Y)\mathrm{d}\psi = \int \pi(\theta_2^*|Y)p(\theta_1, X|Y, \theta_2^*)p(\theta_2^*, \theta_2|Y, \theta_1, X)\mathrm{d}\psi.$$

From this last equation and the definitions of $p(\theta_2, \theta_2^*|Y, \theta_1, X)$ and $p(\theta_2^*, \theta_2|Y, \theta_1, X)$ it follows that

$$p(\theta_2^*|Y) = \frac{\mathbb{E}_1\left[\alpha(\theta_2, \theta_2^*|Y, \theta_1, X)q(\theta_2, \theta_2^*|Y, \theta_1, X)\right]}{\mathbb{E}_2\left[\alpha(\theta_2^*, \theta_2|Y, \theta_1, X)\right]}, \tag{IA.15}$$

where the expectation in the numerator is with respect to the distribution $p(\theta_1, \theta_2, X|Y)$,

and the expectation in the denominator with respect to the distribution $p(\theta_1, X|Y, \theta_2^*)q(\theta_2^*, \theta_2|Y, \theta_1, X)$.

The MCMC algorithm we use to estimate our model provides draws $\{\theta_1^{(g)}, \theta_2^{(g)}, X^{(g)}\}_{g=1}^G$ from the full posterior density $p(\theta_1, \theta_2, X|Y)$. Taking these draws and averaging $\alpha(\theta_2, \theta_2^*|Y, \theta_1, X)q(\theta_2, \theta_2^*|Y, \theta_1, X)$ over these draws thus provides an estimate of the numerator of (IA.15). For the denominator, we fix $\theta_2^*$ and continue the MCMC simulation for an additional $J$ iterations with the two full conditional densities $p(\theta_1|Y, \theta_2^*, X)$ and $p(X|Y, \theta_1, \theta_2^*)$. At each iteration we get values $(\theta_1^{(j)}, X^{(j)})$ and we use these to generate $\theta_2^{(j)} \sim q(\theta_2^*, \theta_2|Y, \theta_1^{(j)}, X^{(j)})$. For every iteration we obtain a draw $(\theta_1^{(j)}, \theta_2^{(j)}, X^{(j)})$ from the distribution $p(\theta_1, X|Y, \theta_2^*)q(\theta_2^*, \theta_2|Y, \theta_1, X)$. We can now estimate $p(\theta_2^*|Y)$ as

$$\hat{p}(\theta_2^*|Y) = \frac{G^{-1}\sum_{g=1}^G \alpha(\theta_2^{(g)}, \theta_2^*|Y, \theta_1^{(g)}, X^{(g)})q(\theta_2^{(g)}, \theta_2^*|Y, \theta_1^{(g)}, X^{(g)})}{J^{-1}\sum_{j=1}^J \alpha(\theta_2^*, \theta_2^{(j)}|Y, \theta_1^{(j)}, X^{(j)})}. \tag{IA.16}$$