Journal of Banking & Finance 36 (2012) 2216–2232

Contents lists available at SciVerse ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Are corporate bond market returns predictable?

Yongmiao Hong [a,b], Hai Lin [c,d], Chunchi Wu [e,*]

[a] Department of Economics, Cornell University, Ithaca, NY 14853, USA
[b] Wang Yanan Institute for Studies in Economics and MOE Key Laboratory in Econometrics, Xiamen University, Xiamen 361005, China
[c] Department of Accountancy and Finance, University of Otago, Dunedin 9054, New Zealand
[d] School of Economics and Finance, Victoria University of Wellington, Wellington 6140, New Zealand
[e] School of Management, State University of New York at Buffalo, Buffalo, NY 14260, USA

ARTICLE INFO

Article history:
Received 8 July 2011
Accepted 1 April 2012
Available online 6 April 2012

JEL classification:
G12
G14
G17

Keywords:
Return predictability
Generalized spectrum
Autocorrelation
Causality
Nonlinearity
Bond pricing
Market efficiency

ABSTRACT

This paper examines the predictability of corporate bond returns using the transaction-based index data for the period from October 1, 2002 to December 31, 2010. We find evidence of significant serial and cross-serial dependence in daily investment-grade and high-yield bond returns. The serial dependence exhibits a complex nonlinear structure. Both investment-grade and high-yield bond returns can be predicted by past stock market returns in-sample and out-of-sample, and the predictive relation is much stronger between stocks and high-yield bonds. By contrast, there is little evidence that stock returns can be predicted by past bond returns. These findings are robust to various model specifications and test methods, and provide important implications for modeling the term structure of defaultable bonds.

1. Introduction

One of the most enduring issues in finance and economics is the question of whether returns on risky assets are predictable. This important issue has been the focus of an extensive literature on asset prices dating back more than a century. Despite an enormous amount of past efforts, whether future asset price changes can be meaningfully predicted is still a subject of ongoing debates and intensive empirical research (see, for example, Ang and Bekaert, 2007; Campbell and Thompson, 2008; Welch and Goyal, 2008; Rapach et al., 2010; Sekkel, 2011).[1]

The literature of asset return predictability has focused on the stock market. There is substantial evidence that stock returns are predictable, either by past price changes or economic variables (see Campbell et al., 1997; Ang and Bekaert, 2007; Campbell and Thompson, 2008; Rapach et al., 2010). Recent efforts have been directed to identifying the predictive components of asset returns at different return horizons, evaluating the predictive power of predictors using more robust tests, and determining how much predictability is compatible with efficiency consistent with risk-based asset pricing models.

Notwithstanding extensive research on equity return predictability, there are only a few studies on corporate bond return predictability (see Keim and Stambaugh, 1986; Kwan, 1996; Hotchkiss and Ronen, 2002; Downing et al., 2009) and empirical evidence is inconclusive. Kwan (1996) shows that significant negative contemporaneous correlation exists between returns of individual stocks and yield changes of bonds issued by the same firm, and that stock returns predict future bond yield changes. Unlike Kwan (1996) and Hotchkiss and Ronen (2002) find that corporate bond returns cannot be predicted by past stock returns based on a sample of 20 high-yield bonds from the National Association of Securities Dealers (NASD). By contrast, Downing et al. (2009) show that stock returns predict convertible bond returns in all rating categories but predict returns of only BBB- and junk-rated nonconvertible bonds.

In this paper, we examine the predictability of corporate bond returns in a narrow sense by focusing on serial dependence and causality tests. Similar to mainstream equity premium studies, we examine return predictability at the aggregate level. We employ bond market index data constructed from transaction prices, instead of dealer quotes used in a number of studies (see,

* Corresponding author. Tel.: +1 716 645 0448; fax: +1 716 645 3823.
E-mail addresses: yh20@cornell.edu (Y. Hong), hai.lin@otago.ac.nz (H. Lin), chunchiw@buffalo.edu (C. Wu).

[1] See also Kessler and Scherer (2009) and Fan et al. (2012).

0378-4266/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jbankfin.2012.04.001

for example, Kwan, 1996; Gebhardt et al., 2005). Our empirical analysis draws heavily on the rich literature in random walk and causality tests (Granger, 1969; Campbell et al., 1997). Similar to Chen and Maringer (2011), we account for nonlinearity in corporate bond index returns. Standard methods of return predictability tests are not robust to nonlinear dependence. To overcome this problem, we employ an advanced generalized spectral method (Hong and Lee, 2005) to detect nonlinear dependencies in returns and to perform robust tests. Furthermore, we conduct causality tests on bond and stock returns by taking into account heteroskedasticity in the error term and potential nonlinearity in the causal relationship.

Knowledge of bond price dynamics is important for formulating optimal strategies for asset allocation and ~~hedging~~. Corporate bonds account for a significant portion of investors' wealth, with a market size near 6 trillion dollars (see Abhyankar and Gonzales, 2009), so understanding corporate bond price dynamics is essential for academics and practitioners. This paper, to the best of our knowledge, is the first that provides comprehensive time-series analysis on serial and cross-serial dependencies in transaction-based corporate bond index returns.

We find strong evidence of serial and cross-serial dependence in corporate bond returns. Empirical analysis reveals a complicated nonlinear structure of serial dependence in corporate bond returns. Investment-grade and high-yield bond returns can be predicted by past stock returns both in-sample and out-of-sample, and the predictive relation is much stronger between stocks and high-yield bonds. By contrast, there is little evidence that stock returns can be predicted by past bond returns. These findings persist even after controlling effects of conditional heteroskedasticity, volatility-induced mean return changes, and time-varying interest rates.

The remainder of the paper is organized as follows. In Section 2, we describe the hypotheses and methodology for testing linear and nonlinear serial dependence in returns. In Section 3, we propose vector autoregressive regression models (VAR) and Granger causality tests with homoskedastic and heteroskedastic returns. In Section 4, we present test results for serial and cross-serial dependence in stock and bond market returns and examine the robustness of results to different model specifications and return measures. In Section 5, we examine the sensitivity of corporate bond returns to concurrent and lagged stock and government bond returns. In Section 6, we conduct out-of-sample tests on return predictability. Finally, we summarize our findings and conclude the paper in Section 7.

2. Tests of serial dependence in returns

A fundamental issue in asset pricing is whether future returns can be predicted by past price changes. In this section, we propose tests on predictive models with past returns. Tests of serial dependence in returns serve a number of purposes. First, by restricting the future return forecast to be a function of past price changes, these tests provide profound insights into the behavior of bond prices and yield important implications for the modeling of term structure of defaultable bonds. Second, an analysis of the nature of serial dependence in returns is important for understanding the structure of return dependence and designing robust statistical tests to accommodate more complicated dependence structure. Third, autocorrelation tests on return series provide essential information for correct model specification. For example, if returns of securities are serially correlated, one must control for this effect in the causality test to avoid spurious relations. In our empirical investigation, we are interested in the lead–lag relation between stock and bond market returns for various reasons, such as assessing information efficiency and understanding the nature of information flow that induces the causal relation. If individual stock returns are serially correlated, the leading and lagged stock returns may be spuriously related with the current change in bond prices even though stock and bond returns are only contemporaneously but not cross-serially correlated. Scrupulous tests of serial dependence can detect such spurious relations and provide critical information for a correct specification of the model.

Past studies on the predictability of corporate bond returns have typically examined the simple autocorrelation pattern in stock and bond returns (see, for example, Kwan, 1996). The standard tests on autocorrelation adopted by these studies lack power in finite sample size and are not robust to nonlinear serial dependence in returns. As a consequence, they may not be able to detect a more complicated dependence structure and to reject the martingale hypothesis correctly. In this paper, we perform not only the standard autocorrelation tests but also advanced tests that are robust to heteroskedasticity and other forms of nonlinearity in return series.

In what follows, we first set forth the hypotheses on serial dependence in conditional mean of bond returns and discuss various tests on serial correlation and the spectral test on the martingale difference sequence (MDS) in returns. Following this, we present empirical test methods and the estimation procedure.

2.1. Test hypothesis

Let $\{X_t\}$ be a weakly stationary return process with $E(X_t) = \mu$. The hypotheses of interest are

$$H_0: \quad E(X_t|I_{t-1}) = \mu$$

against

$$H_A: \quad E(X_t|I_{t-1}) \neq \mu.$$

The test above deals primarily with the question of whether there exists a dependence structure in the conditional mean. It does not impose any assumption on higher-order moments. To the extent that the conditional variance $h_t = \mathrm{var}(X_t|I_{t-1})$ or other higher-order conditional moments are time-varying, higher-moment properties could affect the test statistic for H_0. On the other hand, as no model parameter estimation is involved here, there is no need to consider the potential impact of uncertainty in parameter estimation on the test statistic. The information set I_{t-1} in the conditional mean test may contain only the past history of X_t or the past history of both X_t and other variables. When the information set contains only the history of the own variable, $I_{t-1} = \{X_{t-1}, X_{t-2}, \ldots\}$, it is a test of serial dependence in conditional mean. By contrast, when the information set includes the history of another variable, $I_{t-1} = \{X_{t-1}, Y_{t-1}, \ldots\}$, the test involves cross dependence in conditional mean.

Given that the information set contains only the own history, $I_{t-1} = \{X_{t-1}, X_{t-2}, \ldots\}$, under the null hypothesis H_0 of $E(X_t - \mu|I_{t-1}) = 0$, the martingale difference sequence (after demeaning) implies that

(i) $\{X_t\}$ is serially uncorrelated or white noise (WN),

$$\gamma(j) = \mathrm{cov}(X_t, X_{t-j}) = 0, \quad \text{for all } j > 0$$

or equivalently,

(ii) $\{X_t\}$ has a flat spectrum

$$h(\omega) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty} \mathrm{cov}(\varepsilon_t, \varepsilon_{t-|j|})e^{-ij\omega} = \frac{1}{2\pi}\gamma(0) \quad \text{for all } \omega \in [-\pi, \pi].$$

Thus, we can test H_0 by investigating whether $\gamma(j) = 0$ for all $j > 0$, or alternatively we can examine whether $\{X_t\}$ possesses a flat spectrum.

The hypothesis can be tested using standard autocorrelation tests, such as the Box–Pierce–Ljung tests (see Campbell et al., 1997). Additional tests include variance ratio tests of Lo and MacKinlay (1988) and the spectral density test of Hong (1996). Note that in empirical tests, there is a subtle difference between a white noise and an MDS. Although an MDS is a white noise, a white noise may not be an MDS. This can be illustrated by a simple example involving a nonlinear moving average process $X_t = \alpha\varepsilon_{t-1}\varepsilon_{t-2} + \varepsilon_t$, where $\varepsilon_t \sim i.i.d.(0,\sigma^2)$. For this process, $\mathrm{cov}(X_t, X_{t-j}) = 0$ for all $j > 0$, but $E(X_t|I_{t-1}) = \alpha\varepsilon_{t-1}\varepsilon_{t-2} \neq 0$. Thus, this process is a white noise but not an MDS. A common problem for the tests based on the autocovariance function $\gamma(j)$ or the power spectrum is that they cannot detect such non-MDS alternatives that exhibit zero autocorrelation.

