The Valuation of Corporate Coupon Bonds[1]

Jens Hilscher[2], Robert A. Jarrow[3], and Donald R. van Deventer[4]

Abstract

This paper proposes and estimates a tractable, arbitrage-free valuation model for corporate coupon bonds that includes a more realistic recovery rate process. Most existing studies use a recovery rate process that is misspecified because it includes recovery for coupons due after default. Misspecification errors from assuming recovery on all coupons can be substantial; they increase with recovery rates, coupons, maturity, and default probabilities. For a large sample of market transactions: (i) our model has lower pricing errors than one assuming recovery on all coupons, and (ii) the magnitude of our model's outperformance is linked to misspecification errors from assuming recovery on coupons.

[1]Helpful comments from the editor (Hendrik Bessembinder), two anonymous referees, John Y. Campbell and seminar participants at the Bank of England, the Bank of Finland, the Board of Governors of the Federal Reserve, the European Bank for Reconstruction and Development, the International Monetary Fund, the World Bank, Brandeis University, the 2019 GEA meetings at the Frankfurt School of Finance and Management, and the University of Technology, Sydney are gratefully acknowledged. We thank Tengda Gong for providing excellent research assistance.

[2]University of California, Davis, California 95616. email: jhilscher@ucdavis.edu.

[3]Samuel Curtis Johnson Graduate School of Management, Cornell University, Ithaca, N.Y. 14853. email: raj15@cornell.edu.

[4]SAS Institute Inc, Honolulu, Hawaii 96815. email: donald.vandeventer@sas.com.

stream of the literature prices bonds or related securities using a reduced form model (see Duffee (1999), Duffie, Pedersen, and Singleton (2003), Driessen (2005), Bakshi, Madan, and Zhang (2006)). A careful reading of these papers shows that they all explicitly or implicitly assume that a single credit spread or spread term structure can be used to value risky debt.

The underlying assumption is that a coupon bond is equivalent to a portfolio of risky zero-coupon bonds that can be valued using a single spread or spread term structure. The number of zero-coupon bonds in the portfolio corresponds to the promised coupons and principal with their maturities equal to the payment dates (see expression (3) in the text). Importantly, both promised coupons and principal are discounted using the same spread. For the credit spread estimation literature, this implicit assumption follows because all promised coupons and principal are included when computing a bond's credit spread. In the reduced form model literature, the recovery rate process utilized is the "recovery of market value (RMV)" due to Lando (1998) and Duffie and Singleton (1999), which implies this result. This pricing approach assumes that, when discounting, coupon and principal cash flows are treated the same, and, therefore, that both promised payments entitle the holder to a recovery in default. For subsequent discussion, we call this the "full-coupon recovery" model.

As shown by Jarrow (2004), a single term structure of risky zero-coupon bonds used for valuing coupon bonds is valid if and only if all of the risky zero-coupon bonds are of equal seniority and all have the same recovery rate in the event of default. However, this assumption is inconsistent with industry practice. After default, as evidenced by financial restructurings and default proceedings, only the bond's principal becomes due, and no additional coupon payments are made on or after the default date. This implies that coupon and principal payments cannot be valued using the same (single) credit spread or spread term structure and that basing a bond valuation model on this erroneous assumption of equal seniority will generate model prices with misspecification errors.

Industry practice has been confirmed in the recovery rate estimation literature which finds that alternative recovery rate processes,[2] either the "recovery of face value (RFV)" or the "recovery of Treasuries (RTV)" formulations, provide a better approximation to realized recovery rates than does RMV (Guha and Sbuelz (2005), Guo, Jarrow and Lin (2008), Bakshi, Gao, and Zhong (2022)).[3] And, it is well known that both the RFV and RTV recovery rate processes are consistent with a zero recovery on coupons promised after default. Therefore, these recovery rate processes do not imply the full-coupon recovery model. See Jarrow and Turnbull (2000), Longstaff, Mithal, and Neis (2005), Bielecki and Rutkowski (2002), chapter 13, Collin-Dufresne and Goldstein (2001), and Huang and Huang (2012) for models with zero recovery on coupons promised after default.[4]

The purpose of this paper is to explore, both theoretically and empirically, the effect on bond prices of assuming zero recovery on coupons after default. We refer to such a "no-coupon recovery" model to differentiate it from the "full-coupon recovery" model. We derive an intuitive and straightforward-to-implement no-coupon recovery pricing model that depends only on the risk-free term structure, risk neutral default probabilities,[5] recovery rate, and an illiquidity parameter. We also present a clear and easy-to-calculate measure, the misspecification error, that identifies the effect of using a misspecified full-coupon recovery rate assumption to price bonds. These misspecification errors are due to the full-coupon recovery model's erroneous assumption of positive recovery for coupons after default.

We show theoretically that these misspecification errors are larger if recovery rates, de-

[2]See Bielecki and Rutkowski (2002), Chapter 8 for a discussion of these different recovery rate processes.

[3]See also Guha, Sbuelz and Tarelli (2020), who provide evidence in support of RFV when studying high-yield bond duration.

[4]Huang and Huang (2012) propose a model with no recovery on coupons and a constant recovery rate. Bakshi, Madan and Zhang (2006) use the Lehman Bond price data set to compare different recovery assumptions for a sample of 25 BBB-rated bonds over a nine-year period. They find that pricing errors decline when choosing the RTV or RFV rather than RMV specification. Our paper uses a much larger data set, explores the drivers of model misspecification errors on pricing both theoretically and empirically, and estimates the effect of illiquidity on prices. We also provide direct evidence of prices reflecting no recovery on coupons by looking at prices of bonds immediately after default.

[5]The use of risk neutral default probabilities is essential because we are creating valuation formulas, which require a risk premium. Later, under an additional assumption that default risk is diversifiable, risk neutral and actual probabilities are empirically equivalent.

fault probabilities, maturity, or coupon payments are larger. For example, given a 10-year bond with a face value of \$100, a recovery rate of 50%, a coupon of 2.51%, and an annual default probability of 1%, the full-coupon recovery model will assign a price that is \$0.50 too large. If it is a 30-year bond, the price error is \$4.33, a substantial difference relative to the correct price, which is equal to par in both cases. We calculate exact misspecification errors and also provide an approximate formula that can be used to estimate the misspecification error's magnitudes. In this approximation, misspecification errors are proportional to the recovery rate, the coupon size, the default probability, and the square of the number of coupon payments - which is closely related to maturity. Finally, we provide a comprehensive analysis of the empirical implications of the different pricing models for a large data set of coupon bond transaction prices.

Before this analysis, we present direct evidence of the different payment seniority between principal and coupons. We provide three examples of issuers that have filed for bankruptcy: Lehman Brothers, Pacific Gas and Electric (PG&E), and Weatherford International. We use both the full-coupon recovery and no-coupon recovery models to price the bonds. We find that pricing errors from using the misspecified full-coupon recovery model are between five to ten times larger than the no-coupon recovery model's pricing errors. This evidence is consistent with market prices reflecting zero recovery on coupons promised after default.

Our main empirical investigation performs a comparative analysis of the no-coupon and the full-coupon recovery models using a sample consisting of daily market prices for a collection of liquidly traded bonds from September 2017 - August 2022. This sample contains close to 168 thousand bond price observations. We separately fit both models. If market prices reflect zero recovery for coupons promised after default, the no-coupon recovery model will outperform the full-coupon recovery model. To test this hypothesis, we compute the average outperformance. However, this comparison is less informative if done in isolation. The reason is that our model predicts that the outperformance's magnitude is directly related to the size of the misspecification error – the pricing error from assuming recovery for coupons promised after default. And, a small average outperformance may simply result from a sample in which these misspecification errors are small.

Instead, a more relevant test is whether the misspecification error can explain the *variation* in the magnitude of the no-coupon recovery model's outperformance. If a bond has short maturity with only a few coupons and a small default probability, the two models will predict nearly the same price (the misspecification error is close to zero) and the no-coupon recovery model will outperform only slightly. But, if the maturity is long and the default probability is substantial (and zero recovery of coupons promised after default is reflected in the data) the no-coupon recovery model's outperformance will be large.

Given these insights, our empirical investigation proceeds in two steps. First, we calculate the misspecification errors from assuming recovery for coupon payments after default. Second, we study the performance of both the no-coupon and full-coupon recovery models separately, and analyze whether any outperformance of the no-coupon recovery model depends on the misspecification errors. In this empirical investigation, we fit both models to data obtaining prices, and then compare pricing errors between the two models.

The evidence from the first step shows that the misspecification errors are often quite large (in our data, the 95th percentiles are between \$0.97 and \$2.18 per \$100 face value). However, the median misspecification error is small and between 5 - 13 cents. Thus, although the no-coupon recovery model outperforms the full-coupon recovery model, for some bonds the difference is highly relevant, while for other bonds it is not.

The second step amounts to a horse race between the models, but one that not only tests average outperformance but also tests whether our model's predictions regarding relative outperformance is consistent with the data. We find evidence of the no-coupon recovery model's outperformance in the full sample. More importantly, we show that the no-coupon model's outperformance is larger when the default probability, the recovery rate, the maturity, and the coupons are larger. Thus our approach accurately forecasts when zero recovery of coupons promised after default is important for pricing.

We find that the no-coupon recovery model outperformance is robust to different model implementation choices. We estimate two versions of the two models. Model 1 assumes a fixed recovery rate and no illiquidity effect. The single free parameter is the default probability, which we estimate implicitly. This model has the benefit of being stable and not requiring any additional data apart from bond prices, characteristics, and Treasury rates. In Model 2 we also implicitly estimate the recovery rate.

We fit the model at the issuer-day level, and price a collection of bonds using both models. This allows the full-coupon recovery model to adjust its parameters. Therefore, what matters when comparing model fit is not the average level of the misspecification error. Indeed, if it were the same, biased inputs could result in a low pricing error. Instead, the within issuer-day misspecification error's standard deviation is what is relevant. When it is large, the misspecified model will have difficulty adjusting and is more likely to underperform. We find exactly this pattern in the data. For example, when focusing attention on the top quartile of misspecification error standard deviation observations, the pricing error difference increases from 7.2 to 22.4 cents (model 1).

One implication of our results is that default probability and recovery rate have distinct effects on the bond's price. As a result, it is possible to back out implied recovery rates from observed bond prices, something that is not possible in the full -coupon recovery model.[6] In terms of spreads, recovery rate and default probability have different impacts on principal and coupon-specific spreads (coupon spreads are unaffected by changes in the recovery rate because they have zero recovery). We exploit this pattern in model 3, where we use an external estimate of the default probability and fit a parameter for illiquidity.

The outline of the paper is as follows. Section 2 presents the model for valuing risky

[6]Reflecting this implication of the pricing model, we see instability in fitted recovery rates for the full-coupon recovery model.

coupon bonds. Section 3 quantifies how bond-specific characteristics affect full-coupon model misspecification errors. Section 4 discusses the data and model estimation procedures, while Section 5 presents some illustrative pricing results for three companies that filed for bankruptcy. Section 6 presents a comparative analysis of the two alternative pricing models, discusses variation in model fit and parameters over time, and presents out-of-sample model performance statistics. Section 7 concludes.

II. The Pricing Model

This section presents the pricing model, which is based on the reduced form model of Jarrow and Turnbull (1995). We assume that traded in the economy are default-free zero-coupon bonds of all maturities, a default-free money market account, and a risky coupon bond (to be described later). The market is assumed to be frictionless and competitive. Both the frictionless and competitive market assumptions are relaxed, subsequently, when we add an illiquidity discount to the valuation formula (see expression (4) below).

The default-free money market account earns interest continuously at the default-free spot rate of interest, r_t. The money market account's time t value is denoted by

$$B_t = e^{\int_0^t r_s ds} \tag{1}$$

with $B_0 = 1$. We let the time t value of a default-free zero-coupon bond paying a dollar at time T be strictly positive and denoted by $p(t, T) > 0$.

We consider a firm that issues a bond with a coupon of C dollars, a face value equal to L dollars, and a maturity date T. The bond pays the C dollar coupons at intermediate dates $\{t_1, ..., t_m = T\}$, but only up to the default time τ. For notational convenience, let the current time $t = t_0$. If default happens in the time interval $(t_{k-1}, t_k]$, then the bond pays a stochastic recovery rate of $\delta_{t_k} \in [0,1]$ at time t_k on the notional of L dollars.[7] It is important to note that default can happen anytime within this interval, but the payment only occurs at the end. If default does not happen, the face value of L dollars is repaid at time T.

A. Risk Neutral Valuation

To value the risky coupon bond, we assume (i) that the markets for both the default-free coupon bonds and the risky coupon bond are arbitrage-free and (ii) that enough credit derivatives trade on the risky firm so that the enlarged market is complete (see Jacod and Protter (2010) for a set of sufficient conditions on an incomplete market such that the expanded market is complete). Given the trading of credit default swaps, this is a reasonable approximation.

With only a minor loss of generality, we introduce a novel *conditional independence* assumption to facilitate analytic tractability. The conditional independence assumption (see the online appendix for the formal definition) is that the default-free spot rate r_t, the default time τ, and the recovery rate process δ_t are independent under the risk neutral probability $\mathbb{Q}$ given the information at time t. This is a weak assumption on the evolutions of the default-free spot rate, the default time, and the recovery rate because it imposes very little structure on their evolutions under the statistical probabilities. Under the statistical probabilities, these processes need not be independent. Hence, nonzero pairwise correlations under the statistical probabilities between the observed default-free spot rate, the default time, and the recovery rate processes, are not excluded by this assumption. And, it is well known that non-zero correlations across the default-free spot rate, default times, and recovery rates have been observed in historical data.

[7]In practice, a portion of the next coupon payment after default represents some accrued interest earned, but not yet paid. This accrued interest has a recovery rate associated with it. With a slight loss of generality we exclude this accrued interest payment in the stochastic recovery rate δ_{t_k} defined above. We appreciate the comments from a law firm, Morrison & Foerster, in this regard.

Denote the time $t \leq t_1$ value of the coupon bond as v_t. Under the conditional independence assumption, we show in the online appendix that the coupon bond's price is

$$v_t = \sum_{k=1}^{m} C \times z(t, t_k) + L \times z(t, T) + L \times d_t \sum_{k=1}^{m} x(t, t_k) \tag{2}$$

where

$$d_t := E^{\mathbb{Q}} [\delta_\tau | \mathcal{F}_t]$$

$$z(t, t_k) := p(t, t_k)[1 - Q(t, t_k)]$$

$$x(t, t_k) := p(t, t_k)[Q(t, t_{k+1}) - Q(t, t_k)]$$

$$Q(t, t_i) := Prob^{\mathbb{Q}} [\tau \leq t_i | \mathcal{F}_t] .$$

In this expression:

(i) $Q(t, t_i)$ is the time t conditional risk neutral probability of default before t_i given no default at time t,

(ii) d_t is the time t futures recovery rate (for a futures contract receiving the recovery rate at time T^*, see the online appendix for the details). As a futures price, the recovery rate in our valuation formula is a Q - martingale. This is an important implication of the conditional independence assumption underlying expression (2). Because it is a futures price, it is expected to be slightly larger than the recovery rate if paid on the debt at time t, δ_t (see the online appendix for a proof).

(iii) $z(t, t_k)$ is a *survival digital*, which pays \$1 at time t_k only if default occurs after t_k, 0 otherwise.

(iv) $x(t, t_k)$ is a *default digital*, which pays \$1 at time t_k if default occurs within $(t_{k-1}, t_k]$, 0 otherwise.

We refer to this expression as the "no-coupon recovery" model to emphasize that it has no recovery on the promised coupons after default. In this form it is easy to see that the value of this coupon bond is not equal to the sum of the coupons and principal times the value of a collection of risky zero-coupon bonds. Indeed, let $D(t, t_k)$ denote the time t value of such a risky zero-coupon bond promising to pay a dollar at time t_k for $k = 1, \ldots, m$ with recovery rate δ_t in default. Then, it can be shown that

$$\begin{aligned} v_t^{full\ coupon} &= \textstyle\sum_{k=1}^m C \times D(t, t_k) + L \times D(t, T) \\ &= \textstyle\sum_{k=1}^m C \times z(t, t_k) + L \times z(t, T) + L \times d_t \sum_{k=1}^m x(t, t_k) \\ &\quad + \textstyle\sum_{k=1}^m C \times (m + 1 - k) \times x(t, t_k). \end{aligned} \tag{3}$$

This expression is called the "full-coupon recovery model." The difference between this model and expression (2) is the term $\sum_{k=1}^m C \times (m + 1 - k) \times d_t \times x(t, t_k)$,[8] which represents the present value of the recovery on the coupons promised after default.

