Machine Learning Predictions of Credit and Equity Risk Premia

By Arben Kita *

The emergence of algorithmic high-frequency trading in the market for credit risk affords accurate inference of new risk measures. When combined with machine learning predictive methods, these measures forecast substantial future changes in firms' credit and equity risk premiums in out-of-sample. Parallel measures estimated from firms' stocks fail to predict risk premiums, indicating that credit-market-based risk measures contain valuable information for forecasting firms' risk premia in both markets. The innovative high-volume high-frequency trading has not alleviated short-horizon pricing deviations across firms' equity and credit markets, an epitome of latent arbitrage in the market for credit risk.

JEL: C52, C55, C58, G0, G12, G14, G17

Keywords: Stochastic Volatility, Jumps, Latent Arbitrage, Machine Learning, High Frequency Data

* Kita: University of Southampton, Highfield Campus, Southampton, SO17 1BJ, UK, A.kita@soton.ac.uk.

I. Introduction

The premise of the consumption-based asset pricing models is that markets are fully integrated and that the price of risk is rationally determined and fluctuates countercyclically.[1] The market integration assumption is also fundamental in Merton's (1974) structural model of credit risk. In this model, variations in stock prices and credit spreads must relate precisely to prevent arbitrage. Yet, an active and persistent trading theme in the markets exploits the pricing discrepancies between firms' equity and credit markets. However, the presence of these active arbitrageurs, combined with the market integration hypothesis, should, in theory, ensure that the movements between firms' equity and debt markets are closely related. Failure to find these close links has substantial implications for asset pricing theory and poses practical challenges for risk management.

An important innovation in everyday trading activities is the advent of algorithmic high-frequency trading (HFT). The main advantage of these machine learning (ML) trading methods is that they create high liquidity and mitigate market fragmentation effects. They also assure against human bias, and because they are based on high volumes, they contribute to the price discovery and price formation processes. While the stock market's liquidity is commonly understood, a recent phenomenon in the market for credit risk is a significant increase in "naked trading."[2] This trading allows investors unaffected by the short-selling constraints to express their views on firms' future credit soundness. This practice has enabled a significant increase in trading activity and created a wealth of data.[3]

Therefore, this context provides a beneficial setting to pursue two main inquiry lines not previously studied in the literature. First, can the availability of extensive data and machine learning techniques afford robust predictions of future changes in credit risk premiums? Second, has the advent of automated ML trading and the active capital structure arbitrage eased the market fragmentation? The first question has significant practical risk-management consequences and contributes to our understanding of credit behavior, while

[1] Campbell and Cochrane (1999) present a habit formation process of time-varying risk aversion, while Chen, Collin-Dufresne and Goldstein (2008) discuss how habit formation can explain the low level of defaults if defaults are countercyclical. Bansal and Yaron (2004) look at time-varying consumption volatility, and Barro (2006) and Gabaix (2012) study the time-varying consumption disaster risk.

[2] The naked trading of credit default swap (CDS) consists of speculators buying protection against a default of the firm without owning the underlying credit or bond. These types of positions, who are unlimited, are also known as synthetic.

[3] On February the 4th, 2016– which coincides with the end date of this sample– Financial Times reported that "a record $15.7 bn in gross notional outstanding positions of single name CDS was cleared by investors during Janurary according to the Intercontinental Exchange, the largest credit derivative clearing house". See https://www.ft.com/content/c47dce8e-ca9f-11e5-be0b-b7ece4e953a0.

the second question bears implications for capital theory and policymaking.

Traditional credit risk literature proposes structural models that price corporate securities as claims on a firm's assets.[4] Such models' success is measured by their ability to produce credit spreads that closely mimic the market-observed ones.[5] Recently, Du, Elkamhi and Ericsson (2019), via simulations, have made significant progress in explaining the levels and the changes in the credit spreads.[6] However, such studies cannot offer insights into predictability or provide accurate estimates of the underlying asset value dynamics, crucial for credit risk management and asset pricing, which is the focus of this article.

The importance of modeling the dynamics of credit risk factors is crucial for risk management. For example, in hedging portfolio credit derivatives, the risk is characterized in terms of sensitivities to shifts in risk factors. Typical hedging practices based on such measures of sensitivity have been to "delta-hedge" the changes due to spread fluctuations. However, the lack of well-defined dynamics for risk factors in such static models has proven inefficient for credit risk management during the recent financial crisis.[7] Besides, traditional hedging strategies of credit risk ignore jump risks in the spreads, critical for risk management during the distressed credit market in 2008.

The pricing dynamics of spreads are also important for asset pricing. Models for pricing credit derivatives have traditionally used log-normal approaches to describe changes in credit spreads. Such models' implicit assumption is that changes in spreads are proportional to spreads and that relative spread changes are normally distributed. The presence of (extreme) price movements in the risk premium dynamics reported in this study evinces these models' limitations. Time-varying realized volatility risk is a state variable that drives risk premium dynamics and predicts much of the credit spread changes in out-of-sample. The jump risk measures' contribution further enhances the model's performance in capturing the spreads' pricing dynamics.

The dependence of the asset growth process on state dependencies introduces stochastic volatility in the credit spreads. Empirically, to capture the pricing dynamics of credit

[4] See Collin-Dufresne, Goldstein and Martin (2001), Acharya and Carpenter (2002) for early studies and others thereafter.

[5] See Cremers, Driessen and Maenhout (2008), Huang and Huang (2012), among others.

[6] Christoffersen, Du and Elkamhi (2017) and McQuade (2018) have recently studied structural models with stochastic volatility and jumps and Ait-Sahalia and Kimmel (2007) use polynomial expansions. In contrast, earlier studies looked at option-implied jumps (Cremers, Driessen and Maenhout 2008), equity market implied jumps (Zhang, Zhou and Zhu 2009), and jumps in the US Treasury market (Tauchen and Zhou 2011). None of these studies focuses on the firm-level market for credit risk or relies on this market's high-frequency data to imply the risk parameters, allowing for another innovation of this article to add to this body of work.

[7] In fact, delta-hedging of spread risk is not basd on theory of derivative replication.

spreads, the literature has traditionally used square root factors that impose parametric restrictions (admissibility conditions) to obtain positive conditional variance over a range of state-vectors. However, this method affects the correlations among the factors and worsens the cross-sectional fit of the model. In contrast, constant-volatility Gaussian models with no square root factors do not restrict the signs and magnitude of the conditional and unconditional correlations among the factors. Still, neither do they accommodate the pronounced and persistent volatility fluctuations observed in credit spreads.

This study contributes to the literature by proposing a statistical analysis based on nonparametric estimates of volatility and jumps-risk measures in the asset growth process implied from the credit market information. These risk factors are then used to predict future changes in firms' credit and equity risk premia. A similar approach is used to recover the stock risk measures of the same firm. While this article focuses on predicting future credit risk variations, it also seeks to understand the proposed risk factors' economic importance. Traditional testing of the contribution of risk factors in explaining the asset prices is straightforward. Once the estimated loadings of the stochastic discount factors (SDF) on the risk factors and a set of control variables are evaluated, a test of whether the loadings of risk factors are different from zero reveals their significance for the asset prices. However, given the plethora of factors proposed in the literature due to the curse of dimensionality, evaluating the proposed risk measures' contribution via standard statistical methods is infeasible and results in unreliable estimates and invalid inference.

This article uses the least absolute shrinkage and selection operator (LASSO) to predict credit spread changes. Under certain conditions, this approach allows for selecting correct variables and helps reduce the dimensionality problem. Although this method determines the factors that most contribute to out-of-sample predictability, relying merely on LASSO for model selection, as pointed by Chernozhukov, Hansen and Spindler (2015), produces a poor approximation of estimators' finite-sample distributions and often results in omitted variable bias. Belloni, Chernozhukov and Hansen (2014b) proposed the double-selection LASSO to overcome this issue and obtain robust inference. Therefore, this study blends the ability of ML techniques to get the best out-of-sample predictions in the first instance. In the second step, the double ML includes a battery of theoretically important control variables presented in the credit literature to assess the statistical significance of the proposed risk measures. Finally, the economic importance of the risk factors is assessed in portfolio sorts.

This econometric approach is congruent with this paper's framework. The advent of automated ML trading has allowed for the emergence of a wealth of data. The expected effects of this abundance of high-frequency data are two-fold. On the one hand, they allow for more reliable estimates of volatility and jump risk factors and thus identify the market state variables useful in forecasting;[8] on the other hand, they should lessen the market fragmentation effects. Finally, ML approaches for prediction and inference are robust to nonlinearities and ensure that the results reported in this article are entirely data-driven.

The remainder of this article is organized as follows: Section II summarizes the data set. Section III describes the empirical methodology. Empirical results are presented in Section IV. Finally, Section V concludes. The theoretical motivation of the empirical analysis is reported in the Internet Appendix to save space. This appendix also provides the technical implementation details of the risk factors' estimation method and other results and robustness checks.

II. Data Set

The credit default swap (CDS) data are from ICE Credit Market Analysis Ltd (CMA) DataVision via the Bloomberg data service. The data set consists of five-minute intraday CDS prices for the five-year CDX NA IG index constituents. This CMA database is the most widely used database among financial market participants and the principal data source for Bloomberg-disseminated CDS prices; thus, arguably, these data contain fewer errors.[9] Mayordomo, Peña and Schwartz (2014) note that CMA collects its data from around 40 institutions that are active participants in the CDS market. These participants receive tens of thousands of Bloomberg-formatted pricing messages that are included in the CMA database. Consequently, these prices are very likely to be tradeable. They also find that these prices lead the price discovery process. The time period spans July 9, 2012, to July 9, 2016.

Since the accuracy of the risk estimates is paramount for the empirical investigation, the following steps help mitigate the market microstructure noise (e.g., concerns such as stale prices, discreteness of prices, and bid-ask spreads). Consistent with the literature first, only

[8]See Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2001).

[9]CMA quotes provided by Bloomberg have been extensively used in the literature (e.g., Das, Hanouna and Sarin (2009); Saretto and Tookes (2013); Das, Kalimipalli and Nayak (2014); Boehmer, Chava and Tookes (2015)).

the CMA data, which are aggregated across individual dealers, is used, thereby averaging out some of the noise. Second, sampling at the five-minute frequency helps to manage the effect of microstructure noise on volatility estimates (Hansen and Lunde 2006). Bandi and Russell (2006) calculate optimal sampling intervals in the presence of microstructure noise and find that “since many optimal sampling intervals are near 5 min, the loss is not substantial when a 5-min interval is used.” Similarly, Ait-Sahalia and Mykland (2005) show that more data do not necessarily lead to better estimates of the realized volatility, exactly because of the presence of market microstructure noise. Third, the bipower measures proposed in Huang and Tauchen (2005) ensure against the presence of autocorrelation in adjacent returns, a problem triggered by market microstructure noise. Finally, trading days with fewer than 40 observations of five-minute returns are removed. Repeated observations and obviously misreported prices are also deleted.

The empirical analysis controls whether the proposed credit risk factors help to predict future variations in the firms credit (equity) risk premia above and beyond the plethora of the factors presented in the credit risk literature. In response, the analysis includes a wide range of covariates. The data are divided into the following nine main categories, each containing several control factors.[10]

Credit and equity risk factors (RVC, RJV, $(+/-)RJV$) These are the main variables of interest. The firm-level realized volatility and realized jump risks based on credit and equity market information are estimated from the high-frequency data. We want to know whether these new risk factors can predict firms’ future changes of credit spread and equity risk premium above and beyond other factors presented in the credit literature. The intraday data come from Bloomberg.