Economic behavior and the investor's attitude toward risk and returns can be quite complicated. As these factors affect returns, the return process may not be characterized by a simple linear function. Several methods have been proposed to test if a return series is a martingale difference sequence. These include development of the indicator function test (Domínguez and Lobato, 2003) and the generalized spectral test (Hong and Lee, 2005). In the following, we discuss different methods for serial dependence tests and their implications. The technical details of these methods are included in Appendix A.

2.2. Test methods

The most direct and intuitive test of the martingale hypothesis is to check for serial correlation in the time series of asset returns. Box and Pierce (BP) propose a *Q*-statistic for autocorrelation test (see Campbell et al., 1997). This is an asymptotical test, in that the test statistic $T\sum_{j=1}^{p}\hat{\rho}^2(j)$ follows a χ^2 distribution where T is the number of observations in the return series, and p is the maximum lag order for the sample autocorrelation coefficient $\hat{\rho}$. Ljung and Box (LB) provide the finite-sample correction for this test, which increases the size of the BP test (Campbell et al., 1997). Although the LB test statistic, $LB(p) = T(T+2)\sum_{j=1}^{p}(T-j)^{-1}\hat{\rho}^2(j)$, has a better fit with the χ^2 distribution for small sample size, it lacks power in the presence of conditional heteroskedasticity, which is unfortunately quite common in financial markets. Nevertheless, the LB test remains a useful diagnostic tool because it is conceptually appealing and straightforward to apply to financial data, as statistical packages are readily available.

Variance ratio tests have been used widely in random walk tests of stock returns. The basic principle behind this test is that variance of the increments of a random walk is linear in the sampling interval. For example, if asset prices follow a random walk process, the variance of monthly returns will be four times as large as the variance of weekly returns. In empirical tests, we construct the variance ratio statistic $VR(p) = \left[p\hat{\gamma}(0) + 2p\sum_{j=1}^{p}(1-j/p)\hat{\gamma}(j)\right] \Big/ p\hat{\gamma}(0)$, where p is number of days over a period and $\hat{\gamma}(j)$ is the sample autocovariance function at lag j. Lo and MacKinlay (1988) show the limiting distribution of the test statistic under both homoskedasticity and heteroskedasticity (see Appendix A). Although the standard variance ratio test is convenient, it is not robust to the presence of other forms of nonlinear dependence.

Hong and Lee (2005) propose a nonlinear method to test the MDS hypothesis based on the generalized spectral density, which can accommodate the nonlinear return dependence. The generalized spectral test has several advantages. First, one can use this method to test whether there exists serial dependence in mean or not, where the dimension of the information set can be infinite. Second, the generalized spectral test method can be used conveniently to check dependence in the conditional mean $E(X_t|X_{t-j}) = E(X_t)$ for all $j > 0$, and so it is a more general test on the MDS than either autocorrelation or variance ratio test. Third, since the first-order derivative of the generalized spectral density function is always flat no matter whether higher-order conditional moments are changing over time or not, one can construct a test for the MDS that is robust to time-varying higher-order moments of any unknown form based on this unique property. Lastly, the method can be used to check a large number of lags and it naturally discounts higher-order lags, which is consistent with the stylized fact that the economic behavior is more strongly affected by recent events than remote past events.[2]

Hong and Lee (2005) develop a robust test to check whether $\varepsilon_t = X_t - g(I_{t-1}, \theta)$ satisfies the hypothesis that $E(\varepsilon_t|I_{t-1}) = 0$, where $g(I_{t-1}, \theta) = \mu$ and θ is a parameter. If the null is rejected, the MDS hypothesis is rejected and we conclude that there is a serial dependence in bond returns. The technical detail and implementation of this test procedure are presented in Appendix A.

3. Models for causality tests

Both bonds and stocks are different contingent claims issued by the same firm on the cash flow of the same underlying asset. Information about the expected value of the firm's asset should therefore affect prices of bonds and stocks in the same direction. For example, favorable earnings news increases firm value and stock price. According to the structural model, expected default probability depends on the firm asset value. As firm value appreciates, default risk decreases and bond price increases. If stock and bond markets are equally efficient, this will induce a positive contemporaneous correlation between corporate bond and stock returns. On the other hand, volatility affects bond and stock prices in an opposite direction. A stock can be viewed as a call option on the firm's underlying asset value, and a bond can be viewed as a portfolio long in default-free asset and short in a put option on the firm value. Information that increases asset return volatility but not the mean value of the firm should increase stock price and decrease bond price, resulting in a negative contemporaneous correlation between changes in stock and bond prices. Thus, the direction of the contemporaneous correlation between stock and bond returns depends on the nature of the information signal.

Conversely, if stock and bond markets are not equally efficient, due to either market frictions or other reasons, one market will impound information into prices faster than the other. If the stock price responds to new information faster than the bond price, stock returns will lead bond returns.

Previous studies on the efficiency of stock and bond markets have all relied on the standard causality test or predictive regression test (see Kwan, 1996; Hotchkiss and Ronen, 2002; Downing et al., 2009). This is not surprising given that these test methods are well established and relatively easy to implement. A potential drawback of these methods is that they rely on the linearity of the models, which ignores heteroskedasticity and nonlinear relationships. In this paper, we conduct thorough causality tests based on models with both linear and nonlinear structures. These tests provide unbiased results and avoid spurious causal relations. In addition, we propose causality tests that can detect the sources of causality between stock and bond market returns. In contrast to the traditional causality tests, these tests differentiate the causality induced by macroeconomic factors, such as interest rates, from that by the firm-specific fundamental factors, such as cash flows. In what follows, we present tests on whether the past return in one market can predict the return in another market.

Conventional tests on the lead–lag relationship rely on the VAR model and the Granger causality method. Let $I^{Z}_{t-1} = \left\{I^{Z_1}_{t-1}, I^{Z_2}_{t-1}, \ldots, I^{Z_k}_{t-1}\right\}$, where $I^{Z_1}_{t-1} = \{Z_{1,t-1}, Z_{1,t-2}, \ldots\}$ to $I^{Z_k}_{t-1} = \{Z_{k,t-1}, Z_{k,t-2}, \ldots\}$ rep-

[2] For the advantages and details of the properties of this test method, see Hong and Lee (2005).

resent the historical information sets for k variables. The variable $\{Z_{1t}\}$ is Granger-caused by $\{Z_{2t}, Z_{3t}, \ldots, Z_{kt}\}$ with respect to the information set I_{t-1}^{Z} if $E\left(Z_{1t}|I_{t-1}^{Z}\right) \neq E\left(Z_{1t}|I_{t-1}^{Z_1}\right)$. The dynamic relation between variables can be conveniently cast in a vector autoregressive model,

$$Z_t = \delta + \Phi_1 Z_{t-1} + \cdots + \Phi_p Z_{t-p} + \varepsilon_t,$$

where $Z_t = (Z_{1t}, \ldots, Z_{kt})^T$, $t = 1,2,\ldots,N$, is a k-dimensional vector of variables of interest, p is the lag order, $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{kt})^T$ is a vector of error terms with $E(\varepsilon_t) = 0$, $E\left(\varepsilon_t \varepsilon_t^T\right) = \Sigma$ and $E\left(\varepsilon_t \varepsilon_s^T\right) = 0$ for $s \neq t$, $\delta = (\delta_1, \ldots, \delta_k)^T$ is a constant vector, and Φ_j, $j = 1,\ldots,p$ is a $k \times k$ matrix of response coefficients. This VAR system serves as the basic framework for our causality tests.

3.1. Linear Granger causality test

Consider the bivariate case ($k = 2$), where $Z_t = (Z_{1t}, Z_{2t})^T$ are two stationary time series. Standard linear causality tests can be performed using the bivariate VAR system:

$$Z_{1t} = a_0 + \sum_{j=1}^{p} a_j Z_{1t-j} + \sum_{j=1}^{p} b_j Z_{2t-j} + \varepsilon_t,$$

$$Z_{2t} = c_0 + \sum_{j=1}^{p} c_j Z_{1t-j} + \sum_{j=1}^{p} d_j Z_{2t-j} + v_t.$$

Z_{1t} is Granger-caused by Z_{2t} if some b_j are not zero. We can test whether $b_j = 0$ jointly using the F test. This test assumes conditional homoskedasticity, or var $(\varepsilon_t|I_{t-1}) = \sigma^2$, asymptotically.

3.2. Nonlinear Granger causality test

Aside from the standard linear causality test, we account for conditional heteroskedasticity in the error term using the GARCH model and the effect of volatility on the conditional mean return using the GARCH-M model. Considering the nonlinear structure in the return-generating process yields more robust tests for the dynamic relationship between the two markets. The specification of different models and procedures for causality tests are described in Appendix A.

4. Data and empirical results

4.1. Data

Data include daily returns of the high-yield corporate bond index (*r_nbbh*), the investment-grade corporate bond index (*r_nbbi*), the S&P 500 stock index (*r_sp*500), and the S&P 500 index futures (*r_futures*). Corporate bond return data are based on NASD-Bloomberg US High-Yield and Investment-Grade Bond Indices, which are downloaded from the Bloomberg system. These indices are constructed from actual transaction prices of the active fixed-coupon bonds represented by the Trade Reporting and Compliance Engine (TRACE) system of the NASD that disseminates over-the-counter trades for all publicly traded corporate bonds. The index price is the volume-weighted average price generated from TRACE transactions.[3] The index basket excludes zero-coupon and convertible bonds, and bonds set to mature before the last day of the month for which index rebalance occurs. The sample period of the data is from October 1, 2002 to December 31, 2010 with 2079 daily observations.

Before turning to empirical estimation, it is useful to compare our sample with those of the previous studies closely related to our work, as some of the discrepancies between previous findings and ours are attributable to the differences in the sample selection and study periods. Kwan (1996) uses weekly yield data obtained from Merrill Lynch for 702 corporate bonds issued by 327 firms for the period from January 1986 to December 1990. His sample consists of bid yields from the dealer's quotes. To the extent that these bid yields are not necessarily associated with actual trades, they are subject to the problems of matrix prices and stale quotes. Hotchkiss and Ronen (2002) study transaction prices for 20 high-yield bonds included in the fixed-income pricing system (FIPS) of the NASD for the period between January 3 and October 1, 1995. Because of the limited sample size and the short time frame (9 months), their empirical findings are subject to small sample bias. Downing et al. (2009) analyze transaction data of stocks and bonds for the period from October 1, 2004 to December 31, 2005 (315 days). Their bond data sample is collected from the NASD TRACE system and includes transaction prices for 3000 bonds issued by 439 firms. Although their sample size is much larger than previous studies, the sample period is still relatively short. Tests of return predictability require a long period of data to produce reliable inferences. As such, any empirical finding based on a short sample period can be tenuous and time dependent. Unlike these studies, we use a data sample with a much longer time span (October 1, 2002–December 31, 2010) and conduct subperiod analysis to ensure that our results are robust to different sample periods. In addition, we use transaction-based index data for investment- and speculative-grade bonds that avoid problems of stale quotes and matrix prices.

Using the corporate bond index data has several advantages. First, the index is broad-based, well representing the whole corporate bond market. Second, the index is generated from transaction prices, instead of dealers' quotes or matrix prices, which are not representative of actual transactions (see Gebhardt et al., 2005). Third, the index consists of most liquid bonds. This mitigates the infrequent trading problem and provides a lower bound for inferring bond return predictability. Finally, using the index data bypasses the aggregation problem when summarizing test results across individual bonds to draw an unbiased statistical inference.