B. An Illiquidity Discount

Corporate bond markets are illiquid relative to Treasury bonds or exchange traded equities. This illiquidity implies that corporate bond prices may reflect an illiquidity discount (see Jarrow and Turnbull (1997), Duffie and Singleton (1999), Cherian, Jacquier, and Jarrow (2004)). An illiquidity discount modifies the previous valuation formula to implicitly incorporate the impact on pricing due to transaction costs and trading constraints.

It is important to note that transactions costs (including bid/ask spreads) are a special case of an illiquidity cost paid when trading, which are implicitly included via an illiquidity discount (see Cetin, Jarrow and Protter (2004) for the theoretical justification of this statement). Similarly, taxes paid on coupons and capital gains can also be interpreted as a type

[8] This term follows because if default occurs during the time interval $(t_{k-1}, t_k]$, the remaining future coupons are $\sum_{j=k}^m C = (m + 1 - k)C$. In the full coupon recovery model, one gets a recovery payment on all the remaining coupons.

of transaction cost, and hence they too are implicitly included in the illiquidity discount as well.[9]

We apply the illiquidity discount function $e^{\alpha_t(T-t)}$ symmetrically to all the cash flows promised to the coupon bond. This symmetry enables similar illiquidity discount impacts across different coupon bonds issued by the same credit entity. Given this, we can rewrite the coupon bond's value as

$$(4) \quad v_t^{liq} = \sum_{k=1}^{m} C \times z(t,t_k)e^{\alpha_t(t_k-t)} + L \times z(t,T)e^{\alpha_t(T-t)} + L \times d_t \sum_{k=1}^{m} x(t,t_k)e^{\alpha_t(t_k-t)}.$$

As we discuss below, we fit different versions of this model to the data. When the recovery rate and illiquidity discount are included in the estimation, both the recovery rate d_t and the illiquidity parameter α_t are stochastic, hence, they can vary randomly across time due to changing market conditions. Our estimation procedure allows for these estimated parameter values to reflect this randomness.[10] Expression (4) is the valuation model estimated in the empirical analysis.

III. Misspecification Errors

This section builds intuition for misspecification errors when using the full-coupon recovery model, expression (3) instead of the no-coupon recovery model, expression (2). Recall that the misspecification error, the difference between the full-coupon and no-coupon recovery

[9]The complication of explicitly including illiquidity costs (transaction, taxes) into the model is that different traders face different taxes and transaction costs based on their trading activities. Consequently, to determine a market price, an equilibrium model is needed. Equilibrium models are notoriously ladened with unrealistic assumptions. Furthermore, an argument can be made that the marginal trader, who determines the market price, is the lowest illiquidity cost trader. Here, we note that many institutions pay small transaction costs and there do exist non-taxable institutions that purchase corporate debt.

[10]We use implicit estimation at a fixed time t allowing α_t to depend on the information available at time t.

model prices, is equal to $\sum_{k=1}^{m} C \times (m+1-k) \times d_t \times x(t,t_k)$. Note that these misspecification errors are always positive.

We next quantify the magnitudes of these misspecification errors and provide a simple approximation that allows us to relate the misspecification errors to the model's inputs. Later, we relate the predicted misspecification errors to patterns in the data.

A. Misspecification Error Determinants

For illustrative purposes we make the following simplifying assumptions: (1) coupon bonds are priced on coupon dates, (2) the risk-free term structure of interest rates and the term structure of risk neutral default probabilities are flat,[11] (3) the coupon is set so that the no-coupon recovery model's bond price is equal to par, and (4) there is no illiquidity discount ($\alpha_t = 0$), though we relax this last assumption when we consider the effect of model parameters on spreads.[12] Combined, these imply that the misspecification error is fully determined by the maturity, default probability, recovery rate, and risk-free rate. We note the use of risk neutral default probabilities is essential because we are creating valuation formulas, which require a risk premium. Later, under an additional assumption that default risk is diversifiable, the distinction between risk neutral and actual probabilities disappear because under this assumption they are empirically equivalent.

Our data, which we describe in more detail in Section 4, consists of more than 168 thousand observations from September 2017 to August 2022 for a total of 197 issuers. More than ninety percent of the data have investment grade level ratings equal to BBB- or above;[13]

[11] In the empirical implementation (Model 3) we use a term structure of risk neutral default probabilities, which is not assumed to be flat.

[12] In this case, the full coupon recovery model including a liquidity discount is $v_t^{full\ coupon} = \sum_{k=1}^{m} C \times z(t,t_k)e^{\alpha_t(t_k-t)} + L \times z(t,T)e^{\alpha_t(T-t)} + L \times d_t \sum_{k=1}^{m} x(t,t_k)e^{\alpha_t(t_k-t)} + \sum_{k=1}^{m} C \times (m+1-k) \times x(t,t_k)e^{\alpha_t(t_k-t)}$. The last term is the misspecification error.

[13] The sample consists primarily of investment grade bonds since many high yield bonds have call features, all of which are excluded. An analysis of callable bonds goes beyond the scope of this paper.

average maturity is equal to 3.1 years, the average coupon equals 3.1 percent. We also consider issuer-day-level statistics since we estimate the model at that level: the average issuer-day has five observations; ninety five percent of issuer days have eleven or fewer observations; the average different between the shortest and longest maturity within each issuer-day ('maturity range') is 3.9 years, and ninety percent of observations have maturity ranges between 0.9 and 8.6 years. Our sample is therefore appropriate to study how default risky coupon bonds should be priced.

Table 1 reports misspecification errors across different inputs, assuming that the risk-free term structure is flat at 2%. The par value of the bond is set to 100 and the recovery rate is equal to 50%, a level close to the mean recovery rate we estimate (see below). As expected, misspecification errors increase with the bond's maturity and the issuer's default probability. For short maturity 2-year bonds, the misspecification error is equal to 0.07 if the annual default probability is 2%, while the misspecification error is equal to 1.39 for a 10-year bond with the same default probability. For 30-year bonds the misspecification error can be much larger, reaching a level of 9.62 for a 2% default probability bond, close to 10% of that bond's price.

We now propose a simple approximation for the misspecification error. In the event of default, the present value of the payoff for the first coupon is equal to the discounted value of the product of the coupon rate, the recovery value, and the probability of default, i.e. $C \times d_t \times p(t, t_1)Q(t, t_1)$. The approximate total error is equal to $C \times d_t \times p(t, t_1)Q(t, t_1)m(m+1)/2$ (see the online appendix for additional detail).

We later use the misspecification error to identify portfolios of bonds that are likely to be mispriced by the full-coupon recovery model. We note that the misspecification error is zero if the recovery rate, the default probability, or the coupon payment is zero. The error grows approximately with the square of the number of coupon payments and is exactly proportional to the product of the coupon payment and the recovery rate. Thus, bonds with significant recovery values, default probabilities, and with intermediate to long maturities will have significant misspecification errors.

B. Pricing with Two Credit Spread Curves

If coupons have a zero recovery after default, while the principal payment has a positive recovery, both cash flows will not have the same discount rate. Using the same credit spread for both will result in an inability to price bonds with different maturities and coupons. However, a priori it is not clear if the effect we are focusing on is empirically large or small. Spreads appropriate for discounting coupons and principal may be similar.[14] Before proceeding with our full model estimation, we examine the difference in the two pricing approaches by examining seniority-specific spreads. If there is a misspecification error using this full-coupon recovery model to price bonds, then the two curves will be different. The valuation formula using different credit spreads for coupon and principal payments is

$$v_t^{spread} = \sum_{k=1}^{m} C \times p(t,t_k)e^{-s_C(t,t_k)(t_k-t)} + L \times p(t,T)e^{-s_L(t,T)(T-t)} \tag{5}$$

where $s_C(t,t_k)$ and $s_L(t,T)$ are credit spreads at time t for the coupon and principal cash flows at times t_k and T, respectively, above the default-free rates implicit in the zero-coupon bond prices $p(t,t_k)$.[15]

Table 2 provides some illustrative examples of credit spread curves. We use the same methodology as in Table 1. The only difference is that here we introduce the effect of an illiquidity discount. Panel A reports principal spreads, Panel B reports coupon spreads.

[14] We note that we are interested in pricing multiple bonds simultaneously. It is, of course, possible to calculate a bond-specific yield to maturity and therefore a bond specific credit spread. This, however, does not provide a pricing methodology, but it is simply a transformation of the price into another quantity.

[15] This is the same as defining $p(t,t_k)e^{-s_C(t,t_k)(t_k-t)} = D_S(t,t_k)$ and $p(t,T)e^{-s_C(t,T)(T-t)} = D_L(t,T)$, which correspond to distinct risky zero-coupon bond price term structures for discounting coupons and principal cash flows.

As long as there is a positive recovery, coupon spreads lie above principal spreads since the latter will be worth more and thus are discounted less. The difference between coupon and principal spreads is close to the product of the default probability and the recovery rate, which follows from the misspecification error relation given above, where, for the first coupon, the misspecification error is equal to $C \times d_t \times p(t, t_1)Q(t, t_1)$. A larger default probability makes all spreads higher. If there is no illiquidity discount, coupon spreads are approximately equal to the default probability, and since differences relative to principal spreads depend on the default probability, frictionless spreads are approximately proportional to the default probability. The effect of the illiquidity discount is seen to be symmetric, affecting all cash flows equally. Indeed, both credit spreads increase by the amount of the illiquidity discount.

The results imply that principal payments are safer because they deliver potentially large recovery values in the event of default. Coupon payments, in contrast, do not pay off in default and therefore need a larger discount rate. It is useful to note that using a single spread is not suitable to discount both cash flows with zero or positive recovery. For the former (the coupons), the spread will be too low and for the latter (the principal) it will be too high. Thus, using a single spread (or spread curve) to price a new bond with a different maturity or coupon will result in misspecification errors. In addition, using this 'standard' spread calculation to assess the market's implied risk pricing is not possible.

IV. Data and Estimation

The details of the estimation procedures are as follows. To fit the valuation model to market prices, we obtain traded coupon bond prices for the 1,248 trading days from the beginning of September 2017 until the end of August 2022 using the TRACE system.

The pricing model is for senior unsecured fixed-rate coupon bonds with no embedded options. For each firm, we therefore eliminate from the sample any subordinated bonds, callable and putable bonds, structured bonds, bonds with "death puts" or a "survivor option," and floating rate bonds. Survivor option bonds distort bond prices both because they are issued in small amounts (typically $20 million or less per tranche) and because the value of the embedded put option is significant. The survivor option feature has become more common in recent years.[16] In addition, to be included in our sample, the bond issue's daily trade volume had to exceed $50,000 (in almost every case, volume was much larger) and with at least two separate bonds traded (to ensure model convergence, for issuer-days with only two observations we also require that the maturities are at least half a year apart). We further excluded some bonds of European issuers subject to a 2014 EU regulation allowing regulators to demand an exchange of senior debt securities into equity. Because data assembly and cleaning costs are substantial,[17] we restrict our attention to the sample starting in 2017. Finally, we restrict attention to bond price observations with prices above risk-free bond prices with the same set of cash flows.[18] The resulting sample consists of more than 168 thousand observations for 197 issuers and more than 35 thousand issuer days.

Table 3 presents summary statistics. The average coupon is equal to 3.1%, average

[16]The largest issuers of survivor option bonds as of 2016 included General Electric, Goldman Sachs, Bank of America, Wells Fargo, Ford Motor, HSBC Holdings, National Rural Utilities Cooperative Finance Corporation, Dow Chemical, Prospect Capital, and Barclays PLC. A typical survivor option bond's terms are described as follows in a recent prospectus supplement from General Electric Capital Corporation: "Specific notes may contain a provision permitting the optional repayment of those notes prior to stated maturity, if requested by the authorized representative of the beneficial owner of those notes, following the death of the beneficial owner of the notes, so long as the notes were owned by the beneficial owner or his or her estate at least six months prior to the request. This feature is referred to as a 'Survivor's Option.' Your notes will not be repaid in this manner unless the pricing supplement for your notes provides for the Survivor's Option. The right to exercise the Survivor's Option is subject to limits set by us on (1) the permitted dollar amount of total exercises by all holders of notes in any calendar year, and (2) the permitted dollar amount of an individual exercise by a holder of a note in any calendar year."

[17]It is necessary to screen out callables and survivor options, data on which is only available in the pricing supplement. The SEC and FINRA do not maintain public access to prospectus data for more than about 5 years in easily accessible form. Thus including, for example, data from the financial crisis is not feasible. In addition, the TABB group finds a very high frequency of errors "TABB Group analysis shows reconciliation differences in more than 20% of new issues." (http://www.finregalert.com/an-sec-mandated-corporate-bond-data-monopoly-will-not-help-quality/). There are also non–trivial computational costs.

[18]Some observations have prices above risk-free bond prices (i.e. negative implied credit spreads), perhaps due to data errors. A negative credit spread could signal a potential arbitrage opportunity. However, it may be difficult to capitalize on such mispricing because of illiquidity. We leave further exploration of these patterns to future research.

maturity is equal to 3.1 years,[19] and the mean credit spread is 80 bps. There is quite a bit of variation in the data – 5th to 95th percentile ranges of coupons, maturity and spreads are 4.2, 7.9 years, and 201 bps, respectively. This variation is important for our ability to identify differences in the no-coupon and full-coupon recovery models. Misspecification errors are small if the coupon is low and the maturity is short, while they are high if the maturity is long and the coupon is large. If there is little variation in misspecification errors, the full-coupon recovery model may produce biased estimates, but the pricing errors may be similar to the no-coupon recovery model. However, taking a look at the issuer-day level statistics we find a lot of variation. Average maturity range is almost four years and the average issuer day has five bond price observations in it. There is also variation in credit ratings. The average rating is A- and 7.6% of observations are for non-investment grade (BB+ and below) issuers.

Table 3 Panel B reports additional firm characteristics across rating groups. In order to fit the bond pricing model to data, we require at least two and ideally more observations for each issuer day. This restriction naturally focuses attention on issuers with a lot of outstanding debt, in particular financial institutions that tend to issue a lot of bonds. We note that across the four rating groups average book leverage declines as rating increases (i.e. rating quality declines), no doubt because choice of leverage is endogenous and financial institutions often have low-risk ratings and high leverage. The pattern in stock return volatility is as we would expect; as rating increases, volatility increases from 24% for AA and above to 46% for non-investment grade issuers. While there is little variation in leverage within rating group, the variation in volatility is much larger and increases with rating. Credit spreads exhibit a similar pattern, ranging from 46 bps for AA and above to 210 bps for non-investment grade.

To these data, we add U.S. Treasury yields reported daily by the U.S. Department of the Treasury[20] and derive the maximum smoothness Treasury forward rate curves from these

[19]Model 3 (discussed below) implementation is based on default probabilities that extend to a maturity of ten years. We therefore restrict attention to observations with that maximum maturity.

[20]https://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yield

data (see Adams and van Deventer (1994)). Using these historical forward rate curves, we compute the term structure of default-free zero coupon bond prices on all dates.

Finally, we assemble data on coupon bond prices. The price in the TRACE system does not represent the full amount paid for the bond. The full amount paid is the price plus accrued interest.

In our estimation, we compare the full amount paid when acquiring the bond (the present value of the bond purchase) with the valuation model in expression (4). Specifically, each issuer day, we use the risk-free term structure, coupon payments and payment dates as inputs. We then use non-linear least squares estimation, calculated on a volume-weighted basis, to solve for the best fitting parameter values (recovery rate, default probability, illiquidity). We estimate the no-coupon and the full-coupon recovery models separately and compare pricing errors as well as parameter estimates.

A. Three Empirical Model Implementations

Each issuer day we fit the data to three different empirical implementations of the model.

