Journal of Asset Management (2020) 21:119–134
https://doi.org/10.1057/s41260-020-00151-8

ORIGINAL ARTICLE

A common risk factor and the correlation between equity and corporate bond returns

Amer Demirovic[1] · Ali Kabiri[2,3] · David Tuckett[4] · Rickard Nyman[4]

Revised: 10 February 2020 / Published online: 11 February 2020

Abstract
A growing body of literature documents that security prices within and across asset classes behave similarly highlighting the importance of investors' common expectations about future risk and returns in the asset pricing. Consequently, variations in the common expectations of investors have a major role in determining the correlation among asset prices. We examine the role of these common expectations in determining the relationship between firm-level equity and bond returns. We use a novel measure of the common expectations defined as the difference in relative frequencies of words signalling excitement and anxiety in a large dataset of articles published by Reuters. Further, we also consider the VIX index and the indices of Baker and Wurgler (J Finance 61(4):1645–1680, 2006) and Huang et al. (Rev Financ Stud 28(3):791–837, 2015) as potential common factors. The results show that changes in common expectations, proxied by our index and the VIX, are significant in predicting variations in the correlation between equity and bond returns. An improvement in investors' optimism about future risk and returns causes a weaker correlation. The effect is stronger for the riskiest firms and flattens as firms' credit risk improves. By decomposing our index into the excitement and anxiety components, we find that this predictive power is due to changes in the anxiety components.

Keywords Investor sentiment · Equity–bond correlation · Credit risk · Systematic risk

Introduction

In a recent paper, Bollerslev et al. (2018) show that the volatility patterns within and across asset classes are virtually identical. This suggests that the behaviour of investors is strikingly similar across financial markets and in turn that asset prices can be viewed as being driven by common expectations about future risk and returns, which trigger the collective actions of investors. There are well-documented episodes of both flight-to-quality and flight-to-liquidity effects, which occur when investors' preferences for the highest quality and most liquid assets increase markedly. These episodes provide vivid illustrations of the importance of changes in the common expectations of investors for the pricing of assets during these specific situations.

A growing number of new studies devise measures to capture variations in the common expectations of investors and examine whether they are driven by changes in a common risk premium or alternatively orginate from temporary deviations from rational pricing. The large majority of these studies, however, focus exclusively on the equity markets. We contribute to this growing literature by examining how variations in the common expectations of investors affect the correlations between matched firm-level equity and bond returns. The aim of the paper is to refine our ability to identify this common factor using new techniques in computer-based algorithimic text analysis and also to be able to compare our measures to a suite of other potential drivers of the common factor.

In general, asset correlations are a cornerstone of portfolio management. As asset correlations are stochastic, understanding them is important as tools to mitigate portfolio risk. Financial crises, such as the crisis of 2007–2008,

✉ Amer Demirovic
amer.demirovic@utb.edu.bn

1 Universiti Teknologi Brunei, Jalan Tungku Link, Gadong BE1410, Brunei Darussalam

2 University of Buckingham, Hunter Street, Buckingham MK18 1EG, UK

3 FMG, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK

4 University College London, Gower Street, London WC1E 6BT, UK

are characterised by dramatic changes in the correlation of assets and, consequently, the risk of portfolios. In this particular episode, sharp rises in asset correlations in many cases left financial institutions and investors vulnerable to unforeseen risks. Such correlation risk persisted post-crisis and has been linked to several exacerbating factors such as market liquidity changes where the health and stability of the financial system act to increase asset correlation (IMF 2015). While market liquidity effects are one potential aggravating factor in the asset correlation changes we test for, we restrict our analysis to the identification of the drivers of the common factor from our expectation-based measures from financial news data. The nature of those data currently does not allow us to delineate their exact origins but gives impetus to research to test these factors in the future given the nature of our results.

The ability to understand the underlying determinants of the behaviour of these correlations is essential for asset managers and also for the measurement and setting of regulatory capital. The research therefore can inform policymakers towards a more efficient way of measuring regulatory capital as well as for financial institutions control of risk.

A major impetus for the research is that standard finance models have a relatively poor record in predicting the changes in such correlations, and especially in times of extreme outcomes in the financial markets. The relationship between firm-level equity and corporate bond returns is particularly useful for examining the importance of common pricing factors because of the robust theoretical framework of Merton (1974), which implies the determinants of the correlations as well as the functional relationship between the correlation itself and its determinants. Using this benchmark, we are able to isolate the drivers of the correlations in a robust fashion.

To our knowledge, this is the first study to examine how variations in the common expectations of investors affect the correlation between firm-level equity and bond returns. We add to the literature by showing that changes in the common expectations of investors, which could be captured by an index of expected volatility such as VIX, may also be captured by employing new techniques which use textual analysis of digitised financial and economic news stories. In our study, we use a very large database of news in the Thomson Reuters News Archive database. This database is used in conjunction with a new metric (the relative sentiment shift—RSS) for measuring emotion in text-based articles (Nyman et al. 2018) that has not been previously used in finance studies.

Our study is able to examine the performance of other popular measures of variations in the common expectation of investors (Chicago Board Options Exchange's Volatility Index—VIX, Baker and Wurgler (2006)—BW, and Huang et al. (2015)—HJTZ) in order to be able to more clearly differentiate the information in these indexes. These innovative measures have become prominent in explaining changes in asset prices but have not yet been used to analyse the correlation between firm-level equity and bond returns. More importantly, these innovative indexes focus on different aspects of common expectations—BW and HJTZ look at economic indicators reflecting positive/optimistic expectations, and in contrast, VIX tends to look at negative/pessimistic expectations. By comparing their relative performances to the performance of the RSS, we also shed new light on different types of common expectations that may act upon asset prices. Taken collectively, the research offers further insight into the role of these common expectations as a common factor in the pricing of financial assets.

Most importantly, we show that RSS captures variations in the common risk premium and does so with greater power than previous measures. RSS is strongly correlated with VIX—the 1-month option-implied volatility on the S&P500 stock index, and conversely, we show that RSS is only weakly correlated with the existing indices of BW and HJTZ. The findings, all told, suggest that the *absence* of investors' optimism, when the basis on which BW and HJTZ are constructed is to show the *presence* of optimism, may not necessarily be linked to uncertainty or pessimistic expectations. In other words, when imagining a scale of optimism and pessimism or excitement and anxiety in expectations, the common factor's drivers may not be operating on the same spectrum. Instead, we suggest from our findings that it may be the function of two separate mechanisms. Simply put, the absence of the VIX need not imply optimism in expectations, and hence, VIX or RSS may better proxy for the expectation of negative outcomes and increased asset correlation. The indices of BW and HJTZ are not significant in explaining the stock–bond correlation.

Our results imply that the collective actions of investors are driven by changes in the common risk premium and we uncover a causal correlation between the common risk premium and the correlation between equity and corporate returns. An increase in RSS causes a weakening of the correlation in the subsequent month. This effect is itself dependent on firms' credit risk profiles and weakens in a monotonic fashion as firms move away from the default point. The VIX performs similarly overall to RSS but with some noticeable and practically very important differences. Most prominent of these advantages of RSS is that RSS has strong predictive power over the entire spectrum of credit risk, whereas VIX's effect is muted for some levels of credit risk. We are therefore able to offer a more fine-tuned measure to capture the causal links from expectations to correlations across all credit risk grades of firms.

The remainder of this paper is divided into five sections: In the following section, we review the relevant literature. In Sect. 3, we describe the data and present the methodology for estimating the RSS, the conditional correlation between equity and bond returns, a measure of credit risk, and the empirical model specifications. The results are presented in Sect. 4, and the robustness of the results is examined in Sect. 5. The main conclusions are summarised in Sect. 6.

Literature review

According to the structural model of Merton (1974), all securities issued by a firm may be considered as claims on the firm's assets. Consequently, the correlation between equity and debt securities should be primarily determined by the value and volatility of the firm's assets. The structural model implies that factors impacting the value of firm assets have a positive impact on the correlation, whereas factors affecting the volatility of firm assets induce a negative correlation. Furthermore, the model predicts a nonlinear relationship between the correlation and the asset value and volatility. Hence, the effect of all explanatory variables is strongest when firms are close to the default point and weakens as firms move away from the default point or as credit risk improves. The only common variable which is explicitly accounted for is the risk-free rate.

As Anderson and Sundaresan (2000) note, the structural model is attractive on theoretical grounds because it links the valuation of different financial claims on firms' assets with economic fundamentals. Despite its great appeal and precise formulation, the empirical evidence in support of the structural model has been mixed at best. While relationships between key variables of the structural model and credit spread are generally empirically confirmed as predicted by the model (e.g. Campbell and Taksler 2003), studies usually find that the structural model generates much lower credit spreads than those observed in reality. The inability of the structural model to predict highly realistic credit spreads is generally attributed to the simplifying assumptions used to derive the model and difficulties in estimation of parameters required for implementation. As far as our investigation is concerned, the crucial drawback of the Merton (1974) model stems from the critical assumption that the values of assets evolve as geometric Brownian motions. As a direct result of this assumption, the default probability and hence the credit spread are therefore implied by the volatility of assets and the difference between the asset value and the debt value. This difference, when divided mathematically by the volatility, is usually referred to as the 'distance to default'. Since under the assumptions of a geometric Brownian motion process, the continuously evolving value of assets needs time to change significantly, the default probability over short periods of time is close to zero. The structural model by relying on this assumption ignores the risk of substantial changes or 'jumps' in the values of assets. This is one of the main reasons why the structural model generates a lower credit spread than those observed.