Illiquidity risk (λ) This risk factor includes the firm-specific and market-wide illiquidity shocks based on credit and equity market information. Therefore, this measure includes firm-level credit and stock bid-ask spreads, Barclay’s liquidity cost scores, CDX index bid-ask spreads, Roll (1984) measures, and their expected shocks. The estimation procedure for the illiquidity risk is discussed in the internet appendix. Intraday and daily data come from Bloomberg.

Historical risk factors (ι) Traditionally the credit risk literature has inferred the

[10]The sampled entities, the data source, and definitions are discussed in the Internet Appendix B. Table IB.1 reports the firms’ names and provides their summary statistics. Table IB.2 discusses the economic intuition on the relation between the covariates employed in this article, and the credit spreads and equity.

firm-level volatility and jump risks from daily non-overlapping standard deviation, skewness, and kurtosis (-$VOLA_t$, -$KURT_t$, -$SKEW_t$) of credit (equity) returns. The credit risk measures have a C- prefix, and the equity counterparts have an E- prefix. The daily data come from Bloomberg.

OPTIONS MARKET (π) These include two control measures namely, the deep-out-of-the-money (DOTM) Put options of the underlying stock and the VIX index. While VIX, the fear index, captures markets' downside risk expectations and has been widely used in the credit literature, put options due to their similar contingent-claim payoffs as the CDS are studied in Cao, Yu and Zhong (2010). The daily data come from Bloomberg.

FIRM-CHARACTERISTICS (ϕ) The firm-level characteristics control variables containing the firm leverage ratio ($LEVERAGE$) estimated as the sum of current- and long-term debt divided by the sum of total equity and current- and long-term debt, and the firm's size, proxied as the log of market capitalization (SIZE), which controls for firm-size effects. In the structural models for credit risk, the default occurs when the leverage ratio nears unity while the size is commonly assumed as an inverse proxy for the expected cost of bankruptcy (Rajan and Zingales 1995). To capture the firm's value process, for robustness, regressions also control for individual firm's EQUITY RETURN. The daily and quarterly data come from Bloomberg.

ASSET RISK FACTORS (α) Asset risk factors include the asset volatility (σ_{ASSET}) estimated as in Kelly, Manzo and Palhares (2017) and the asset beta (β_{ASSET}), estimated similar to Schwert and Strebulaev (2014). These risk factors capture assets total and systematic risk. Within the structural framework, changes in the firm's asset value process affect its credit and equity pricing dynamics. Inferential tests account for this interaction. The daily and quarterly data come from Bloomberg.

BUSINESS RISK FACTORS ($\mathbf{B}$) These include, TOBIN'S Q, TANGIBILITY, calculated as a ratio of tangible equity over tangible assets, and, PROFITABILITY estimated as a ratio of net income over net revenue. Monthly data come from Bloomberg.

COMMON RISK FACTORS (κ) They include Fama-French three-factors MKT, SMB and HML; the excess return on the market portfolio and the return on two long/short portfolios that capture the size and book-to-market effects, respectively. Further controls include the default risk premia in the credit market, DEF (The difference between the 10-year Bloomberg Barclays US BBB Corporate Bond Index and the 10-year US Treasury rate), and the TERM (the difference between the returns on 10-year- and 3-month Treasury securities). The daily and monthly data come from Bloomberg and the Federal Reserve Economic database (FRED).

Macroeconomic factors (μ) Monthly S&P 500 returns, a proxy for the economy's overall state and captures the business climate changes. LEVEL is the 10-year Benchmark Treasury yields. The term structure's squared and cube level ($LEVEL^2, LEVEL^3$, respectively) accounts for potential nonlinear effects due to convexity. SLOPE is the difference between the 10-year and 2-year Treasury yields. The term structure of the interest rates is affected by the term structure's level and its slope. Since the interest rates affect the firm value, an increase in the Treasury curve slope increases the expected future short rates; therefore, it affects the spreads. The final macroeconomic variable is the SWAP rate, the difference between the 10-year US dollar swap rates and the 10-year US Treasury rate. This proxy for changes in the overall risk in the economy.

Table 1, Panel A, reports summary statistics for the monthly CDS spreads, volatility, jump and illiquidity risk measures implied from credit and equity markets, and the respective first-order autocorrelations divided into four rating groups, as well as for the whole sample. Risk measures based on the credit market information have the prefix C-, and those from the equity market have the prefix E-. Volatility series exhibit pronounced serial dependencies. In untabulated results, the first 10 autocorrelations of realized volatility are all highly significant with the gradual but very slow decay suggestive of long-memory features. This feature is also evident from the time series plots of the realized volatility series reported in Panel (A) on the left-hand side of Figure 1. The average monthly spread realized volatilities are 3.59%, 3.66%, and 3.68% for the investment-grade entities and peak at 4.75% for the noninvestment-grade entities. The right side of Panel (A) in Figure 1 reports the corresponding realized volatility measures of the same entities implied from the equity market. The average realized jump volatility, on the other hand, is more homogenous across the rating groups and drops to less than 1%.

An interesting result from Table 1 is that, while the credit implied volatility measures are lower than their equity counterparts, the credit market-based jump risk measures are higher than their corresponding equity market-based jump risk measures. This difference is also evinced by Figure 1, Panle (B) which plots a horse race comparison of the time series of realized jump volatility estimates implied from credit returns (left side) and equity returns (right side). This result provides initial insight into the different pricing patterns between the credit and equity market-based risk measures and warrants further investigation.

III. Empirical Method

The theoretical links between the asset value process and the credit and equity price processes and their variations are discussed in the Internet Appendix A. This discussion points to an essential aspect of a close relation between firms' assets volatility and the credit spread and stock pricing behavior.

A. *Credit Spread Return*

In a simple form, a corporate bond can be viewed as a combination of a risk-free bond and a risky asset that pays out an annual coupon and demands payment (net of recovery) if a credit event occurs. This risky asset accurately reflects the CDS contract; its payoff can be synthesized by going long on a corporate bond and shorting the riskless bond. Consider an investor who sells protection using a CDS contract j at time i-1, where i is the i^{th} intraday observation within trading day t at a CDS spread of $CDS_{j,t,(i-1)}$, paid quarterly. At time i within trading day t, the investor buys an offsetting contract at a spread $CDS_{j,t,i}$. The net cash flow generated prior to maturity or default is therefore $(CDS_{j,t,i} - CDS_{j,t,(i-1)})$. Since, as from the discussion in the previous section, spreads follow relative rather than absolute changes, the i^{th} within-day excess returns of contract j are modeled as

$$R_{j,t,i} = \frac{CDS_{j,t,i}}{250} - D_{CDS_{j,t,i}}(CDS_{j,t,i} - CDS_{j,t,(i-1)}), \tag{1}$$

where $D_{CDS_{j,t,i}}$ is the risky duration. These are returns to a CDS seller, and represent insurance sellers' accrued risk premia.[11] These returns are then combined with the risk-free and equity returns to estimate firm's daily asset returns. The assets monthly riskiness measures are used in the empirical analysis.[12]

[11] These returns are similar to Kelly, Manzo and Palhares (2017). When returns are estimated as absolute change in spreads as in Bongaerts, De Jong and Driessen (2011), results remain generally consistent.

[12] Asset returns are estimated as in Kelly, Manzo and Palhares (2017). The daily standard deviations of asset returns are aggregated into monthly frequency.

B. Realized Volatility and Jumps

The importance of extreme price movements and their source (diffusion- or jump-induced) for asset pricing has long been recognized (Merton 1976), though their estimation is not trivial when low-frequency data are employed.[13] Filtering out the jumps from volatility based on low-frequency discrete returns remains empirically challenging, rendering jump diffusion models less attractive in practical risk management. An important issue for hedging in practice is understanding the dynamics of credit spreads under the real-world measure. In hedging strategies based on pricing models, for example, it is not clear how well those models can capture the pricing dynamics of spreads under this measure.

Realized volatility estimations and the increase in HF trading in the market for credit risk afford a more precise estimation of volatility and jump risks. Bipower variation measures allow to separate realized volatility into continuous (RVC) and jump (RJV) components.[14]

Barndorff-Nielsen and Shephard (2004) propose two general measures of the quadratic variation process: realized variance, RV_t, and realized bipower variation, BV_t. Both RV_t and BV_t converge uniformly, as the intraday sampling frequency, Δ, $\Delta \to 0$ or $m = 1/\Delta \to \infty$, to different components of the underlying jump diffusion process

$$RV_{j,t,i} \equiv \sum_{i=1}^{m} (R_{j,t,i}^2) \to \int_{t-1}^{t} \sigma_s^2 ds + \sum_{i=1}^{m} (J_{j,t,i}^s)^2, \tag{2}$$

where the right side is the quadratic variation of the price over the time interval $[t, t-1]$. The term $R_{j,t,i}^2$ is the squared spread (equity) returns estimated in Equation (1). For increasingly fine-sampled increments ($m = 1/\Delta \to \infty$), realized volatility consistently estimates the total ex post variation of the price process. By decomposing the summation of the squared increments (of the quadratic variation) into separate summations of small and large price changes, it is possible to estimate the variation in the continuous sample price path

$$BV_{j,t,i} \equiv \frac{\pi}{2} \sum_{i=2}^{m} |R_{j,t,i}| \cdot |R_{j,t,i-1}| \to \int_{t-1}^{t} \sigma_s^2 ds. \tag{3}$$

The difference between the realized variance and the bipower variation is zero when there

[13] See Andersen, Benzoni and Lund (2002) and Ait-Sahalia (2004), among others.

[14] See Barndorff-Nielsen and Shephard (2004), Huang and Tauchen (2005), Barndorff-Nielsen (2006) and Andersen, Bollerslev and Diebold (2007). RVC or volatility and RJV or jump risks are used interchangeably throughout the paper to denote the short-run firm-specific volatility and jump risks.

is no jump and is strictly positive with a jump. The realized jump measure RJ_t is then

$$RJ_{j,t,i} \equiv \frac{RV_{j,t,i} - BV_{j,t,i}}{RV_{j,t,i}}. \tag{4}$$

To identify the jumps, this study uses Tauchen and Zhou (2011) extension of the "significant jump" approach of Andersen, Bollerslev and Diebold (2007), based on the signed square root of the significant jump:

$$J^s_{j,t,i} = sign(R^s_{j,t,i}) \times \sqrt{RV_{j,t,i} - BV_{j,t,i}} \times I(z > \Phi^{-1}_{\alpha}), \tag{5}$$

where Φ is the cumulative standard normal distribution function, α is the significance level of 0.99 of the z-test, and $I(z > \Phi^{-1}_{\alpha})$ is the indicator function pertaining to whether there is a jump during day t.[15] The approach proposed herein relies on the economic intuition that jumps in financial markets, and particularly in the credit market, are rare and large. This method, therefore, allows for credit market-based sentiment measure to be based on "observed" realized jumps without any assumptions on the jump distribution.

Once the realized jump volatilities have been estimated, based on the assumption that rare jumps have a dominant effect on returns on the day they occur, the direction of the jump, up or down ($(+)RJV$) or $(-)RJV$, respectively) is assigned based on the return direction of that day. Formally,

$$\begin{aligned} (+)RJV &= I\{r_t \geq 0\}RJV \\ (-)RJV &= I\{r_t < 0\}RJV. \end{aligned} \tag{6}$$

C. Empirical Analysis

Credit risk literature has proposed numerous factors, discussed in the data section, with varying theoretical justification to explain the credit spreads. Empirical works usually select ad hoc a handful of control factors to assess the contribution of the proposed factor.