1. Implied Default Probabilities (Models 1 and 2)

Model 1 starts with a restricted version of the model, which allows us to clearly trace the effect of estimating both the full and no-coupon recovery models on misspecification errors, pricing errors, and parameter estimates. We first fit our model, expression (4), assuming a flat term structure of default probabilities estimated implicitly. We also assume no liquidity discount,[21] and a fixed recovery rate futures price (recovery rate for short) equal to 50%, which

[21] In the model, a change in the default probability affects the present value of all cash flows, both coupons and principal. The same is true for a change in the illiquidity parameter. From expression (4) and the spread curve examples in Table 2 we know that the effect is not exactly the same, and so it is possible

is close to the average recovery rate estimated in the less restrictive model implementation given below. Importantly, this model has only one parameter, the default probability. Model 1 allows us to see the direct impact of misspecification on estimated default probabilities as well as model outperformance relative to the full-coupon recovery model.

In Model 2, we relax the restriction on recovery rates and instead allow it to vary between 0.1 and 0.8. This model has two parameters and it is significantly more flexible. We restrict default probabilities to be greater than 0.1% – close to the first percentile of the distribution when using an historically estimated default probability, discussed below. These bounds on the default probability and recovery rates reduce excessive model flexibility, in particular for the full-coupon recovery model, and allow us to trace the effect of misspecification errors on pricing errors more easily. Both models 1 and 2 have the advantage that they require only bond prices and the term structure of risk-free Treasury rates as inputs.

2. Historically Estimated Default Probabilities (Model 3)

Model 3, instead of estimating default probabilities implicitly using bond price data, uses a historically estimated default probability. That is, we employ independently estimated default probabilities from a proportional hazard rate model. Because there is then one less parameter to fit, in this version of the model we include an illiquidity discount parameter together with the recovery rate.

To facilitate the estimation of the default intensity process, we assume that the default time τ corresponds to the first jump time of a Cox process with intensity $\lambda_t = \lambda_t(\Gamma_t) \geq 0$ where $\Gamma_t = (\Gamma_1(t), \ldots, \Gamma_m(t))' \in \mathbb{R}^m$ are a collection of stochastic processes characterizing the state of the firm and the market at time t. In addition, we assume that default risk is diversifiable in the sense of Jarrow, Lando, and Yu (2005).[22] This assumption enables the

to estimate both separately. Nevertheless, to guard against unstable estimates or overfitting, we set the illiquidity discount equal to zero and estimate only a default intensity. In Model 2, we also fit the recovery rate. When using historically-estimated default probability (Model 3) we also add an illiquidity parameter (see below).

[22] For additional detail, see the online appendix.

estimation of default intensities without the need to adjust the intensity process for a default jump risk premium. In conjunction, these two assumptions imply that we can estimate the default probabilities using a proportional hazard rate model (see Fleming and Harrington (1991), p. 126), i.e.

$$\lambda_t(\Gamma_t) = \theta e^{\phi \Gamma_t}$$

where θ is a constant and where ϕ is a vector of constants. For an application of such a hazard rate model applied to corporate default probabilities see Chava and Jarrow (2004).

As discussed in Jarrow, Lando, and Yu (2005), this assumption does not imply that risky coupon bonds earn no risk premium. Quite the contrary. If the state variables Γ_t driving the default process represent systematic risk, which is the most likely case, then risky coupon bond prices necessarily earn a risk premium due to the bond price's correlation to Γ_t. The diversifiable risk assumption just states that the timing of the default event itself, after conditioning on Γ_t, is diversifiable in a large portfolio. Alternatively stated, in a poor economy all firms are more likely to default. But, the timing of which firms actually default depends on the idiosyncratic risks of the firm's management and operations.

The default process parameters (θ, ϕ) from the proportional hazard rate model were provided by the Kamakura Risk Information Services (KRIS) division of SAS Institute, Inc.[23] KRIS uses a refinement of the approach employed by Chava and Jarrow (2004) to estimate these parameters that are then used to construct the full term structure of cumulative default probabilities.[24] Specifically, for each issuer-day we obtain cumulative default probabilities from the 10-year term structure of monthly marginal default probabilities (the monthly probability of default conditional on no prior default). The state variables used in KRIS's hazard rate estimation include both firm specific and macroeconomic variables. Importantly,

[23] See www.kamakuraco.com.

[24] The model underlying the default probability calculations is similar to the one used in Campbell, Hilscher and Szilagyi (2008, 2011), who extend Chava and Jarrow (2004) and Shumway (2001). Campbell et al. show that the default probability measure is a more accurate predictor of failure than Moody's EDF numbers, data that have been widely used in academic studies, e.g. Berndt, Douglas, Duffie and Ferguson (2018).

the default probabilities do not use traded bond or CDS prices as inputs. Default probabilities are therefore separate inputs relative to the observed bond prices that we fit using model 3.

We restrict the recovery rate to lie between 0.1 and 0.8 and the illiquidity discount to lie between zero and -5%. Doing so will reduce the influence of observed bond price errors on the estimates. We report robustness checks below.

V. Illustrations: Coupon and Principal Seniority in Default

Before moving to the full sample estimation, this section provides evidence that market prices reflect the difference in seniority between principal and coupons in default. We consider three companies that filed for bankruptcy: Lehman, PG&E, and Weatherford International. Lehman is chosen because of the size and importance of its bankruptcy. The latter two firms are in our sample because each firm has a sufficient number of bonds traded. In each case we focus on senior bonds, including callable bonds, because on the day bankruptcy is announced the call option is worthless and can be ignored. We fit the no-coupon and full-coupon recovery models to the data.

The key reason for analyzing issuer bonds after they file for bankruptcy is that the default probability equals 100%.[25] The recovery amount for the no-coupon recovery model is the recovery rate times the notional of $100 (par value) for each of the bonds. In contrast, the recovery amount for the full-coupon recovery model is $100 plus the dollar coupon times the number of remaining payments on each bond, a different amount for each issue. For each

[25]Since the bonds are in default, we can include both non-callable as well as callable bonds in this analysis.

issuer day and for both of the models, referring back to expression (4),

$$v_t = d_t \times L, \tag{6}$$

because: (i) default has occurred and there are no more coupon payments after time t, (ii) the principal is immediately due, and (iii) the liquidity discount is assumed to be zero because the bond has defaulted. We run a regression on this formula to derive the recovery rate and the present values (price plus accrued interest) for each bond. As noted earlier, expression (4) ignores the interest earned since the last coupon payment date. This would be included in the market prices.

Figures 1-3 depict the pricing errors. We order the bonds by maturity. Pricing errors when using the no-coupon recovery model (in blue) are substantially lower than those resulting from the full-coupon recovery model (in red). Mean absolute errors are more than five times as large for Lehman (2.0 vs. 11.0) and almost ten times as large for PG&E (2.1 vs. 19.7) and Weatherford International (2.6 vs. 21.0).

We see that the full-coupon recovery model results in prices that are too large, especially for bonds of longer maturities that have more coupons, which if they were of equal seniority, would entitle the bondholder to a recovery value. However, in default those coupons are worthless and so any coupon paying bond would have pricing errors that are positive as long as the model was using unbiased inputs. However, in an attempt to fit the data, the model tries to reduce the average pricing error resulting in bonds with short maturities being underpriced and bonds with long maturities being overpriced. The maximum errors lie between 19.6 and 37.4. It is worth noting that average market prices are equal to 32.4 (Lehman), 78.2 (PG&E), and 65.0 (Weatherford International) so that the maximum errors are around one half the market price. The (negative) minimum errors are similar in size lying between -37.2 and -14.2.

In contrast, the no-coupon recovery model's maximum and minimum pricing errors are much smaller. They lie between 3.9 and 9.1 and -5 and -2.7, and so are approximately one quarter of the full-coupon recovery model pricing errors. Importantly, and in direct support of the no-coupon recovery model, its pricing errors have no clear pattern relative to the bond's maturity.

To summarize, Lehman, PG&E and Weatherford International's bond prices provide direct evidence in support of the no-coupon relative to the full-coupon recovery model. Failing to take into account the different seniority of coupons and principal results in substantial pricing errors, which have a predictable pattern consistent with our model.

VI. The Pricing Model Comparison

This section provides a comparative analysis of the no-coupon and full-coupon recovery models.[26] The full-coupon recovery model produces different prices only if misspecification errors (see section 3) are non-zero. We therefore investigate *both* if the no-coupon recovery model has better fit on average, and also if it has better fit when misspecification errors are larger. Both of these predictions are implications of our model because if bonds are priced according to the no-coupon recovery model, then a necessary condition for model outperformance is the presence of misspecification errors.

A. Misspecification Errors

[26] In a previous version of the paper we also compared the no-coupon recovery model to one based on ratings. In that model coupons are assumed to have full recovery and the credit spread is assumed to depend only on the rating. The ratings-based valuation model is consistent with numerous pronouncements from the Basel Committee on Banking Supervision (2010, 2017). It performs poorly primarily because of the erroneous assumption that all firms that have the same rating have the same risk; analyzing it is therefore less relevant when comparing no-coupon and full-coupon recovery models.

Before analyzing model fit and outperformance, we first consider the distribution of the misspecification errors. We use the unbiased estimates from the no-coupon recovery model as inputs to compute the misspecification error as the difference in prices between the full-coupon and the no-coupon recovery model. Recall that the misspecification error is approximately equal to $C \times d_t \times p(t, t_1) \times Q(t, t_1) \times m(m+1)/2$.

The full-coupon recovery model can only exhibit a worse fit if misspecification errors are present. If they are zero, the two models are the same. The misspecification error's approximation formula is multiplicative in coupon rate, recovery rate, default probability, and maturity. It is useful to note that the approximation is quite accurate in capturing the variation in misspecification errors. When we regress actual on predicted misspecification errors, the R^2 lies between 93% and 95%.

Table 4 reports summary statistics of model fit, misspecification errors, model outperformance, and parameter estimates. The 25th percentile issuer-day misspecification error is 4 cents (the median is 13 cents). Thus, as expected, a large fraction of the data is not greatly affected by the pricing differences between the two models. However, the mean misspecification error is more than twice as large and equal to 30 cents, which means that there are many bonds with large implied price differences across the two models (when input parameters are held constant). The 95th percentile of the misspecification error distribution is 1.02, which is substantial, and the 75th percentile (unreported) is 0.30, which is also quite large. The median observed bond price is equal to 101.33. As a result, these numbers are directly comparable to those in Table 1. Of course, when estimating the two models separately, which is what we do next, the full-coupon recovery model may adjust parameters, resulting in possible biases. However, as we saw in Section 4, when discussing the bond prices of defaulted companies, an incorrect model not only produces biased parameter estimates, but a worse fit.

B. Model Performance

Following Bakshi, Madan and Zhang (2006) we report mean absolute pricing error to compare model fit.[27] Another alternative could be to measure performance comparing yields or credit spreads. We do not choose this route because our analysis of model misspecification errors and our measure of outperformance focuses specifically on dollar pricing errors, not yields. In addition, one contribution of our paper is to point out that yields and credit spreads should not be used to price bonds.

As noted previously, if bonds are priced according to the no-coupon recovery model and not the full-coupon recovery model, this has two implications. First, in a sample of bonds that have large default risk, we should detect that the no-coupon recovery model has a better fit. Second, the outperformance will be larger when the two models disagree by more (i.e. when the misspecification errors are larger and more variable). We next provide evidence supporting both of these implications. We note that there is nothing mechanical about the relation between misspecification errors and model outperformance. If bonds were priced according to the full-coupon recovery model, we would find that it outperforms the no-coupon recovery model, and that it does so by more when differences between model implied prices are larger.

We also note that average model outperformance must, necessarily, be a function of sample characteristics. What this implies is that we expect a sample with higher misspecification errors to have higher no-coupon model outperformance. In addition, it implies that we may find low average outperformance for other sample characteristics. This is one reason to focus on explaining variation in model outperformance: to be consistent with the theory, we should find evidence that the no-coupon model outperforms where we expect it to outperform.

1. Model 1: A Fixed Recovery Rate

[27] Eom, Helwege and Huang (2004) calculate percentage pricing errors, and Bakshi, Madan and Zhang (2006) also report these. In our sample, 90 percent of prices lie between 95.70 and 110.11; results are therefore robust to using percentage errors instead.

We compare the two models after fitting them independently to the data. For each issuer-day, we estimate both the no-coupon and full-coupon recovery models and calculate the volume-weighted mean absolute error. These average error statistics are reported in Table 4 Panel A.

As expected, we find that the no-coupon recovery model fits the data better than the full-coupon recovery model. The mean model outperformance is equal to 0.07. However, also as expected, there are many observations for which the error difference is very small (median outperformance is equal to 0.02). At the same time, there are also issuer days with larger pricing error differences. The 95th percentile of the outperformance distribution is 0.27. We also report overall model fit – the no-coupon recovery model has an average dollar pricing error of 0.35.

What is more relevant for the performance comparison is how the models perform when model prices differ, resulting in misspecification errors. In fact, having a large misspecification error standard deviation within the data sample is crucial to validating the no-coupon recovery model. When predicted misspecification errors are small, for example because the default probabilities are low, pricing differences will also be small. Reflecting the large portion of the data with low misspecification errors, we see that a large fraction of the data also has a low standard deviation of those errors. The average standard deviation is 26 cents, the median is 12 cents, and the 25th percentile is 3 cents.

But, even if predicted misspecification errors are large, the full-coupon recovery model may still generate low pricing errors, for example, if the bonds in the specific issuer-day sample have very similar predicted misspecification errors. In this case the misspecified model will be able to adjust by changing parameters and the resulting pricing error differences relative to the no-coupon recovery model can be small, albeit at the cost of biased model parameter estimates. But, if there is a high variability in the misspecification errors, the incorrect model will fail to fit prices as well as the no-coupon recovery model.

As predicted, we find that the no-coupon recovery model outperformance is large when the misspecification error standard deviation is large. The second set of results in Panel A, first line, is for the subsample of issuer days in the top quartile of the misspecification error standard deviation distribution. In that group, the minimum misspecification error standard deviation is 30 cents and the mean is 75 cents. Correspondingly, the outperformance (difference in mean absolute errors across the two models) is much larger, exactly as our model predicts. The average outperformance more than triples to 0.22. Even though we are considering only 25% of issuer days, this sample reflects prices from 51,133 observations, which is 30.4% of observations, or 5.7 observations per issuer-day, compared to 4.7 for the full sample.

Breaking up the sample across credit rating groups, we find, as expected, that the misspecification error increases with rating. The misspecification error increases from 0.10 (AA and above) to 0.94 (BB and below). Correspondingly, model outperformance increases with rating, from 0.03 (AA and above) to 0.23 (BB and below). The main driver for larger misspecification errors is a larger annual default probability, which in model 1 is the parameter estimated. It increases from 1.1% (AA and above) to 4.5% (BB and below). These are the parameters for the no-coupon recovery model.

When estimating the full-coupon recovery model, estimated default probabilities are larger. If they were not, the full-coupon recovery model would produce prices that are too large on average. We therefore can see the bias introduced when using the misspecified model. Of course, all models are incomplete approximations of reality and we can also compare the no-coupon recovery model estimates to independently-estimated default probabilities based on historical data, summary statistics for which we report below. Those numbers are lower, mainly because Model 1 assumes no variation in illiquidity.

2. Model 2: A Variable Recovery Rate

We next estimate Model 2, which relaxes the constraint on the recovery rate, but continues to estimate implied default probabilities from bond prices. Adding a degree of freedom, we expect the overall fit to increase and the no-coupon recovery model outperformance to decrease, because there are now two parameters also in the full-coupon recovery model, both of which can be biased. Indeed, average outperformance is equal to 0.03, but that number approximately triples to 0.10 for the high misspecification standard deviation sample. As before, outperformance increases with credit rating. We also note that average estimated recovery rate is close to 50% (restricted value in Model 1) both for the full sample, as well as for rating subsamples.

C. Determinants of the No-coupon Recovery Model's Outperformance

We now explore the determinants of the no-coupon recovery model's outperformance (the difference between the no-coupon and full-coupon model volume-weighted absolute pricing errors). As discussed in Section III, outperformance should be related to the misspecification error – if this error is larger, we expect outperformance of the no-coupon recovery model to be larger as well. Table 5 Panel A reports outperformance for the full sample and several subsamples. In Panel B we regress outperformance on different sets of explanatory variables in order to explore determinants of no-coupon recovery model outperformace. On average, the no-coupon recovery model provides a better fit (also see Table 4). For Model 1, pricing errors are 7.2 cents larger when using the full-coupon recovery model and the difference is statistically significant. We have already seen that the (in)ability of the full-coupon recovery model to fit the data reflects its misspecified assumption. Thus, we expect a strong relationship between the no-coupon recovery model outperformance and the misspecification error's standard deviation.