As large shifts and sudden changes in the values of securities are rare events and usually confined to crisis and boom periods, it is near impossible to estimate the probability of their occurrence and the magnitude of their effect on asset prices with any high degree of precision. In other words, investors have a limited ability to quantify how likely rare events are ex ante and the scale of any impact upon asset prices if and when they occur. On this basis, Caballero and Krishnamurthy (2008) argue that these events are more about uncertainty, where outcomes do not occur on a known distribution with probabilities attached, rather than risk. Barro (2006) also emphasises the importance of rare events in asset pricing. He shows that accounting for rare events helps explain the high equity premium and other asset pricing puzzles.

There is much empirical support for the idea that the consequences of changes in the common expectations of investors are statistically and economically significant. Longstaff (2004) estimates that the liquidity premium, which is linked to sharp declines in the investors' expectations, accounts for as much as 15% of the value of risk-free bonds. Nozawa (2017) reports that variations in a common risk factor can explain about 50% of the credit spread. Collin-Dufresne et al. (2001) show that changes in corporate credit spreads are mostly driven by a single common factor rather than firm-level proxies for the theoretical variables. Demirovic et al. (2017) find that the relationship between equity and bond returns breaks down during periods of high uncertainty, measured by using the VIX index. They conclude that common factors overshadow the theoretical firm-level factors during financial crises, which is consistent with the flight-to-quality/flight-to-liquidity phenomenon and associated spikes in correlations between different assets classes. Durand et al. (2011) find that changes in the VIX, which proxies for investors' expectations of market volatility, affect the Fama–French risk factors (Fama and French 1993). On a more general note, as highlighted in introduction, Bollerslev et al. (2018) show that the volatility patterns in equity, bond, commodity, and currency markets are virtually identical.

Since variations in the common expectations of investors have been well examined in the literature and appear to have a significant role in asset pricing, it is primarily important to understand what drives them and also how best to take advantage of modern econometric and computer science techniques to measure them accurately.

As already noted, collective actions of investors are triggered by changes in expectations of future returns and risk. If investors' expectations of future market volatility affect risk factors as Durand et al. (2011) argue, these expectations will be relevant in explaining and predicting asset returns and correlations among asset returns. In line with this prediction, Lee et al. (2002) find that excess returns and volatility are contemporaneously correlated with changes in the ratio of 'bullish' to 'bearish' recommendations by investment advisory services. Likewise, Brown and Cliff (2005) find that shifts in the ratio of 'bullish' to 'bearish' recommendations are positively correlated with the pricing errors of models containing Fama–French and other widely used risk factors.

Baker and Wurgler (2006) construct a novel measure of investors' optimism by extracting the common components of six market-based variables (the closed-end fund discount, NYSE share turnover, the number of IPOs, the average first-day returns on IPOs, the equity share in new issues, and the dividend premium). They report that the cross section of future equity returns is conditional on investors' optimism at the period beginning. When investor optimism is high, securities that are attractive to speculators and unattractive to arbitrageurs tend to earn low returns in subsequent periods. This result implies that investor optimism is a noise driving prices away from their intrinsic values rather than a priced risk factor.

Huang et al. (2015) note that, by design, the principal component analysis used to isolate the common component of the variables used by Baker and Wurgler (2006) is unable to distinguish variations in the investors' optimism relevant for asset pricing from the approximation error of investors' optimism. They filter out the irrelevant approximation error by utilising the two-stage partial least squares estimation method. In line with Baker and Wurgler, they find that the first-day returns on IPOs and the equity share in new issues are the most significant proxies for investors' optimism. On the other hand, the number of IPOs and the dividend premium are not statistically significant components of their index.

Another approach to derive measures of investors' expectations is to use a survey of consumer confidence as a proxy or to survey investors directly. Another more recent alternative method infers changes in investors' expectations from the word content of digitised news media articles. One incarnation of this approach can be found in the combination of algorithmic data techniques to scan large samples of digitised news data. An excellent example of this approach is found in Manela and Moreira (2017) who examine the relationship between words appearing in the front page of the Wall Street Journal from 1890 to 2010 and changes in the VIX. They construct a measure which is significant in predicting equity returns and show that this return predictability can be attributed to changes in investors' concern about rare disasters.

Earlier work in this area has also produced other significant results. Cutler et al. (1989) examine the relationship between news related to fundamental economic values and changes in equity prices. They report that the relevant economic news cannot explain a significant portion of variations in equity prices. In other words, a large part of fluctuations in equity prices seems to be driven by news not related to systematic risks. Tetlock (2007) constructs a measure of investors' expectations from the Wall Street Journal's 'Abreast of the Market' daily column. He finds that the measure based on the number of words classified as pessimistic or optimistic in the daily column can predict future returns and trading volumes. Negative words are reported to be much more relevant than other words. Tetlock et al. (2008) examine the relevance of negative words in stories about S&P 500 firms published in the Wall Street Journal and the Dow Jones News Service. They find that the fraction of negative words forecasts low earnings and that stock prices temporarily underreact to the information embedded in negative words. Garcia (2013) adopts a similar approach and conducts the text analysis of financial columns published in the New York Times. He reports that the predictive power of news' content is concentrated in periods of elevated market uncertainty during recessions. In a similar analysis, Kräussl and Mirgorodskaya (2017) find that a measure based on the text analysis of financial columns predicts market returns and volatilities months in advance.

Da et al. (2015) use daily Internet search keywords and volume to capture changes in investors' expectations. They report that the number of queries for keywords such as recession and bankruptcy predicts aggregated market returns. Smales (2014) notes that a measure of investors' expectations based on the news for the constituents of the S&P 500 Index is negatively related to changes in the VIX. In other words, affirmative news coverages are related to a decrease in the VIX. He reports that this relationship is much stronger during the period of market turbulence in 2007–2009. A compelling rejoinder to these papers is that an absence of news coverage may also predict returns in asset markets. Fang and Peress (2009) find that media coverage affects equity returns. Securities with no media coverage significantly outperform securities regularly featured in media after controlling for common risk factors.

Methodology

This section describes our measure of investors' expectations, the methodology for estimating the conditional correlation between equity and bond returns, control variables, and the data set. Lastly, we specify the empirical model.

Relative sentiment shift (RSS)

Our approach to deriving the RSS index is similar to the approaches used by Tetlock (2007), Tetlock et al. (2008), Garcia (2013) and Kräussl and Mirgorodskaya (2017).

Generally, these measures are based on the frequency of words classified as positive or negative. Tetlock (2007) derives his measure by counting words in predetermined categories of the Harvard Psychosocial Dictionary appearing in the Wall Street Journal's daily column. Tetlock et al. (2008) consider the frequency of words appearing in all Wall Street Journal and Dow Jones News Service articles about S&P 500 firms. Garcia (2013) counts positive and negative words in two financial columns appearing in the New York Times. Kräussl and Mirgorodskaya (2017) obtain their measure from Wall Street Journal, Financial Times, and New York Times articles classified under the category Banking and Finance in the *LexisNexis* database.

The RSS is based on the frequency of words in news stories from the Thomson Reuters News Archive Database. Unlike other measures, the RSS is based on the frequency of preselected ordinary English emotional words (e.g. unease, perils, distrusted, brilliant, attracts, energising). This potentially makes the RSS orthogonal to any economic news in the text and at least partially mitigates the problem of reverse causality. In other words, the RSS is not directly based on the frequency of words which may be a reaction to the market movements. The approach directs the word search towards just two particular groups of emotion thought to encourage or inhibit action in conditions of uncertainty. Specifically, the emotion groups explored are those associated with excitement about gain (evoking approach) and anxiety about loss (evoking avoidance). A sample of words from both groups of words is presented in Nyman et al. (2018).

Given the word list, we define measures of anxiety and excitement as the sum of relative frequencies of words from the lists, while the RSS is defined as the difference between the measures of excitement and anxiety.

$$\mathrm{RSS}_t = \frac{\mathrm{Excitement}_t - \mathrm{Anxiety}_t}{N_t}. \tag{1}$$

where $\mathrm{Excitement}_t$ is the total number of occurrences of words from the excitement list in a period t, $\mathrm{Anxiety}_t$ is the total number of incidents of words from the anxiety list, and N_t is the total number of words.