Fig. 1 plots the return series and histograms of the data. Returns of corporate bonds are substantially less volatile than stock and index futures returns. Volatility of corporate bonds is relatively low with an exception for the recent financial crisis period. Daily returns are centered on zero with occasional spikes. There are significant volatility clusterings, and distributions of returns clearly deviate from normal for both stocks and bonds.

Panel A of Table 1 provides summary statistics for all return series. Over the sample period, the high-yield bond index has the highest mean daily return (0.043%), and the S&P 500 stock index futures has the lowest mean return. The investment-grade bond index return has the lowest daily volatility (0.250%), while the stock index futures return has the highest volatility (1.360%). The return-risk trade-off measured by the ratio of mean return to standard deviation is highest for investment-grade bonds (0.088), followed by high-yield bonds (0.074), the stock index (0.015), and index futures (0.014). Historically, returns of corporate bonds vary from one period to another. Over our sample period, corporate bond returns are higher than stock returns, partly due to the equity market slump in 2002 and the severe downturn in the recent financial crisis. High-yield bond index returns are negatively skewed and have high kurtosis. Investment-grade bond index returns are also negatively skewed and have kurtosis, but the magnitude is milder compared with high-yield bonds. Consistent with

[3] The NASD-Bloomberg indices reflect actual transactions throughout the most actively traded portion of the corporate bond market where 65% of that activity occurs at the retail level. All bonds included in the basket must have traded on average at least three times per day, with at least one trade on 80% of the 60 trading days prior to the rebalance date and have a total issued amount of outstanding reported publicly. The index values are calculated at 5:15 p.m. each day using TRACE transactions. The indices are rebalanced on a monthly basis.

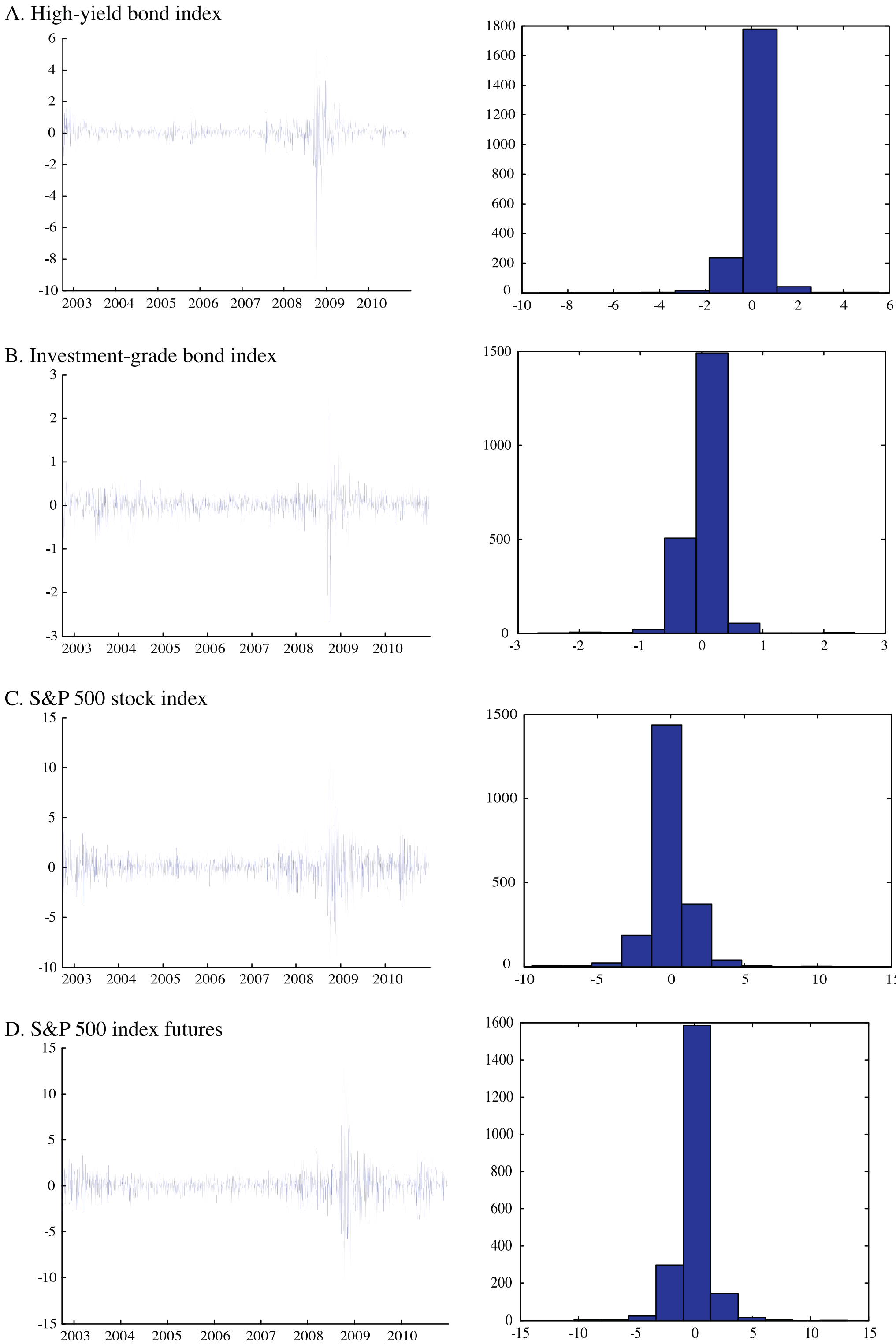

Fig. 1. Daily returns and histograms.

previous findings, daily returns of S&P 500 spot and futures indices exhibit kurtosis. Excessive kurtosis is partly attributed to extreme movements in stock and bond prices during the recent financial market turmoil.

Panel B of Table 1 shows that returns are positively contemporaneously correlated across securities. The positive correlation between corporate bond and S&P 500 index returns is consistent with the negative contemporary correlation between yield changes of corporate bonds and stock returns documented by Kwan (1996). Correlation between investment-grade bond and stock index returns is low (0.016), compared with that between high-yield bond and stock index returns (0.316). Cash flows of investment-grade bonds are relatively stable, and thus their prices are less sensitive to firm-specific earnings news. On the other hand, speculative bond prices are sensitive to news about firms' earnings, similar to stocks, because default risk is high. This may explain the differ-

Table 1
Summary statistics.

	Mean (%)	Standard deviation (%)	Skewness	Kurtosis
Panel A: Summary statistics				
r_nbbh	0.043	0.584	−1.622	44.253
r_nbbi	0.022	0.250	−0.711	23.219
r_sp500	0.021	1.355	−0.201	10.189
r_futures	0.019	1.360	0.031	13.468
	r_nbbh	*r_nbbi*	*r_sp500*	*r_futures*
Panel B: Contemporaneous correlation				
r_nbbh	1.000			
r_nbbi	0.453	1.000		
r_sp500	0.316	0.016	1.000	
r_futures	0.348	0.050	0.976	1.000
Lag	*r_nbbh*	*r_nbbi*	*r_sp500*	*r_futures*
Panel C: Autocorrelation				
$P = 1$	0.396***	0.284***	−0.120***	−0.090***
	(326.13)	(167.26)	(30.146)	(16.984)
$P = 2$	0.141***	−0.009***	−0.070***	−0.099***
	(367.68)	(167.42)	(40.290)	(37.161)
$P = 3$	0.139***	0.011***	0.069***	0.072***
	(407.88)	(167.65)	(50.039)	(48.041)
$P = 4$	0.181***	0.036***	−0.024***	−0.021***
	(476.33)	(170.37)	(51.083)	(48.970)
$P = 5$	0.129***	0.033***	−0.029***	−0.029***
	(510.60)	(172.65)	(52.776)	(50.320)
$P = 6$	0.040***	0.073***	0.015***	0.011***
	(513.87)	(183.63)	(53.256)	(50.580)
$P = 7$	0.054***	0.082***	−0.046***	−0.044***
	(519.95)	(197.38)	(57.567)	(54.616)
$P = 8$	0.014***	0.050***	0.022***	0.027***
	(520.35)	(202.53)	(58.538)	(56.178)
$P = 9$	0.051***	0.050***	−0.007***	−0.030***
	(525.66)	(207.58)	(58.648)	(58.029)
$P = 10$	0.054***	0.110***	0.025***	0.016***
	(531.66)	(232.56)	(59.875)	(58.524)

		r_nbbh_t	r_nbbi_t	r_sp500_t	$r\_futures_t$
Panel D: Cross correlation					
Υ_1	r_nbbh_{t-1}	0.396	0.190	−0.068	−0.075
	r_nbbi_{t-1}	0.166	0.284	−0.067	−0.089
	r_sp500_{t-1}	0.396	0.230	−0.123	−0.086
	$r\_futures_{t-1}$	0.406	0.233	−0.109	−0.094
Υ_2	r_nbbh_{t-2}	0.141	0.061	0.004	0.003
	r_nbbi_{t-2}	0.009	−0.009	−0.053	−0.059
	r_sp500_{t-2}	0.117	0.074	−0.071	−0.078
	$r\_futures_{t-2}$	0.103	0.068	−0.092	−0.100
Υ_3	r_nbbh_{t-3}	0.139	0.042	0.029	0.027
	r_nbbi_{t-3}	0.050	0.011	0.002	−0.005
	r_sp500_{t-3}	0.043	0.052	0.066	0.074
	$r\_futures_{t-3}$	0.017	0.045	0.063	0.070

The Ljung–Box statistics are included in parentheses. *r_nbbh*, *r_nbbi*, *r_sp500*, and *r_futures* represent high-yield, investment-grade, S&P 500 index and index futures returns, respectively.
** Denotes significance at the 5% level.
*** Denotes significance at the 1% level.

ence between these contemporaneous correlation coefficients. The bottom row reports the correlation between returns of stock index futures and other indices. Correlations between bond returns and stock index future returns are somewhat higher than those between bond returns and stock index returns. Not surprisingly, the contemporary correlation between S&P 500 stock index returns and index futures returns is close to one.

4.2. Tests of serial dependence in conditional mean

Panel C of Table 1 reports autocorrelations of returns. The Ljung–Box (LB) statistics are significant up to a lag order of $p = 10$ for all return series. Bond index returns tend to have positive autocorrelations. The first-order autocorrelation of bond returns is strongly positive but that of stock returns is negative, and the former is much higher than the latter in terms of absolute magnitude. In addition, autocorrelation of high-yield bond returns (*r_nbbh*) is higher than that of investment-grade bond returns (*r_nbbi*). These results imply that past returns would have higher predictive ability for future returns of corporate bonds than for stocks, and the predictability would also be higher for future returns of high-yield bonds than for future returns of investment-grade bonds.

Panel D of Table 1 shows cross correlations among returns at different lags. Cross autocorrelation concentrates on the lower left-hand corner of the correlation matrix, indicating that cross autocorrelations from stocks to bonds are much higher than from bonds to stocks. Furthermore, the cross autocorrelation from stocks to high-yield bonds is higher than that from stocks to investment-grade bonds. Positive stock returns lead to positive bond returns on the next day. Investment-grade bond returns are also positively cross-serially correlated with high-yield bond returns on the next day.

Table 2 reports results of variance ratio tests and generalized spectral tests. Variance ratio tests account for conditional heteroskedasticity. The upper panel shows results based on the entire sample. As shown, all return series fail the variance ratio test (see the left side), consistent with the finding in Table 1 that returns are significantly serially correlated. The significance level of rejection associated with the variance ratio test is, however, lower than that of the LB test. One possible reason for the higher rejection level in the LB test is that conditional heteroskedasticity is unaccounted for in this test. Variance ratio test statistics are much higher for high-yield bonds, suggesting the existence of a more complicated serial dependence structure for this bond market segment.