We first focus on the subsample with the top 25% of default probabilities. For that subsample, the average outperformance is 19.1 cents, almost three times as large as in the full sample. This outperformance is even larger when considering the subsample with the largest 25% average misspecification error issuer days. Here the outperformance is 21.8 cents on average. As expected, the variable finding the largest outperformance is the misspecification error's standard deviation. It may be large because of a large dispersion in maturities combined with large default probabilities. The misspecified full-coupon recovery model does not have sufficient degrees of freedom to match the data well. The no-coupon recovery model has, on average, a 22.4 cent lower pricing error than the full-coupon recovery model.

The pattern is the same, only stronger, when examining the pricing errors in the top deciles. For large default probability issuer days the outperformance is equal to 29.5 cents, for the misspecification error it is 35.4 cents and for the standard deviation it is equal to 37.3 cents. Pricing error differences are large when the model predicts them to be large. A similar pattern emerges for Model 2, though pricing error differences are smaller throughout, reflecting the additional degree of freedom (two instead of one fitted model parameter), and resulting potential bias in that model. For the top decile of misspecification error standard deviations, outperformance is on average equal to 16 cents, which is highly statistically significant.

We next explore in more detail what determines the size of the no-coupon recovery model outperformance. Specifically, we regress pricing error differences on maturity, coupon, and default probabilities. Together with the recovery rate, these four variables are the main determinants of the misspecification error. The recovery rate is fixed in Model 1 and is therefore not included, but we do include it when explaining variation in the outperformance of Model 2. Maturity and the default probability are highly significant with a positive sign; the coupon rate also has a positive sign but is insignificant.

All variables enter the misspecification error, but they do so in a specific way. We next include the actual misspecification error and the misspecification error standard deviation into the regression. We expect the dispersion of the misspecification errors to determine model outperformance. When we include the misspecification error's standard deviation, the coefficient on it is highly significant, while the coefficients on all the other variables becomes either indistinguishable from zero or switches sign. The R^2 of the regression increases from 37.1% to 60.6%. Dropping all the other variables and keeping only the misspecification error's standard deviation in the regression results in a similar fit of 58.4%.

The ability of the misspecification error's standard deviation to explain the variation in model outperformance is direct evidence supporting our hypothesis that bond market prices are consistent with the no-coupon recovery model instead of the full-coupon recovery model. When fitting Model 2, the results are very similar. All four variables individually enter with the expected sign. When the misspecification error's standard deviation is included along with the average misspecification error, the coefficients drop in size by more than half or become negative. As before, the regression R^2 increases dramatically from 18.7% to 41.4% and it is only slightly smaller at 38.2% when only the misspecification error's standard deviation remains in the regression.

To summarize, our analysis provides strong evidence that the pricing error differences between the two models are statistically significant for the full sample. Importantly, variation in model outperformance occurs exactly when the model predicts it. This evidence is for a sample consisting of 93.5% investment grade debt (98.7% with a rating of BB+ or above) and thus one where market participants perceive default is not imminent.

D. Outperformance Over Time

We have documented the no-coupon recovery model's outperformance in the full sample and in sub-samples. We next consider the time variation in its outperformance. Each week we calculate average outperformance and the average misspecification error's standard deviation for all issuer days. Figure 4 plots the time series of the no-coupon recovery model's outperformance based on Model 1. There is some noticeable time variation. Toward the end of 2018 and in the beginning of 2019 the model's outperformance increases, reaching a local peak of about 11 cents. This episode happened contemporaneously with a stock market downturn and a corresponding increase in volatility and default probabilities. Then, at the beginning of the pandemic in early 2020 we see a large increase in model outperformance reaching a weekly average of 37 cents. Toward the end of the sample, in 2022, the average outperformance reaches 24 cents. Note also that average outperformance of the no-coupon recovery model is positive throughout the sample.

As reported in Table 5, the misspecification error's standard deviation explains 58% of the variation in no-coupon recovery model outperformance. Using weekly averages – the data from the figure – results in an R^2 of 79%. We note that, during the pandemic, there is a lot of variation in both outperformance and the misspecification error's standard deviation; during that time the relationship is weaker (R^2 of 64%). Figure 5 plots weekly averages outside of 2020. We find a clear linear relation between outperformance and the misspecification error's standard deviation; the R^2 of this relationship is 93%. In short, model outperformance is large when we expect it to be large.

For Model 2 the pattern outside of 2020 is very similar even though the average outperformance is lower (see Tables 4 and 5). However, during the height of pandemic (March 2020) the relationship between outperformance and the misspecification error's standard deviation is no longer present when using Model 2. This may be due to the additional degree of freedom in that model and reflects the lower overall no-coupon recovery model outperformance. Using a more constrained model with a historically-estimated default probability and in which we estimate the recovery rate (discussed below) results in a strong relationship between model outperformance and the misspecification error's standard deviation in 2020.

E. Model 3: Using a Historically Estimated Default Probability and an Illiquidity Discount

In Models 1 and 2, the default probability is estimated implicitly. We now use historically-estimated default probabilities from a proportional hazard rate model as an input (as discussed in Section 4.1.2). The data are from the Kamakura Risk Information Services (KRIS) division of SAS Institute, Inc. We also allow for an effect of illiquidity (as discussed in Section 2.2).

Table 6 reports estimation results. The no-coupon recovery model's mean average pricing error is equal to 0.29, slightly lower than Model 2. The average misspecification error is equal to 0.21 and the average no-coupon recovery model's outperformance is 0.02. When focusing on the large misspecification error's standard deviation subsample the no-coupon recovery model's outperformance more than triples and is equal to 0.07.

The median annual default probability used as an input to the model is 0.7% and the mean is equal to 1.3%.[28] We estimate the mean recovery rate as 49%, in line with the restriction in Model 1 and the mean recovery rate in Model 2. The mean illiquidity parameter is -0.4%. It is useful to note that both the recovery rate and illiquidity estimates are reasonable. This is, of course, not guaranteed given that our estimates are implicitly estimated using the traded bond prices and the no-coupon recovery model. Jankowitsch, Nagler, and Subrahmanyam (2014) report an average recovery rate value of 0.38. Since our recovery rate is in fact the recovery rate futures price, as shown in the online appendix, it is expected that our estimate should be slightly larger than these estimates. In Panel A we report full-sample illiquidity estimates of 0.26% (median) and 0.42% (mean). Though somewhat lower, these are broadly consistent with, for example, the spread between Aaa-rated corporate bonds and Treasury

[28]These probabilities lie below estimated default probabilities in Models 1 and 2. Those measures are higher because we do not include an illiquidity discount in those models.

debt. We next discuss variation in the estimated illiquidity parameter over time and note that it also spikes in March of 2020, similar to the Aaa-Treasury spread.

F. Time Variation in Parameter Estimates

We next study the illiquidity and recovery rate parameter estimates over time. Before proceeding, it is useful to get a sense of what kind of variation is present in the default probabilities (which are an input to the model). Figure 6 reports monthly average default probabilities over the sample period. Two important drivers of default probabilities are volatility and leverage (see, e.g., Merton (1974), Jarrow (2009), Guha, Sbuelz, and Tarelli (2020)), which we also plot. Variation in default probabilities over time are dominated by the effect of the pandemic in 2020. It is also notable that variation in default probabilities is tracked by variation in stock return volatility, while book leverage remains close to constant throughout the sample period.

Figure 7 plots weekly average illiquidity, based on the full estimation sample (Table 6). We notice a slight increase in late 2018 and early 2019. The more striking and larger increase in the estimated illiquidity parameter occurs at the beginning of the pandemic. Average illiquidity increases to just under 3%. This increase is matched by an increase in the Aaa-Treasury spread, which we also plot in the figure. We interpret the Aaa-Treasury spread as a related measure of illiquidity. Indeed, as is evident from the figure, the two series move together; the correlation is equal to 65% (in levels, 79% in changes). The close relationship between the Aaa-Treasury spread provides independent validation of our model and its implied corporate bond illiquidity measure.

There is also considerable variation in the illiquidity parameter across ratings. Figure 8 shows the weekly average illiquidity discounts at a monthly frequency across rating groups, again based on the full estimation sample. The three investment grade groups have somewhat similar illiquidity levels, while the non-investment grade's group has noticeably larger illiquidity and is also less correlated with the other three groups.[29]

The average recovery rates over time are plotted in Figure 9. The fact that there is no recovery of coupons in the event of default means that recovery rates can be inferred directly from bond prices. This is an important empirical novelty compared to an inability to disentangle recovery rates and default probabilities if equal seniority is assumed for coupons and principal, and there is a single spread used to discount all cash flows. There is not much variation over time, supporting the assumption of a fixed recovery rate in Model 1. Indeed, there is also little variation in average recovery rate across rating groups (Table 6 Panel B).

We can use these estimates to interpret what happened during the height of the COVID pandemic. Figure 10 plots weekly averages of recovery rates and illiquidity parameters in 2020. From Figure 6 we already know that the historically estimated default probabilities increased dramatically in early 2020. We can now see that there was also a dramatic increase in illiquidity, from an average of 0.18% in January to a maximum of 2.8% in the second half of March. There is a modest increase in recovery rates from close to 55% in January to 68% in March, perhaps due to an increased market focus on short-term liquidity-induced defaults, which may have been perceived to result in slightly lower levels of loss in the event of default. Bond prices can decline either because of higher default probabilities, lower recovery rates, or higher illiquidity. Our estimates suggest that, at least during March of 2020, the two main effects were a loss of liquidity and an increase in default probabilities.

G. Out-of-sample Model Performance

Our analysis so far is based on fitting the pricing model in-sample. A concern is that in-sample model fit statistics can be biased by overfitting noise in the data. To address this issue,

[29] Investment grade illiquidity averages lie below the Aaa-Treasury spread (see Table 6) suggesting that, although the two measures are highly correlated in levels and changes, they do not capture exactly the same market friction.

our focus was on relative performance. We measured if the no-coupon model had a lower overall error as compared to the full-coupon recovery model. Both models have the same number of parameters and a similar structure. Unless we believe that there is a differential bias in the model fit statistics, our measure of outperformance should not be affected by the in-sample methodology.

Nevertheless, it is useful to check relative model performance using an out-of-sample pricing approach as well. Each issuer day, we split the sample of bonds into two groups of equal or close to equal size (if there are an odd number of observations, one group will be slightly larger than the other). One group is the estimation sample that is used to calculate the models' parameters – default probability, recovery rate, and illiquidity. We take these estimated parameters and compute fitted prices for the other group's observations, the out-of-sample group. We fit both the full and the no-coupon recovery model in this way. In order to ensure that we can estimate the parameters in this manner, we restrict attention to issuer days with at least four observations; that way each group always contains at least two observations. We repeat the process but switching the observation groups so that the second group of bonds on that issuer day (that we previously used for out-of-sample fitting) is now used for estimation purposes. In this way, when calculating each bond price, we are using parameters that are estimated using a different sample (on the same day, but including different bonds), while at the same time ensuring that we can work with a large data set of out-of-sample model prices. To ensure that the samples are comparable, groups are assigned based on maturity rank.

We estimate all three model implementations, Models 1, 2, and 3. We report full-sample no-coupon recovery model outperformance as well as outperformance for the sample with large misspecification error standard deviations. Table 7 Panel A reports summary statistics. All three models show the no-coupon recovery model outperformance and in each case the magnitude of the outperformance increases substantially for the sample with large misspecification error standard deviation observations. These are the main two patterns previously identified, and both are present when using this out-of-sample estimation approach. We find that average outperformance is highly statistically different from zero (Panel B) for all three models and for the full sample as well as the large misspecification error standard deviation subsamples. Panel C reports average outperformance across rating groups. As the credit rating increases, the no-coupon recovery model outperformance increases also, consistent with the in-sample estimation results.

We have performed several additional robustness tests, including a quasi simulation regarding biased parameter estimates in the full-coupon recovery model, larger daily observation cutoffs, grouping observations by month, and using alternative recovery rate estimation bounds. We discuss these robustness checks in the online appendix.

VII. Conclusion

This paper presents evidence that the common corporate bond pricing assumption of equal seniority of principal and coupon payments is not supported by market transaction prices. We propose a tractable coupon bond valuation model, which includes a more realistic recovery rate process that distinguishes between coupon payments received before and after default. This setup has important advantages that support our empirical investigation. (1) The model implies that a single spread or spread term structure cannot be used to discount all cash flows. Instead, seniority-specific discount rates reflect different recovery rates for principal and coupons. (2) The model has a clear prediction about the importance of modeling the market practice of zero recovery paid on coupons after default. We calculate misspecification errors – those resulting from using the full-coupon recovery model rather than the no-coupon recovery model. When these errors are large, differences in the no-coupon and the full-coupon recovery model's predictions are larger. Misspecification errors are shown to depend directly on the coupon, recovery rate, default probability, and time to maturity, and they can be substantial in size.

We find that our no-coupon recovery model's predictions are reflected in a large data set of bond transaction prices, evidence that market prices of risky coupon bonds reflect zero coupon recovery after default. The model has a clear prediction when no recovery on coupons after default is relevant for pricing and when it is less important. Indeed, if default probabilities, coupons, recovery rate, or maturity are small, then the effect of the differing coupon recovery assumptions has only a very small impact. We find evidence supporting this prediction, while also identifying our model's outperformance in the full sample.

We document that model outperformance is closely related to the misspecification error's standard deviation within the estimation sample (generally an issuer-day). When that standard deviation is large, our model predicts that bond prices will be the most affected by the erroneous assumption of full-coupon recovery after default. The fact that the misspecification error's standard deviation explains model outperformance well is thus direct evidence supporting the no-coupon recovery model. Separately, we find that the no-coupon recovery outperformance is evident when considering bond prices of companies in bankruptcy. Model outperformance is also higher for non-investment grade issuers.

Finally, our model allows for direct estimation of implied recovery rates and an illiquidity parameter's effects on bond prices. Average recovery rates, though a little higher, are generally in line and consistent with the previous literature. When the Covid-19 pandemic hit markets in March of 2020, both the Aaa-Treasury and our illiquidity parameter spiked (also see Kargar et al. (2021)). In the crisis, default probabilities increased, bond prices dropped, and illiquidity increased markedly. Indeed, we find that variation in our illiquidity parameter has a strikingly close relationship to the Aaa-Treasury spread, a standard measure of corporate bond illiquidity. The close relationship between the two measures, neither of which uses the same data nor methodology to compute, provides independent validation for our pricing model.

References

Adams, K. and D. van Deventer. "Fitting Yield Curves and Forward Rate Curves with Maximum Smoothness." *Journal of Fixed Income*, June (1994), 52 - 62.

Bakshi, G., X. Gao, and Z. Zhong. "Decoding default risk: A review of modeling approaches, findings, and estimation methods." *Annual Review of Financial Economics*, forthcoming (2022).

Bakshi, G., D. Madan, and F. Zhang. "Investigating the Role of Systematic and Firm-Specific Factors in Default Risk: Lessons from Empirically Evaluating Credit Risk Models." *The Journal of Business*, 79 (4) (2006), 1955 - 1987.

Basel Committee on Banking Supervision. Basel III: A global regulatory framework for more resilient banks and banking systems. December, Bank for International Settlements (2010).

Basel Committee on Banking Supervision. Basel III: Finalizing the post-crisis reforms. December, Bank for International Settlements (2017).

Berndt, A., R. Douglas, D. Duffie, and M. Ferguson. "Corporate Credit Risk Premia." *Review of Finance*, 22(2) (2018), 419–454.

Bielecki, T. and M. Rutkowski. *Credit Risk: Modeling, Valuation and Hedging*, Springer Verlag, Berlin (2002).

Bremaud, P. *Point Processes and Queues: Martingale Dynamics*, Springer Verlag, Berlin (1980).

Campbell, J. Y., J. Hilscher, and J Szilagyi. “In Search of Distress Risk.“ *Journal of Finance*, 63(6) (2008), 2899–2939.

Campbell, J. Y., J. Hilscher, and J. Szilagyi. “Predicting Financial Distress and the Performance of Distressed Stocks.” *Journal of Investment Management*, 9(2) (2011), 14–34.