Conditional correlation between equity and bond returns

Equity returns are calculated in the usual manner. Define $P^E_{i,t}$ as the share price of firm i at time t and $D_{i,t}$ as dividends paid from time $t-1$ to time t. The rate of return is defined as:

$$R^E_{i,t} = \ln\frac{P^E_{i,t} + D_{i,t}}{P^E_{i,t-1}}. \tag{2}$$

The holding period returns for bonds are calculated in a similar manner. Define $P^B_{i,t}$ as the bond price of firm i at time t, $C_{i,t}$ as the coupon payments, and $AC_{i,t}$ as the accrued interest on bond i from time $t-1$ to time t. The rate of return is defined as:

$$R^B_{i,t} = \ln\frac{P^B_{i,t} + C_{i,t} + AC_{i,t}}{P^B_{i,t-1} + AC_{i,t-1}} \tag{3}$$

The conditional correlation between equity and bond returns is obtained from a bivariate generalised autoregressive conditional heteroscedasticity (GARCH) process. The mean equations are given by:

$$R^E_t = c_1 + \varepsilon_{E,t} \quad \text{and} \quad R^B_t = c_2 + \varepsilon_{B,t}, \tag{4}$$

where R^E_t, R^B_t, $\varepsilon_{E,t}$, and $\varepsilon_{B,t}$ are equity and bond returns, and the disturbance terms, respectively. One of the most popular models for estimating the conditional covariance is Bollerslev et al. (1988). We use the parsimonious version referred to as the diagonal VECH (1,1). In order to guarantee that the conditional covariance matrix is positive semi-definite, we follow Engle and Kroner (1995) and Ding and Engle (2001) and restrict the coefficient matrices to rank 1 matrices. This gives the following variance/covariance equations:

$$\begin{aligned} h_{E,t} &= c_1 + a_1\varepsilon^2_{E,t-1} + b_1 h_{E,t-1} \\ h_{B,t} &= c_2 + a_2\varepsilon^2_{B,t-1} + b_2 h_{B,t-1} \\ h_{EB,t} &= c_1c_2 + a_1a_2\varepsilon^2_{E,t-1}\varepsilon^2_{B,t-1} + b_1b_2 h_{EB,t-1}, \end{aligned} \tag{5}$$

where $h_{E,t}$, $h_{B,t}$, and $h_{EB,t}$ are, respectively, the equity variance, the bond variance, and the equity–bond covariance. This specification is widely utilised in empirical studies (e.g. Bekaert and Wu 2000; Ang and Chen 2002; Belke 2011).

Control variables

The correlation between equity and bond returns depends on firms' credit risk exposure. The correlation of returns on securities issued by the riskiest firms is expected to the strongest. Therefore, it is critical to control for credit risk.

The distance to default (DD) is the difference between the market value of the assets and the book value of debt relative to the volatility of the market value of the assets (Merton 1974). DD follows directly from the Black and Scholes (1973) call option pricing equation:

$$E = AN(d_1) - De^{-rT}N(d_2), \tag{6}$$

where

$$d_1 = \frac{\ln\left(\frac{A}{D}\right) + \left(r + \frac{\sigma_A^2}{2}\right)T}{\sigma_A\sqrt{T}},$$

$d_2 = d_1 - \sigma_A\sqrt{T}$, E, and A are the market values of the firm's equity and assets, σ_A is the volatility of the market value of the firm's assets, D is the book value of the firm's debt, r is the risk-free rate, T is the time horizon in years, and $N(.)$ is the cumulative density of the standard normal distribution.

The market value of the firm's assets is assumed to follow a geometric Brownian motion process. Assuming that the firm's equity value follows the same process, its dynamics under the risk-neutral probability measure can be described by:

$$\mathrm{d}E = rE\,\mathrm{d}t + \sigma_E E\mathrm{d}X, \tag{7}$$

where σ_E is the volatility of the market value of the firm's equity and $\mathrm{d}X_t$ is the standard Wiener process. Since the equity value is a function of the asset value and time, Itô's lemma can be applied to give:

$$\mathrm{d}E = \left[\frac{\partial E}{\partial t} + rA\frac{\partial E}{\partial A} + \frac{1}{2}\left(\sigma_A A\right)^2\frac{\partial^2 E}{\partial A^2}\right]\mathrm{d}t + \frac{\partial E}{\partial A}\sigma_A A\mathrm{d}X. \tag{8}$$

A comparison of the coefficient multiplying the stochastic components in the two preceding equations gives the following identity:

$$\sigma_E E = \frac{\partial E}{\partial A}A\sigma_A. \tag{9}$$

The unobservable market value and volatility of the firm's assets are estimated by simultaneously solving Eqs. (6) and (9). This approach is widely used in empirical studies (e.g. Ronn and Verma 1986; Hillegeist et al. 2004; Campbell et al. 2008). The equity volatility is estimated using a GARCH(1,1) model (Bollerslev 1986). Once the asset value and volatility are estimated, the distance to default is calculated as follows:

$$DD = \frac{\ln\left(\frac{A}{D}\right) + \left(r - \frac{\sigma_A^2}{2}\right)T}{\sigma_A\sqrt{T}}. \tag{10}$$

To control for the maturity of bonds, the duration is calculated according to the following formula:

$$d = \frac{1}{B_d}\sum_{t=1}^{N}\frac{\mathrm{CF}_t}{(1+Y)^t}t, \tag{11}$$

where B_d is the dirty bond price (clean price + accrued interest), CF_t is the cash flow in period t, N is the number of periods to maturity, and Y is the per-period yield to maturity. The control variable for the size of the bond issue is the natural logarithm of the bond's market price multiplied by the number of outstanding bonds.

Panel data analysis

The data set consists of the conditional correlation between equity and bond returns, and a set of independent variables for n firms over T consecutive time periods. Because of the possible common factors influencing the correlation, we use a panel data model with period fixed effects:

$$C_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it}, \quad \varepsilon_{it} \sim i.i.d.\left(0, \sigma^2\right), \tag{12}$$

where C_{it} is the conditional correlation between the equity and bond returns of firm i at time t, α_i is the firm effect, β is the $k \times 1$ parameter vector, x_{it} is the vector of explanatory variables, and ε_{it} is the disturbance term.

We conduct our analysis by regressing the conditional correlation between equity and bond returns, C_{it}, on measures of the common factor and control variables. The common factor's measures are our index (RSS) and other popular indices (VIX, BW, and HJTZ). The control variables are DD, firm's size, bond duration, and bond issue size.

As discussed above, the most important control variable is DD which is a measure of credit risk. As implied by the structural model, the impact of a change in DD is expected to depend on the level of credit risk. In other words, a small change in a large DD should have only a limited impact on the correlation between equity and bond returns, while the magnitude of impact should grow as DD falls. To account for this non-linearity, we add the squared DD variable to the model.

Furthermore, to control for a potential discrete form of nonlinear impact of changes in credit risk, we include DD dummies, i.e. $\mathrm{DD}_{it}^s = I(\tau_s \leq \mathrm{DD}_{it} < \tau_{s+1})$, where $I(.)$ is the indicator function and τ_s are thresholds. The thresholds are commonly selected arbitrarily. Campbell and Taksler (2003), for example, use 5, 10, and 20 as the thresholds for the interest coverage variable and 10, 25, and 35% as the threshold values for the debt ratio. On the other hand, Cremers et al. (2008) treat financial leverage as a continuous variable. Since controlling for credit risk is crucial in this study, we select the number and values of thresholds as suggested by Hansen (2000). They are determined by estimating ms with different sets of dummies and threshold values and selecting the model on the basis of a minimum Akaike information criterion (AIC). Specifically, for each number of dummies ($s = 1, \ldots, n$), we estimate models for all threshold combinations (with discrete increment steps). For the DD variable, we estimated 4753 models. The lowest AIC is achieved with 15 dummies, but 94% of the improvement in AIC is achieved by a set of four dummies. Therefore, in order to have as parsimonious a model as possible, we use the optimal set of four dummy variables: 0.8, 1.8, 2.8, and 3.8. The coefficient of DD_{it} then captures the average effect of the DD, while the dummy coefficients capture the additional effect of the DD for predefined risk classes. The threshold selection procedure is described in 'Appendix'.

The common factor may have different effects for different firms. To account for this possibility, we include an interaction between the common factor and firms' credit risk (DD). Overall, the full regression is expressed as follows:

$$C_{it} = \alpha_i + \alpha_1 \mathrm{SENT}_t + \sum_{s=1}^{4} \delta_s \mathrm{SENT}_t \mathrm{DD}_{it}^s + \alpha_2 \mathrm{DD}_{it} + \alpha_3 \mathrm{DD}_{it}^2 + \sum_{s=1}^{4} \gamma_s \mathrm{DD}_{it}^s + \varepsilon_{it}. \tag{13}$$

As emphasised by Petersen (2009), the OLS standard errors in Eq. (13) may be biased and underestimate the true variability in the estimated coefficients due to the correlation of the residuals across time for a given firm (time series dependence) and/or across different firms (cross-sectional dependence). The standard errors are corrected to account for the cross-sectional dependence (i.e. common factors affecting all firms at the same time), while the serial dependence (i.e. firm-specific factors affecting individual firms) is addressed by adding dummy variables.

Data

We use the Thomson Reuters News Archive database which contains several million articles on political, financial, and economic news. We specifically focus on news articles 'tagged' (originating) in Washington DC and New York as those stories are most directly related to the US economy. These data average 10,832 articles per month over the period from 1996 to 2011. We exclude articles tagged as Sport, Weather, and Human Interest. Tetlock (2007) uses a single daily column, and Kräussl and Mirgorodskaya (2017) have 3135 articles per month which are less than a third of our sample. An extensive database ensures that our algorithm can produce high-quality results due to the volume of matches to our algorithm using a monthly aggregate.