[15] Internet Appendix B reports the implementation details. An alternative strategy to the "significant jumps" approach is the method proposed by Lee and Mykland (2008). Their model allows filtering out multiple jumps and has the desirable property to distinguish two jumps a day, with low and high variance. This approach is readily adaptable in the context of the equity market. However, the credit risk market's structural and technical factors, the asymmetric and intermittent nature of credit that display large peaks and lows, higher volatility, and illiquidity, make the one significant jump a day approach more suitable and provides for a reasonable comparison between the two markets. Also, recently popularized Hawkes self-excited processes are challenging to fit data.

The implicit assumption of such approaches is that the chosen model is the true data generating process (DGP) and that the model is robust to misspecification. Besides the doubtfulness of these assumptions, recently, credit risk literature employs highly-stylized models that explain the spreads' levels and variations via simulations. This literature is mostly silent regarding these models' ability to predict the changes of spreads, especially in out-of-sample. Besides the difficulty in predict *changes* in spreads as opposed to *levels*, the credit literature emphasizes the importance of changes in spreads since they represent the market's altered assessment of the credit risk.

LASSO proposed by Tibshirani (1996) is commonly used for prediction. This method minimizes the prediction error by including a penalty function in the least-squares optimization. This penalty increases as a function of the model's increased complexity, thereby imposing a sparsity condition and omitting variables. The least-square minimization is obtained by including a penalty value on the parameter estimates and choosing one of these solutions as the best solution based on entirely different criteria, that of the out-of-sample estimate error.[16] Due to regularization issues, Belloni et al. (2012) and Belloni and Chernozhukov (2013) propose a post-LASSO estimation, which runs LASSO for model selection and then refits the least-squares by including only the non-zero coefficients from the first step. The penalty function is based ("tuned") on the cross-validation (CV) method that minimizes estimates of the out-of-sample prediction error.[17]

In response, predictive regressions include all proposed factors in the literature. The aim is to use machine learning to evaluate whether the proposed risk factors based on credit market information alone matter for out-of-sample predictions above and beyond

[16] Formally, the optimization problem solved by linear lasso of point estimates $\widehat{\beta}$ is: $\widehat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \left\{ \frac{1}{2n} \sum_{i=1}^{n} \left(y_i - \mathbf{x}_i \boldsymbol{\beta}' \right)^2 + \lambda \sum_{j=1}^{p} \omega_j \left| \beta_j \right| \right\}$. Here, y is the outcome variable, β is the vector of coefficients of x, λ is the penalty parameter, ω_j are the penalty loadings, β_j is the j-th element of β, and n is the sample size. The first term is the least-squares and the last term is the penalty term. λ and ω_j are the "tuning" parameters that specify the weights applied to the penalty parameter. With $\lambda = 0$ LASSO reduces to OLS. As λ decreases from λ_{max} (for which all estimated coefficients are zero) the number of non-zero coefficients increases. Lasso excludes the covariates with zero coefficients.

[17] CV splits the sample into ten random K folds for a given penalty function. Once a fold is chosen a linear regression fits the other K-1 folds using the variables in the model for that penalty function. The prediction is then computed using these new estimates for the data of the chosen fold K. This process is repeated for the other K-1 folds. Finally, the average mean squared errors of the prediction across the folds are computed. For robustness, we also use the adaptive LASSO method. Adaptive LASSO runs multiple LASSOs. The first LASSO's chosen penalty parameter is used to construct the penalty weights from these coefficient estimates, which are then later used in the second LASSO where another penalty function is selected. The sample is split 50% training and 50% validation.

the plethora of factors already presented in the literature. The full prediction model is,

$$\begin{aligned}\delta y_{t,j} = c &+ \beta_1 \delta RVC^{'}_{t,j} + \beta_2 \delta RJV^{'}_{t,j} + \beta_3 \delta \lambda^{'} + \beta_4 \delta \iota^{'}_{t,j} + \beta_5 \delta \pi_{t,j} \\ &+ \beta_6 \delta \phi^{'}_{t,j} + \beta_7 \delta \alpha^{'}_{t,j} + \beta_8 \delta B^{'}_{t,j} + \beta_9 \delta \kappa^{'}_{t} + \beta_{10} \mu^{'}_{t} \\ &+ Firm\, FE + Month\, FE + Ratings\, FE + Sector\, FE + \epsilon_{t,j}. \end{aligned} \tag{7}$$

The dependent variable $y_{t,j}$ is the changes in credit or equity risk premia. The variables of interest, the volatility and jump measures from high-frequency CDSs and equity prices, are estimated as detailed in Section III.B and internet appendix. They also include the up-and-down jumps estimated in Equation (6). Other covariates, also discussed in data section II, include the firm-level illiquidity risk factor ($\lambda^{'}$), historical volatility and jump risk factors ($\iota^{'}$), the equity deep-out-of-the-money put option (π), firm characteristics ($\phi^{'}$), firm asset risk factors ($\alpha^{'}$), business risk factors ($\mathbf{B}$), common risk factors ($\kappa^{'}$), and finally macroeconomic factors ($\mu^{'}$). All regressions include firm (*Firm*), time (*Month*), credit rating (*Rating*), and industry (*Sector*) fixed effects. δ denotes the changes in the variables.

However, ML techniques are designed for prediction, and as discussed in Feng, Giglio and Xiu (2020), LASSO can omit critical economic factors when recovering the SDF. Furthermore, LASSO does not provide standard errors to assess the proposed risk factors' statistical validity. In response, this work uses the cross-fit partialing-out, also known as double machine learning method proposed by Chernozhukov et al. (2018). This method splits the sample into roughly two equal sizes; coefficients obtained from one sample are used in another, and therefore, is robust to model selection mistakes that LASSO makes. The inferential analysis thus uses the cross-fit partialing-out of the same specification, Equation (7).

By including a large set of covariates, this article also addresses the omitted factor critique. The double-ML approach adopted here also benefits from reporting estimates of values from the true model that generated the data. Empirical results in all tables report the out-of-sample R^2 of post-LASSO prediction regressions, and coefficients note the robust standard errors of the double-ML that account for within-cluster correlations that are not captured by the fixed effects.

IV. Empirical Results

In traditional empirical contingent claim analysis, most of the explanatory variables are jointly determined with credit spreads. Consequently, initial investigation focuses on the role of credit volatility and jump risk measures to predict future credit spread *levels* in pooled OLS regressions. The preliminary results are reported in the Internet Appendix Table AI. 1. Regressions use only lagged explanatory variables to account for the high persistence of credit spreads and avoid simultaneity problems that would artificially increase predictive power. Results evince the utility of the volatility and jump risk measures based on the credit market information for predicting future spread levels.

A. Comparing the Models: CDS and Equity Implied Volatility and Jump Risks

While the advantage of the proposed volatility and jump risk for predicting future spread levels is well founded in structural models theory (see discussions in Internet Appendix), three considerations can be made. First, while the credit literature recognizes the importance of predicting credit spreads, it also emphasizes the much more difficult task of forecasting *changes* in spreads. Changes in spreads are essential because they represent returns that investors earn from holding a position on a particular portfolio of spreads. Hence, they capture changes in the market's assessment of the credit risk of the underlying entity. Second, the proposed risk measures' successful predictive ability does not necessarily imply that they would perform equally well out of the sample. The highly stylized models commonly employed in credit literature, or the recent simulation approach, suffer from overfitting problems, which is especially critical in a nonlinear world of abrupt changes in spreads.[18] Third, relying on structural models, the literature so far has implied volatility and jump risks from equity returns or equity options; in contrast, the innovation of this work is to propose risk measures based on the emergence of HF credit market information alone. Thus, the credit market risk-based measures are theoretically well-founded. However, investors need to compare the performance of proposed credit risk measures with the traditional ones, including the HF stock and option-implied risk measures.[19] The

[18]Recent studies reported in Du, Elkamhi and Ericsson (2019) rely on simulations and invoke strong assumptions on the data generating process. Culp, Nozawa and Veronesi (2018) show via the use of the "pseudo bonds" (the difference between Treasury bonds and put options on firm assets) that the excessive tails risk and firm-specific asset risks are critical factors of corporate spreads. These works, however, do not focus on predictability.

[19]Cao, Yu and Zhong (2010) and Cremers, Driessen and Maenhout (2008) study the usefulness of stock options for credit spreads. Equity market-implied jumps are studied in Zhang, Zhou and Zhu (2009).

ability to predict future changes in credit spreads is undoubtedly essential for practical risk-management and investment purposes. On the other hand, the advent of high-volume high-frequency machine trading allows for this work's additional contribution: to shed new light on the market integration hypothesis postulated in asset pricing and the structural models, essential for the asset pricing models and with policy-making implications as second line of inquiry.

Preliminary analysis show that RVC, as implied from credit and equity returns, has a positive but weak correlation of 0.14. RJV, on the other hand, has a negative correlation of -0.24. It, therefore, appears that these two markets entail differing movements. A horse race exercise can compare the empirical performance of the two models.

The first line of inquiry of this work is to systematically study if the proposed risk measures outperform the risk measures proposed in the literature in forecasting future *changes* in credit risk premia. Asset pricing models with Epstein-Zin preferences predict that, if risk factors are market state variables, then they should be valuable in predicting future (returns) changes in spreads. Furthermore, due to the high-dimensional setting adopted in this study, traditional statistical methods used in asset pricing are ineffective and provide poor estimation and, moreover they cannot highlight the potential factors' marginal contribution. In response, this study uses the LASSO machine learning methods for prediction and inference.

Table 2 reports the results of the LASSO predictive regressions defined as in Equation (7). Models (1) and (2) use risk measures based only on equity market information. In contrast, Models (3) and (4) report results with risk measures implied from the credit market. Finally, Model (5) combines both equity and credit market-based risk measures. Since the leaps' direction is crucial for risk management, to better understand the direction of the realized jump volatility, the jump risk measure is divided into up (or positive) and down (or negative) jumps based on the sign of the return on the day of the jump (Equation 6). Furthermore, due to the assumed jump size distribution, the correlations of the up jumps with conditional volatility are more robust than the correlation of the down jumps. Consequently, the analysis accounts for the up and down jumps.

Results reported in Table 2 show that credit-market-based risk measures predict a significant portion of future changes in credit spreads. These measures report a striking out-of-sample R^2 of 47% (Model 4). In contrast, parallel estimates based on equity market information in Model (2) predict only 11% of changes in credit spreads in out-of-sample. Results show that puts alone are the main drivers of this predictability. In combining information from both markets, Model (5) reports, in fact, a slightly smaller R^2 of 46%, indicative of the curse of dimensionality.