The right side in the upper panel of Table 2 reports the results of MDS tests using the generalized spectral test of Hong and Lee (2005). Results overwhelmingly reject the hypothesis of no serial dependence in corporate bond and stock returns. The generalized spectral test can detect a wide range of model misspecifications in mean and are robust to conditional heteroskedasticity and higher-order time-varying moments of unknown form. A distinct advantage is its ability to uncover both linear and nonlinear serial dependence in the moments of any order. Results in Table 2 confirm that the generalized spectral test indeed has more power than the variance ratio test, and rejects the hypothesis of serial independence more strongly. The generalized spectral test rejects the independence hypothesis for index futures returns even at long lags. Test statistics are the highest for high-yield bonds, followed by investment-grade bonds, stocks, and index futures. Results suggest the existence of a more complicated dependence structure for bond returns.

The test results in the upper panel of Table 2 are based on the whole sample, which includes the period of the recent financial crisis. Extreme observations could have a dominating effect and affect the inferences. To examine the impacts of including extreme observations, we redo the tests based on the subsample which excludes the data in the recent credit crunch and subsequent periods (after June 2007).[4] The results are reported in the lower panel of Table 2. Results show that extreme observations reduce the significance of the tests for stock market and index futures returns, particularly for the variance ratio test. On the other hand, the impacts on the test results for investment- and speculative-bond indices, which are our major focus, are small. Results show that our tests for the corporate bond markets are robust to the inclusion of the financial crisis period. We next examine if there exists a component of nonlinear dependence in returns.

[4] Friewald et al. (2012) also use July 2007 as the beginning time of the subprime financial crisis.

Table 2
Variance ratio and MDS tests for corporate bond, stock, and index futures returns.

Lag (p)	Variance ratio test				Martingale difference series test			
	r_nbbh	r_nbbi	r_sp500	r_futures	r_nbbh	r_nbbi	r_sp500	r_futures
Whole sample period: 10/2002–12/2010								
1					74.500***	32.819***	6.231***	2.671***
2	5.404***	3.718***	−2.945***	−2.077**	69.700***	30.649***	5.686***	2.659***
3	5.322***	3.239***	−3.189***	−2.617***	65.958***	29.151***	5.370***	2.734***
4	5.359***	2.954***	−2.606***	−2.155**	62.622***	27.814***	5.209***	2.816***
5	5.601***	2.851***	−2.371***	−1.983**	59.485***	26.644***	5.115***	2.899***
6	5.831***	2.805***	−2.256**	−1.901*	56.727***	25.816***	5.055***	2.969***
7	5.927***	2.847***	−2.117**	−1.803*	54.192***	25.151***	5.012***	3.035***
8	5.983***	2.952***	−2.068**	−1.776*	51.853***	24.556***	4.981***	3.108***
9	5.998***	3.076***	−1.971**	−1.694*	49.829***	23.984***	4.949***	3.176***
10	6.018***	3.207***	−1.895*	−1.656*	48.061***	23.454***	4.908***	3.231***
Subsample: 10/2002–06/2007								
1					86.505***	16.013***	2.935***	2.126**
2	7.913***	6.293***	−2.334***	−1.841*	83.363***	14.975***	2.695***	2.048**
3	9.219***	5.791***	−1.675*	−1.055	79.434***	14.007***	2.468***	2.017**
4	9.764***	5.374***	−1.511	−1.025	76.515***	13.038***	2.331***	1.954**
5	9.817***	5.378***	−1.387	−0.985	74.037***	12.249***	2.224**	1.890**
6	9.700***	5.474***	−1.427	−1.095	71.779***	11.627***	2.143**	1.853**
7	9.408***	5.442***	−1.515	−1.224	69.352***	11.185***	2.053**	1.811**
8	9.182***	5.466***	−1.698*	−1.460	67.042***	10.834***	1.956**	1.760**
9	9.031***	5.602***	−1.771*	−1.568	64.964***	10.551***	1.858**	1.707**
10	8.918***	5.701***	−1.861*	−1.699*	62.890***	10.313***	1.763**	1.651*

p is the lag order used in computing test statistics. Test statistics follow the normal distribution.
* Denotes significance at the 10% level.
** Denotes significance at the 5% level.
*** Denotes significance at the 1% level.

Although tests in Table 2 consider the effect of conditional heteroskedasticity, it remains unclear whether serial dependence is caused by linear and nonlinear factors. To check whether serial dependence contains nonlinearity, we can remove the linear component of returns from an ARMA model and examine whether the adjusted return can pass the variance ratio and generalized spectral tests. If the adjusted return passes the variance ratio test, this implies that all linear relations have been purged. The generalized spectral test will then tell us more about the possibility of nonlinear dependence. A rejection by the generalized spectral test points to the existence of nonlinearity in return dependence. To remove the linear component of serial dependence, we estimate the following ARMA model:

$$r_t = \alpha + \sum_{j=1}^{p} \beta_p r_{t-p} + \sum_{k=1}^{q} \gamma_q \varepsilon_{t-q} + \varepsilon_t,$$

where lag orders p and q are determined by the BIC information criterion. The BIC criterion provides a consistent order selection for a weakly stationary linear process. After estimating the model, we remove the AR component from returns[5] and apply the variance ratio test and the generalized spectral test to the linearly adjusted return to check if there is still significant serial dependence.

Table 3 reports results of tests for the adjusted return. The upper panel reports the results based on the whole sample. After removing the linear return component, we find that all residual returns pass the variance ratio test (see the left side of the table). This finding suggests that the linear dependence in returns has been successfully purged. However, all adjusted returns fail the generalized spectral test reported on the right side of the table. Thus, although there is no evidence against the white noise hypothesis, the generalized spectral test rejects the MDS hypothesis. Results strongly suggest that nonlinear dependence exists in returns, which cannot be detected by the variance ratio test. The lower panel of Table 3 reports the results excluding the recent credit crunch and subsequent periods. Although excluding the extreme observations weakens significance of the MDS tests for stock market and index futures returns, results are not affected as much for corporate bonds. Results continue to show the existence of nonlinear dependence in corporate bond returns.

The findings above demonstrate the power of the generalized spectral test in detecting more complicated serial dependence structure in returns. The structure of serial dependence appears to be more complex for high-yield bonds, as indicated by the relatively high test statistics for this group.

4.3. Linear Granger causality tests

We next turn to Granger causality tests for the lead–lag relationship between the two markets. We first estimate the VAR model that includes returns of high-yield and investment-grade bonds and stocks where the lag order is determined by the BIC criterion. Table 4 reports the results for the vector autoregressive model. The return of the high-yield bond index is significantly related to the lagged-one return (0.112) of the investment-grade bond at the 5% level and the lagged-one return (0.140) of the stock index at the 1% level. The return of the investment-grade bond is significantly related to the lagged-one returns of the high-yield bond (−0.024) and stock (0.045) at the 5% and 1% levels, respectively. The coefficient of the lagged-one stock return in the investment-grade bond return equation is much lower than that in the high-yield bond equation, indicating that high-yield bond returns are more closely related to lagged stock returns. On the other hand, the return of the S&P 500 index is only significantly related to its own return at lag one (−0.142). However, it is significantly related to both investment-grade (0.145) and high-yield bond returns (−0.444) at lag two. Intercept estimates are 0.027, 0.016, and 0.031 for the high-yield bond, investment-grade bond, and S&P 500 index return equations, respectively.

Results of linear Granger causality tests are reported in Table 5. The left side of the table reports the results based on the whole

[5] The BIC criterion suggests an AR (2) process for the high-yield bond and an AR (1) process for the investment-grade bond, S&P 500 stock index, and index futures return series.

Table 3
Variance ratio and MDS tests after adjusting for the linear relationship.

Lag (p)	Variance ratio test				Martingale difference series test			
	r_nbbh AR(2)	*r_nbbi* AR(1)	*r_sp500* AR(1)	*r_futures* AR(1)	*r_nbbh* AR(2)	*r_nbbi* AR(1)	*r_sp500* AR(1)	*r_futures* AR(1)
Whole sample: 10/2002–12/2010								
1					4.588***	4.048***	0.384	0.401
2	0.029	0.365	−0.209	−0.169	4.786***	4.736***	0.725	0.996
3	−0.151	−0.260	−0.997	−1.109	4.858***	5.230***	1.181	1.527*
4	−0.067	−0.409	−0.761	−0.907	4.867***	5.591***	1.546*	1.915**
5	0.193	−0.384	−0.723	−0.868	4.909***	5.857***	1.850**	2.233**
6	0.423	−0.365	−0.763	−0.888	4.998***	6.081***	2.091**	2.473***
7	0.528	−0.283	−0.763	−0.881	5.095***	6.259***	2.287**	2.666***
8	0.645	−0.156	−0.815	−0.920	5.186***	6.396***	2.451***	2.833***
9	0.700	−0.035	−0.808	−0.898	5.263***	6.502***	2.586***	2.972***
10	0.764	0.064	−0.796	−0.904	5.329***	6.580***	2.693***	3.085***
Subsample: 10/2002–06/2007								
1					2.345***	0.881	0.045	0.795
2	−0.040	0.431	0.277	0.396	2.485***	1.178	0.209	0.976
3	−0.120	−0.375	0.483	0.780	2.511***	1.379*	0.438	1.192
4	0.011	−0.572	0.383	0.582	2.452***	1.507*	0.579	1.290*
5	−0.177	−0.300	0.284	0.434	2.428***	1.601*	0.681	1.356*
6	−0.240	0.036	0.070	0.179	2.442***	1.672**	0.767	1.413*
7	−0.247	0.174	−0.177	−0.080	2.476***	1.733**	0.830	1.449*
8	−0.153	0.371	−0.488	−0.426	2.516***	1.784**	0.857	1.461*
9	−0.112	0.651	−0.671	−0.625	2.549***	1.841**	0.854	1.454*
10	−0.045	0.847	−0.845	−0.823	2.572***	1.904**	0.834	1.433*

p is the lag order used in computing test statistics. Test statistics follow the normal distribution.
* Denotes significance at the 10% level.
** Denotes significance at the 5%level.
*** Denotes significance at the 1% level.

Table 4
VAR estimation of corporate bond and stock returns.

	$Z_t = [r\_nbbh_t, r\_nbbi_t, r\_sp500_t]^T$		
δ^T	0.027***	0.016***	0.031
	(2.40)	(3.11)	(1.04)
Φ_1	0.239***	0.112**	0.140***
	(9.00)	(2.14)	(15.06)
	−0.024**	0.314***	0.045***
	(−1.97)	(13.11)	(10.67)
	0.062	−0.219	−0.142***
	(0.89)	(−1.60)	(−5.83)
Φ_2	0.077***	−0.134***	0.012
	(3.20)	(−2.57)	(1.26)
	0.034***	−0.109***	0.006
	(3.12)	(−4.53)	(1.26)
	0.145**	−0.444***	−0.108***
	(2.30)	(−3.22)	(−4.29)

The estimated model is $Z_t = \delta + \Phi_1 Z_{t-1} + \Phi_2 Z_{t-2} + \varepsilon_t$, where $\delta = [\delta_1, \delta_2, \delta_3]^T$, Φ_1 and Φ_2 are 3×3 matrices that include the first- and second-lag response coefficients, respectively. The t-statistics are included in the parentheses.
* Denotes significance at the 10% level.
** Denotes significance at the 5% level.
*** Denotes significance at the 1% level.

sample period. Results show that the high-yield bond returns are significantly Granger-caused by the investment-grade bond returns, and the investment-grade bond returns are also significantly Granger-caused by the high-yield bond returns. Both high-yield and investment-grade bond returns are Granger-caused by stock returns at the 1% level. Results also show that stock returns are Granger-caused by both high-yield and investment-grade bond returns, but test statistics are much weaker.