Following Scruggs and Glabadanidis (2003), the correlation between equity and bond is estimated at the monthly level as noise in the returns at higher frequencies makes it challenging to determine the true relation between the returns. We use firm-level equity and bond data. Since bond data are relatively scarce compared to equity data, we start our sample selection with all straight corporate bonds issued by non-financial companies in the US market available in the Thompson Reuters Datastream database. When multiple bonds are available from the same issuer, the bond with the maximum number of observations is used. This is preferred to averaging the data of different bonds with a common issuer as bonds have different characteristics, such as duration and issue size. Bonds with less than 36 monthly observations, asset-backed bonds, bonds with any sort of collateral, or with an average market value of less than \$10 million are excluded from the sample. Once the bond data are collected, they are matched with the equity data, also obtained from the Datastream database. The matched sample consists of 351 firms and 33,870 firm-month observations.

All other variables (distance to default, firm asset value, bond duration, and bond issue value) are estimated at the daily level and then converted into monthly series by averaging daily observations. The accounting data required for the estimation of DD are obtained from Compustat, and the risk-free interest rate data are obtained from the Datastream database. VIX Index data are obtained from the Chicago Board Options Exchange. Finally, Baker and Wurgler (2006) and Huang et al. (2015) indices data are obtained from Jeffrey Wurgler's and Guofu Zhou's websites. The descriptive statistics for the variables used in the empirical model are presented in Table 1.

The sample covers the period from August 1996 to February 2011. Not all series cover the entire sample period, so our panel is unbalanced. It should be noted that the number of observations available at the beginning of the sample period (1996–2000) is much lower than that later in the sample period (2001–2011). However, the earlier dataset is still large (1519 observations for 33 firms) compared to other studies dealing with bond data (Table 2).

Results

We first examine the relationship between the RSS, VIX, BW (Baker and Wurgler 2006), and HJTZ (Huang et al. 2015) indices. The RSS and VIX indices are strongly correlated ($\rho = -64\%$). Similarly, a strong correlation is observed between the BW and the HJTZ. This implies that the RSS/VIX and BW/HJTZ capture different effects. The correlation coefficients are presented in Table 3 while the normalised time series are illustrated in Fig. 1.

To examine the relationships among the RSS, VIX and other indices, we estimate a VAR(2) model. The first lag of RSS is significant at the 5% level in explaining changes in the VIX, while the second lag of RSS is significant at the 10% level in explaining changes in the BW index. The lags of HJTZ are significant in explaining changes in the RSS and the 5% level and other indices at the 1% level, while BW's lags explain changes in the HJTZ. The model explains over 95% variations in the BW and HJTZ, 78% of fluctuations in the VIX, and only 46% of changes in the RSS.

Table 1 Descriptive statistics

	Mean	Maximum	Minimum	SD	Skewness	Kurtosis
Equity–bond returns correlation	0.09	1.00	−0.93	0.28	0.25	3.12
RSS	−0.10	−0.01	−0.27	0.05	−0.97	4.25
BW (Baker and Wurgler 2006, Eq. 2)	0.17	2.84	−0.93	0.77	1.34	4.91
HJTZ (Huang et al. 2015)	−0.04	3.03	−1.03	0.90	1.61	4.93
VIX	22.37	62.64	10.82	8.45	1.73	8.17
Equity return	0	1.64	−1	0.11	−0.73	16.62
Bond period return	0.01	1.19	−0.75	0.04	−0.03	83.62
Distance to default	5.43	28.05	−2.42	2.73	0.87	4.87
Asset value	25,449.62	832,438.00	66.69	46,157.04	5.73	61.39
Bond duration	6.91	18.24	0	3.61	0.19	2.08
Bond issue value	238.06	4567.08	2.12	268.3	5.1	54.3

The RSS is obtained from Eq. (1). Asset value and bond issue value are in USD millions. Equity and bond returns are logarithmic as specified in Eqs. (2) and (3); the distance to default (DD) is the difference between the market value of the assets and the book value of debt relative to the volatility of the market value of the assets (Eq. 10); the duration is in years (Eq. 11). The indices, equity, and bond returns are calculated at monthly frequency, while other variables are calculated at the daily frequency and are converted into monthly series as the average of daily observations within the given months

Table 2 Correlation matrix

	RSS	VIX	BW	HJTZ
RSS	1.00			
VIX	−0.64	1.00		
BW	0.05	−0.04	1.00	
HJTZ	−0.11	0.35	0.66	1.00

Table 3 VAR model of the indices

	RSS_t	VIX_t	BW_t	$HJTZ_t$
RSS_{t-1}	0.64***	−0.16**	0.01	0.01
RSS_{t-2}	0.12	0.06	0.04	−0.02
VIX_{t-1}	0.1	0.87***	−0.05	0
VIX_{t-2}	−0.03	−0.1	−0.01	−0.02
BW_{t-1}	−0.29	−0.01	0.86***	0.2***
BW_{t-2}	0.35	−0.01	0.04	−0.21***
$HJTZ_{t-1}$	0.56**	0.34*	0.15*	1.45***
$HJTZ_{t-2}$	−0.65**	−0.25	−0.04	−0.48***
C	0.00	0	−0.01	0
Adjusted R^2	0.46	0.78	0.96	0.97

(−1) and (−2) indicate the first and the second lags of the variables

***,**,*Significance at the 1%/5%/10%

Table 4 depicts an estimate of the model in Eq. (13). The model regresses the correlation between equity and bond returns on the investors' expectations and controls for credit risk. The results reveal that the RSS and the three expectation–credit risk interaction variables are significant in explaining the correlation. The relationship is negative. In other words, an increase in the RSS is associated with a decrease in the correlation. The effect is the strongest for the riskiest firms and monotonically weakens as firms move away from the default point. The equation with the VIX yields similar results, while the model with the BW and HJTZ indices provides mixed results. In the BW model, only one interaction variable is significant, while the index and two interaction coefficients are significant in the HJTZ models.

Re-estimating the model presented in Table 4 with the first lags of the explanatory variables reveals that the RSS is able to predict the correlation. The coefficients imply that an increase in the RSS leads to a decrease in the correlation in a subsequent month. The strength of this effect depends on a firm's credit risk. The first lag of the index coefficient and the first three interaction coefficients are significant while the fourth (i.e. the lowest credit risk) interaction coefficient is not significant at any of the conventional levels. The VIX model yields similar results. In the HJTZ model, the index and two interaction coefficients are significant, while only one interaction coefficient is significant in the BW model.

As depicted by Eq. (1), the RSS is the difference between the frequency of words indicating excitement and the frequency of words indicating anxiety. Investors may respond to positive and negative news differently (e.g. Brown et al. 1988; Tetlock 2007). To examine this hypothesis, we replace the RSS with the excitement and anxiety components and estimate the models presented in Tables 4 and 5. In the contemporaneous model, the coefficients of the RSS components and most of the interaction variables are significant. The coefficient signs and magnitude are as expected—an increase in excitement lowers the correlation while an increase in anxiety strengthens the correlation. The effect is more substantial for riskier firms. Surprisingly, the

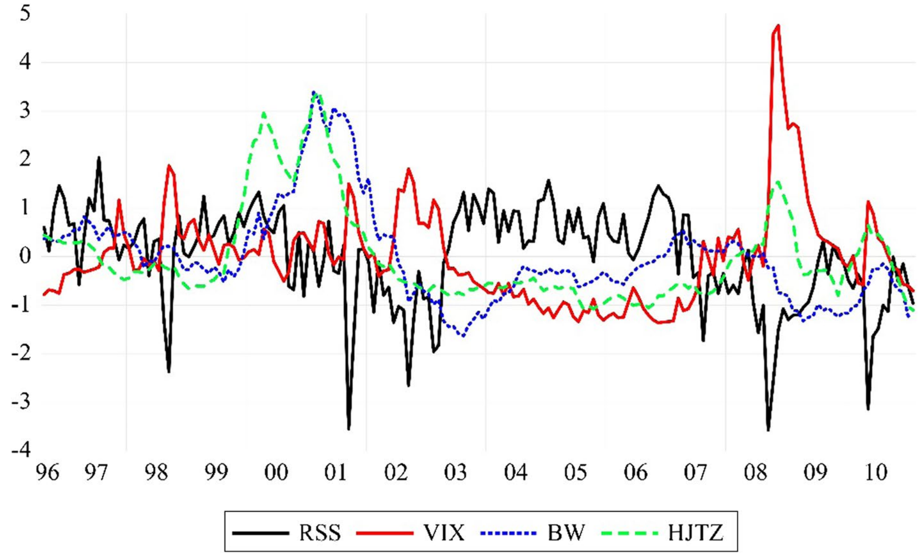

Fig. 1 Normalised values of VIX, RSS, BW, and HJTZ

Table 4 Cotemporaneous relationship between the indices and the conditional correlation between equity and bond returns

Variable	RSS	VIX	BW	HJTZ
DD	−0.01**	−0.01***	0.00	0.00
DD^2	0.00***	0.00**	0.00	0.00
$I(DD<0.8)$	−0.13***	−0.15**	0.07**	0.01
$I(DD\geq 0.8$ AND $DD<1.8)$	−0.1**	−0.15***	0.03*	0.02
$I(DD\geq 1.8$ AND $DD<2.8)$	−0.03**	−0.05**	0.02*	0.02**
$I(DD\geq 2.8$ AND $DD<3.8)$	−0.01	−0.01	0.00	0.00
INDEX	25.47***	0.00***	0.00	−0.01**
INDEX * $I(DD<0.8)$	−109.17**	0.01***	0.06**	0.07**
INDEX * $I(DD\geq 0.8$ AND $DD<1.8)$	−88.66***	0.01***	−0.01	0.05**
INDEX * $I(DD\geq 1.8$ AND $DD<2.8)$	−35.23***	0.00***	−0.01	0.01
INDEX * $I(DD\geq 2.8$ AND $DD<3.8)$	−9.23	0.00	−0.01	0.00
C	0.14***	0.16	0.09*	0.09***
Adjusted R^2	61.1%	61.3%	61.0%	61.1%

The models are estimated using fixed firm effects with correction for cross-sectional error dependence. The estimated model is Eq. (13). The dependent variable, C_{it}, is the conditional correlation estimated from the VECH (1,1) from Eq. (5). The distance to default, DD_{it}, is obtained from Eq. (10), and $I(.)$ is an indicator function that equals 1 if the argument is true and zero otherwise

***,**,*Significance at the 1%/5%/10% level

magnitude of coefficients implies that a change in excitement is associated with a larger change in the correlation than a change in anxiety. For firms with the distance to default from 0.8 to 1.8, a 0.001 increase in excitement is associated with a decrease in the correlation of 18% while a similar increase in anxiety is associated with an increase in the correlation of 5%.