LASSO predictive regressions suggest that only few factors matter for predicting future changes in spreads. Figure 2 provides a visual representation of the out-of-sample R^2 of the univariate regressions for the LASSO selected variables. Results suggest that equity-realized jumps bear greater empirical relevance in forecasting changes in spreads than the equity-realized volatility measure. On the other hand, put options have superior predictive power compared to E-RVC and E-RJV. Given the similar payoff structure of put options and the credit default swaps, this result supports models relying on the contingent-claim options market's information content to predict changes in spreads. More to the point, the proposed risk measures based on the credit market information alone in this study are the driving force in predicting the significant portion changes in spreads. C-RVC forecasts substantial changes in spreads, 42%, while C-RJV predicts 14%, only a percentage point more than puts. It is interesting to note that E-RJV, puts, C-RJV, and C-RVC risk measures, in addition to their theoretical importance, are also consistently selected by all LASSO models (1 to 5). These results, in addition to providing empirical support to Culp, Nozawa and Veronesi (2018) that tail risk is the critical driver of firm spreads, they also further our understanding of firms' spread behavior by identifying the key risk factors that generate the data process in firms' credit spreads, and thus drive the predictability.

Next, in the structural model's framework, asset value drives firms' stock and credit pricing dynamics. Therefore, it is crucial to see whether the credit- and equity- market-based risk factors also bear empirical relevance for predicting the equity risk premium changes. Table 3 reports similar results to those reported in Table 2 with $y_{t,j}$ replaced by the equity risk premium. Interestingly, there is some evidence of out-of-sample predictability of credit market risk factors (16%) for future changes in the equity risk premia. This predictability is substantially higher than what risk factors based on stock market information can forecast changes in credit spreads: 1.4%. Models (1) and (2) report that equity market risk factors predict 17% in future changes in equity risk premia. However, untabulated univariate regressions show that it is the put options that drive entirely this predictability, 17%. Puts empirical relevance remains robust even when all variables are pooled together (Model (5)).

This result reinforces the notion of tail risk measures' importance for predicting credit and equity risk premiums changes. Puts are in many aspects similar to CDS. They both have similar contingent-claim payoffs and offer protection against downside risk. Thus, finding that these factors can predict risk premiums in both markets affords a better understanding of the importance of assets' with a contingent-claim payoff to predict the risk premiums' pricing dynamics. However, while credit market-based risk measure alone predicts a similar portion of future changes in equity risk premia as put options (16% and 17%, repsecitvely), puts predict significantly less of future changes in credit premia, 9%, than the credit risk factors, 47%. These results highlight the importance of the credit-market information in predicting firms risk premia.

The first part of the analysis uses ML to predict changes in credit and equity risk premia. However, predictive LASSO regressions do not allow for robust inferential analysis. Therefore, the statistical significance of the risk measures is obtained via the double-machine learning LASSO proposed by Chernozhukov et al. (2018). Regression specifications are like (7), with HF credit and equity risk factors being the variables of interest. Results in Tables 2 and 3 report robust standard errors of these cross-fit partialing-out inferences. The double-machine learning approach is also important in that it recovers the true data-generating process.[20] Regressions use a battery of control variables to assess the statistical significance of the risk factors. These controls are used in two distinctive ways. Models (1) and (3) report results from a sparse model specification that uses the variables of interest, *i.e.*, credit and equity volatility and jump risk measures (-RVC and -RJV, up/down RJV, respectively), illiquidity (λ), and firm characteristics (ϕ), and controls only for fixed effects. Models (2) and (4) include the variables of interest and control for all covariates derived from their respective equity and credit markets. Model (5) includes the full model as in Equation (7). Since in structural models, the asset value process drives firms' stock and credit pricing dynamics, Models (2), (4) and (5) also include interaction terms of asset risk factors (α), the asset volatility and asset beta, with the equity and credit volatility risk factors.[21] Including a large number of control variables and the interaction terms affords a two-fold contribution. First, this approach allows for a comprehensive investigation on the marginal importance of the proposed HF credit market-based risk measures concerning a

[20]See Belloni, Chernozhukov and Hansen (2014) for using LASSO for inference and making causal interpretations.
[21]This specification results in 166 controls, of which cross-fit partialing-out selects 60.

complete list of covariates presented in the credit literature. Furthermore, it highlights the limitations of the traditional statistical measures used in asset pricing literature favoring the machine-learning approach.

Results reported in Table 2 show that, when all control variables are included in Model (5), E-RJV is the only equity risk factor that is statistically significant. The credit market-based risk factors proposed in this article are consistently statistically highly significant across all model specifications. They are consumed by neither the other risk factors presented in the literature nor by including their interactions with assets total and systematic risks. Likewise, the credit illiquidity risk measure (λ_C) and the put options are also consistently statistically highly significant in all model specifications. These factors also result important in all predictive LASSO model selections discussed before.

Regarding the equity risk premium, results reported in Table 3 show that only the DOTM puts are statistically highly significant across three out of four models. Credit risk measures, although they predict an equal portion of future changes in equity risk premia as puts (16% vs. 17%), they are not statistically significant and appear to be subsumed by puts in the kitchen sink regressions in Model (5).

Next, firms' leverage and size both result highly significant across all models reported in Table 2. In the structural framework, leverage represents the moneyness of the implicit put option on firms' debt. Because in this model, defaults occur when leverage nears unity, the leverage should interact with risk measures. To see the importance of the HF credit and equity risk measures for predicting credit and equity risk premia changes across different moneyness levels, Tables 4 and 5 sort firms into four equally-weighted portfolios based on the firm's leverage. Results are mostly consistent with those reported in Tables 2 and 3, and indicative of the notion that the credit risk factors are a persistent data feature. Another implication of studying predictability by leverage groups is to demonstrate the robustness of the results across different samples. Moreover, this strategy enables to relate the impact of asset pricing models (risk factors and return predictability) with the structural models (the importance of leverage). Furthermore, since the risk factors are implied at the firm level result reported here also addresses Lucas' critique.

These results imply potentially unspanned volatility and jump risks in equity markets, which are idiosyncratic to the credit market. Credit markets differ from other asset classes by the magnitude and clarity with which these factors are revealed. Assets' risk-neutral distribution implied from credit market information provides insights into the tail risk regions, which are not revealed by the equity, and to a large extent the options market, or macro-financial factors. These results are important for risk management, but they also point to market fragmentation, which is addressed in the following section.

B. Market Integration

The previous section showed how carefully-estimated risk parameters based on credit market information and machine-learning prediction techniques significantly outperform the equity risk counterparts in forecasting credit and equity risk premia. However, results also indicate a divergent movement in these markets, contrary to the asset pricing and structural models' predictions, warranting further investigation.

Consequently, the second line of inquiry of this article explores the market integration premise. The advent of automated machine learning trading and the active arbitrageurs have created an abundance of data. Accordingly, the expectations are that an econometrician should see only marginal pricing discrepancies between these markets, as predicted by asset pricing and structural models and the slow-moving capital theory's premise.

This article adopts the test of Kapadia and Pu (2012) for market integration for two reasons. First, by moving this analysis as close to theirs, we can see whether and how results change. Although the differences between these two studies are expected, they can still provide useful information. While their research relies on daily data spanning 2001 to 2009, thereby including the start, and thus nonconsolidated credit markets, their sample also includes the major shock in this market, the 2008 credit crunch. Pricing divergences in their results are, therefore, to be expected. Yet, finding pricing discrepancies in this work, and moreover, of a similar magnitude as in their sample, can shed light on the effects of the innovative algorithmic HF trading for the market integration and efficiency and afford for new and robust testing of asset pricing and structural model predictions. Second, Kapadia and Pu (2012) test is based on the structural models' forecast of a close relationship between the firm's bond and stock prices. The test recognizes these two closely related assets' pricing discrepancy by implying a concordance level between these two assets' price changes. The concordance measure is directly linked to pricing discrepancy and thus has a straightforward interpretation.[22]

[22]The test is defined over k=1,..,M nonoverlapping intervals of observations where the changes (first differences

Results collated by credit rating and for the entire sample are reported in Table 6. Results suggest that the pricing discrepancy is a persistent feature in the data when the analysis is performed daily (Panel A) or monthly (Panel B) frequency. For the whole sample, equity and credit spread co-move only in 48% and 49% of the time, for daily and monthly data. This persistent occurrence of pricing divergences challenges the slow-moving capital theory's premise. This theory predicts that by the time that the arbitrage capital is deployed, the pricing discrepancy will diminish. Thus, results reported here show that even with HF trading, the pricing divergence is still persistent. Results also show that for about 3% of the time, there is a pricing anomaly where there are no changes in equity and credit spreads.

At the credit-rating levels, there are some important cross-sectional differences. The pricing divergences between credit and equity markets, at daily and monthly frequency, are remarkably more pronounced for the AA-rated (49.5% and 37.5% for daily and monthly data, respectively) entities than for the speculative-grade ones, about 48% at both daily and monthly frequency.

There are several interesting points to note concerning the results reported here and those in Kapadia and Pu. While in both studies, the pricing divergence is more pronounced for the investment-grade entities than the high-yield ones, the monthly co-movements in the correct direction ($\Delta CDS/\Delta P/P < 0$) between the markets have decreased from about 63% in their study to 49% reported here. Furthermore, to the extent that daily anomalies (the zero price changes) represent higher trading costs, the advent of high-frequency trading seems to have increased these costs from 1.9% to 3.8% and 2.8% for daily and monthly data reported here.

These results pose a challenge to traditional credit risk models, with 48% of daily co-movements representing arbitrage opportunities. They are also informative for policymaking. The reported increased costs are likely an epitome of the "latency arbitrage" effects.[23] Furthermore, result reported in Tables 2 and 4 show that the illiquidity risk is statistically significant, indicative of the "ghost liquidity" notion.[24] While results reported in Kapa-

Δ) of CDS and equity prices represent arbitrage if $\Delta CDS/\Delta P/P > 0$ and the market integration test is based on the frequency of occurrence of such arbitrages $\hat{\kappa}_i = \sum_{\tau=1}^{M-1} \sum_{k=1}^{M-\tau} 1_{\left[\Delta CDS_{i,k}^t \Delta P_{ik}^t > 0\right]}$. Markets are more integrated if $\hat{\kappa}_i < \hat{\kappa}_j$.

[23]Latency arbitrage occurs when front runners hike the price. On the other hand, underpriced latency is a symptom of the limitations of HFT in where analysts' skills and availability of capital are crucial for a profitable trade, hence increasing costs.

[24]A significant criticism of the HFT disputes their claim as liquidity providers since the assets are held only for a

dia and Pu (2012) spans period with young credit markets and include the 2008 financial shock; therefore, pricing discrepancies reported in their work is largely not surprising; the pricing inconsistencies reported here are implied by HFT and a mature credit market. Consequently, they are puzzling and symptomatic of the notion of latency arbitrage exacerbating the market fragmentation effects.

Next, for the investors and analysts, it is important to recognize the lead-lag relationship between the volatility time series in these two markets and how the trading patterns emerge. In response, this study conducts a wavelet coherence analysis between the credit and equity markets. Wavelet coherence allows for analysis of the coherence (correlation) and the phase lag between time series as a function of both frequency and time.[25] Results are reported in Figure 3. Panel (A) reports the monthly wavelet correlations results and shows initially a lead of the credit market. On the other hand, panel (B) reports results of the daily wavelet cross-correlation sequence (a scale-localized version of the usual cross-correlation) of the daily realized volatility. Results suggest that after the initial negative correlation with credit-realized volatility leading the equity, the correlation turns positive by the end of the week.

These results suggest that in addition to the market segmentation and frictions, missing state variables also can cause price discrepancy between the two markets. Tail risk measures based on credit market data bear more meaningful information for predicting risk premiums in both credit and equity markets. This appears a robust feature of the data, not explained by firm heterogeneity or subsumed by other factors presented in the literature. However, if these risk factors are market state variables, then they should also command a market premium, as examined in the following section.