The whole sample period covers the most recent financial crisis period. To test whether there is a change in the causal relationship due to the financial crisis, we exclude the data after June 30, 2007 and redo the Granger causality test. The right side of Table 5 reports the results. The Granger causality effects from stock returns to high-yield bond returns and investment-grade bond returns continue to be quite significant, whereas the Granger causality effects from high-yield bond returns and investment-grade bond returns to stock returns become insignificant. Thus, the causal relation from bonds to stocks on the left side of Table 5 can be attributed to the effect of extreme observations. The result shows that stock returns Granger-cause bond returns, and this causal relationship is robust to extreme observations.

4.4. Nonlinear Granger causality tests with heteroskedasticity

Standard Granger causality tests assume that the residual terms in the VAR model are conditionally homoskedastic. If this assumption is violated, causality tests are biased and one may obtain spurious correlations between variables in fitting supposedly uncorrelated data with conditional heteroskedasticity to the linear VAR model.

To check the robustness of standard Granger causality tests, we conduct the VAR test by accounting for conditional heteroskedasticity. We estimate the following bivariate VAR(1)-GARCH (1,1) model for each pair of security returns:

$$Z_t = \delta + \Phi_1 Z_{t-1} + \varepsilon_t,$$

where $\delta = [\delta_1, \delta_2]^T$, $\Phi_1 = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}$, $\varepsilon_t | I_{t-1} \sim N(0, H_t)$, and the conditional variance- covariance matrix $H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{bmatrix}$ is assumed to have a BEKK representation (see Baba et al., 1990).

Table 6 reports results of the bivariate VAR model with conditional heteroskedasticity. The effect of the lagged stock return on the high-yield bond return remains significant, even after accounting for the GARCH effect. By contrast, there is no evidence that corporate bond market returns can predict stock market returns. For the high-yield bond regression, the lagged-one coefficient of the stock index return is 0.071 and significant at the 1% level, whereas for the investment-grade bond regression,

Table 5
Linear Granger causality tests between corporate bond and stock returns.

Hypothesis	Whole sample: 10/2002–12/2010 Granger causality test	Subsample: 10/2002–06/2007 Granger causality test
High-yield bond return vs. investment-grade bond return		
Hypothesis 1	3.82**	1.90
Hypothesis 2	6.87***	3.03**
High-yield bond return vs. stock return		
Hypothesis 3	62.01***	14.95***
Hypothesis 4	3.56***	1.45
Investment-grade bond return vs. stock return		
Hypothesis 5	59.55***	4.88***
Hypothesis 6	4.84***	0.49

Reported are the *F* statistics for Granger causality tests. The lag orders (p) are selected using the BIC information criterion.
Hypothesis 1: high-yield bond returns are not Granger-caused by investment-grade bond returns.
Hypothesis 2: investment-grade bond returns are not Granger-caused by high yield bond returns.
Hypothesis 3: high-yield bond returns are not Granger-caused by stock returns.
Hypothesis 4: stock returns are not Granger-caused by high-yield bond returns.
Hypothesis 5: investment-grade bond returns are not Granger-caused by stock returns.
Hypothesis 6: stock returns are not Granger-caused by investment-grade bond returns.
* Denotes significance at the 10% level.
** Denotes significance at the 5% level.
*** Denotes significance at the 1% level.

the lagged-one coefficient of the stock index return is only 0.024 (significant at the 1% level). Results again show that high-yield bonds are more closely related to stocks.

Results of the causality test under conditional heteroskedasticity are reported in Table 7 where the lag order (p) of the VAR is determined by the BIC for each pair of return series. The left side of the table reports the results based on the whole sample period. During the whole sample period, high-yield bond returns are not Granger-caused by investment-grade bond returns, and investment-grade bond returns are Granger-caused by high-yield bond returns at the 1% level. Both investment-grade and high-yield bond returns are Granger-caused by stock returns at the 1% level. But stock returns are neither Granger-caused by high-yield nor by investment- grade bond returns. This finding contrasts with the result on the left side of Table 5 based on the whole sample, which assumes conditional homoskedasticity. It appears that results on the left side of Table 5 are spurious due to ignorance of heteroskedasticity. More robust results in Table 7 suggest that the stock market return leads the bond market return, but not vice versa.

The right side of Table 7 reports the results excluding the recent credit crunch and subsequent periods. The results show that the Granger causality relation between high-yield and investment-grade bond returns is sensitive to the sample period. By contrast, the Granger causality relations from stock returns to high-yield bond returns and investment-grade bond returns are robust to the exclusion of the financial crisis and subsequent periods. Results continue to show that stock returns Granger-cause bond returns, regardless of whether we include the extreme observations or not.

4.5. The GARCH-M effect

Previous studies have documented an existence of the GARCH-M effect in the stock market (see Hamao et al., 1990). We next examine whether there exists a similar GARCH-M effect in the corporate bond market. Table 8 reports results for the bivariate VAR(1)-GARCH (1,1)-M model. For convenience, the estimate of the GARCH-M response coefficient matrix Ψ_1 is transposed. For instance, in the high-yield bond index return equation, the coefficient of its own variance is 0.054, the coefficient of the covariance term between high-yield and investment- grade bond index returns is 0.012, and the coefficient of the variance of investment-grade bond index returns is 0.008; none is statistically significant.

Table 6
Estimation of bivariate VAR (1)-GARCH (1,1) models.

	$Z_t = [r\_nbbh_t, r\_nbbi_t]^T$		$Z_t = [r\_nbbh_t, r\_sp500_t]^T$		$Z_t = [r\_nbbi_t, r\_sp500_t]^T$	
δ^T	0.043***	0.016***	0.040***	0.074***	0.020**	0.037
	(8.27)	(3.49)	(7.51)	(3.67)	(2.62)	(0.96)
Φ	0.355***	0.049	0.301***	0.071***	0.171***	0.024***
	(14.50)	(1.13)	(12.58)	(10.05)	(3.70)	(5.60)
	0.035***	0.183***	−0.015	−0.116***	0.149	−0.051
	(4.00)	(6.32)	(−0.19)	(−4.91)	(0.25)	(−1.50)
c	0.041***	0.009	0.046***	0.025	0.033	0.016
	(6.62)	(1.13)	(4.28)	(0.70)	(1.09)	(0.06)
		−0.016***		−0.093***		−0.095***
		(−4.63)		(−5.40)		(−3.34)
a	0.432***	0.052***	−0.398***	0.047	−0.232*	−0.307
	(10.80)	(2.91)	(−9.03)	(0.56)	(−1.90)	(−0.18)
	−0.090	0.151***	−0.024	−0.207***	0.002	0.224***
	(−1.65)	(6.14)	(−1.34)	(−9.47)	(0.03)	(2.42)
g	−0.904***	0.020***	−0.902***	−0.048	−0.958***	0.005
	(−56.88)	(3.20)	(−34.04)	(−1.02)	(−12.04)	(0.01)
	−0.022	−0.986***	0.003	−0.969***	0.005	−0.969***
	(−1.42)	(−178.96)	(0.70)	(−138.10)	(0.15)	(−18.77)

The estimated models are $Z_t = \delta + \Phi_1 Z_{t-1} + \varepsilon_t$, where $\delta = [\delta_1, \delta_2]^T$ is the intercept vector, $\Phi_1 = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}$ is the VAR response coefficient matrix, $\varepsilon_t | I_{t-1} \sim N(0, H_t)$, and the conditional variance–covariance matrix $H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{bmatrix}$ is assumed to have a BEKK representation, i.e., $H_t = c^T c + a^T \varepsilon_{t-1} \varepsilon_{t-1}^T a + g^T H_{t-1} g$, where $c = \begin{bmatrix} c_{11} & c_{12} \\ 0 & c_{21} \end{bmatrix}$, $a = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $g = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$, $\varepsilon_t = \begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \end{bmatrix}$. The t-statistics are included in parentheses. The intercept vector δ is transposed in the report where the first (second) element is associated with the first (second) dependent variable in the bivariate VAR system.
* Denotes significance at the 10% level.
** Denotes significance at the 5% level.
*** Denotes significance at the 1% level.

Table 7
Nonlinear Granger causality tests between corporate bond and stock returns with conditional heteroskedasticity.

Hypothesis	Whole sample: 10/2002–12/2010 Granger causality test	Subsample: 10/2002–06/2007 Granger causality test
High-yield bond return vs. investment-grade bond return		
Hypothesis 1	2.21	4.73***
Hypothesis 2	14.67***	1.34
High-yield bond return vs. stock return		
Hypothesis 3	43.04***	19.85***
Hypothesis 4	1.70	1.34
Investment-grade bond return vs. stock return		
Hypothesis 5	29.46***	4.00**
Hypothesis 6	0.10	0.21

Reported are the F statistics for Granger causality tests. The lag orders (p) are determined by using the BIC information criterion.
Hypothesis 1: high-yield bond returns are not Granger-caused by investment-grade bond returns.
Hypothesis 2: investment-grade bond returns are not Granger-caused by high-yield bond returns.
Hypothesis 3: high-yield bond returns are not Granger-caused by stock returns.
Hypothesis 4: stock returns are not Granger-caused by high-yield bond returns.
Hypothesis 5: investment-grade bond returns are not Granger-caused by stock returns.
Hypothesis 6: stock returns are not Granger-caused by investment-grade bond returns.
* Denotes significance at the 10% level.
** Denotes significance at the 5% level.
*** Denotes significance at the 1% level.

Overall, we find that the GARCH-M effect is modest. There is some evidence that stock return volatility may affect bond returns but no evidence that volatility of corporate bond returns affects stock returns. The cross dependence of conditional mean corporate bond returns on past stock returns remains unchanged. Results continue to show that stock returns predict corporate bond returns.

4.6. Cross dependence in excess returns

The total return of a risky asset can be divided into a risk-free interest component, which represents a compensation for the investor's time preference, and an excess return component, a reward for bearing risk. If the stock market is more efficient in impounding the firm's cash flow information than the corporate bond market, the excess stock return should lead the excess corporate bond return. Focusing on excess returns allows us to net out the effect of interest rates. As such, causality tests on excess returns are expected to be more revealing. We next examine the relationship between the excess returns of corporate bonds and stocks.

The risk-free rate data were downloaded from Ken French's Web site. We subtract the risk-free rate from the return of risky securities and conduct the Granger-causality test on excess returns. Table 9 reports test results that take into account conditional heteroskedasticity and the lag orders are determined by BIC. The left side of the table reports the results based on the whole sample. Results show that investment-grade bond excess returns are Granger-caused by high-yield bond excess returns, but not vice versa.

Table 8
Estimation of bivariate VAR (1)-GARCH (1,1)-M models.