The statistical significance and magnitude of coefficients in the causal models reveal that the predictive power of the RSS is due to its anxiety component. All anxiety coefficients are statistically significant and show that an increase in excitement causes an increase in the correlation. The effects monotonically decrease as firms move away from the default point. On the other hand, only two interaction coefficients of the excitement components are significant.

As noted above, the effect of changes in the investors' expectations depends on the riskiness of firms. The effect weakens as firms become safer and reverses the sign for the safest group of firms. In other firms, a negative change in the expectations of investors is associated with a steep increase in the correlation between equity and bond returns of the riskiest firms, whereas it is related to a decrease in the correlation of the safest firms (Fig. 2).

A similar dependence on credit risk is present in the causal model too. A decline in the investors' expectations

Table 5 Causal relationship between the indices and the conditional correlation between equity and bond returns

Variable	RSS_t	VIX_t	BW_t	$HJTZ_t$
DD_{t-1}	−0.01**	−0.01***	0.00	−0.004
DD^2_{t-1}	0.00*	0.00**	0.00	0.000*
$I(DD_{-1} < 0.8)$	−0.25**	−0.14**	0.03	0.000
$I(DD_{t-1} \geq 0.8 \text{ AND } DD_{t-1} < 1.8)$	−0.13***	−0.13***	0.02	0.010
$I(DD_{t-1} \geq 1.8 \text{ AND } DD_{t-1} < 2.8)$	−0.05**	−0.04*	0.01	0.008
$I(DD_{t-1} \geq 2.8 \text{ AND } DD_{t-1} < 3.8)$	−0.03**	−0.02	0.01	0.007
$INDEX_{t-1}$	18.18**	0.00***	0.00	−0.006*
$INDEX_{t-1}*(DD_{t-1} < 0.8)$	−162.09**	0.00**	0.05**	0.032
$INDEX_{t-1} * I(DD_{t-1} \geq 0.8 \text{ AND } DD_{t-1} < 1.8)$	−94.71***	0.00***	−0.01	0.031*
$INDEX_{t-1} * I(DD_{t-1} \geq 1.8 \text{ AND } DD_{t-1} < 2.8)$	−38.52***	0.00*	−0.01	0.000
$INDEX_{t-1} * I(DD_{t-1} \geq 2.8 \text{ AND } DD_{t-1} < 3.8)$	−30.26**	0.00	0.00	0.003
Intercept	0.13***	0.17***	0.10***	0.1***
Adjusted R^2	61.1%	61.2%	61.0%	61.0%

The models are estimated using fixed firm effects with correction for cross-sectional error dependence. The estimated model is Eq. (13). The dependent variable, C_{it}, is the conditional correlation estimated from the VECH (1,1) from Eq. (5). The distance to default, DD_{it}, is obtained from Eq. (10), and $I(.)$ is an indicator function that equals 1 if the argument is true and zero otherwise. $(_{t-1})$ indicates the first lag of the variable

***,**,*Significance at the 1%/5%/10% level

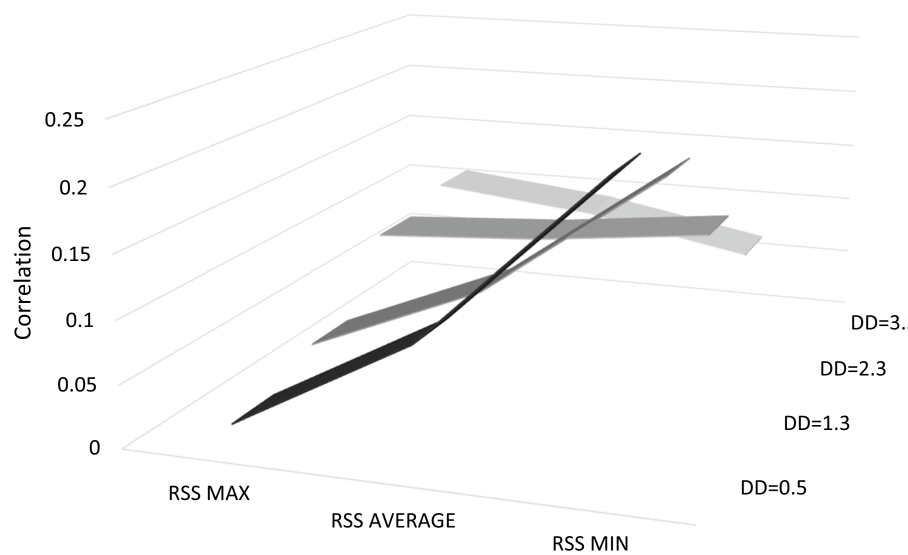

Fig. 2 Contemporaneous correlation between equity and bond returns conditioned to the maximum, average, and minimum RSS value

causes a steep increase in the correlation of risky firms. The magnitude of the effect decreases as firms move away from the default point and becomes essentially zero for the safest firms (Fig. 3).

Robustness checks

Firm size, bond duration, and bond issue size are important characteristics that can potentially influence the correlation between equity and bond returns. Credit risk exposure may be related to firm size. The relation between the duration and the risk inherent in a bond is straightforward: a longer duration indicates higher risk, *ceteris paribus*. Therefore, the returns on long-term bonds should behave more like equity returns than the returns on short-term bonds. The size of a bond issue may affect the correlation through the liquidity mechanism. Large bond issues are more liquid, and therefore, their values should react more quickly to shocks in the value of the issuing firm's equity.

To examine whether the results are sensitive to changes in firm size, the maturity, and liquidity of bonds in the sample, the models presented in Tables 4, 5 and 6 are augmented with three sets of dummy variables to control for firm size, bond duration, and issue size. The largest firms, the largest issue size, and the longest duration are the benchmarks for which we have no dummies in order to avoid the dummy variable trap. In each case, the number of dummies and

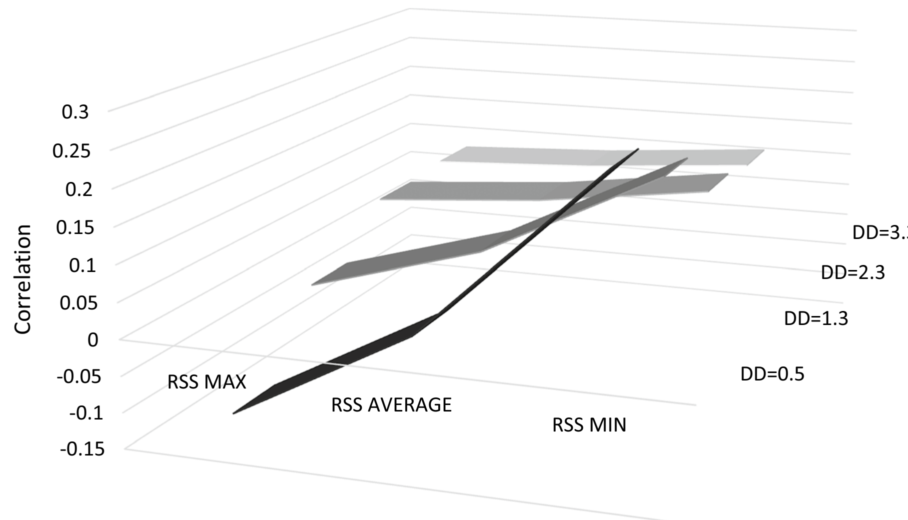

Fig. 3 Causal correlation between equity and bond returns conditioned to the maximum, average, and minimum RSS value

Table 6 Cotemporaneous and causal relationship between the RSS's components and the conditional correlation between equity and bond returns

	Cotemporaneous	Causal
DD	−0.006**	−0.005**
DD^2	0.000*	0.000*
I(DD < 0.8)	0.148	−0.62***
I(DD ≥ 0.8 AND DD < 1.8)	0.119	0.075
I(DD ≥ 1.8 AND DD < 2.8)	0.093	0.012
I(DD ≥ 2.8 AND DD < 3.8)	−0.024	−0.111**
RSS Anxiety	−24.208***	−17.886**
RSS Anxiety * I(DD < 0.8)	82.708	198.743***
RSS Anxiety * I(DD ≥ 0.8 AND DD < 1.8)	72.918**	80.342***
RSS Anxiety * I(DD ≥ 1.8 AND DD < 2.8)	27.465**	34.878***
RSS Anxiety * I(DD ≥ 2.8 AND DD < 3.8)	10.137	34.933***
RSS Excitement	36.769**	20.026
RSS Excitement * I(DD < 0.8)	−255.61*	33.263
RSS Excitement * I(DD ≥ 0.8 AND DD < 1.8)	−215.776***	−212.409***
RSS Excitement * I(DD ≥ 1.8 AND DD < 2.8)	−109.679***	−74.027*
RSS Excitement * I(DD ≥ 2.8 AND DD < 3.8)	1.147	18.248
Intercept	0.117***	0.124***
Adjusted R^2	61.2%	61.2%

The models are estimated using fixed firm effects with correction for cross-sectional error dependence. The estimated model is Eq. (13). The dependent variable, C_{it}, is the conditional correlation estimated from the VECH (1,1) from Eq. (5). The distance to default, DD_{it}, is obtained from Eq. (10), and I(.) is an indicator function that equals 1 if the argument is true and zero otherwise. The first lags of the variables are used in the causal model

***,**,*Significance at the 1%/5%/10% level

their associated thresholds were selected using the approach detailed in 'Appendix'.