C. Volatility and Risk Premia

Asset pricing models with Epstein-Zin preferences suggest that SDF's are dependent not only on current consumption growth but also on the news about future volatility. These models predict that investors are willing to pay a premium to hedge the volatility risk. Consequently, if these risk factors are market state variables valuable for predicting future returns, they should also be market priced sources of risk and, therefore, correlate with

short time (seconds). Therefore, the liquidity is not real.

[25] Wavelet allows to derive from parameter p its characteristic frequency $f(p)$ and characteristic time $t(p)$ and, therefore, informs about the temporal extent of the signal as well as the frequency spectrum signal.

the market risk premia. Furthermore, in Merton (1974) model, the leverage is the default boundary for firms and represents the implicit put option's moneyness in a firm's debt. This put option gives a mechanical loading on shocks to asset volatility and earns a variance risk premium.

Delta-hedged equally-weighted quartile portfolios sorted on firms' leverage levels reveal the spreads' exposure that stems from bearing variance risk. The hedge ratio is estimated from a full sample regression of CDS returns for portfolio p onto the underlying equity return. Formally,

$$R_{H_{p,t}} = R_{CDS_{p,t}} - \Delta_{p,t} \times R_{EQUITY_{p,t}}. \tag{8}$$

However, in the event of a large negative jump in firm value, the appropriate hedging tool for corporate debt may not be the firm's equity; rather the DOTM puts on the firm's equity. Therefore, results depicted in Figure 4 also report DOTM put delta-hedged portfolios.[26] Results show that Sharp ratios move with moneyness and decline with an increase in leverage; as the spreads move "at-the-money." Results suggest that the variance risk premium is an important factor in the market for credit risk.

V. Conclusions

The first line of inquiry of this study showed how volatility and jump risk factors implied from high-frequency credit spreads combined with machine learning predictive methods forecast a substantial part of future changes in firms' credit premiums in out-of-sample. Strikingly, this predictive ability also extends for the equity risk premiums, albeit their statistical significance is not confirmed. Results indicate that tail risk measures based on credit market data alone contain valuable information for predicting firms' risk premia in both the credit and equity market, while parallel estimates based on a firm's high-frequency stock data do not. This result is reinforced by deep-out-of-the-money put options' ability to predict credit and equity risk premia in out-of-sample. Put options' contingent-claim payoff is similar to credit spreads and provide cheap and convenient tail risk protection. However, puts predictive power is significantly lower for credit risk premia than the credit risk measures. Credit market-based risk factors represent the most significant determinant

[26] To address potential concerns about the statistical significance of the reported Sharpe ratios, Figure 4 also report bars of the 95% confidence intervals of 10,000 studentized bootstrap repetitions. The confidence intervals do not include zero suggesting that the Sharpe ratios are significantly different from zero. Put delta-hedged Sharpe ratios are plotted with dotted lines while the stock delta-hedged Sharpe ratios are reported with the continuous line. The same figure also reports the unhedged Sharpe ratios depicted with a dashed line.

of credit spread changes across all firms and leverage groups, indicating that credit spread volatility is a state variable that drives the risk premium pricing dynamics. This appears a robust feature of the data, not explained by firm heterogeneity or subsumed by other factors presented in the literature. Results suggest that, even with stochastic volatility and jumps taken into account, the credit market information still bears greater empirical relevance for firm risk premiums.

The second line of inquiry showed how the emergence of the HFT in the credit market and the presence of active arbitrageurs have not lessened the market fragmentation effects, contrary to the expectations posed in asset pricing and structural models and slow-moving-capital theory's premise.

The credit and equity market's divergent movements suggest that, initially, credit markets lead the equity, providing some intuition on how the trading patterns arise and supporting the findings of the credit risk factors' superior predictive ability. This fact, coupled with the increased market fragmentation and trading costs, is an epitome of latent arbitrage in the market for credit risk. Results suggest that, while the front runner has become more sophisticated, their intentions remain predictable and with uncertain welfare implications, as evidenced by higher trading costs and increased market fragmentation.

These results are of interest to investors, academicians, and policymakers alike. The volatility and jump risk measures proposed in this paper are based on model-free unconditional variance estimates extracted from diverse firms' price data. These measures can provide analysts with an easy-to-compute method of the credit market-based risk measures on any day and a useful instrument for their risk management decisions. In their efforts to gauge the credit market's dynamics and soundness, regulators can also find these new easily quantifiable risk measures informative in their policy decisions.

REFERENCES

Acharya, Viral V., and Jennifer N. Carpenter. 2002. "Corporate Bond Valuation and Hedging with Stochastic Interest Rates and Endogenous Bankruptcy." *Review of Financial Studies*, 15(5): 1355–1383.

Ait-Sahalia, Yacine. 2004. "Disentangling diffusion from jumps." *Journal of Financial Economics*, 74(3): 487–528.

Ait-Sahalia, Yacine, and Per A. Mykland. 2005. "How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise." *Review of Financial Studies*, 18(9611): 351–416.

Ait-Sahalia, Yacine, and Robert Kimmel. 2007. "Maximum likelihood estimation of stochastic volatility models." *Journal of Financial Economics*, 83(2): 413 – 452.

Andersen, Torben G., Luca Benzoni, and Jesper Lund. 2002. "An Empirical Investigation of Continuous-Time Equity Return Models." *Journal of Finance*, 57(8510): 1239–1284.

Andersen, Torben G., Tim Bollerslev, and Francis X. Diebold. 2007. "Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility." *The Review of Economics and Statistics*, 89(4): 701–720.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens. 2001. "The distribution of realized stock return volatility." *Journal of Financial Economics*, 61(1): 43–76.

Bandi, Federico M., and Jeffrey R. Russell. 2006. "Separating microstructure noise from volatility." *Journal of Financial Economics*, 79(3): 655 – 692.

Bansal, Ravi, and Amir Yaron. 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *The Journal of Finance*, 59(4): 1481–1509.

Barndorff-Nielsen, O E. 2006. "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation." *Journal of Financial Econometrics*, 4(1): 1–30.

Barndorff-Nielsen, Ole, and Neil Shephard. 2001. "Non-Gaussian OU Based Models and Some of Their Uses in Financial Economics." *Journal of the Royal Statistical Society Series B*, 63: 167–241.

Barndorff-Nielsen, Ole E., and Neil Shephard. 2004. "Power and Bipower Variation with Stochastic Volatility and Jumps." *Journal of Financial Econometrics*, 2: 1–37.

Barro, Robert J. 2006. "Rare Disasters and Asset Markets in the Twentieth Century*." *The Quarterly Journal of Economics*, 121(3): 823–866.

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen. 2012. "Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain." *Econometrica*, 80(6): 2369–2429.

Belloni, Alexandre, and Victor Chernozhukov. 2013. "Least squares after model selection in high-dimensional sparse models." *Bernoulli*, 19(2): 521–547.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives*, 28(2): 29–50.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014b. "Inference on Treatment Effects after Selection among High-Dimensional Controls." *The Review of Economic Studies*, 81(2): 608–650.

Boehmer, Ekkehart, Sudheer Chava, and Heather E. Tookes. 2015. "Related Securities and Equity Market Quality: The Case of CDS." *Journal of Financial and Quantitative Analysis*, 50(03): 509–541.

Bongaerts, Dion, Frank De Jong, and Joost Driessen. 2011. "Derivative Pricing with Liquidity Risk: Theory and Evidence from the Credit Default Swap Market." *Journal of Finance*, 66(1): 203–240.

Campbell, John Y., and John H. Cochrane. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy*, 107(2): 205–251.

Cao, Charles, Fan Yu, and Zhaodong Zhong. 2010. "The information content of option-implied volatility for credit default swap valuation." *Journal of Financial Markets*, 13(3): 321 – 343.

Chen, Long, Pierre Collin-Dufresne, and Robert S. Goldstein. 2008. "On the Relation Between the Credit Spread Puzzle and the Equity Premium Puzzle." *The Review of Financial Studies*, 22(9): 3367–3409.

Chernozhukov, Victor, Christian Hansen, and Martin Spindler. 2015. "Valid Post-Selection and Post-Regularization Inference: An Elementary, General Approach." *Annual Review of Economics*, 7(1): 649–688.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. "Double/debiased machine learning for treatment and structural parameters." *The Econometrics Journal*, 21(1): C1–C68.

Christoffersen, Peter, Du Du, and Redouane Elkamhi. 2017. "Rare Disasters, Credit, and Option Market Puzzles." *Management Science*, 63(5): 1341–1364.

Collin-Dufresne, Pierre, Robert S. Goldstein, and J. Spencer Martin. 2001. "The Determinants of Credit Spread Changes." *The Journal of Finance*, 56(6): 2177–2207.

Cremers, K. J. Martijn, Joost Driessen, and Pascal J. Maenhout. 2008. "Explaining the Level of Credit Spreads: Option-Implied Jump Risk Premia in a Firm Value Model." *Review of Financial Studies*, 21: 2209–42.

Culp, Christopher L., Yoshio Nozawa, and Pietro Veronesi. 2018. "Option-Based Credit Spreads." *American Economic Review*, 108(2): 454–88.

Das, Sanjiv, Madhu Kalimipalli, and Subhankar Nayak. 2014. "Did CDS trading improve the market for corporate bonds?" *Journal of Financial Economics*, 111(2): 495 – 525.

Das, Sanjiv R., Paul Hanouna, and Atulya Sarin. 2009. "Accounting-based versus market-based cross-sectional models of CDS spreads." *Journal of Banking & Finance*, 33(4): 719 – 730.

Du, Du, Redoune Elkamhi, and Jan Ericsson. 2019. "Time-Varying Asset Volatility and the Credit Spread Puzzle." *The Journal of Finance*, 74(4): 1841–1885.

Feng, Guanhao, Stefano Giglio, and Dancheng Xiu. 2020. "Taming the Factor Zoo: A Test of New Factors." *The Journal of Finance*, 75(3): 1327–1370.

Gabaix, Xavier. 2012. " Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance *." *The Quarterly Journal of Economics*, 127(2): 645–700.

Hansen, Peter R., and Asger Lunde. 2006. "Realized Variance and Market Microstructure Noise." *Journal of Business & Economic Statistics*, 24(2): 127–161.

Huang, Jing-Zhi, and Ming Huang. 2012. "How Much of the Corporate-Treasury Yield Spread Is Due to Credit Risk?" *The Review of Asset Pricing Studies*, 2(2): 153.

Huang, Xin, and George E. Tauchen. 2005. "The Relative Contribution of Jumps to Total Price Variance." *Journal of Financial Econometrics*, 3: 456–99.

Kapadia, Nikunj, and Xiaoling Pu. 2012. "Limited arbitrage between equity and credit markets." *Journal of Financial Economics*, 105(3): 542 – 564.

Kelly, Bryan T., Gerardo Manzo, and Diogo Palhares. 2017. "Credit-Implied Volatility." *Chicago Booth Research Paper No. 17-12.*

Lee, Suzanne S., and Per A. Mykland. 2008. "Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics." *The Review of Financial Studies*, 21(6): 2535–2563.

Mayordomo, Sergio, Juan Ignacio Peña, and Eduardo S. Schwartz. 2014. "Are All Credit Default Swap Databases Equal?" *European Financial Management*, 20(4): 677–713.