	$Z_t = [r\_nbbh_t, r\_nbbi_t]^T$		$Z_t = [r\_nbbh_t, r\_sp500_t]^T$		$Z_t = [r\_nbbi_t, r\_sp500_t]^T$	
δ^T	0.039***	0.012	0.016	0.051*	−0.002	0.055
	(2.99)	(1.34)	(1.48)	(1.91)	(−0.11)	(0.94)
Φ_1	0.354***	0.033	0.289***	0.071***	0.173***	0.022***
	(14.53)	(1.07)	(12.07)	(9.86)	(5.27)	(5.05)
	0.050***	0.182***	−0.026	−0.118***	−0.149	−0.084***
	(4.17)	(6.31)	(−0.33)	(−5.03)	(−1.23)	(−3.60)
Ψ_1^T	0.054	0.032	0.112	−0.098	0.050	−0.010
	(0.68)	(0.94)	(1.04)	(−0.35)	(0.12)	(−0.01)
	0.012	−0.008	−0.289**	−0.048	0.244	−0.004
	(0.02)	(−0.05)	(−2.21)	(−0.14)	(1.60)	(−0.01)
	0.008	0.001	0.050**	0.047	0.022***	−0.001
	(0.02)	(0.01)	(2.24)	(1.09)	(3.87)	(−0.02)
c	0.042***	0.010	−0.047***	−0.027	0.087***	0.082***
	(6.52)	(1.17)	(−3.99)	(−0.79)	(5.71)	(2.70)
		−0.016***		0.093***		−0.001
		(−4.39)		(5.52)		(−0.72)
a	0.432***	0.052***	−0.388***	0.071	−0.340***	−0.772***
	(10.52)	(2.91)	(−8.36)	(0.77)	(−4.09)	(−5.54)
	−0.091	0.152***	−0.030	−0.214***	−0.053***	0.156***
	(−1.61)	(5.94)	(−1.49)	(−10.09)	(−4.00)	(4.23)
g	−0.904***	0.020***	−0.902***	−0.051	−0.799***	0.479***
	(−55.04)	(3.20)	(−29.46)	(−1.08)	(−12.30)	(3.62)
	−0.022	−0.986***	0.004	−0.968***	−0.003***	−0.962***
	(−1.35)	(−168.63)	(0.95)	(−138.03)	(−0.99)	(−157.01)

The estimated model is $Z_t = \delta + \Phi_1 + Z_{t-1} + \Psi_1 \widetilde{H}_t + \varepsilon_t$, where $\delta = [\delta_1, \delta_2]^T$ is the intercept vector, $\Phi_1 = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}$ is the response coefficient matrix, $\Psi_1 = \begin{bmatrix} \Psi_{11} & \Psi_{12} & \Psi_{13} \\ \Psi_{21} & \Psi_{22} & \Psi_{23} \end{bmatrix}$ is the GARCH-M matrix, $\widetilde{H}_t = [h_{11,t}, h_{12,t}, h_{22,t}]^T$ is the conditional variance–covariance vector, $\varepsilon_t | I_{t-1} \sim N(0, H_t)$, and the conditional variance–covariance matrix $H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{bmatrix}$ is assumed to have a BEKK representation, i.e., $H_t = c^T c + a^T \varepsilon_{t-1} \varepsilon_{t-1}^T a + g^T H_{t-1} g$, where $c = \begin{bmatrix} c_{11} & c_{12} \\ 0 & c_{21} \end{bmatrix}$, $a = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $g = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$, $\varepsilon_t = \begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \end{bmatrix}$. The t-statistics are included in parentheses. The intercept vector δ and GARCH-M matrix Ψ_1 are transposed in the report for convenience.
* Denotes significance at the 10% level.
** Denotes significance at the 5% level.
*** Denotes significance at the 1% level.

Table 9
Nonlinear Granger causality tests between corporate bond and stock excess returns with conditional heteroskedasticity.

Hypothesis	Whole sample: 10/2002–12/2010 Granger causality test	Subsample: 10/2002–06/2007 Granger causality test
High-yield bond excess return vs. investment-grade bond excess return		
Hypothesis 1	2.25	4.74***
Hypothesis 2	15.20***	1.40
High-yield bond excess return vs. stock excess return		
Hypothesis 3	42.83***	19.71***
Hypothesis 4	1.74	1.35
Investment-grade bond excess return vs. stock excess return		
Hypothesis 5	29.67***	4.00**
Hypothesis 6	0.12	0.19

Reported in the table are the F statistics for Granger causality tests. The lag orders (p) are selected by using the BIC information criterion.
Hypothesis 1: high-yield bond excess returns are not Granger-caused by investment-grade bond excess returns.
Hypothesis 2: investment-grade bond excess returns are not Granger-caused by high-yield bond excess returns.
Hypothesis 3: high-yield bond excess returns are not Granger-caused by stock excess returns.
Hypothesis 4: stock excess returns are not Granger-caused by high-yield bond excess returns.
Hypothesis 5: investment-grade bond excess returns are not Granger-caused by stock excess returns.
Hypothesis 6: stock excess returns are not Granger-caused by investment-grade bond excess returns.
* Denotes significance at the 10% level.
** Denotes significance at the 5% level.
*** Denotes significance at the 1% level.

Table 10
Nonlinear Granger causality test between corporate bond markets and S&P 500 futures markets with conditional heteroskedasticity.

	Whole sample: 10/2002–12/2010 Granger causality test	Subsample: 10/2002–06/2007 Granger causality test
Panel A: Causality tests using returns		
High-yield bond return vs. futures return		
Hypothesis 1	42.97***	20.05***
Hypothesis 2	1.26	0.82
Investment-grade bond return vs. futures return		
Hypothesis 3	28.20***	3.05**
Hypothesis 4	0.20	0.34
Panel B: Causality tests using excess returns		
High-yield bond excess return vs. futures excess return		
Hypothesis 5	4.27***	19.98***
Hypothesis 6	1.29	0.82
Investment-grade bond excess return vs. futures excess return		
Hypothesis 7	28.41***	3.04**
Hypothesis 8	0.19	0.33

Reported in the table are F statistics for Granger causality tests. The lag orders (p) in VAR models are selected using the BIC information criterion.
Hypothesis 1: high-yield bond returns are not Granger-caused by futures returns.
Hypothesis 2: futures returns are not Granger caused by high-yield bond returns.
Hypothesis 3: investment-grade bond returns are not Granger-caused by futures returns.
Hypothesis 4: futures returns are not Granger-caused by investment-grade bond returns.
Hypothesis 5: high-yield bond excess returns are not Granger-caused by futures excess returns.
Hypothesis 6: futures excess returns are not Granger caused by high-yield bond excess returns.
Hypothesis 7: investment-grade bond excess returns are not Granger-caused by futures excess returns.
Hypothesis 8: futures excess returns are not Granger-caused by investment-grade bond excess returns.
* Denotes significance at the 10% level.
** Denotes significance at the 5% level.
*** Denotes significance at the 1% level.

High-yield bond excess returns and investment-grade bond excess returns are Granger-caused by stock excess returns at the 1% significance level.

The right side of Table 9 reports the results excluding the recent credit crunch and subsequent periods. Results show that high-yield bond excess returns are Granger-caused by investment-grade bond excess returns, but not vice versa. Stock excess returns are neither Granger-caused by high-yield bond nor by investment-grade bond excess returns. By contrast, both high-yield bond and investment-grade bond excess returns are Granger-caused by stock excess returns at the 1% and 5% significance level, respectively. Results support the hypothesis that the stock market leads the corporate bond market even after controlling for the effect of extreme observations. Thus, the lead–lag relation between stock and bond market returns is robust to the exclusion of the financial crisis period.

4.7. Tests using stock index futures data

Previous studies have found that the futures market leads the spot market (see, for example, Wu et al., 2005). We next replace the S&P 500 stock index return by the S&P 500 index futures return in the causality test. The S&P 500 futures return is driven by similar fundamental factors that affect the stock index return. Nevertheless, if the index futures market impounds information more quickly, we should observe a stronger predictive effect on bond returns.

Panel A of Table 10 reports results of causality tests using the S&P 500 futures and corporate bond returns. These tests account for conditional heteroskedasticity in returns. The left side of Panel A shows the results based on the whole sample period, whereas the right side reports the results excluding the most recent credit crunch and subsequent periods. Results show that both high-yield and investment-grade bond returns are significantly Granger-caused by the S&P 500 futures return, whereas the S&P 500 futures return is neither Granger-caused by high-yield nor by investment-grade bond returns. These results hold for both the whole sample period and the subsample period excluding the recent financial crisis. Results strongly support the hypothesis that corporate bond returns are predicted by past stock index futures returns.[6]

We next extend the test on the excess returns by using index futures returns. Panel B of Table 10 reports results of causality tests. Both high-yield and investment-grade bond excess returns are Granger-caused by stock index futures excess returns. By contrast, index futures excess returns are not Granger-caused by either investment-grade or high-yield bond excess returns. These results are again robust to the exclusion of the recent credit crunch and subsequent periods. Results suggest that the causal relation is primarily driven by the response to the flow of fundamental information rather than the discount rate.

5. Effects of government bond and stock returns on corporate bond returns

Corporate bonds can be viewed as a hybrid of riskless bonds and stocks, and so their returns are expected to be related to both government bond and stock returns. These relations depend on the risk of corporate bonds. High-yield bonds have high default risk, which makes their expected cash flows tied more closely to firm value changes. As such, returns of high-yield bonds are more sensitive to news about firms' earnings and economic fundamentals.

[6] For a robustness check, we also estimate the bivariate VAR-GARCH (1,1)-M model using futures data. The results are consistent with the Granger causality test. These results are available upon request.

As variations of stock returns are closely linked to changes in firms' expected future cash flow, the correlation will be high between high-yield bond and stock returns. Thus, high-yield bonds would be more like stocks than riskless bonds. On the other hand, investment-grade bonds have relatively stable expected cash flows because their default risk is small. This makes returns of investment-grade bonds less sensitive to news about firms' future cash flows. Instead, these safe bonds are more sensitive to interest rate changes and behave more like government bonds. Therefore, one would expect higher contemporaneous and lagged correlations between investment-grade bond and government bond returns.

We test these hypotheses using the following time-series regression:

$$r_t = \alpha + \sum_{j=1}^{q_1} \beta_j^{AR} r_{t-j} + \sum_{j=0}^{q_2} \beta_j^{S} r\_sp500_{t-j} + \sum_{j=0}^{q_3} \beta_j^{L} r\_LEHMAN_{t-j} + \varepsilon_t,$$

where the dependent variable r_t is either the high-yield bond return r_nbbh_t or the investment-grade bond return r_nbbi_t, β^{AR} is the autoregressive coefficient, β^S is the response coefficient associated with the S&P 500 index return, β^L is the coefficient of the Lehman intermediate government bond index return, and q_1, q_2, q_3 denote the lag orders for corporate bond, stock, and government bond returns, respectively.[7] We incorporate current and lagged stock index returns and Lehman intermediate government bond index returns to capture both contemporaneous and lagged effects on corporate bond returns.