The augmented cotemporaneous models are presented in Table 7. The coefficients of the bond value and duration control variables are mostly significant, but the results on the effect of changes in the investor's expectations are not altered. The last RSS–credit risk interaction variable, which is insignificant in Table 4, is now significant at the ten per cent level.

The results presented in Table 8 show that the results on the causal relationship between changes in the investors' expectations and the correlation hold after controlling for the firm size, bond issue size, and duration. An improvement in the investors' expectations leads to a weaker correlation between equity and bond returns.

Finally, the models presented in Table 6 are augmented with the control variables. The first excitement–credit risk

Table 7 Contemporaneous relationship between the indices and the conditional correlation between equity and bond returns—robustness checks

	RSS	VIX	BW	HJTZ
DD	−0.008***	−0.008***	−0.003	−0.004
DD^2	0.000**	0.000**	0.000	0.000*
I(DD < 0.8)	−0.086	−0.112*	0.088***	0.030
I(DD ≥ 0.8 AND DD < 1.8)	−0.095**	−0.141***	0.031**	0.020*
I(DD ≥ 1.8 AND DD < 2.8)	−0.036**	−0.045***	0.014*	0.014*
I(DD ≥ 2.8 AND DD < 3.8)	−0.012	−0.009	0.001	0.001
INDEX	29.849***	−0.002***	−0.001	−0.005*
INDEX * I(DD < 0.8)	−89.841**	0.004***	0.063**	0.058**
INDEX * I(DD ≥ 0.8 AND DD < 1.8)	−84.093***	0.005***	−0.009	0.043**
INDEX * I(DD ≥ 1.8 AND DD < 2.8)	−37.014***	0.002***	−0.007	0.004
INDEX * I(DD ≥ 2.8 AND DD < 3.8)	−10.979*	0.00	−0.006	−0.002
Intercept	0.213***	0.223***	0.169***	0.155***
Asset Value Dummy 1 (smallest)	−0.277***	−0.257***	−0.277***	−0.259***
Asset Value Dummy 2	−0.029*	−0.015	−0.026	−0.018
Asset Value Dummy 3	−0.020*	−0.010	−0.016	−0.012
Asset Value Dummy 4	0.008	0.015*	0.014	0.016*
Asset Value Dummy 5	0.002	0.007	0.007	0.008
Asset Value Dummy 6	−0.001	0.001	0.003	0.002
Bond Value Dummy 1 (smallest)	−0.223***	−0.220***	−0.234***	−0.212***
Bond Value Dummy 2	−0.122***	−0.121***	−0.123***	−0.103***
Bond Value Dummy 3	−0.051	−0.045	−0.058	−0.040
Bond Value Dummy 4	−0.073**	−0.068**	−0.081**	−0.066*
Bond Value Dummy 5	−0.094***	−0.092***	−0.1***	−0.086***
Bond Value Dummy 6	−0.075**	−0.074**	−0.08***	−0.067**
Bond Value Dummy 7	−0.099***	−0.097***	−0.104***	−0.09***
Bond Value Dummy 8	−0.078***	−0.073***	−0.085***	−0.074***
Bond Duration Dummy 1 (shortest)	0.063***	0.051***	0.047***	0.049***
Bond Duration Dummy 2	0.048***	0.039***	0.036***	0.038***
Bond Duration Dummy 3	0.016***	0.011*	0.010*	0.010*

The models are estimated using fixed firm effects with correction for cross−sectional error dependence. The estimated model is Eq. (13). The dependent variable, C_{it}, is the conditional correlation estimated from the VECH (1,1) from Eq. (5). The distance to default, DD_{it}, is obtained from Eq. (10), and I(.) is an indicator function that equals 1 if the argument is true and zero otherwise

The asset value Dummies 1–6 take the value of 1 if the log of firm's asset value in millions of US dollars is less than 6 or between the two thresholds of 6, 7, 8, 9, 10, 11 (i.e. 403, 1097, 2981, 8103, 22,026, 59,874 million). The Bond Value Dummies 1–8 take the value of 1 if the log of bond issue value in millions of US dollars is less than 3.1 or between the two thresholds of 3.1, 3.6, 4.1, 4.6, 5.1, 5.6, 6.1, and 6.6 (i.e. 22, 37, 60, 99, 164, 270, 446, and 735 million). The Bond Duration Dummies 1–3 take the value of 1 if the bond duration is less than 3.7 or between the two thresholds of 3.7, 7.2, and 10.5 years. The selection of dummy variable sets is described in 'Appendix'

***,**,*Significance at the 1%/5%/10% level

interaction variables in the contemporaneous model, which is significant at the ten per cent level in the un-augmented model, are now insignificant, in which the last anxiety–credit risk interaction variable becomes significant at the ten per cent level. However, the obtained results, including the one that changes in the anxiety component cause changes in the correlation, are unaltered (Table 9).

Conclusions

It is well documented that investors' preference for assets of the highest quality and liquidity sharply increases when the markets are unstable, such as crises periods like the one witnessed in 2008. At these times, shifts in the common expectations about risk and returns are the primary drivers of asset prices. Moreover, a recent study by Bollerslev et al.

Table 8 Causal relationship between the indices and the conditional correlation between equity and bond returns—robustness checks

	RSS_t	VIX_t	BW_t	$HJTZ_t$
DD_{t-1}	−0.006**	−0.009***	−0.003	−0.004
DD^2_{t-1}	0.000**	0.000**	0.000	0.000*
$I(DD_{t-1} < 0.8)$	−0.231**	−0.108	0.045	0.015
$I(DD_{t-1} \geq 0.8$ AND $DD_{t-1} < 1.8)$	−0.127***	−0.123***	0.016	0.009
$I(DD_{t-1} \geq 1.8$ AND $DD_{t-1} < 2.8)$	−0.052***	−0.043*	0.005	0.005
$I(DD_{t-1} \geq 2.8$ AND $DD_{t-1} < 3.8)$	−0.036***	−0.026	0.004	0.004
$INDEX_{t-1}$	22.326***	−0.002***	−0.002	−0.005*
$INDEX_{t-1} * I(DD-_{t-1} < 0.8)$	−157.403***	0.004**	0.05**	0.023
$INDEX_{t-1} * I(DD_{t-1} \geq 0.8$ AND $DD_{t-1} < 1.8)$	−93.133***	0.004***	−0.004	0.028*
$INDEX_{t-1} * I(DD_{t-1} \geq 1.8$ AND $DD_{t-1} < 2.8)$	−40.553***	0.002*	−0.006	0.000
$INDEX_{t-1} * I(DD_{t-1} \geq 2.8$ AND $DD_{t-1} < 3.8)$	−31.897***	0.001	−0.004	0.003
Intercept	0.182***	0.207***	0.143***	0.14***
Asset Value Dummy 1 (smallest)	−0.256***	−0.239***	−0.249***	−0.245***
Asset Value Dummy 2	−0.017	−0.006	−0.014	−0.012
Asset Value Dummy 3	−0.013	−0.006	−0.011	−0.009
Asset Value Dummy 4	0.013	0.017*	0.016*	0.017*
Asset Value Dummy 5	0.006	0.008	0.009	0.009
Asset Value Dummy 6	0.001	0.001	0.003	0.003
Bond Value Dummy 1 (smallest)	−0.193***	−0.179***	−0.184***	−0.18***
Bond Value Dummy 2	−0.096**	−0.083*	−0.08*	−0.079*
Bond Value Dummy 3	−0.029	−0.019	−0.023	−0.019
Bond Value Dummy 4	−0.055	−0.048	−0.053	−0.048
Bond Value Dummy 5	−0.079**	−0.073**	−0.075**	−0.072**
Bond Value Dummy 6	−0.065**	−0.061*	−0.061*	−0.057*
Bond Value Dummy 7	−0.091***	−0.088***	−0.09***	−0.085***
Bond Value Dummy 8	−0.074***	−0.07***	−0.079***	−0.074***
Bond Duration Dummy 1 (shortest)	0.063***	0.058***	0.053***	0.056***
Bond Duration Dummy 2	0.049***	0.044***	0.04***	0.043***
Bond Duration Dummy 3	0.016***	0.014**	0.012*	0.013**