McQuade, Timothy. 2018. "Stochastic Volatility and Asset Pricing Puzzles." *Working Paper, Stanford Univeristy.*

Merton, Robert C. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *Journal of Finance*, 29(2): 449–70.

Merton, Robert C. 1976. "Option pricing when underlying stock returns are discontinuous." *Journal of Financial Economics*, 3: 125–144.

Rajan, Raghuram G., and Luigi Zingales. 1995. "What Do We Know about Capital Structure? Some Evidence from International Data." *Journal of Finance*, 50(5): 1421–1460.

Roll, Richard. 1984. "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market." *Journal of Finance*, 39(4): 1127–1139.

Saretto, Alessio, and Heather E. Tookes. 2013. "Corporate Leverage, Debt Maturity, and Credit Supply: The Role of Credit Default Swaps." *Review of Financial Studies*, 26(5): 1190.

Schwert, Michael, and Ilya A. Strebulaev. 2014. "Capital Structure and Systematic Risk." *Working Paper, Rock Center for Corporate Governance at Stanford University Working Paper No. 178.*

Tauchen, George, and Hao Zhou. 2011. "Realized jumps on financial markets and predicting credit spreads." *Journal of Econometrics*, 160(1): 102–118.

Tibshirani, Robert. 1996. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1): 267–288.

Zhang, Benjamin Yi-Bin, Hao Zhou, and Haibin Zhu. 2009. "Explaining Credit Default Swap Spreads With the Equity Volatility and Jump Risks of Individual Firms." *Review of Financial Studies*, 22: 5099–5131.

VI. Tables and Figures

Table 1—Summary Statistics

This table reports the monthly summary statistics of the risk factors by the rating groups. Panel A reports the average monthly credit spreads, the firm-level continuous realized volatility of CDS and equity returns (C-RVC and E-RVC, respectively), their corresponding jump risk measures (C-RJV and E-RJV), and their respective autocorrelation functions. Volatility risk measures are reported in percentage points and annualized. λ denotes the illiquidity risk measure implied from both the credit (λ_C) and equity (λ_E) markets. Risk measures based on the credit market information have the prefix C- and those from the equity market have the prefix E-. N reports the number of firms for each S&P (Moody's) last-quarter available S&P (Moody's) credit ratings. Panel B reports the number of firms for each of the nine industry sectors.

Panel A: Risk Factors divided on S&P (Moody's) Credit Rating

	AA		A		BBB		BB-B		All	
	Mean	Std. dev	Mean	Std. dev	Mean	Std. dev	Mean	Std. dev	Mean	Std. dev
CDS	90.79	110.96	82.26	90.90	81.66	85.58	90.20	115.20	83.11	91.86
C-RVC	3.59	1.25	3.66	1.88	3.68	1.17	4.75	1.99	3.79	1.47
AR(1)	0.26		0.21		0.20		0.29		0.28	
C-RJV	1.67	1.25	1.69	1.10	1.68	1.07	1.66	1.14	1.68	1.48
AR(1)	0.02		0.10		0.02		0.18		0.11	
λ_C	0.23	0.18	0.24	0.17	0.23	0.16	0.23	0.14	0.23	0.16
AR(1)	0.15		0.12		0.30		0.15		0.18	
E-RVC	14.60	4.68	15.12	1.71	15.18	1.40	14.00	2.26	15.01	9.38
AR(1)	0.12		0.13		0.17		0.20		0.17	
E-RJV	1.48	1.45	1.58	1.10	1.57	1.11	1.59	1.18	1.57	1.88
AR(1)	0.05		0.10		0.10		0.06		0.13	
λ_E	0.35	0.18	0.34	0.27	0.34	0.18	0.33	0.18	0.34	0.21
AR(1)	0.11		0.10		0.25		0.21		0.16	
N	3		25		50		10		88	
Leverage	0.53		0.50		0.48		0.50		0.50	

Panel B: Industry Sector

	Cons. Non-cyclical —	Cons. Cyclical —	Industrial —	Financial —	Basic Materials —	Utilities —	Technology —	Communications —	Energy —	All
N	17	19	15	15	5	2	4	7	4	88
Leverage	0.48	0.49	0.50	0.51	0.51	0.47	0.51	0.48	0.51	0.50

Table 2—Forecasting changes in monthly Credit Risk Premia

This table reports results of LASSO inferential analysis with LASSO-selected coefficients derived from cross-fit partialing-out regressions (double-machine learning) of Chernozhukov et al. (2018). Models (1) and (3) take the form

$$\delta y_{t,j} = c + \beta_1 \delta RVC^{'}_{t,j} + \beta_2 \delta RJV^{'}_{t,j} + \beta_3 \delta \lambda^{'} + \beta_4 \delta \iota^{'}_{t,j} + \beta_5 \delta \pi_{t,j}$$
$$+ \beta_6 \delta \phi^{'}_{t,j} + (Firm\, FE + Month\, FE + Ratings\, FE + Sector\, FE) + \epsilon_{t,j}.$$

Model (3) considers only credit market information i.e., excludes λ_E, equity returns, and puts. Models (2) and (4) take the form:

$$\delta y_{t,j} = c + \beta_1 \delta RVC^{'}_{t,j} + \beta_2 \delta RJV^{'}_{t,j} + \beta_3 \delta \lambda^{'} + \beta_4 \delta \iota^{'}_{t,j} + \beta_5 \delta \pi_{t,j}$$
$$+ \beta_6 \delta \phi^{'}_{t,j} + (\beta_7 \delta \alpha^{'}_{t,j} + \beta_8 \delta B^{'}_{t,j} + \beta_9 \delta \kappa^{'}_{t} + \beta_{10} \mu^{'}_{t}$$
$$+ Firm\, FE + Month\, FE + Ratings\, FE + Sector\, FE) + \epsilon_{t,j},$$

where asset risk factors α interact with the credit and equity RVC, RJV and up/down Jumps. Model (5) contains all controls. The variables inside parenthesis denote the control variables. The table reports the out-of-sample R^2 of the predictive LASSO regressions of the same model specifications. The dependent variable $y_{t,j}$ is the frim's changes in credit spreads. δ denotes the first differences in the variables. The sample period spans July 2012 to July 2016. The superscripts $***$, $**$, and $*$ indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
E-RVC	3.719	3.774*			1.321
	(2.291)	(2.253)			(1.781)
E-RJV	131.404**	127.303**			102.935***
	(55.845)	(56.130)			(38.298)
E-(-)RJV	5.904	-3.512			67.275**
	(50.860)	(50.793)			(33.609)
E-(+)RJV	-45.373	-46.528			-52.095
	(48.712)	(49.157)			(37.137)
C-RVC			6.585***	6.009***	6.359***
			(0.301)	(0.268)	(0.275)
C-RJV			3.369***	3.666***	3.879***
			(0.754)	(0.722)	(0.745)
C-(+)RJV			5.319***	5.399***	5.838***
			(1.815)	(1.690)	(1.735)
C-(-)RJV			-8.102***	-7.828***	-8.295***
			(1.815)	(1.743)	(1.806)
λ_E	1.490**	1.391**			-0.810
	(0.624)	(0.628)			(0.824)
λ_C			45.852***	30.819***	29.874***
			(7.583)	(6.627)	(6.847)
EQUITY RETURN	7.272	2.692		-13.158	-5.107
	(30.544)	(30.554)		(23.907)	(24.630)
Eh-VOLATILITY	2.440	1.816			4.459**
	(2.299)	(2.282)			(1.984)
Eh-SKEWNESS	-0.001	-0.002			-0.004
	(0.006)	(0.006)			(0.004)
Eh-KURTOSIS	0.001	0.001			-0.000
	(0.001)	(0.001)			(0.000)
Ch-VOLATILITY			-0.163	-0.104	-0.127
			(0.105)	(0.103)	(0.110)
Ch-SKEWNESS			3.544***	2.749**	2.319*
			(1.184)	(1.160)	(1.199)
Ch-KURTOSIS			0.178	0.151	0.092
			(0.128)	(0.127)	(0.135)
DOTM Put	3.354***	3.199***		1.342***	1.965***
	(0.294)	(0.299)		(0.144)	(0.188)
LEVERAGE	25.583***	24.686***	13.626**	16.584***	16.716***
	(7.460)	(7.450)	(5.302)	(5.134)	(5.349)
SIZE	-8.533***	-9.171***	-3.841***	-5.886***	-5.846***
	(1.389)	(1.473)	(1.088)	(1.076)	(1.223)
σ_{ASSET}	1.818	2.257	1.898**	2.571**	2.346
	(1.186)	(1.385)	(0.955)	(1.004)	(1.563)
β_{ASSET}	0.274	0.058	2.606***	0.582	1.438*
	(1.355)	(1.421)	(0.542)	(1.173)	(0.837)
Out-of-Sample R^2	0.10	0.11	0.44	0.47	0.46
N	3602	3602	3965	3965	3587

Table 3—Forecasting changes in monthly Equity Risk Premia

This table reports results of LASSO inferential analysis with LASSO-selected coefficients derived from cross-fit partialing-out regressions (double-machine learning) of Chernozhukov et al. (2018). Regressions are same as in Table 2 with the dependent variable $y_{t,j}$ replaced with changes in equity risk premia. δ denotes the first differences in the variables. The sample period spans July 2012 to July 2016. The superscripts $***$, $**$, and $*$ indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
E-RVC	0.080	0.076			0.118
	(0.701)	(0.699)			(0.701)
E-RJV	8.359	9.539			8.174
	(11.991)	(11.969)			(12.030)
E-(-)RJV	6.393	6.222			8.026
	(9.501)	(9.503)			(9.611)
E-(+)RJV	3.539	3.589			1.724
	(11.193)	(11.189)			(11.163)
C-RVC			0.000	0.000	0.000
			(0.001)	(0.001)	(0.001)
C-RJV			-0.001	-0.001	-0.002
			(0.002)	(0.002)	(0.002)
C-(+)RJV			-0.001	-0.001	-0.001
			(0.005)	(0.005)	(0.005)
C-(-)RJV			-0.005	-0.005	-0.006
			(0.004)	(0.004)	(0.005)
λ_E	0.253	0.243			0.183
	(0.318)	(0.315)			(0.327)
λ_C			0.016	0.017	0.018
			(0.027)	(0.027)	(0.029)
Eh-VOLATILITY	1.715	1.664			1.751
	(1.108)	(1.108)			(1.121)
Eh-SKEWNESS	0.000	0.000			0.000
	(0.002)	(0.002)			(0.002)
Eh-KURTOSIS	-0.000	-0.000			-0.000
	(0.000)	(0.000)			(0.000)
Ch-VOLATILITY			0.000	0.000	-0.000
			(0.000)	(0.000)	(0.000)
Ch-SKEWNESS			0.004	0.004	0.001
			(0.005)	(0.005)	(0.005)
Ch-KURTOSIS			0.001*	0.001*	0.001
			(0.001)	(0.001)	(0.001)
DOTM Put	-0.001***	-0.001***		-0.000	-0.001***
	(0.000)	(0.000)		(0.000)	(0.000)
LEVERAGE	0.024	0.025	0.010	0.010	0.025
	(0.021)	(0.021)	(0.020)	(0.020)	(0.021)
SIZE	-0.000	-0.001	-0.002	-0.002	-0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
σ_{ASSET}	0.001	0.001	0.001	0.001	0.002
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
β_{ASSET}	0.003	0.003	0.004	0.004	0.003
	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)
Out-of-Sample R^2	0.17	0.17	0.16	0.16	0.17
N	3602	3602	3965	3965	3587