The above regression ignores the fact that investment- and speculative-grade bond returns are cross autocorrelated (see Panel D of Tables 1 and 4). To account for this dynamic relation, we estimate the following time-series regression:

$$r_t = \alpha + \sum_{j=1}^{q_1} \beta_j^{AR} r_{t-j} + \sum_{j=0}^{q_2} \beta_j^{S} r\_sp500_{t-j} + \sum_{j=0}^{q_3} \beta_j^{L} r\_LEHMAN_{t-j} + \sum_{j=1}^{q_4} \beta_j^{C} r_{t-j}^{C} + \varepsilon_t,$$

where r_{t-j}^C is the other corporate bond index return (i.e., if the dependent variable $r_t = r/nbbh_t$, then $r_t^C = r\_nbbi_t$, and vice versa) and q_4 is the lag order of the cross serial correlation term of corporate bond returns. Here we add the lagged investment-grade bond returns in the high-yield bond return regression and the lagged high-yield bond returns in the investment-grade bond return regression to capture the cross autocorrelation between these two bond market segments.

Table 11 reports results for these regressions. The BIC tests reveal that a lag order of 4 is adequate to capture the lagged effects of regressors, but we also examine the sensitivity of results to different lag orders. Conditional heteroskedasticity is taken into account in all regressions. Results show a significant positive effect of government bond returns on both high-yield and investment-grade bond returns. The effect of government bond returns, as measured by the sum of response coefficients $\sum \beta_j^L$, is significant at the 1% level for both high-yield and investment- grade bonds. But the magnitude of the effect of government bond returns on investment-grade bonds is four times as large compared with that on high-yield bonds. The result shows that investment-grade bonds are much more sensitive to risk-free rate changes.

The effect of past investment-grade bond returns on high-yield bond returns is significant at the 5% level and the effect of past high-yield bond returns on investment-grade bond returns is significant at the 1% level, after controlling for the effects of government bond, stock index and own past returns. The effect of stock index returns on high-yield bond returns is significantly positive, confirming that high-yield bond returns are affected by risk factors similar to those driving stock returns. This positive relationship remains quite significant even after including lagged investment-grade bond returns in the regression. In comparison, the effect of stock index returns on investment-grade bond returns is much smaller, although it is still significantly positive. Tests reject the hypothesis of no stock return effect regardless of whether we include the contemporaneous stock index returns or not (see the test results in the last two columns). In sum, stock returns predict corporate bond returns, and this predictive relationship is stronger for high-yield bonds. Moreover, investment-grade bonds behave more like government bonds, whereas high-yield bonds behave more like stocks.

Empirical results are not sensitive to the lag order for the time-series regression. Panel B increases the lag order by one for each explanatory variable, while Panel C reports the results using the differential lag structures for different variables and including only the current return of government bonds in the regression. Results again show that high-yield bonds are more closely associated with stocks and investment-grade bonds are more like government bonds. Both high-yield and investment-grade bond returns are significantly related to concurrent and lagged stock market returns.

Our results for investment-grade bonds are consistent with the finding of Kwan (1996), while our results for high-yield bonds are more in line with the finding of Downing et al. (2009) based on individual bond data. Using investment-grade bond data, Kwan (1996) finds that lagged stock returns have explanatory power for current bond returns. We find that the predictive power of stock returns remains strong even after controlling for the effects of interest rates and the cross dependence and nonlinearity in bond returns. Downing et al. show that the lead–lag relation holds only for individual convertible bonds with conversion options deep in the money and nonconvertible bonds with severe financial distress. By contrast, using the broad-based market index data, we find that stock market returns lead both investment- and speculative-grade straight (nonconvertible) bond returns. On the other hand, our finding contrasts sharply with Hotchkiss and Ronen's (2002) results, which show no evidence that stock returns lead high-yield bond returns. This discrepancy could be due to differences in the sample and the study period. Their sample size is small (20 high-yield bonds) and sample period is short (9 months). Our bond indexes cover a large number of speculative- and investment-grade bonds. As our sample is broadly based, results are supposedly more representative. In addition, the longer sample period in our study increases the efficiency of parameter estimates and the power of statistical tests.

Our results show that the lead–lag relation between stocks and bonds is driven mainly by the fundamental information related to cash flow, instead of the interest rate factor. As stock returns have predictive power for bond returns, our empirical results are consistent with the contention that the stock market is more efficient than the bond market in impounding the information related to cash flow. The literature has suggested that informed traders prefer to trade in a market with high liquidity and low trading cost. As such, private information is likely to be embedded faster in securities in a market with higher liquidity and lower trading cost. This argument has been used to explain the lead–lag relation between the derivative and stock markets documented in the literature. To the extent that the corporate bond market is much less liquid than the equity market and cost of trading bonds is higher, informed traders would prefer to trade in the stock market than in the corporate bond market. Our empirical finding is consistent with this informed trading hypothesis and suggests that the stock market leads the bond market in conveying the fundamental information.

[7] The Lehman intermediate government bond index is downloaded from Datastream. The name changed to Barclay after Lehman went bankrupt in September 2008.

Table 11
Relations between corporate bond returns and government bond and stock index returns.

		$\sum_{j=1}^{q_1}\beta_j^{AR}$	$\sum_{j=0}^{q_2}\beta_j^{S}$	$\sum_{j=0}^{q_3}\beta_j^{L}$	$\sum_{j=1}^{q_4}\beta_j^{C}$	Test1: $\beta_j^S=0,\ \forall j$	Test2: $\beta_j^S=0,\ j>1$
Panel A: Lag order: $q_1 = q_2 = q_3 = q_4 = 4$							
r_nbbh_t	Model 1	0.374***	0.186***	0.211***			
		(0.032)	(0.016)	(0.060)		113.71***	71.60***
	Model 2	0.351***	0.188***	0.130*	0.120**		
		(0.034)	(0.016)	(0.078)	(0.064)	113.89***	68.85***
r_nbbi_t	Model 1	0.204**	0.064***	0.808***		48.55***	35.27***
		(0.037)	(0.007)	(0.049)			
	Model 2	0.103	0.049***	0.889***	0.072***	41.27***	28.17***
		(0.061)	(0.007)	(0.051)	(0.014)		
Panel B: Lag order: $q_1 = q_2 = q_3 = q_4 = 5$							
r_nbbh_t	Model 1	0.293***	0.228***	0.290***		98.48***	61.54***
		(0.035)	(0.018)	(0.067)			
	Model 2	0.283***	0.221***	0.061	0.243***		
		(0.039)	(0.017)	(0.085)	(0.070)	95.81***	60.23***
r_nbbi_t	Model 1	0.233***	0.066***	0.823***		46.25***	33.48***
		(0.037)	(0.007)	(0.051)			
	Model 2	0.104	0.049**	0.882***	0.076***	34.35***	19.55***
		(0.070)	(0.008)	(0.057)	(0.015)		
Panel C: Lag order: $q_1 = 3$, $q_2 = 4$, and $q_3 = 0$							
r_nbbh_t	Model 1	0.387***	0.177***	0.083***		110.56***	70.82***
		(0.030)	(0.014)	(0.022)			
	Model2 ($q_4 = 3$)	0.393***	0.174***	0.108***	0.136**	118.18***	75.37***
		(0.032)	(0.015)	(0.022)	(0.039)		
r_nbbi_t	Model 1	0.225***	0.060***	0.730***		37.83***	22.07***
		(0.020)	(0.005)	(0.011)			
	Model 2 ($q_4 = 3$)	0.189***	0.042***	0.723***	0.052***	33.39***	13.95***
		(0.039)	(0.011)	(0.006)	(0.011)		

The table reports the regression results of corporate bond market return on stock market return and intermediate government bond returns. The co-movement between the high-yield bond market and investment-grade bond markets is excluded in Model 1 but considered in Model 2. Standard errors are reported in parentheses. In Test 1, we test whether $\beta_j^S = 0$ for all j including the contemporary return, whereas in Test 2 we test whether $\beta_j^S = 0$ for $j > 0$. The heteroskedasticity in the error term is captured by the GARCH (1,1) process. The lag orders are q_1, q_2 and q_3 for returns of corporate bonds, stocks and government bonds, respectively. Panels A–C impose different lag orders for these returns.

* Denotes significance at the 10% level.

** Denotes significance at the 5% level.

*** Denotes significance at the 1% level.

6. Out-of-sample results

The results above are from in-sample estimation. We next conduct the out-of-sample forecasts. The out-of-sample performance is a useful model diagnostic for in-sample regressions. A good model should not only be able to explain what happened in the past but also to forecast the future. In addition, understanding the out-of-sample performance is of interest for an investor who seeks to use the predictive model for market timing or hedging purposes.

We compare the out-of-sample performance of two predictive regression models for future return forecasts. The first model uses only the security's own historical return information (i.e., the autoregressive model), and the second model uses both the security's own past returns and another security's past returns to forecast the future return.[8] At day t, we use the return data available up to t to estimate the parameters and predict the return at day $t + 1$ using each model. We then calculate the prediction error at time $t + 1$, which is the difference between the actual and predicted returns. This rolling regression estimation and forecasting continue until the last date of our sample. After obtaining the forecast errors, we calculate the out-of-sample R^2 and the root-mean-square error (RMSE) to evaluate the performance of the predictive regression models.

The out-of-sample R^2 is defined as $1-(MSE_R/MSE_N)$, where MSE_R is the mean-square-error of the out-of-sample forecast by the predictive regression model and MSE_N is the mean-square-error of the naive forecast. The naive forecast is represented by the historical average return. The predictive regression is represented by the univariate model including only the own lagged returns or the bivariate model including either lagged stock or another bond returns. This out-of-sample R^2 is commonly used in the literature to measure the forecasting performance (see, for example, Campbell and Thompson, 2008). The out-of-sample R^2 can be directly compared with the in-sample R^2. The out-of-sample R^2 is positive when the predictive regression has lower mean square prediction errors than the forecast based on the historical average return.

We report results of out-of-sample forecasts in Table 12. The out-of-sample prediction period begins in January 2005. Panel A compares the out-of-sample R^2 with the in-sample R^2 from regressions for the whole sample period and the subperiod that excludes the recent credit crunch and subsequent periods. For each security, we report the forecasts of three predictive regressions that include own past returns and/or anther security's past returns. For example, for high-yield bonds (r_nbbh), the first predictive regression includes only own past returns, the second regression includes both own past returns and past investment-grade bond returns, and the third includes own past returns and past stock returns. Similar arrangements are made for the forecasts of investment-grade bond and stock returns. The left and right sides of Panel A report the results for the models with conditional homoskedasticity and heteroskedasticity, respectively.

The out-of-sample R^2 values are all positive, indicating that the predictive regression model performs better than the forecast based on the historical average. The out-of-sample R^2 behaves much like the in-sample R^2. In fact, for the model with heteroske-

[8] The lag orders of past returns are again determined by BIC.

Table 12
Out-of-sample forecasts.