The models are estimated using fixed firm effects with correction for cross-sectional error dependence. The estimated model is Eq. (13). The dependent variable, C_{it}, is the conditional correlation estimated from the VECH (1,1) from Eq. (5). The distance to default, DD_{it}, is obtained from Eq. (10), and $I(.)$ is an indicator function that equals 1 if the argument is true and zero otherwise. (−1) indicates the first lag of the variable

The asset value Dummies 1–6 take the value of 1 if the log of firm's asset value in millions of US dollars is less than 6 or between the two thresholds of 6, 7, 8, 9, 10, 11 (i.e. 403, 1097, 2981, 8103, 22,026, 59,874 million). The Bond Value Dummies 1–8 take the value of 1 if the log of bond issue value in millions of US dollars is less than 3.1 or between the two thresholds of 3.1, 3.6, 4.1, 4.6, 5.1, 5.6, 6.1, and 6.6 (i.e. 22, 37, 60, 99, 164, 270, 446, and 735 million). The Bond Duration Dummies 1–3 take the value of 1 if the bond duration is less than 3.7 or between the two thresholds of 3.7, 7.2, and 10.5 years. The selection of dummy variable sets is described in 'Appendix'

***,**,*Significance at the 1%/5%/10% level

(2018) shows that volatility patterns across the markets are highly similar most of the time. This implies that the importance of changes in the common expectations of investors extends beyond the relatively short and infrequent episodes of flights to quality and liquidity.

In line with other studies, authors examining the pricing of corporate debt find that the firm-level variables implied by the structural model of Merton (1974) cannot explain a large portion of the credit spread. Consequently, they emphasise the importance of a common factor. We contribute to the literature by examining how changes in the common expectations of investors affect the correlation between equity and corporate bond returns. We use a measure of investor's expectations, termed the relative sentiment shift (RSS), not previously used in financial research. The RSS is defined as the difference in relative frequencies of words signalling excitement and anxiety. We also consider alternative measures of the investors' expectations, namely the VIX, BW (Baker and Wurgler 2006), and HJTZ (Huang et al. 2015) indices.

Table 9 Cotemporaneous and causal relationship between the RSS's components and the conditional correlation between equity and bond returns—robustness checks

	Cotemporaneous	Causal
DD	−0.007**	−0.006**
DD^2	0**	0**
I(DD < 0.8)	0.131	−0.708***
I(DD ≥ 0.8 AND DD < 1.8)	0.065	0.038
I(DD ≥ 1.8 AND DD < 2.8)	0.089	0.006
I(DD ≥ 2.8 AND DD < 3.8)	−0.024	−0.105**
RSS Anxiety	−29.276***	−22.554**
RSS Anxiety * I(DD < 0.8)	69.076	203.868***
RSS Anxiety * I(DD ≥ 0.8 AND DD < 1.8)	72.476**	81.18***
RSS Anxiety * I(DD ≥ 1.8 AND DD < 2.8)	29.066***	36.864***
RSS Anxiety * I(DD ≥ 2.8 AND DD < 3.8)	11.489*	35.875***
RSS Excitement	35.815***	19.419
RSS Excitement * I(DD < 0.8)	−205.182	94.015
RSS Excitement * I(DD ≥ 0.8 AND DD < 1.8)	−177.396**	−189.25**
RSS Excitement * I(DD ≥ 1.8 AND DD < 2.8)	−111.977***	−75.01*
RSS Excitement * I(DD ≥ 2.8 AND DD < 3.8)	−3.088	10.14
Intercept	0.202***	0.186***
Asset Value Dummy 1 (smallest)	−0.268***	−0.258***
Asset Value Dummy 2	−0.027	−0.017
Asset Value Dummy 3	−0.02*	−0.013
Asset Value Dummy 4	0.008	0.013
Asset Value Dummy 5	0.002	0.006
Asset Value Dummy 6	−0.001	0.001
Bond Value Dummy 1 (smallest)	−0.224***	−0.191***
Bond Value Dummy 2	−0.122***	−0.093**
Bond Value Dummy 3	−0.052	−0.027
Bond Value Dummy 4	−0.074**	−0.055
Bond Value Dummy 5	−0.095***	−0.078**
Bond Value Dummy 6	−0.076**	−0.064**
Bond Value Dummy 7	−0.1***	−0.09***
Bond Value Dummy 8	−0.079***	−0.074***
Bond Duration Dummy 1 (shortest)	0.064***	0.063***
Bond Duration Dummy 2	0.048***	0.048***
Bond Duration Dummy 3	0.016***	0.016***

The models are estimated using fixed firm effects with correction for cross-sectional error dependence. The estimated model is Eq. (13). The dependent variable, C_{it}, is the conditional correlation estimated from the VECH (1,1) from Eq. (5). The distance to default, DD_{it}, is obtained from Eq. (10), and I(.) is an indicator function that equals 1 if the argument is true and zero otherwise. The first lags of the variables are used in the causal model

The asset value Dummies 1–6 take the value of 1 if the log of firm's asset value in millions of US dollars is less than 6 or between the two thresholds of 6, 7, 8, 9, 10, 11 (i.e. 403, 1097, 2981, 8103, 22,026, 59,874 million). The Bond Value Dummies 1–8 take the value of 1 if the log of bond issue value in millions of US dollars is less than 3.1 or between the two thresholds of 3.1, 3.6, 4.1, 4.6, 5.1, 5.6, 6.1, and 6.6 (i.e. 22, 37, 60, 99, 164, 270, 446, and 735 million). The Bond Duration Dummies 1–3 take the value of 1 if the bond duration is less than 3.7 or between the two thresholds of 3.7, 7.2, and 10.5 years. The selection of dummy variable sets is described in 'Appendix'

***,**,*Significance at the 1%/5%/10% level

The RSS is strongly correlated with VIX ($\rho = -64\%$) and weakly correlated with the BW and HJTZ indices. This implies that the RSS and VIX capture the same effect, whereas the other two indices proxy for another effect. In a VAR(2) model of all four indices, the RSS equation has the lowest explanatory power ($R^2 = 48\%$) which suggests that the largest portion of changes in the RSS is due to orthogonal factors. The first lag of RSS is significant in explaining VIX while past changes in VIX do not explain variations in RSS.

After controlling for credit risk, we find that the RSS and VIX perform similarly in explaining the contemporaneous correlation between equity and bond returns. An improvement in the investors' expectations is associated with a weaker correlation. The magnitude of the effect decreases as firms move away from the default point. In other words, a change in the investors' expectations will have an outsized impact on the correlation between equity and bond returns of the riskiest firms and little or no impact on the correlation of the safest firms.

Moreover, shifts in the investors' expectations, proxied by the RSS and VIX indices, are significant in *predicting* changes the correlation between equity and bond returns. An improvement in the investors' expectations causes a weaker correlation. Here we also discover that RSS is much more effective than VIX at predicting correlation across the credit risk spectrum. The effect is stronger for the riskiest firms and flattens as firms' credit risk improves. As we are able to decompose RSS into the excitement/optimism and anxiety/pessimism components, we find that this predictive power stems primarily from changes to the anxiety components. In other words, shifts in investors' anxiety lead to changes in the correlation between equity and bond returns of risky firms across the full credit spectrum.

Appendix

The threshold values for the dummy variables are determined in the spirit of Hansen (2000). We basically regress the conditional correlation between equity and bond returns on all potential combinations of models using predetermined threshold increments and select the optimal model based on the Akaike information criterion (AIC).

We assign $D_{it}^{\mathrm{k}} = I(\tau_{\mathrm{k}-1} \leq X_{it} < \tau_{\mathrm{k}})$ as the dummy variables, where X_{it} is the value of a variable for firm i at time t, τ_{k} are thresholds, and $I(.)$ is the indicator function. τ_0 is equal to the variable's sample minimum, the first threshold, τ_1, is equal to the lower limit K_{L}, and the last threshold, τ_{n}, is equal to the upper limit K_{U}. The difference between the lower and upper limits covers the large majority of observations. The first threshold, τ_1, increases by an increment of 0.1, and the difference between two thresholds,s, starts at 0.5 and increases by an increment of 0.5, i.e. $\delta = 0.5$, 1.0, 1.5,…. The threshold selection procedure involves estimation of models with all possible combinations of the number or thresholds (n), the starting value of τ_1, and the differences between two thresholds (δ), which covers the range from K_{L} to K_{U}.

In the case of one threshold, the procedure simplifies to estimating the models with one dummy variable $D_{it}^{1} = I(X_{it} < \tau_1)$ with $\tau_1 = K_{\mathrm{L}}, K_{\mathrm{L}} + 0.1, K_{\mathrm{L}} + 0.2, \ldots, K_{\mathrm{U}}$.

In the case of two thresholds, where $\tau_1 = K_{\mathrm{L}}, K_{\mathrm{L}} + 0.1, K_{\mathrm{L}} + 0.2, \ldots, K_{\mathrm{U}} - \delta$ and $\tau_2 = \tau_1 + \delta$.

In the case of n thresholds, $D_{it}^{1} = I(X_{it} < \tau_1)$ and $D_{it}^{\mathrm{k}} = I\left(\tau_{\mathrm{k}-1} \leq X_{it} < \tau_{\mathrm{k}}\right)$, where $\tau_1 = K_{\mathrm{L}}, K_{\mathrm{L}} + 0.1, K_{\mathrm{L}} + 0.2, \ldots, K_{\mathrm{U}} - (n-1)\delta$ and $\tau_{\mathrm{k}} = \tau_1 + (k-1)\delta$.