Table 4—Forecasting changes in monthly Credit Risk Premia by Quartile Portfolios

This table reports results of LASSO inferential analysis on credit returns by equally-weighted leverage quartiles. Firms with the lowest average level are in quartile < 25%, and 25%-50% is the quartile of firms with the second lowest average level. The second highest quartile are 50%-75%, and finally, > 75% is the fourth quartile of firms with the highest average level of leverage. Models (1) and (2) are similar to Models (1) and (3) in Table 2. The sample period spans July 2012 to July 2016. The superscripts ∗ ∗ ∗, ∗∗, and ∗ indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Leverage Groups					
	<25%			25-50%		
	(1)	(2)	(3)	(1)	(2)	(3)
E-RVC	4.023		5.556*	1.827		-3.254
	(4.471)		(3.312)	(5.626)		(4.018)
E-RJV	214.863**		208.435***	257.292**		272.498***
	(107.281)		(78.629)	(107.688)		(78.361)
E-(-)RJV	158.834		115.213**	-40.884		26.286
	(100.784)		(57.834)	(110.065)		(70.024)
E-(+)RJV	-8.977		-2.759	14.640		-6.403
	(89.163)		(65.228)	(85.426)		(72.874)
C-RVC		6.408***	6.372***		6.085***	6.300***
		(0.632)	(0.548)		(0.593)	(0.577)
C-RJV		2.492	3.618**		4.245***	3.519**
		(1.681)	(1.493)		(1.558)	(1.508)
C-(+)RJV		2.256	3.220		0.822	1.991
		(2.731)	(2.736)		(3.328)	(3.128)
C-(-)RJV		-2.454	-3.919*		-4.725	-4.461
		(2.271)	(2.260)		(4.207)	(4.180)
λ_E	1.030		0.083	-0.746		-8.381**
	(0.869)		(0.484)	(5.417)		(3.841)
λ_C		45.826***	32.366**		39.749**	40.451**
		(16.342)	(13.481)		(18.032)	(16.776)
EQUITY RETURN	-52.333		-58.171*	103.014		17.177
	(42.723)		(34.310)	(63.564)		(43.906)
DOTM Put	3.259***		1.682***	3.396***		2.061***
	(0.513)		(0.314)	(0.616)		(0.392)
LEVERAGE	19.478	3.532	9.770	42.415**	21.223*	28.490**
	(20.117)	(14.377)	(14.762)	(18.828)	(12.847)	(13.672)
SIZE	-4.543	-5.444***	-7.363***	-11.180***	-4.414*	-6.812***
	(2.923)	(1.877)	(2.261)	(2.903)	(2.476)	(2.642)
σ_{ASSET}			12.234*			0.503
			(6.664)			(3.484)
β_{ASSET}			-9.195*			1.966***
			(5.274)			(0.744)
Out-of-Sample R^2	0.12	0.45	0.48	0.13	0.44	0.49
N	922	1022	907	912	1013	899
	50-75%			>75%		
E-RVC	2.150		1.930	1.263		2.288
	(5.034)		(3.193)	(4.273)		(3.265)
E-RJV	43.298		29.672	64.683		49.211
	(123.783)		(67.434)	(100.047)		(78.714)
E-(-)RJV	95.062		106.889	-53.010		21.957
	(105.175)		(70.844)	(89.501)		(69.453)
E-(+)RJV	-103.536		-90.288	-115.943		-91.915
	(84.829)		(62.698)	(114.459)		(88.255)
C-RVC		6.399***	6.006***		7.304***	6.856***
		(0.523)	(0.461)		(0.580)	(0.568)
C-RJV		5.208***	6.853***		1.349	2.712**
		(1.498)	(1.555)		(1.221)	(1.259)
C-(+)RJV		5.286	7.327**		14.709***	14.730***
		(3.504)	(3.352)		(4.858)	(5.170)
C-(-)RJV		-9.764***	-12.077***		-19.748***	-20.270***
		(2.496)	(2.666)		(5.725)	(7.087)
λ_E	2.808		-10.895***	2.329***		-0.636
	(5.079)		(3.440)	(0.786)		(0.933)
λ_C		49.357***	36.007***		49.936***	38.848***
		(14.729)	(12.751)		(11.643)	(11.677)
EQUITY RETURN	43.402		25.086	-53.762		-16.910
	(63.627)		(59.336)	(75.338)		(53.514)
DOTM Put	3.131***		2.004***	3.830***		2.315***
	(0.663)		(0.412)	(0.562)		(0.364)
LEVERAGE	-5.698	17.302	7.352	-7.682	-16.487	-18.773
	(17.021)	(12.287)	(12.668)	(20.076)	(13.850)	(14.795)
SIZE	-5.699**	-1.555	-6.230**	-8.925***	-3.401*	-4.763*
	(2.736)	(2.287)	(2.592)	(2.623)	(1.976)	(2.503)
σ_{ASSET}			0.906			-0.049
			(1.033)			(1.457)
β_{ASSET}			-28.731***			-33.059***
			(6.856)			(7.863)
Out-of-Sample R^2	0.12	0.44	0.48	0.13	0.45	0.47
N	912	1013	891	909	1017	890

Table 5—Forecasting changes in monthly Equity Risk Premia by Quartile Portfolios

This table reports results of LASSO inferential analysis on changes in equity risk premium by equally-weighted leverage quartiles. Quartile portfolios and regression models are similar to Table 4. The sample period spans July 2012 to July 2016. The superscripts $***$, $**$, and $*$ indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

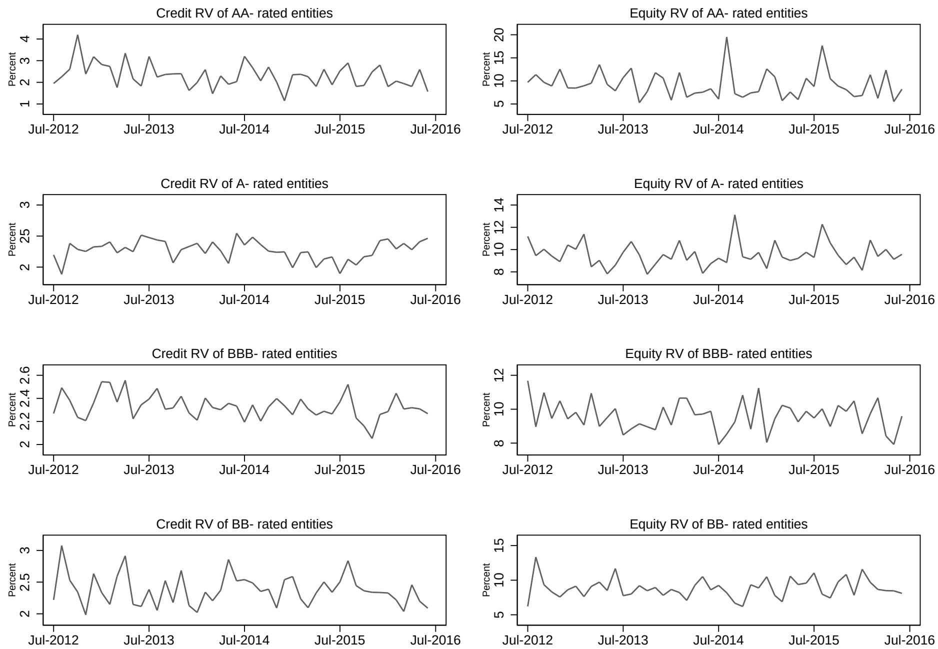
	Leverage Groups					
	<25%			25-50%		
	(1)	(2)	(3)	(1)	(2)	(3)
E-RVC	0.057		0.073	-1.624		-2.019
	(1.708)		(1.738)	(1.905)		(1.904)
E-RJV	6.671		9.019	18.334		4.055
	(21.992)		(21.782)	(22.910)		(23.528)
E-(-)RJV	46.580*		48.504*	4.731		8.318
	(26.078)		(25.854)	(17.182)		(17.885)
E-(+)RJV	-12.652		-2.347	2.332		-2.014
	(26.485)		(27.391)	(19.087)		(18.778)
C-RVC		-0.000	0.000		0.001	0.002
		(0.001)	(0.001)		(0.001)	(0.001)
C-RJV		0.006*	0.003		-0.005	-0.005
		(0.003)	(0.003)		(0.003)	(0.003)
C-(+)RJV		0.001	-0.002		-0.007	-0.010
		(0.006)	(0.006)		(0.008)	(0.007)
C-(-)RJV		-0.007	-0.005		-0.009	-0.004
		(0.007)	(0.007)		(0.010)	(0.009)
λ_E	0.853***		0.817***	-1.188		-1.571
	(0.193)		(0.178)	(1.407)		(1.434)
λ_C		0.034	0.008		0.015	0.055
		(0.062)	(0.066)		(0.055)	(0.050)
DOTM Put	-0.001		-0.001*	-0.002*		-0.001
	(0.001)		(0.001)	(0.001)		(0.001)
LEVERAGE	0.114**	0.109*	0.123**	0.113*	0.079	0.101
	(0.054)	(0.056)	(0.054)	(0.060)	(0.058)	(0.062)
SIZE	0.004	-0.001	0.006	-0.007	0.001	-0.002
	(0.007)	(0.007)	(0.008)	(0.009)	(0.009)	(0.010)
σ_{ASSET}			0.012			-0.014
			(0.010)			(0.015)
β_{ASSET}			0.025			-0.006
			(0.021)			(0.004)
Out-of-Sample R^2	0.17	0.18	0.17	0.18	0.18	0.18
N	922	1022	907	912	1013	899
	50-75%			>75%		
E-RVC	-0.599		-0.065	0.782		1.184
	(1.496)		(1.145)	(1.390)		(1.379)
E-RVJ	4.767		1.897	-1.843		17.477
	(23.651)		(24.765)	(27.727)		(29.742)
E-(-)RVJ	-1.821		-18.755	-19.248		-7.541
	(21.464)		(21.541)	(21.095)		(19.997)
E-(+)RVJ	-27.373		-1.926	3.778		7.157
	(26.673)		(19.759)	(27.068)		(29.342)
C-RVC		-0.001	0.000		0.000	0.000
		(0.001)	(0.001)		(0.001)	(0.001)
C-RJV		0.005	0.003		-0.007**	-0.006*
		(0.003)	(0.003)		(0.003)	(0.004)
C-(+)RJV		0.001	-0.000		0.022	0.029
		(0.011)	(0.011)		(0.019)	(0.023)
C-(-)RJV		-0.012	-0.011		0.019	0.016
		(0.007)	(0.008)		(0.014)	(0.012)
λ_E	1.216		1.150	-0.341		-0.354
	(1.287)		(1.382)	(0.271)		(0.237)
λ_C		0.012	0.027		0.021	-0.008
		(0.053)	(0.057)		(0.064)	(0.072)
DOTM Put	-0.001		-0.002*	-0.000		-0.000
	(0.001)		(0.001)	(0.001)		(0.001)
LEVERAGE	-0.014	-0.028	-0.018	0.126	0.128*	0.064
	(0.060)	(0.056)	(0.061)	(0.077)	(0.073)	(0.078)
SIZE	0.003	-0.006	-0.000	-0.013	-0.006	-0.011
	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)
σ_{ASSET}			0.001			0.008
			(0.001)			(0.010)
β_{ASSET}			0.004			-0.048
			(0.003)			(0.039)
Out-of-Sample R^2	0.19	0.19	0.19	0.20	0.19	0.20
N	912	1013	891	909	1017	890