Campbell, J. Y. and G. B. Taksler. “Equity Volatility and Corporate Bond Yields.” *Journal of Finance*, 58 (2003), 2321-2350.

Cetin, U., R. Jarrow, and P. Protter. “Illiquidity Risk and Arbitrage Pricing Theory.” *Finance and Stochastics*, 8 (3), August (2004), 311 - 341.

Chava, S. and R. Jarrow. “Bankruptcy Prediction with Industry Effects.” *Review of Finance*, 8 (4) (2004), 537 - 569.

Cherian, J., E. Jacquier, and R. Jarrow. “A Model of the Convenience Yields in On the Run Treasuries.” *Review of Derivatives Research,* 7 (2004), 79 - 97.

Collin-Dufresne, P. and R. Goldstein. “Do Credit Spreads Reflect Stationary Leverage Ratios.” *Journal of Financ*e, 56 (6) (2001), 1929 - 1957.

Collin-Dufresne, P., R. Goldstein, and J. Martin. “The Determinants of Credit Spread Changes.” *Journal of Finance*, 56 (5) (2001), 2177 - 2207.

Driessen, J. “Is Default Event Risk Priced in Corporate Bonds?” T*he Review of Financial Studies*, 18 (1) (2005), 165 - 195.

Duffee, G. "Estimating the Price of Default Risk." *The Review of Financial Studies*, 12 (1) (1999), 197 - 226.

Duffie, D., L. Pedersen, and K. Singleton. "Modeling Sovereign Yield Spreads: A Case Study of Russian Debt." *Journal of Finance*, 63 (1) (2003), 119 - 159.

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

Elton, E., M. Gruber, D. Agrawal, and C. Mann. "Explaining the Rate Spread on Corporate Bonds." *Journal of Finance*, 56 (1) (2001), 247 - 277.

Eom, Y. H. J. Helwege, and J. Huang. "Structural Models of Corporate Bond Pricing: An Empirical Analysis." *The Review of Financial Studies*, 17 (2) (2004), 499 - 544.

Fleming, T. and D. Harrington. *Counting Processes and Survival Analysis*, John Wiley & Sons, Inc., New York (1991).

Gilchrist, S. and E. Zakrajšek. "Credit Spreads and Business Cycle Fluctuations." *American Economic Review*, 102 (4) (2012), 1692-1720.

Guha, R. and A. Sbuelz. "Structural Recovery of Face Value at Default." unpublished working paper, University of Verona, Italy (2005).

Guha, R., A. Sbuelz, and A. Tarelli. "Structural recovery of face value at default," *European Journal of Operational Research*, 283 (3) (2020), 1148-117.

Guo, X., R. Jarrow and H. Lin. "Distressed Debt Prices and Recovery Rate Estimation." *Review of Derivatives Research*, 11 (2008), 171 - 204.

Heath, D., R. Jarrow and A. Morton. "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation" *Econometrica*, 60 (1) (1992), 77 - 105.

Hilscher, J. and Y. Nosbusch. “Determinants of Sovereign Risk: Macroeconomic Fundamentals and the Pricing of Sovereign Debt.” *Review of Finance*, 14 (2) (2010), 235–262.

Huang, J. and M. Huang. “How Much of the Corporate-Treasury Yield Spread Is Due to Credit Risk?,” *Review of Asset Pricing Studies*, 2 (2) (2012), 153–202.

Jacod, J. and P. Protter. “Risk Neutral Compatibility with Option Prices.” *Finance and Stochastics*, 14 (2010), 285 - 315.

Jankowitsch, R., F. Nagler, and M. G. Subrahmanyam. “The Determinants of Recovery Rates in the US Corporate Bond Market.” *Journal of Financial Economics*, 114 (1) (2014), 155-177.

Jarrow, R. “Risky Coupon Bonds as a Portfolio of Zero-coupon Bonds,” *Finance Research Letters*, 1 (2004), 100 - 1005.

Jarrow, R. “Credit Risk Models.” *Annual Review of Financial Economics*, 1 (2009), 37 - 68.

Jarrow, R. *Continuous-Time Asset Pricing Theory: A Martingale-Based Approach*, 2nd Edition, Springer (2021).

Jarrow, R., D. Lando and F. Yu. “Default Risk and Diversification: Theory and Empirical Applications,” *Mathematical Finance*, 15 (1) (2005), 1 - 26.

Jarrow, R. and P. Protter. “An Introduction to Financial Asset Pricing.” *Handbooks in OR&MS*, Chap. 1 (J. R. Birge & V. Linetsky eds.), Vol. 15. Philadelphia: Elsevier B.V. (2008).

Jarrow, R. and S. Turnbull. “Pricing Derivatives on Financial Securities Subject to Credit Risk." *Journal of Finance*, 50 (1) (1995), 53 - 85.

Jarrow, R. and S. Turnbull. “An Integrated Approach to the Hedging and Pricing of Eurodollar Derivatives.” *Journal of Risk and Insurance*, 64 (2) (1997), 271 - 299.

Jarrow, R. and S. Turnbull. “The Intersection of Market and Credit Risk.” *Journal of Banking and Finance*, 224 (2000), 271 - 299.

Kargar, M., B. Lester, D. Lindsay, S. Liu, P. O. Weill, and D. Zúñiga. “Corporate Bond illiquidity during the COVID-19 Crisis.” *The Review of Financial Studies*, (2021), 34 (11), 5352–5401.

Lando, D. “On Cox Processes and Credit Risky Securities.” *Review of Derivatives Research*, 2 (1998), 99 - 120.

Longstaff, F., S. Mithal, and E. Neis. “Corporate Yield Spreads: Default Risk or illiquidity? New Evidence from the Credit Default Swap Market.” *Journal of Financ*e, 60 (5) (2005), 2213 - 2253.

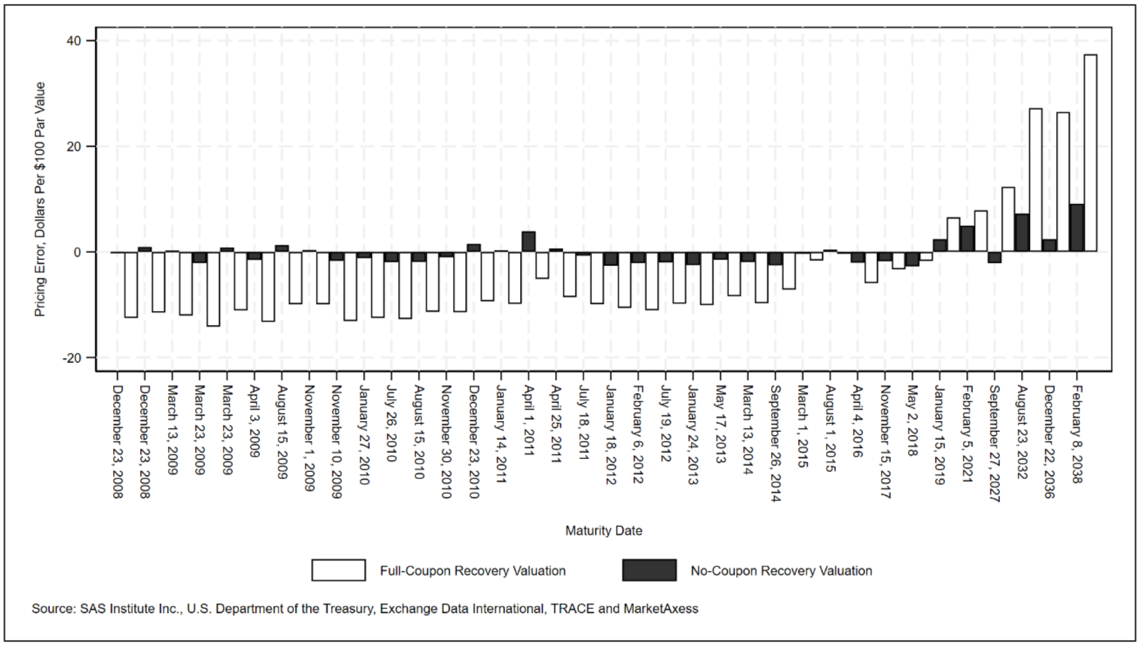
Madan, D. and H. Unal. “Pricing the Risks of Default.” *Review of Derivatives Research*, 2 (1998), 121 - 160.

Merton, R. C. “On the Pricing of Corporate Debt: The Risk Structure of Interest Tates.” *Journal of Finance*, 29(2) (1974), 449-470.

Protter, P. *Stochastic Integration and Differential Equations*, 2nd edition, Springer-Verlag: New York (2005).

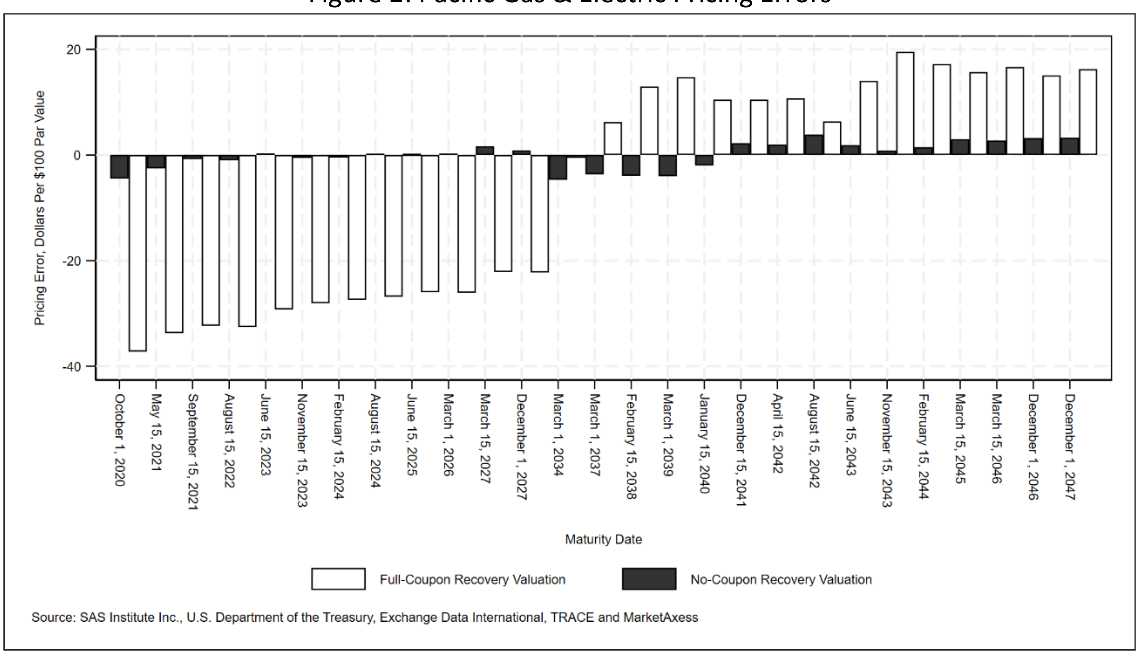
Shumway, T. "Forecasting Bankruptcy More Accurately: A Simple Hazard Model." *Journal of Business*, 74(1) (2001), 101–124.

Figure 1: Lehman Brothers Pricing Errors



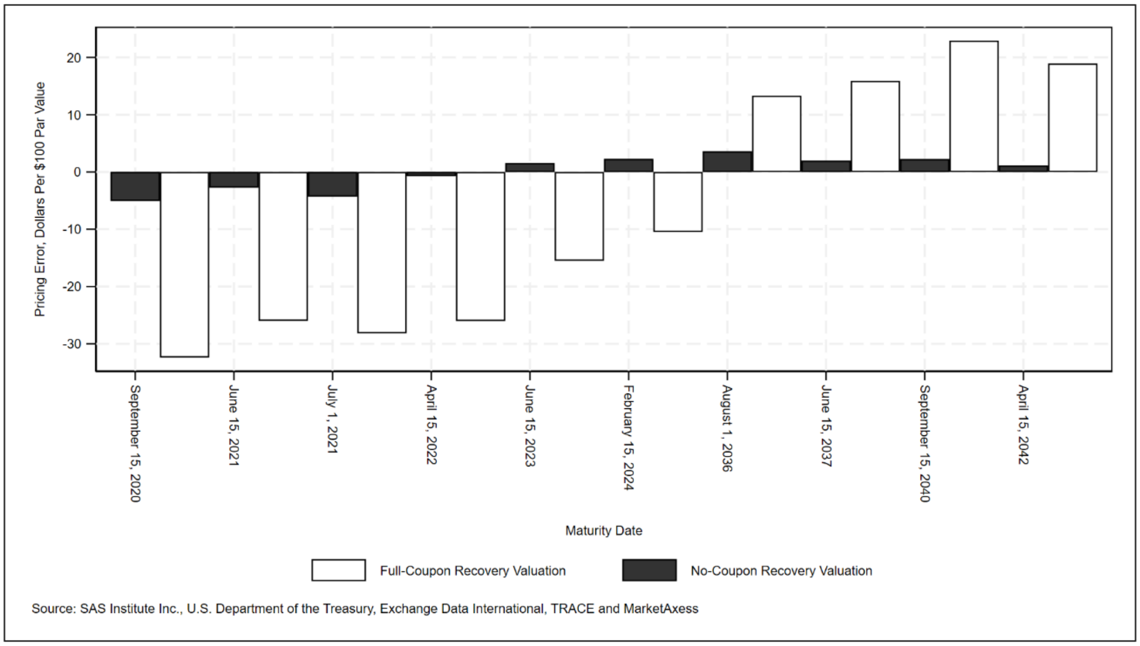
Note: Full-coupon and no-coupon recovery model pricing errors for Lehman Brothers bonds on September 15, 2008.

Figure 2: Pacific Gas & Electric Pricing Errors



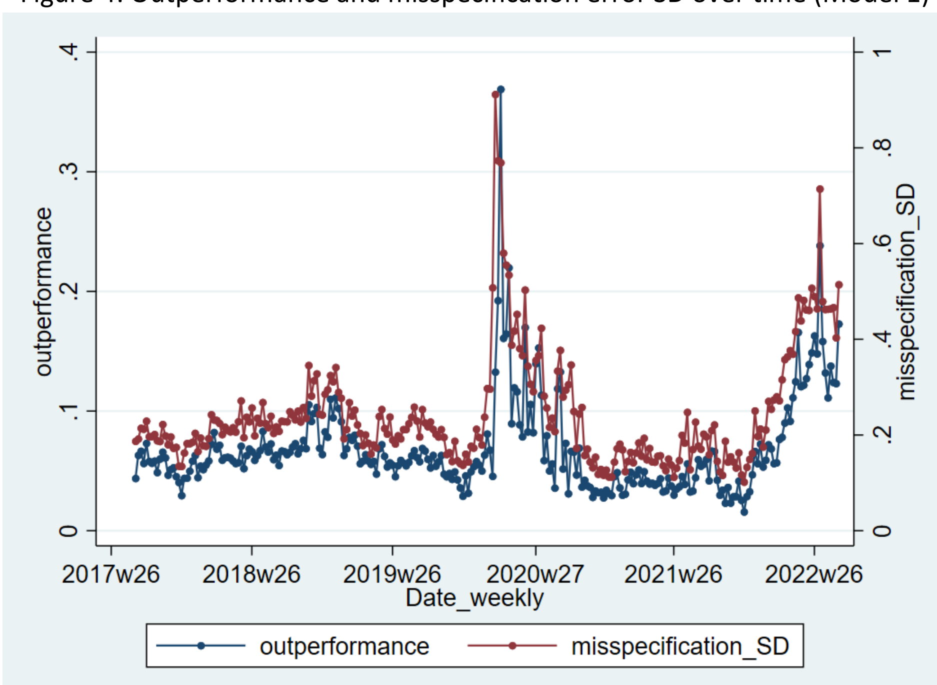
Note: Full-coupon and no-coupon recovery model pricing errors for PG&E January 14, 2019.