Performing the above procedure on the distance to default variable requires estimation of 4753 models with all combinations of the number of dummies (1–26), the starting value of $\tau_1 = 0.5$, and the differences between two thresholds $\delta = 0.5$, 1, 1.5…, which cover the range from $K_{\mathrm{L}} = 0.5$ to $K_{\mathrm{U}} = 13$. The values between these limits cover 98% of the observations. The lowest AIC gives a model with 15 dummy variables or thresholds, and the lowest SSE gives a model with 26 dummy variables. The greatest improvement in AIC (94%) and SSE (90%) is achieved by the best performing model with four dummies. Therefore, we use the best performing four-dummy model in order to present a model that is as parsimonious as possible. Thus, the optimal thresholds for the distance to default dummies are 0.8, 1.8, 2.8, and 3.8. The threshold values for the robustness variables are selected in the same manner.

References

Anderson, R., and S. Sundaresan. 2000. A Comparative Study of Structural Models of Corporate Bond Yields: An Exploratory Investigation. *Journal of Banking & Finance* 24(1/2): 255.

Ang, A., and J. Chen. 2002. Asymmetric Correlations of Equity Portfolios. *Journal of Financial Economics* 63(3): 443–494.

Baker, M., and J. Wurgler. 2006. Investor Sentiment and the Cross-section of Stock Returns. *Journal of Finance* 61(4): 1645–1680.

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

Bekaert, G., and G. Wu. 2000. Asymmetric Volatility and Risk in Equity Markets. *The Review of Financial Studies* 13(1): 1–42.

Belke, A. and Gokus, C. 2011. Volatility Patterns of CDS, Bond and Stock Markets Before and During the Financial Crisis: Evidence From Major Financial Institutions. *Working Paper*. Berlin: German Institute for Economic Research.

Black, F., and M. Scholes. 1973. The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy* 81(3): 637–654.

Bollerslev, et al. 2018. Risk Everywhere: Modeling and Managing Volatility. *Review of Financial Studies* 31(7): 2729–2773.

Bollerslev, T. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31(3): 307–327.

Bollerslev, T., R. Engle, and J. Wooldridge. 1988. A Capital Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy* 96(1): 116–131.

Brown, G.W., and M.T. Cliff. 2005. Investor Sentiment and Asset Valuation. *The Journal of Business* 78(2): 405–440.

Brown, K.C., W.V. Harlow, and M.C. Tinic. 1988. Risk Aversion, Uncertain Information, and Market Efficiency. *Journal of Financial Economics* 22(2): 355–385.

Caballero, L.J., and A. Krishnamurthy. 2008. Collective Risk Management in a Flight to Quality Episode. *Journal of Finance* 63(5): 2195–2230.

Campbell, J.Y., and G. Taksler. 2003. Equity Volatility and Corporate Bond Yields. *The Journal of Finance* 58(6): 2321–2350.

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

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

Cremers, M., J. Driessen, P. Maenhout, and D. Weinbaum. 2008. Individual Stock-Option Prices and Credit Spreads. *Journal of Banking & Finance* 32(12): 2706–2715.

Cutler, D.M., J.M. Poterba, and L.H. Summers. 1989. What Moves Stock Prices? *Journal of Portfolio Management* 15(3): 4–12.

Da, Z., J. Engelberg, and P. Gao. 2015. The Sum of all FEARS: Investor Sentiment and Asset Prices. *Review of Financial Studies* 28(1): 1–32.

Demirovic, A., C. Guermat, and J. Tucker. 2017. The Relationship Between Equity and Bond Returns: An Empirical Investigation. *Journal of Financial Markets* 35(C): 47–64.

Ding, Z., and R. Engle. 2001. Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing. *Journal of Empirical Finance* 1(1): 83–106.

Durand, R.B., D. Lim, and J.K. Zumwalt. 2011. Fear and the Fama-French Factors. *Financial Management* 40(2): 409–426.

Engle, R., and K. Kroner. 1995. Multivariate Simultaneous Generalized ARCH. *Econometric Theory* 11(1): 122–150.

Fama, E.F., and K.R. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33(1): 3–56.

Fang, L., and J. Peress. 2009. Media Coverage and the Cross-section of Stock Returns. *Journal of Finance* 64(5): 2023–2052.

García, D. 2013. Sentiment During Recessions. *Journal of Finance* 68(3): 1267–1300.

Hansen, B.E. 2000. Sample Splitting and Threshold Estimation. *Econometrica* 68(3): 575–603.

Hillegeist, S.A., E.K. Keating, D.P. Cram, and K.G. Lundstedt. 2004. Assessing the Probability of Bankruptcy. *Review of Accounting Studies* 9(1): 5–34.

Huang, D., F. Jiang, J. Tu, and G. Zhou. 2015. Investor Sentiment Aligned: A Powerful Predictor of Stock Returns. *Review of Financial Studies* 28(3): 791–837.

IMF Global Financial Stability Report. 2015. Vulnerabilities, Legacies, and Policy Challenges-Risks Rotating to Emerging Markets.

Kräussl, R., and E. Mirgorodskaya. 2017. Media, Sentiment and Market Performance in the Long Run. *European Journal of Finance* 23(11): 1059–1082.

Lee, W.Y., C.X. Jiang, and D.C. Indro. 2002. Stock Market Volatility, Excess Returns, and the Role of Investor Sentiment. *Journal of Banking & Finance* 26(12): 2277–2299.

Longstaff, F.A. 2004. The Flight-to-Liquidity Premium in U.S. Treasury Bond Prices. *Journal of Business* 77(3): 511–526.

Manela, A., and A. Moreira. 2017. News Implied Volatility and Disaster Concerns. *Journal of Financial Economics* 123(1): 137–162.

Merton, R. 1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance* 29(2): 449–470.

Nozawa, Y. 2017. What Drives the Cross-Section of Credit Spreads?: A Variance Decomposition Approach. *The Journal of Finance* 72(5): 2045–2072.

Nyman, R., S. Kapadia, D. Tuckett, D. Gregory, P. Ormerod, and R. Smith. 2018. News and Narratives in Financial Systems: Exploiting Big Data for Systemic Risk Assessment. *Bank of England Staff Working Paper* No. 704.

Petersen, M. 2009. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *The Review of Financial Studies* 22(1): 435–480.

Ronn, E.I., and A.K. Verma. 1986. Pricing Risk-Adjusted Deposit Insurance: An Option-Based Model. *The Journal of Finance* 41(4): 871–895.

Scruggs, J., and P. Glabadanidis. 2003. Risk Premia and the Dynamic Covariance Between Stock and Bond Returns. *The Journal of Financial & Quantitative Analysis* 38(2): 295–316.

Smales, L.A. 2014. News Sentiment and the Investor Fear Gauge. *Finance Research Letters* 11(2): 122–130.

Tetlock, P.C. 2007. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *Journal of Finance* 62(3): 1139–1168.

Tetlock, P.C., M. Saar-Tsechansky, and S. Macskassy. 2008. More than Words: Quantifying Language to Measure Firms' Fundamentals'. *Journal of Finance* 63(3): 1437–1467.

Amer Demirovic is an assistant professor the Universiti Teknologi Brunei. He currently serves as Deputy Dean of the School of Applied Sciences and Mathematics. Amer earned his PhD in finance from the University of the West of England. His research focuses on the equity - corporate bond markets integration and the interaction between equity and credit risks.

Ali Kabiri is the Head of Economics & International Studies Dept. at the University of Buckingham and a Research Associate at the Financial Markets Group (FMG) at LSE. His current research projects at the CSDU focus on narratives and their underlying emotions to understand several facets of the financial markets and macroeconomy. His research focuses primarily on the Great Depression in the USA and the role of emotions as drivers of the macroeconomy and financial markets of the time. His research on the modern period investigates the role of emotions in Financial Broker reports and their impact on prices, implied volatility and forecast predictability for US Stocks. Further projects investigate the role of emotions as a common risk factor for corporate credit spreads and the conditional correlation between equities and corporate bonds during the GFC in the USA.

David Tuckett trained in Economics, Medical Sociology and Psychoanalysis and is Professor and Director of the Centre for the Study of Decision-Making Uncertainty at UCL in the Faculty of Brain Sciences, as well as a Fellow of the Institute of Psychoanalysis in London. He works part-time in clinical practice but since winning a 2006 Leverhulme Research fellowship for a "psychoanalytic study of investment markets" has been collaborating with a range of colleagues to introduce psychoanalytical understanding to behaviour in the financial markets and the economy more generally. He has published books and articles in sociology, psychoanalysis, economics, and finance and is a former President of the European Psychoanalytic Federation, Editor in Chief of the International Journal of Psychoanalysis and Principal of the Health Education Studies Unit at the University of Cambridge.

Rickard Nyman has a degree in Mathematics from Imperial College London and a degree in Financial Computing jointly from University College London and London Business School. He obtained his PhD in early 2016 at the Centre for Financial Computing, University College London. Rickard is a postdoctoral researcher working for the CSDU on leveraging machine learning and big data to understand and forecast decision making under uncertainty by incorporating ideas from Conviction Narrative Theory.