Table 6—Market Integration

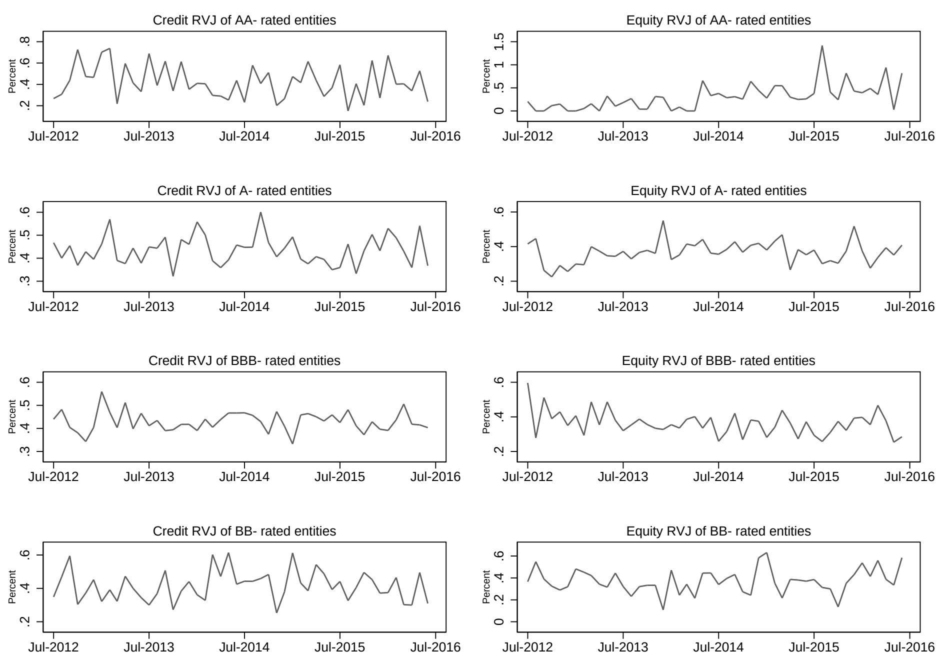
This table reports the results of the market integration test of Kapadia and Pu (2012). The credit and equity markets' co-movements represent the pricing discrepancies within the Merton (1974) model, i.e., $\Delta CDS\Delta P/P > 0$. $\Delta CDS\Delta P/P < 0$, and $\Delta CDS\Delta P/P = 0$ are registered as a share of total observations measured over non-overlapping daily (Panel A) and monthly (Panel B) time intervals. The $|\Delta CDS|$ expressed in bps denotes the mean of absolute spread changes. $|\Delta P/P|$ represents the average absolute stock returns. The sample period spans July 2012 to July 2016.

		AA	A	BBB	BB-B	ALL
Panel A: Daily						
$\Delta CDS\Delta P/P < 0$	Fraction	0.487	0.483	0.452	0.485	0.479
	$\|\Delta CDS\|$	0.985	8.539	10.640	10.849	9.857
	$\|\Delta P/P\|$	6.023	8.419	4.486	6.484	6.772
$\Delta CDS\Delta P/P > 0$	Fraction	0.495	0.478	0.453	0.495	0.483
	$\|\Delta CDS\|$	0.985	8.539	10.640	10.849	9.857
	$\|\Delta P/P\|$	6.023	8.419	4.486	6.484	6.772
$\Delta CDS\Delta P/P = 0$	Fraction	0.018	0.039	0.095	0.020	0.038
Panel B: Monthly						
$\Delta CDS\Delta P/P < 0$	Fraction	0.604	0.499	0.518	0.475	0.492
	$\|\Delta CDS\|$	2.131	7.674	5.419	6.413	5.302
	$\|\Delta P/P\|$	0.403	3.522	1.209	2.704	2.488
$\Delta CDS\Delta P/P > 0$	Fraction	0.375	0.484	0.467	0.486	0.480
	$\|\Delta CDS\|$	2.131	3.723	5.419	6.413	5.302
	$\|\Delta P/P\|$	0.403	2.994	1.209	2.704	2.488
$\Delta CDS\Delta P/P = 0$	Fraction	0.021	0.018	0.015	0.039	0.028

Figure 1. Realized Volatility and Jumps in Credit and Equity Markets



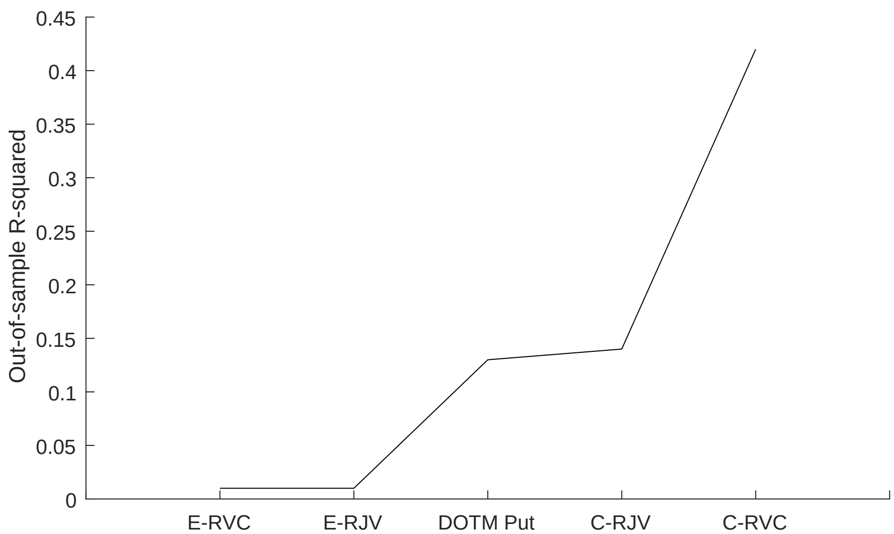
(a) Credit and Equity Implied Realized Volatility



(b) Credit and Equity Implied Realized Jump Volatility

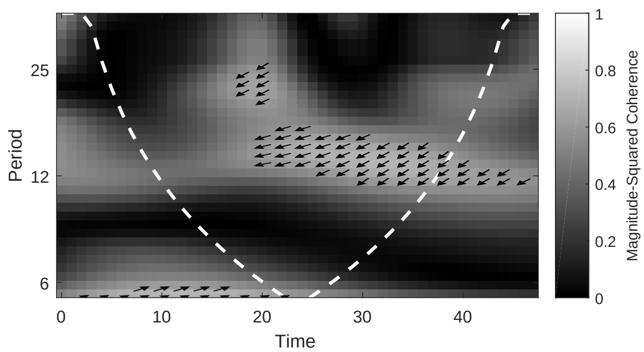
Panel (a) of this figure plots the horse-race comparison of continuous realized volatility estimates implied from credit and equity returns. The left-hand side panels plot by rating group the monthly credit spread implied realized volatility of 88 entities. The right-hand side panels plot the monthly equity implied realized volatility for the same matching entities. Panel (b) plots the corresponding realized jump volatility estimates. Section III and the internet appendix illustrate the method used to estimate realized (jump) volatility. The sample period spans July 2012 to July 2016.

FIGURE 2. OUT-OF-SAMPLE R^2 OF UNIVARIATE LASSO PREDICTIVE REGRESSIONS

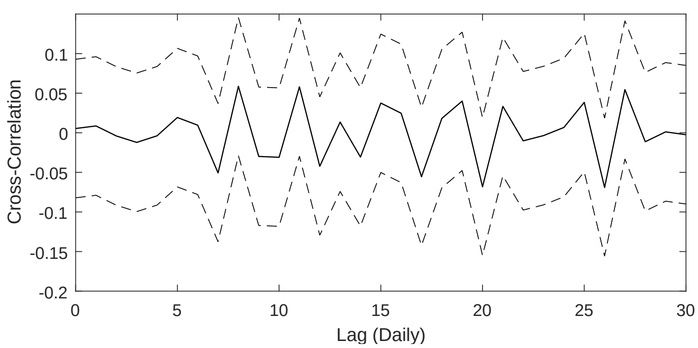


This figure plots the out of sample R^2 of univariate LASSO predictive regressions. RVC and RJV report realized volatility and jump risk measures, $DOTMPut$ is the deep out of the money put options of firms stocks. Equity-market-based risk measures have an E- prefix, and the credit market risk measures are reported with the C- prefix. The sample period spans July 2012 to July 2016.

Figure 3. Wavelet Coherence and Cross-Correlation of Credit and Equity Realized Volatilities



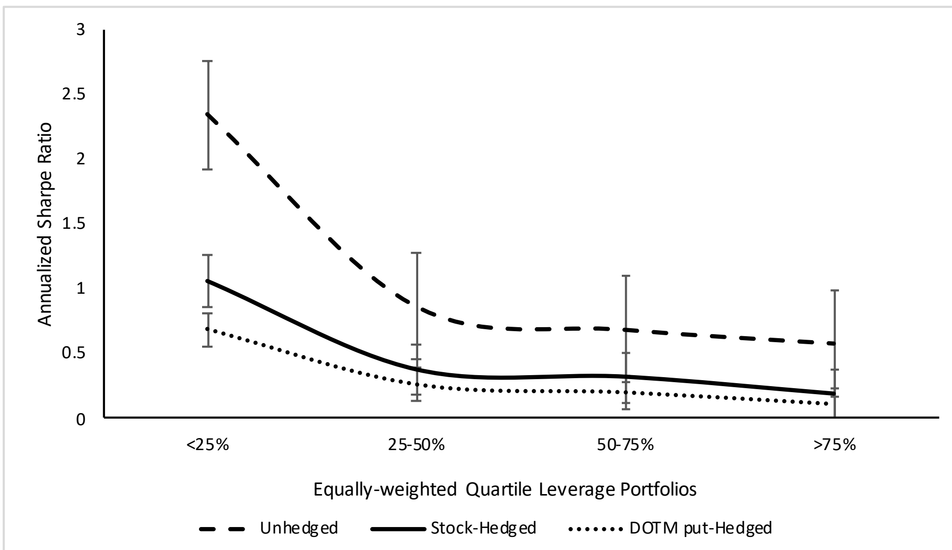
(a) Monthly Wavelet Coherence of Credit and Equity Realized Volatility



(b) Daily Wavelet Cross-Correlation Sequence of Credit and Equity Realized Volatility

This figure in panel (a) plots the monthly wavelet coherence of credit and equity realized volatility. Panel (b) plans the daily wavelet cross-correlation order of credit and equity realized volatility. The sample period spans July 2012 to July 2016.

Figure 4. Sharpe Ratios of Delta-Hedged Portfolios



This figure plots the annualized Sharpe ratios from monthly delta-hedged CDS returns of equally-weighted leverage quartile portfolios. Firms with the lowest average level are in quartile $< 25\%$, and 25%-50% is the quartile of firms with the second lowest average level. The second highest quartile are 50%-75%, and finally, $> 75\%$ is the fourth quartile of firms with the highest average level of leverage. The dotted line plots the deep out of the money (DOTM) put delta-hedged Sharpe ratios. The continuous line plots the equity return delta-hedged Sharpe ratios and the dashed line plots the unhedged Sharpe ratios. The bands indicate the 95% confidence intervals of the 10,000 studentized bootstrapped repetitions. The sample period spans July 2012 to July 2016.