Figure 3: Weatherford International Pricing Errors



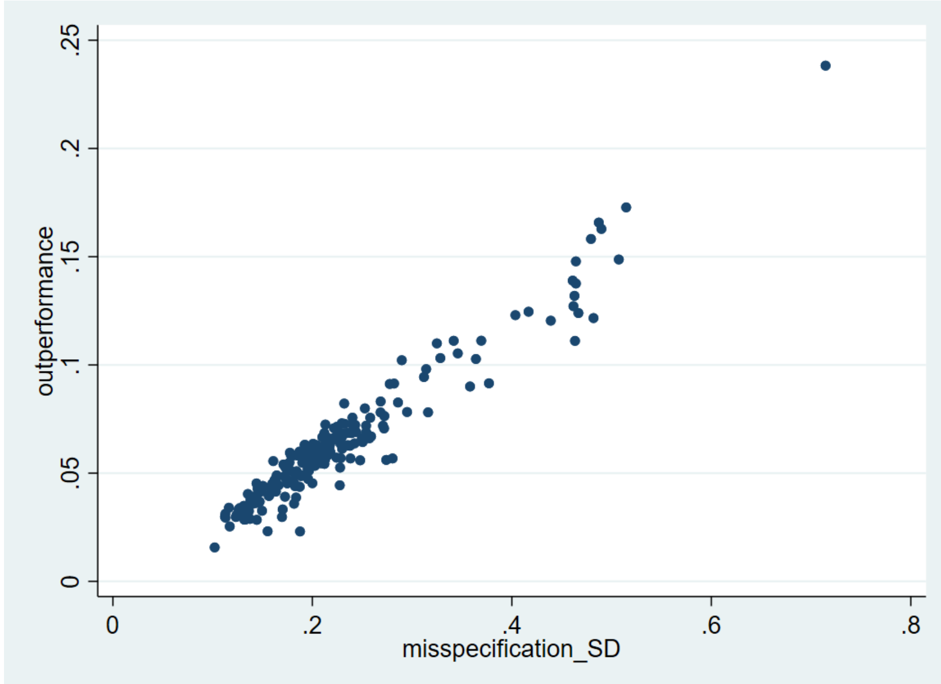
Note: Full-coupon and no-coupon recovery model pricing errors for Weatherford International May 10, 2019.

Figure 4: Outperformance and misspecification error SD over time (Model 1)



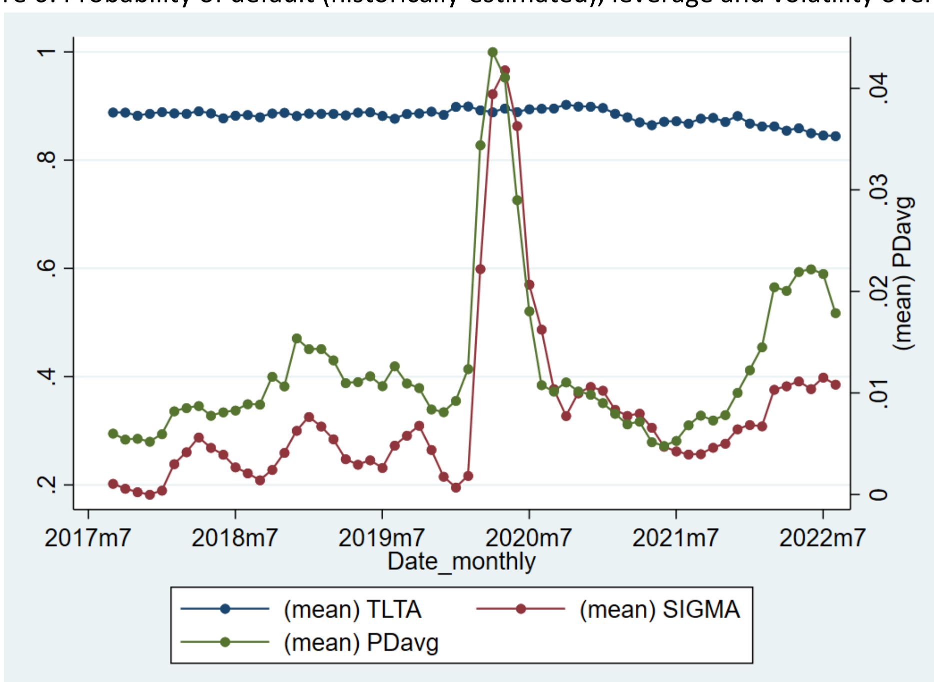
Note: Model 1 weekly averages: outperformance is the difference in volume-weighted absolute pricing errors (in dollars) of the no-coupon recovery relative to the full-coupon recovery model; misspecification_SD is the average of the issuer-day misspecification error standard deviation.

Figure 5: No-coupon recovery model outperformance outside of 2020 (Model 1)



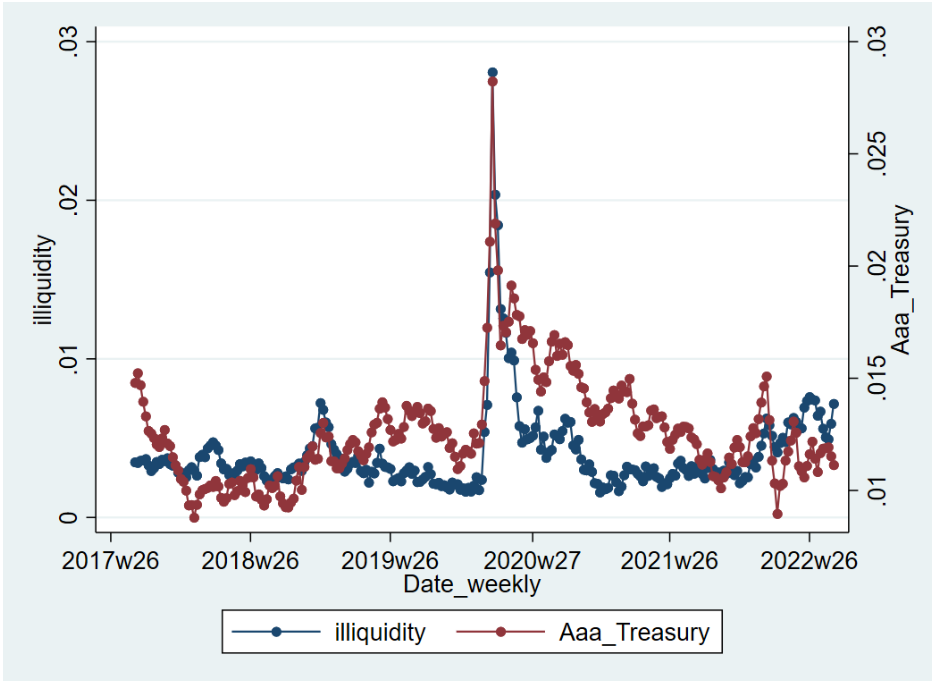
Note: Model 1 weekly averages: outperformance is the difference in volume-weighted absolute pricing errors (in dollars) of the no-coupon recovery relative to the full-coupon recovery model; misspecification_SD is the average of the issuer-day misspecification error standard deviation.

Figure 6: Probability of default (historically-estimated), leverage and volatility over time



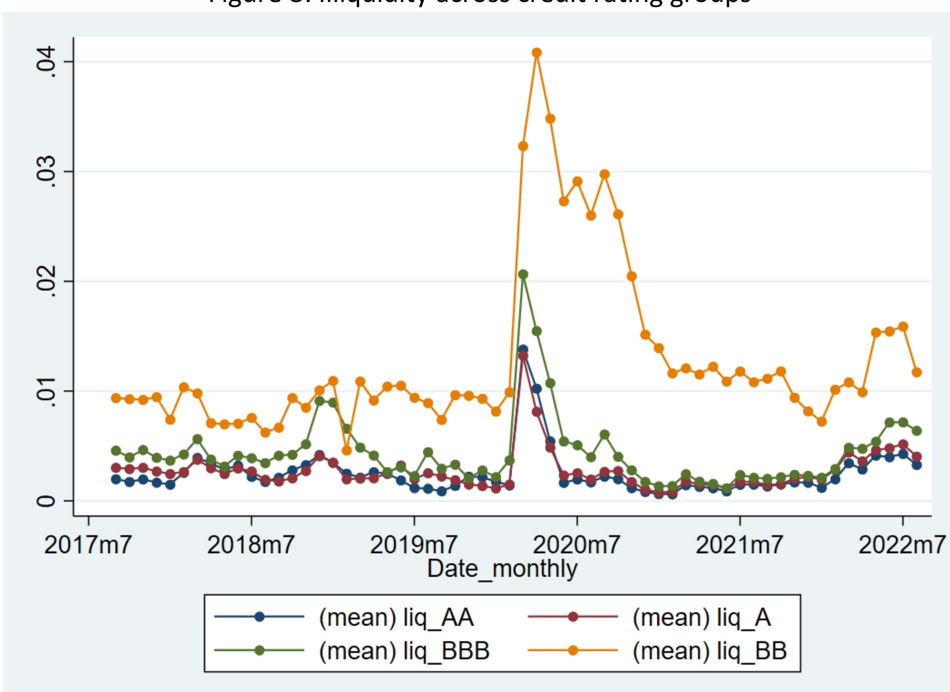
Note: Monthly averages of: historically-estimated annual default probabilities (averaged across maturities), book leverage (total assets divided by total liabilities), and stock return volatility (SIGMA). Data provided by Kamakura Risk Information Services (KRIS).

Figure 7: Model illiquidity and Aaa-Treasury spread



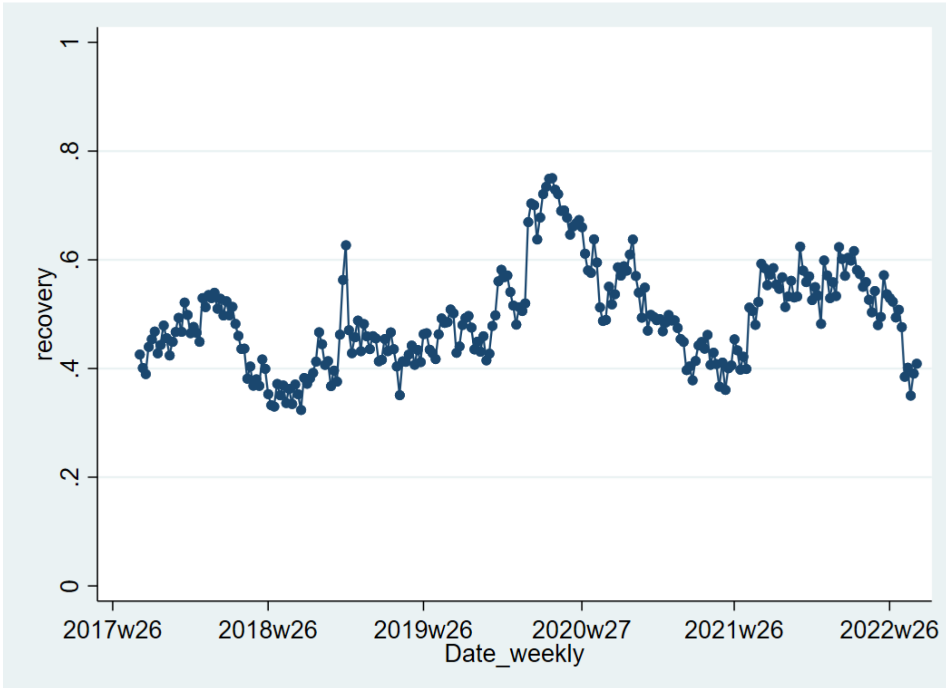
Note: Weekly averages of: illiquidity estimated from Model 3 and the Aaa – Treasury spread from FRED.

Figure 8: Illiquidity across credit rating groups



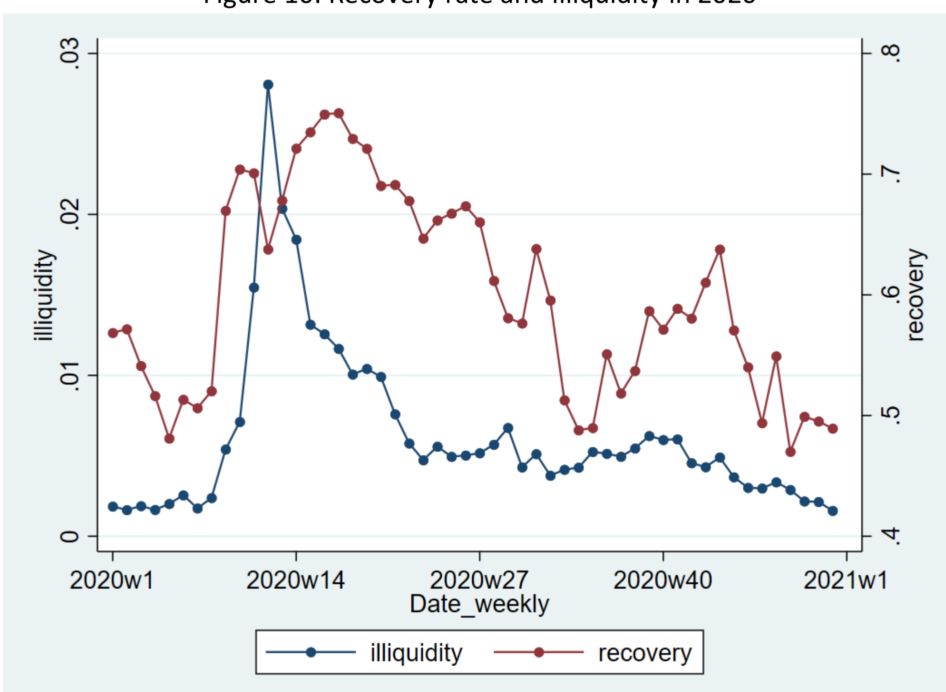
Note: Monthly averages of: illiquidity estimated from Model 3, calculated across rating groups: AA and above, A, BBB, and BB and below.

Figure 9: Average recovery rate over time



Note: Weekly average recovery rate estimated from Model 3.

Figure 10: Recovery rate and Illiquidity in 2020



Note: Weekly averages of: illiquitidy and recovery rate estimated from Model 3.

Table 1: Misspecification error determinants

In this table we calculate prices based on the no-coupon recovery and the full coupon recovery models. We assume a flat risk-free term structure of 2%, a flat default probability term structure and different maturities. Coupons are chosen so that bonds (based on the no-coupon recovery model price) trade at par ($100). We report the misspecification error (in dollars) resulting from using the full-coupon recovery instead of the no-coupon recovery model (column five, missspec. error).

maturity	recovery	default probability (annual)	coupon	misspec. error (in dollars)
2	0.5	1%	2.51%	0.03
2	0.5	2%	3.03%	0.07
5	0.5	1%	2.51%	0.16
5	0.5	2%	3.03%	0.39
10	0.5	1%	2.51%	0.60
10	0.5	2%	3.03%	1.39
30	0.5	1%	2.51%	4.33
30	0.5	2%	3.03%	9.62

Table 2: Coupon and principal spreads

This table reports spreads (in percent) appropriate for discounting coupons and principal (C and P) for various maturities (see equation (5) in the text), annual default probabilities, and illiquidity values. As in the previous table, we assume a flat risk-free term structure of 2% and a flat default probability term structure. Panel A reports spreads appropriate for disounting principal, Panel B reports spreads appropriate for discounting coupons.

Def prob	1%	1%	2%	2%
Recovery	0.5	0.5	0.5	0.5
Illiquidity	0	-0.5%	0	-0.5%
Maturity				
Panel A: Principal spreads				
1	0.50%	1.01%	1.01%	1.51%
3	0.49%	0.99%	0.98%	1.48%
5	0.48%	0.98%	0.94%	1.44%
10	0.44%	0.93%	0.86%	1.34%
Panel B: Coupon spreads				
1	1.02%	1.52%	2.04%	2.55%
3	1.02%	1.52%	2.04%	2.55%
5	1.02%	1.52%	2.04%	2.55%
10	1.02%	1.52%	2.04%	2.55%

Table 3: Summary statistics

This table reports detailed summary statistics for the main sample of bond prices (the sample which we use to perform our empirical analysis). Panel A reports bond characteristics and overall ratings. We report both observation- and issuer-day-level statistics. Panel B contains issuer-day level statistics by rating group. Spread, reported in basis points (bps) is the bond-specific credit spread (standard definition); maturity range is the difference for each issuer-day between the maximum and minimum maturity; maturity SD is the issuer day maturity standard deviation; number of observations counts how many bond prices are in the data set each issuer day (we require a minimum of two), Rating is the S&P issuer credit rating, TLTA is the ratio of COMPUSTAT book value of total liabilities divided by book value of total assets, SIGMA is the stock return standard deviation. To be included in the sample, the spread, coupon, and maturity must be positive, and the maturity range (if there are only two observations) must be at least 0.5. The restrictions, which result in only a small share of the data being dropped, are discussed further in the text.

Panel A: Bond characteristics

	Bond-level stats			Issuer-day level stats			
	Coupon	Maturity	Spread (bps)	Maturity range	Maturity SD	No. of obs.	Rating
Mean	3.07	3.1	81	3.9	1.8	5	A-
SD	1.26	2.3	185	2.3	1.0	3	2
p5	1.25	0.3	14	0.9	0.5	2	AA-
p50	2.85	2.6	58	3.5	1.6	4	A-
p95	5.45	8.2	215	8.6	3.5	11	BB+
No. of obs	168,285	168,285	168,285	35,635	35,635	35,635	35,635

Panel B: Firm characteristics across ratings

	TLTA	SIGMA	Spread (bps)	TLTA	SIGMA	Spread (bps)	
	Averages			Standard deviations			No of obs
AA, above	0.89	0.24	46	0.11	0.16	35	5,650
A	0.89	0.30	58	0.12	0.17	56	17,297
BBB	0.86	0.41	101	0.12	0.26	328	9,981
BB, below	0.84	0.46	210	0.13	0.22	292	2,707

Table 4: No-coupon recovery model outperformance statistics and parameters

This table reports summary statistics of model fit. Panel A reports statistics for a model where recovery rate is set equal to 0.5 (Model 1), Panel B reports results when allowing for a variable recovery rate between 0.1 and 0.8 (Model 2). Both models restrict default probability to lie above 0.1%. Estimation is done at the issuer-day level minimizing the volume-weighted squared pricing error. All data are reported at the issuer-day level. Mean absolute error (MAE) is the no-coupon recovery volume-weighted error (in dollars) for a given issuer day. MAE difference is the difference between the full-coupon recovery and the no-coupon recovery model MAE (in dollars). Avg miss error is the average misspecification error (in dollars), Def prob is the annual fitted default probability, which is reported first for the no-coupon recovery model and next when estimated using the full-coupon recovery model. In each panel we also report model fit statistics for a subsample of issuer-day observations in the top 25 percent of within issuer-day misspecification error standard deviation, as well as for four rating groups.

Panel A: Model 1 (fixed recovery rate, variable default proability)

	Mean abs error (MAE, in dollars)	MAE difference (full-coupon relative to no-coupon rec)	Avg Miss error (in dollars)	Def Prob (no-coupon rec)	Def Prob (full coupon rec)
Mean	0.35	0.07	0.30	1.8%	2.0%
p5	0.02	0.00	0.01	0.4%	0.4%
p50	0.26	0.02	0.13	1.3%	1.4%
p95	0.98	0.27	1.02	4.4%	5.1%
Number of issuer days: 35,635					
Subsamples: high misspecification error standard deviation, rating groups					
Top quartile miss SD	0.66	0.22	0.85	3.3%	3.9%
AA, above	0.25	0.03	0.10	1.1%	1.2%
A	0.30	0.05	0.19	1.3%	1.4%
BBB	0.45	0.10	0.42	2.2%	2.5%
BB, below	0.53	0.23	0.94	4.5%	5.2%

Panel B: Model 2 (variable recovery rate and default probability)

	Mean abs error (MAE, in dollars)	MAE difference (full-coupon relative to no-coupon rec)	Avg Miss error (in dollars)	Def Prob (no-coupon rec)	Recovery rate
Mean	0.32	0.03	0.52	2.6%	0.51
p5	0.01	0.00	0.00	0.4%	0.10
p50	0.24	0.01	0.11	1.8%	0.79
p95	0.90	0.14	2.18	7.7%	0.80
Number of issuer days: 35,635					
Subsamples: high misspecification error standard deviation, rating groups					
Top quartile miss SD	0.46	0.10	1.71	5.7%	0.78
AA, above	0.23	0.01	0.19	1.6%	0.50
A	0.28	0.02	0.34	1.9%	0.49
BBB	0.40	0.05	0.71	3.3%	0.54
BB, below	0.44	0.08	1.72	6.7%	0.56

Table 5: Determinants of model outperformance

This table reports statistics of model outperformance, the volume-weighted average absolute error difference (in dollars) when comparing the full-coupon recovery model and the no-coupon recovery model. Panel A reports averages for different samples; Panel B reports results from regressions of model outperformance (i.e. volume-weighted MAE difference) on different sets of explanatory variables. We report results for both Model 1 (fixed recovery rate, variable default probability) and Model 2 (variable recovery rate and default probability). PD is default probability, Miss error is the average within issuer-day misspecification error and Miss error SD is its standard deviation. Recovery rate is from the no-coupon recovery model; coupon and maturity are averaged at the issuer-day level; Def prob is the fitted default proability. Standard errors, reported below coefficients, are robust and clustered by issuer and date; *** denotes significant at the 1% level, ** at the 5% level and * at the 10% level.

Panel A: No-coupon recovery model outperformance -- full and subsamples

Sample	full	PD top quartile	Miss error top quartile	Miss err SD top quartile	PD top decile	Miss error top decile	Miss err SD top decile
	Model 1 (fixed recovery rate, variable PD) outperformance (in dollars)						
Average	0.072***	0.191***	0.218***	0.224***	0.295***	0.354***	0.373***
	(0.014)	(0.038)	(0.038)	(0.037)	(0.088)	(0.080)	(0.077)
	Model 2 (variable recovery rate and PD) outperformance (in dollars)						
Average	0.032***	0.091***	0.100***	0.100***	0.126***	0.162***	0.160***
	(0.005)	(0.014)	(0.014)	(0.014)	(0.029)	(0.028)	(0.029)
Issuer days	35,635	8,909	8,909	8,909	3,563	3,564	3,564

Panel B: Determinants of variation in outperformance (no-coupon relative to full-coupon recovery)

	Model 1			Model 2		
Maturity	0.031***	-0.011		0.015***	0.003	
	(0.005)	(0.009)		(0.002)	(0.003)	
Coupon	0.009	-0.021**		0.013**	0.006***	
	(0.008)	(0.009)		(0.006)	(0.002)	
Def Prob	6.378***	1.658		0.309	-0.634**	
	(1.867)	(1.028)		(0.197)	(0.282)	
Recovery				0.069***	0.032**	
				(0.011)	(0.015)	
Miss error		0.050*			-0.007	
		(0.028)			(0.018)	
Miss err SD		0.312***	0.373***		0.068***	0.057***
		(0.050)	(0.047)		(0.025)	(0.013)
R-square	0.371	0.606	0.584	0.187	0.415	0.382
Issuer days	35,635	35,635	35,635	35,635	35,635	35,635

Table 6: Fit and parameters when using historically-estimated default proability as an input

This table reports results when using historically-estimated default probability as an input and there is an illiquidity effect on all cash flows (see text for additional detail). As before, recovery rate is constrainted to lie between 0.1 and 0.8; the effect of illiquidity lies between 0 and -5%. Estimation is done minimizing the volume-weighted squared pricing error. All data are reported at the issuer-day level. Mean absolute error (MAE) is the no-coupon recovery volume-weighted error (in dollars) for a given issuer day. MAE diff is the difference between the full-coupon recovery and the no-coupon recovery model MAE. Avg miss error is the average misspecification error, Default probability (annual) is the maturity-weighted historically-estimated default probability (from Kamakura Risk Information Services division of SAS Institute), Illiquidity and Recovery rate are both fitted. In Panel B we report model fit statistics for a subsample of issuer-day observations for the subsample with the top 25 percent of within issuer-day misspecification error standard deviation, as well as for four rating groups.

Panel A: Model 3 (fitted recovery rate and illiquidity, historically-estimated PD)						
	Mean abs error (MAE, in dollars)	MAE difference (in dollars)	Avg Miss error (in dollars)	Default Probability	Illiquidity	Recovery rate
Mean	0.29	0.02	0.21	1.3%	-0.42%	0.49
p5	0.00	-0.03	0.00	0.1%	-1.50%	0.10
p50	0.15	0.00	0.05	0.7%	-0.26%	0.54
p95	1.01	0.15	0.97	4.2%	0.00%	0.80
Number of issuer days: 35,635						
Panel B: Subsamples -- high misspecification error standard deviation, rating groups						
Top quartile miss SD	0.60	0.07	0.70	3.1%	-0.65%	0.70
AA, above	0.17	0.01	0.06	0.6%	-0.27%	0.48
A	0.24	0.02	0.18	1.0%	-0.29%	0.49
BBB	0.42	0.02	0.29	2.0%	-0.48%	0.51
BB, below	0.43	0.02	0.43	2.1%	-1.37%	0.49

Table 7: Model outperformance based on out-of-sample fitting

This table reports results when the three models are fit out-of-sample. Each issuer day, we choose one half of the observations (based on maturity rank) for estimation and calculate out-of-sample prices for the other half. The exercise is then repeated using the other half of the observations so that each issuer-day has a full set of out-of-sample prices, i.e. model prices that are calculated from a sample of observations not including the one for which we compare model to actual price. We require a minimum of four observations for each issuer day ensuring that parameters are calculated based on a minimum of two prices. Panel A reports summary statistics for the full sample and for the top quartile of misspecification error standard deviation. Panel B reports means and standard errors when regressing no-coupon recovery model outperformance (i.e. differences in volume weighted mean absolute errors) on a constant, using standard errors that are robust and double clustered by issuer and date (same as in Table 5). Panel C reports average outperformance across four rating groups.

Panel A: Out-of-sample outperformance (in dollars) of no-coupon recovery model

	Model 1		Model 2		Model 3	
Sample	full	High Miss error SD	full	High Miss error SD	full	High Miss error SD
Mean	0.09	0.24	0.02	0.05	0.04	0.14
median	0.04	0.18	0.01	0.03	0.01	0.10
SD	0.23	0.43	0.14	0.27	0.17	0.31
No. of issuer days	18,708	4,677	18,708	4,677	18,708	4,677

Panel B: Statistical significance of average outperformance (in dollars, full and subsample)

	Model 1		Model 2		Model 3	
Sample	full	High Miss error SD	full	High Miss error SD	full	High Miss error SD
	0.087***	0.236***	0.023***	0.045**	0.043***	0.143***
	(0.016)	(0.038)	(0.006)	(0.019)	(0.015)	(0.036)

Panel C: Average out-of-sample outperformance (in dollars) across rating subsamples

	Model 1	Model 2	Model 3
AA, above	0.04	0.01	0.02
A	0.08	0.01	0.05
BBB	0.10	0.03	0.04
BB, below	0.30	0.10	0.06

Online Appendix to "The Valuation of Corporate Coupon Bonds"

A. Internet Appendix

1. The Bond Valuation Formula

This appendix formalizes the pricing model in Section 3. We consider a continuous trading model on a finite horizon $[0, T^*]$. The uncertainty in the model is characterized by a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T^*]}, \mathbb{P})$ where the filtration $(\mathcal{F}_t)_{t\in[0,T^*]}$ satisfies the usual hypotheses and $\mathcal{F} = \mathcal{F}_{T^*}$. See Protter (2005) for the definitions of these various terms. Here $\mathbb{P}$ is the statistical probability measure.

The default-free money market account earns interest continuously at the default-free spot rate of interest, r_t, which is adapted to $(\mathcal{F}_t)_{t\in[0,T^*]}$. As in the text, money market account's time t value is

$$B_t = e^{\int_0^t r_s ds}.$$

Of course, we assume the necessary measurability and integrability such that the following expression is well-defined. The default-free zero-coupon bond, denoted by $p(t,T) > 0$, is adapted to $(\mathcal{F}_t)_{t\in[0,T^*]}$.

We consider a firm that issues a bond with a coupon of C dollars, a face value equal to L dollars, and a maturity date T. The bond pays the C dollar coupons at intermediate dates $\{t_1, ..., t_m = T\}$, but only up to the default time τ. For notational convenience, let the current time $t = t_0$. If default happens in the time interval $(t_{k-1}, t_k]$, then the bond pays a stochastic recovery rate of $\delta_{t_k} \in [0,1]$ at time t_k on the notional of L dollars. It is important to note that default can happen anytime within this interval, but the payment only occurs at the end. If default does not happen, the face value of L dollars is repaid at time T.

Let $\Gamma_t = (\Gamma_1(t), \ldots, \Gamma_n(t))' \in \mathbb{R}^n$ be a collection of stochastic processes characterizing the state of the firm and the market at time t with $\mathcal{F}_t^{\Gamma}$ representing the filtration generated by the state variables Γ_t up to and including time $t \geq 0$. We assume that these state variables are adapted to $\mathcal{F}_t$, which implies that $\mathcal{F}_t^{\Gamma} \subset \mathcal{F}_t$. We assume that r_t is $\mathcal{F}_t^{\Gamma}$ - measurable.

Let $\lambda : [0, T^*] \times \mathbb{R}^n \longrightarrow [0, \infty)$, denoted $\lambda_t = \lambda_t(\Gamma_t) \geq 0$, be jointly Borel measurable with $\int_0^{T^*} \lambda_t(\mathit{\Gamma}_t)dt < \infty$ a.s. $\mathbb{P}$, and let $N_t \in \{0, 1, 2, \cdots\}$ with $N_0 = 0$ be a Cox process conditioned on $\mathcal{F}_T^{\Gamma}$ with $\lambda_t(\mathit{\Gamma}_t)$ its intensity process (see Lando (1998)). Finally, let the default time $\tau \in [0, T^*]$ be the stopping time adapted to the filtration $\mathcal{F}_t$ defined by

$$\tau \equiv inf\left\{t > 0 : \; N_t = 1\right\}.$$

The function $\lambda_t(\mathit{\Gamma}_t)$ is the firm's default intensity. A Cox process is a point process which, conditional upon the information set generated by the state variables process Γ_t over the entire trading horizon $[0, T^*]$ behaves like a standard Poisson process. In particular,

$$\mathbb{P}(\tau > \mathcal{T} \left| \mathcal{F}_{T^*}^{\Gamma} \vee \mathcal{F}_t \right.) = e^{-\int_t^{\mathcal{T}} \lambda_u du}$$

and

$$\mathbb{P}(\tau > \mathcal{T} \left| \mathcal{F}_t \right.) = E^{\mathbb{P}}\left[e^{-\int_t^{\mathcal{T}} \lambda_u du} \left| \mathcal{F}_t \right.\right].$$

We assume that the markets are arbitrage-free.

Assumption. (Existence of an Equivalent Martingale Measure)

There exists an equivalent probability measure $\mathbb{Q}$ such that

$$\frac{p(t, \mathcal{T})}{B_t} \; for \; all \; \mathcal{T} \in [0, T^*] \quad and \quad \frac{v_t}{B_t}$$

are $\mathbb{Q}$ martingales.

It is well known that this assumption implies that the market is arbitrage-free. See Jarrow and Protter (2008). Given an equivalent martingale measure $\mathbb{Q}$, define $\widetilde{\lambda}_t \equiv \lambda_t \kappa_t$ to be the intensity process of the Cox process under $\mathbb{Q}$ where $\kappa_t(\omega) \geq 0$ is a predictable process with $\int_0^{T^*} \lambda_t(\Gamma_t)\kappa_t dt < \infty$ a.s. $\mathbb{P}$ (see Bremaud (1980), p. 167). The process $\kappa_t(\omega)$ represents a default jump risk premium.

We add the following assumption to simplify the analytic formulas.

Assumption. (Conditional Independence)
The default-free spot rate r_t, the default time τ, and the recovery rate process δ_t are independent under $\mathbb{Q}$ given the filtration $\mathcal{F}_t$ for all t.

Denote the time $t \leq t_1$ value of the coupon bond as v_t, then using risk neutral valuation yields

$$\begin{aligned} v_t = \textstyle\sum_{k=1}^m C E^{\mathbb{Q}}\left[1_{\{\tau > t_k\}} e^{-\int_t^{t_k} r_u du} \,|\mathcal{F}_t\right] + L E^{\mathbb{Q}}\left[1_{\{\tau > T\}} e^{-\int_t^{T} r_u du} \,|\mathcal{F}_t\right] \\ + \textstyle\sum_{k=1}^m L E^{\mathbb{Q}}\left[1_{\{t_{k-1} < \tau \leq t_k \leq T\}} \delta_{t_k} e^{-\int_t^{t_k} r_u du} \,|\mathcal{F}_t\right] \end{aligned}$$

where $\mathbb{Q}$ denotes the risk-neutral probabilities.

Using the independence assumption, this implies

$$E^{\mathbb{Q}}\left[1_{\{\tau > t_k\}} e^{-\int_t^{t_k} r_u du} \,|\mathcal{F}_t\right] = E^{\mathbb{Q}}\left[1_{\{\tau > t_k\}} \,|\mathcal{F}_t\right] E^{\mathbb{Q}}\left[e^{-\int_t^{t_k} r_u du} \,|\mathcal{F}_t\right]$$

and

$$E^{\mathbb{Q}}\left[1_{\{t_{k-1} < \tau \leq t_k \leq T\}} \delta_{t_k} e^{-\int_t^{t_k} r_u du} \,|\mathcal{F}_t\right]$$

$$= E^{\mathbb{Q}}\left[1_{\{t_{k-1} < \tau \leq t_k \leq T\}} \,|\mathcal{F}_t\right] E^{\mathbb{Q}}\left[\delta_{t_k} \,|\mathcal{F}_t\right] E^{\mathbb{Q}}\left[e^{-\int_t^{t_k} r_u du} \,|\mathcal{F}_t\right].$$

Thus, the previous equation simplifies to

$$\begin{aligned} v_t = Lp(t,T)\left[1 - Q(t,T)\right] + (C - Ld_t) \textstyle\sum_{k=1}^m p(t,t_k)\left[1 - Q(t,t_k)\right] \\ + Ld_t \textstyle\sum_{k=1}^m p(t,t_k)\left[1 - Q(t,t_{k-1})\right] \end{aligned}$$

where

$$d_t = E^{\mathbb{Q}}\left[\delta_\tau \left|\mathcal{F}_t\right.\right]$$

Proof. The only difficult term is $E^{\mathbb{Q}}\left[1_{\{t_{k-1}<\tau\leq t_k\leq T\}}\delta_{t_k}e^{-\int_t^{t_k} r_u du}\left|\mathcal{F}_t\right.\right]$. We have by conditional independence that

$$\begin{aligned} &E^{\mathbb{Q}}\left[1_{\{t_{k-1}<\tau\leq t_k\leq T\}}\delta_{t_k}e^{-\int_t^{t_k} r_u du}\left|\mathcal{F}_t\right.\right] \\ &= E^{\mathbb{Q}}\left[1_{\{t_{k-1}<\tau\leq t_k\leq T\}}\left|\mathcal{F}_t\right.\right]E^{\mathbb{Q}}\left[\delta_{t_k}\left|\mathcal{F}_t\right.\right]E^{\mathbb{Q}}\left[e^{-\int_t^{t_k} r_u du}\left|\mathcal{F}_t\right.\right]. \end{aligned}$$

Using the definition of d_t and $p(t, t_k)$ gives

$$\begin{aligned} v_t = &\textstyle\sum_{k=1}^m Cp(t,t_k)E^{\mathbb{Q}}\left[1_{\{\tau>t_k\}}\left|\mathcal{F}_t\right.\right] + Lp(t,T)E^{\mathbb{Q}}\left[1_{\{\tau>T\}}\left|\mathcal{F}_t\right.\right] \\ &+\textstyle\sum_{k=1}^m Lp(t,t_k)d_tE^{\mathbb{Q}}\left[1_{\{t_{k-1}<\tau\leq t_k\leq T\}}\left|\mathcal{F}_t\right.\right]. \end{aligned}$$

□

To simplify expression (**??**), we introduce two simple credit derivatives: a survival and default digital. These building blocks were first discussed in the literature by Madan and Udal (1998).

At time t, consider the time interval $[t_{k-1}, t_k]$ for $k = 0, \ldots, m$. Note that when $k = m$, $t_m = T$. We define a survival and a default digital as follows

- (A Survival Digital) This security pays \$1 at time t_k only if default occurs after t_k. The value of this security at time $t < t_k$ is

$$z(t, t_k) = E^{\mathbb{Q}}\left[1_{\{\tau>t_k\}}e^{-\int_t^{t_k} r_u du}\left|\mathcal{F}_t\right.\right].$$

- (A Default Digital) This security pays \$1 at time t_k if default occurs within $(t_{k-1}, t_k]$. The value of this security at time $t < t_k$ is

$$x(t, t_k) = E^{\mathbb{Q}}\left[1_{\{t_{k-1}<\tau\leq t_k\}}e^{-\int_s^{t_k} r_u du}\left|\mathcal{F}_t\right.\right].$$

Using the conditional independence assumption, we can write these as

$$z(t, t_k) = p(t, t_k)[1 - Q(t, t_k)]$$

and

$$x(t, t_k) = p(t, t_k)[Q(t, t_{k+1}) - Q(t, t_k)].$$

The identity $1_{\{t_{k-1}<\tau\}} = 1_{\{t_{k-1}<\tau\leq t_k\}} + 1_{\{\tau>t_k\}}$ implies

$$E^{\mathbb{Q}}\left[1_{\{t_{k-1}<\tau\}}e^{-\int_s^{t_k} r_u du}|\mathcal{F}_t\right] = E^{\mathbb{Q}}\left[1_{\{\tau\leq t_k\}}e^{-\int_s^{t_k} r_u du}|\mathcal{F}_t\right]$$

$$+E^{\mathbb{Q}}\left[1_{\{\tau>t_k\}}e^{-\int_s^{t_k} r_u du}|\mathcal{F}_t\right],$$

which simplifies to the given expression:

$$p(t, t_k)[1 - Q(t, t_{k-1})] = x(t, t_k) + z(t, t_k).$$

The left side represents the present value of the t_k maturity zero-coupon bond at time t, which is received only if there is no default before or at time t_{k-1}. The right side is the present value of the default digital for the interval $(t_{k-1}, t_k]$ and the survival digital with maturity t_k.

Using these two digitals, the valuation expression can be written as

$$v_t = \sum_{k=1}^m Cz(t, t_k) + Lz(t, T)$$
$$+Ld_t \sum_{k=1}^m x(t, t_k).$$

This expression shows that a risky coupon bond can always be decomposed into a portfolio of survival and default digitals.

2. The Recovery Rate Futures Price

This section explains why d_t can be interpreted as a futures price. Define

$$d_t := E^{\mathbb{Q}} \left[\delta_{T^*} \left| \mathcal{F}_t \right. \right]$$

for $t \in [0, T^*]$ where δ_{T^*} corresponds to the recovery rate if default happens at time T^*. That this can be interpreted as a futures price follows from the commodities pricing literature (for example see Jarrow (2021), chapter 6). Since d_t is a $\mathbb{Q}-$ martingale, the recovery rate futures price is equal to

$$d_t = E^{\mathbb{Q}} \left[\delta_\tau \left| \mathcal{F}_t \right. \right],$$

which is the risk adjusted expected value of the recovery rate at the default time τ. The relation between the recovery rate process δ_t and the futures price d_t is given by the following expression

$$\delta_t = d_t E^{\mathbb{Q}} \left[\frac{B_t}{B_\tau} \left| \mathcal{F}_t \right. \right].$$

Proof.

Define the recovery rate process δ_t on $[0, T^*]$ as the liquidation value of the firm at time t. Using the notation in the text, the following time line documents relevant dates.

$$\begin{array}{ccccccc} 0 \cdots & t \cdots & t_k & \tau \cdots & t_{k+1} \cdots & \tau^* \cdots & T^* \\ & & & default & & payment & \\ (d_0, \delta_0) & (d_t, \delta_t) & & (d_\tau, \delta_\tau) & & (d_{\tau^*}, \delta_{\tau^*}) & (d_{T^*} = \delta_{T^*}) \end{array}$$

Both τ and τ^* are stopping times with $\tau^* > \tau$. τ is the firm's default time. Due to cross-defaulting provisions, if one liability defaults, all default at the same time. τ^* corresponds to the date the payments on all of the firm's liabilities are paid, after liquidation or financial restructuring. It occurs, as indicated, after the default date. Time T^* is when the model ends. The arbitrage free value of the recovery rate payment at time $t \in [0, T^*]$ is

$$\delta_t = E^{\mathbb{Q}}\left[\frac{\delta_{T^*}}{B_{T^*}} \left| \mathcal{F}_t\right.\right] B_t.$$

Using iterated expectations, we obtain

$$\delta_t = E^{\mathbb{Q}}\left[\frac{\delta_\tau}{B_\tau} \left| \mathcal{F}_t\right.\right] B_t.$$

Under the conditional independence of δ_t and r_t, we can rewrite this as

$$\delta_t = E^{\mathbb{Q}}\left[E^{\mathbb{Q}}\left[\delta_\tau \left| \mathcal{F}_t \vee \tau\right.\right] E^{\mathbb{Q}}\left[\frac{B_t}{B_\tau} \left| \mathcal{F}_t \vee \tau\right.\right] \left| \mathcal{F}_t\right.\right]$$

$$= \sum_{i=1}^{M} E^{\mathbb{Q}}\left[\delta_{t_i} \left| \mathcal{F}_t; t_{i-1} < \tau \leq t_i\right.\right] E^{\mathbb{Q}}\left[\frac{B_{t_i}}{B_\tau} \left| \mathcal{F}_t; t_{i-1} < \tau \leq t_i\right.\right] Prob^{\mathbb{Q}}\left[t_{i-1} < \tau \leq t_i \left| \mathcal{F}_t\right.\right]$$

$$= \sum_{i=1}^{M} E^{\mathbb{Q}}\left[\delta_{t_i} \left| \mathcal{F}_t; t_{i-1} < \tau \leq t_i\right.\right] E^{\mathbb{Q}}\left[\frac{B_{t_i}}{B_\tau} \left| \mathcal{F}_t; t_{i-1} < \tau \leq t_i\right.\right] \left[Q(t, t_i) - Q(t, t_{i-1})\right]$$

where $t_M = T^*$ and $Q(t, t_i) = E^{\mathbb{Q}}\left[1_{\tau \leq t_i} \left| \mathcal{F}_t\right.\right] = Prob^{\mathbb{Q}}\left[\tau \leq t_i \left| \mathcal{F}_t\right.\right]$.

Recall that all payments occur at the next payment date after default.

Using the conditional independence of δ_t and r_t from τ, we can rewrite this as

$$\delta_t = \sum_{i=1}^{M} E^{\mathbb{Q}}\left[\delta_{t_i} \left| \mathcal{F}_t\right.\right] E^{\mathbb{Q}}\left[\frac{B_t}{B_{t_i}} \left| \mathcal{F}_t\right.\right] \left[Q(t, t_i) - Q(t, t_{i-1})\right].$$

Last, using the definition of d_t and noting that $E^{\mathbb{Q}}\left[\frac{B_t}{B_{t_i}} \left| \mathcal{F}_t\right.\right] = p(t, t_i)$ we get

$$\delta_t = d_t \sum_{i=1}^{M} p(t, t_i) \left[Q(t, t_i) - Q(t, t_{i-1})\right].$$

We note that this implies

$$\delta_t = d_t E^{\mathbb{Q}} \left[\frac{B_t}{B_\tau} |\mathcal{F}_t \right].$$

End of proof.

Given this expression, it is easily seen that the recovery rate futures price is strictly greater than the recovery rate δ_t since $E^{\mathbb{Q}} \left[\frac{B_t}{B_\tau} |\mathcal{F}_t \right] < 1$. This difference is expected to be small since interest rates are small over our sample period.

3. Misspecification Error Approximation

In the event of default, the present value of the payoff for the first coupon is equal to the discounted value of the product of the coupon rate, the recovery value, and the probability of default, i.e. $Cd_t p(t, t_1) Q(t, t_1)$. For the second coupon, default can occur either in the first or in the second period. The present value of the payoff in period one is the same as for the first coupon. The conditional expectation at time one of the payoff in period two is also the same as for the first coupon. However, this payment still needs to be discounted back to time zero and adjusted to take into account that the firm must have survived to period one. For longer time periods, a similar logic applies. This implies that the misspecification error is exactly proportional to the product of the coupon payment and the recovery value. In contrast, the total misspecification error is only approximately proportional to the error resulting from the first coupon. The reason is that the probability of survival depends on the default probability and because payments farther into the future will be discounted by a larger amount.

Next we consider the effect of the bond's maturity or, equivalently, the number of coupon payments. Ignoring discounting and the adjustment for the probability of survival, the expected recovery payment from the second coupon is twice as large as for the first coupon. The reason is that default on the second coupon can happen either in the first or the second period. For a bond receiving m coupon payments the factor is therefore $m(m+1)/2$. This means that the approximate total error is equal to $Cd_t p(t,t_1)Q(t,t_1)m(m+1)/2$.

4. Robustness Tests

This section provides several robustness tests of the model's outperformance.

1. Errors and Biased Parameter Estimates in the Full-Coupon Recovery model: A Quasi-Simulation

Another way to check that our results are due to the misspecified assumptions underlying the full-coupon recovery model is to calculate pricing errors using a simulation. We do this using the actual data as the basis for a quasi-simulation. Specifically, we assume that the no-coupon recovery model's parameters are unbiased and calculate market prices based on that model. We then use the full-coupon recovery model to price these bonds. This enables us to trace out the effects of using the full-coupon recovery model's misspecified assumptions on model prices and parameters.

We find that the expected biases appear. The misspecification error's standard deviation is closely related to the pricing error. Thus, whenever the full-coupon recovery model has more difficulty using the parameters to fit the data, the no-coupon recovery model's outperformance is larger, just as we observed in the main sample. We also find that the full-coupon recovery parameter estimates are indeed biased relative to the no-coupon recovery model's parameters, which in this exercise are assumed to be the actual underlying parameters.

In Model 1 the default probability has a downward bias; the effect is similar to what we see in the main sample. In Model 3 the recovery rate has a downward bias, as we would expect. The full-coupon recovery model's estimated recovery rate is 3% lower for the full sample and 4.5% lower for the subsample with the highest quartile of the misspecification error's standard deviation. For the illiquidity parameter the effect is much smaller. For the full sample, the bias is equal to 0.02% (full-coupon model-implied illiquidity is more negative), while it is 0.06% for the top 25% of the misspecification error's standard deviation subsample. In short, the patterns we find in the data are directly driven by the full-coupon model's assumptions.

For Model 2, the biases are larger. The recovery rate is 28% too large, while the default probabilities are on average 1.4% too large. This model's instability underscores the need for more restricted models, namely Model 1 (restricted recovery rate) and Model 3 (historically-estimated default probabilities).

2. Daily Observation Cutoffs

In our estimation, we group bonds each issuer day and then fit the two pricing models. We require a minimum of two bond price observations for each issuer day. Another possibility is to choose minimum observation cutoffs, for example requiring at least five or ten observations for each issuer day. Such a restriction has the benefit of reducing noise but shrinks the sample size. We have checked that our results are robust to increasing the minimum required number of observations.[1]

3. Monthly Observations

An alternative way to reduce noise is to pool observations each issuer month rather than each issuer day. This approach significantly increases the number of observations in each group. It also has the benefit of not reducing the overall sample size. Our results are robust to this change too. In particular, we see a similar level of outperformance of the no-coupon recovery model, though both models fit less well since there is now less parameter flexibility. Our findings suggest that outperformance is not directly linked to sample size. This result is consistent with our findings that the misspecification error's variance explains well the variation in model outperformance.

4. Alternative Recovery Rate Estimation Bounds

[1] In the case of Model 3, the no-coupon recovery model outperformance increases substantially when restricting attention to the sub-sample of issuer days with ten or more observations; however it cuts the number of issuers to only 18.

As pointed out, our data set contains a lot of financial institutions, partly because they issue a lot of debt, which is then actively traded in the secondary market. Regulated financial institutions may have higher recovery rates than other firms since the regulator may close or seize the bank before losses increase. It is therefore possible that, for at least a subset of the data, recovery rates lie above the 80% cutoff we impose. We have checked what happens if we increase the upper bound on recovery rate to 95%. This increased flexibility may or may not be useful for Model 2, which, as we have discussed, is quite flexible already. However, we may see an effect for Model 3, and that is indeed what we find. No-coupon recovery model outperformance is higher for both models. This evidence suggests that for a subsample of the data issuers have high expected levels of recovery rates. When using this less restrictive estimation for the out-of-sample estimation, Model 2 outperformance drops while Model 3 outperformance increases. This evidence supports our conjecture that Model 2 is somewhat unstable, while it supports the use of restrictions (Models 1 and 3).