	Conditional homoskedasticity				Conditional heteroskedasticity			
	1/2005–12/2010		1/2005–06/2007		1/2005–12/2010		1/2005–06/2007	
	In-sample (%)	Out-of-sample (%)	In-sample (%)	Out-of-sample (%)	In-sample (%)	Out-of-sample (%)	In-sample (%)	Out-of-sample (%)
Panel A: R^2 for out-of-sample forecasts and in-sample regressions								
r_nbbh								
r_nbbh	15.70	10.25	15.10	10.14	14.55	14.49	14.91	11.90
r_nbbh and *r_nbbi*	16.01	9.83	15.37	9.97	14.63	14.52	15.10	11.44
r_nbbh and *r_sp*500	26.49	19.25	19.35	14.49	21.49	21.86	19.12	16.60
r_nbbi								
r_nbbi	8.90	8.20	5.50	4.62	8.04	9.36	5.43	4.57
r_nbbi and *r_nbbh*	9.50	4.27	5.98	2.70	8.69	8.24	5.83	2.92
r_nbbi and *r_sp*500	13.85	13.11	6.27	4.77	12.89	14.13	6.09	4.61
*r_sp*500								
*r_sp*500	2.51	1.52	0.97	−0.37	1.65	1.85	0.64	−0.01
*r_sp*500 and *r_nbbh*	3.18	−2.14	1.05	−0.97	1.33	0.48	0.96	−0.29
*r_sp*500 and *r_nbbi*	2.65	1.02	1.45	−0.49	1.56	1.72	0.69	−0.16

$\Delta RMSE$: $RMSE_2 - RMSE_1$						*Percentage change: $(RMSE_2 - RMSE_1)/RMSE_1$*					
r_nbbh		*r_nbbi*		*r_sp*500		*r_nbbh*		*r_nbbi*		*r_sp*500	
r_nbbi	*r_sp*500	*r_nbbh*	*r_sp*500	*r_nbbh*	*r_nbbi*	*r_nbbi* (%)	*r_sp*500 (%)	*r_nbbh* (%)	*r_sp*500 (%)	*r_nbbh* (%)	*r_nbbi* (%)
Panel B: Changes and percentage changes in RMSE for out-of-sample forecasts of bond and stock returns											
Conditional homoskedasticity: 1/2005–12/2010											
0.0014	−0.0317	0.0051	−0.0074	0.0214	0.0037	0.24	−5.15	2.12	−3.03	1.47	0.26
Conditional homoskedasticity: 1/2005–06/2007											
0.0003	−0.0067	0.0014	−0.0001	0.0013	0.0004	0.09	−2.53	1.02	−0.08	0.19	0.06
Conditional heteroskedasticity: 1/2005–12/2010											
−0.0001	−0.0279	−0.0015	−0.0074	0.0101	0.0008	−0.01	−4.62	−0.61	−3.06	0.70	0.05
Conditional heteroskedasticity: 1/2005–06/2007											
0.0007	−0.0071	0.0012	−0.0003	0.0009	0.0005	0.26	−2.69	0.86	−0.25	0.14	0.07

This panel reports the R^2 for out-of-sample forecasts and in-sample regressions. The out-of-sample R^2 is $1 - (MSE_R/MSE_N)$ where MSE_R is the mean-squared-error of the out-of-sample regression forecast and MSE_N is the mean-squared-error of naive forecasts. Each naive forecast is the average return preceding the out-of-sample forecast. The in-sample R^2 are from in-sample regressions.
This panel reports the improvement in *RMSE* when the past return of the other index is used in the out-of-sample forecast of high-yield bond (*r_nbbh*), investment-grade bond (*r_nbbi*) and stock index returns (*r_sp*500). The out-of-sample forecast begins from 2005. $RMSE_1$ is the *RMSE* of the out-of-sample forecast using the own past return while $RMSE_2$ is the *RMSE* of the out-of-sample forecast using both the own past return and the other index's past return. The change in *RMSE* and the percentage change in *RMSE* are reported.

dasticity in returns (see the right side of the panel), the out-of-sample R^2 is quite comparable to the in-sample R^2. Since the out-of-sample forecast produces a pattern similar to the in-sample regression test, the out-of-sample tests support our basic inferences about the causal relation. Excluding the recent credit crunch and subsequent periods does not improve the out-of-sample forecast. In both sample periods, the performance of out-of-sample forecasts is much better (or out-of-sample R^2 is higher) when past stock returns are used as an additional predictor to forecast bond returns (see the third row in each set of regressions for bonds). On the other hand, past bond returns are not helpful for forecasting stock returns. Results consistently show that stock returns are useful for forecasting bond returns and strongly support the hypothesis for the predictability of bond returns.

Panel B of Table 12 reports the results based on the *RMSE* when the past return for another security is utilized as a predictor in the out-of-sample forecast. $RMSE_1$ is the *RMSE* of the out-of-sample forecast using only own past returns whereas $RMSE_2$ is the *RMSE* of the out-of-sample forecast using both own past returns and past returns of another bond or stock index. We calculate both changes in the *RMSE* ($\Delta RMSE = RMSE_2 - RMSE_1$) and percentage changes in the *RMSE* ($\Delta RMSE/RMSE_1$). The left side of the panel reports changes in the *RMSE* and the right side reports percentage changes in the *RMSE*. A negative value of $\Delta RMSE$ or $\%\Delta RMSE$ implies an improvement in out-of-sample forecast when including another forecasting variable in the predictive regression model.

The upper (lower) part of Panel B shows results for the predictive regression model assuming conditional homoskedasticity (heteroskedasticity) in returns. Results in the first column show that when forecasting the high-yield bond return (*r_nbbh*), including the information for the past investment-grade bond return (*r_nbbi*) does not reduce the *RMSE*. For the whole sample, the *RMSE* for the high-yield bond increases slightly (0.0014) instead when past investment-grade bond returns are added to the regression which has already included past high-yield bond returns. For the subsample excluding the recent credit crunch and subsequent periods (row 2), the increase in *RMSE* for high-yield bonds is smaller (0.0003) but remains positive. Similar results are obtained for the investment-grade bond return forecast when including past high-yield bond returns as an additional forecasting variable, that is, high-yield bond returns are not helpful for forecasting future investment-grade bond returns. Considering heteroskedasticity (see the lower part of the panel) does not change results materially.

By contrast, there is a significant improvement in the *RMSE* when past stock returns are used to help forecast both the high-yield and investment-grade bond returns. For the model with homoskedasticity, the *RMSE* for the forecast of high-yield bonds decreases by 0.0317 (or 5.15%) when adding lagged stock returns as a forecasting variable in the regression for the whole sample period. The gain in *RMSE* drops to 0.0067 (or 2.53%) for the subperiod analysis excluding the credit crunch and subsequent periods but it is still a meaningful improvement in forecasting performance in percentage terms. A similar pattern is found for the investment-grade bond return forecast using the predictive regression model that includes past stock returns. For the whole sample period, including past stock returns leads to a drop of 0.0074 (or 3.03%) in the *RMSE* of investment-grade bonds when using the model with conditional homoskedasticity. For the subperiod excluding the financial crisis and subsequent periods, the gain in the *RMSE* of investment-grade bonds is reduced but remains positive (i.e., the change in *RMSE* is still negative). Results based on the regression model with conditional heteroskedasticity show a similar pattern.

Results show that historical stock return information is important for predicting corporate bond returns out-of-sample, a finding consistent with in-sample Granger-causality tests. The decrease in the *RMSE* is larger for the high-yield bond than for the investment-grade bond. High-yield bond returns are more sensitive to stock market performance, as these risky bonds are linked more closely to stocks. Including the past stock return information thus improves the forecasting performance of the predictive regression model more. On the other hand, neither the high-yield bond return nor the investment-grade bond return is helpful for predicting the stock return out-of-sample. As shown, the *RMSE* for stocks does not decrease regardless of whether past investment-grade or high-yield bond returns are added in the predictive regression.

Overall, there is strong evidence that stock returns forecast corporate bond returns but not vice versa. Excluding the extreme observations associated with the recent credit crunch and subsequent periods does not change our basic inferences. Results of out-of-sample forecasting are consistent with in-sample tests, confirming that corporate bond returns are predictable.

7. Conclusion

An issue central to financial research is the predictability of asset returns. There is substantial evidence that stock returns are predictable. An important question is whether returns are also predictable for other asset classes. This paper examines this issue for the corporate bond market and employs empirical methodologies that are robust to nonlinearity in serial return dependence and conditional heteroskedasticity.

Empirical evidence strongly suggests that corporate bond market returns are predictable. There is evidence of return predictability for both investment-grade and high-yield bonds. These results are robust to alternative model specifications, return measures, and exclusion of extreme observations. Results show that corporate bond market returns exhibit higher autocorrelation and a more complicated structure of serial dependence than stock market returns. Stock market returns lead both high-yield and investment-grade bond returns, whereas there is little evidence that corporate bond market returns lead stock market returns. This lead–lag relation is stronger between high-yield bond returns and stock returns. The lead–lag relation remains strong when we control the effects of interest rates and serial and cross-serial dependence in bond returns. Moreover, out-of-sample tests show results consistent with in-sample tests. Results show that the past stock return is useful information for predicting both speculative- and investment-grade bond returns out of sample.

Our findings provide important implications for corporate bond modeling and asset pricing tests. Results suggest that tests of the risk-return tradeoff in corporate bonds should take into account the cross-serial dependence between bond and stock returns. Our findings also impose restrictions on the specification of the term structure model. In particular, our results suggest that the term structure model of defaultable bonds should account for the serial and cross-serial dependence in corporate bond price changes in order to provide a more satisfactory explanation for corporate bond price behavior.

The cause of the predictability of corporate bond returns is not immediately clear. Return predictability could be due to bond illiquidity, transaction cost, market structure, or other frictions. An exploration of the cause for our results is an important extension of this paper, and we leave this for a future study.

Acknowledgments

We thank the participants at the 2006 WISE-SMU joint seminar at Xiamen University, 2012 New Zealand Finance Colloquium, and department seminars at Cornell University, University of Otago and Victoria University of Wellington for helpful comments. Hai Lin acknowledges the financial support from University of Otago, National Natural Science Foundation of China Grant Nos. 70971114 and 71101121, and the Humanities & Social Sciences Fund of The Chinese Ministry of Education Grant No. 11YJC790014. All errors are our own.

Appendix A

In this appendix, we describe the empirical methodology and the test procedure for the models used in the paper.

A.1. Autocorrelation tests

Define the sample autocovariance function at lag j as

$$\hat{\gamma}(j) = T^{-1}\sum_{t=|j|+1}^{T}(X_t - \overline{X})(X_{t-|j|} - \overline{X}), \quad j = 0, \pm 1, \ldots, \pm(T-1),$$

where $\overline{X}$ is the sample mean. The sample autocorrelation functions are

$$\hat{\rho}(j) = \hat{\gamma}(j)/\hat{\gamma}(0), \quad j = 0, \pm 1, \ldots, \pm(T-1).$$

Under the *i.i.d.* assumption on $\{X_t\}$, one can employ the Box–Pierce test to detect autocorrelation:

$$BP(p) = T\sum_{j=1}^{p}\hat{\rho}^2(j) = \sum_{j=1}^{p}\left[\sqrt{T}\hat{\rho}(j)\right]^2 \xrightarrow{d} \chi^2_p.$$

To improve the size of the Box–Pierce test in finite samples, one can use the Ljung–Box test:

$$LB(p) = T(T+2)\sum_{j=1}^{p}(T-j)^{-1}\hat{\rho}^2(j) \xrightarrow{d} \chi^2_p.$$

The asymptotic distribution of LB statistics holds with conditional homoskedasticity $\mathrm{var}(X_t|I_{t-1}) = \sigma^2$.

A.2. Variance ratio tests

Let $\sum_{j=1}^{p}X_{t-j}$ be the cumulative return over a period of p days. Then, under the null hypothesis H_0,

$$\frac{\mathrm{var}\left(\sum_{j=1}^{p}X_{t-j}\right)}{p\mathrm{var}(X_t)} = \frac{p\gamma(0) + 2p\sum_{j=1}^{p}(1-j/p)\gamma(j)}{p\gamma(0)} = 1,$$

if the autocovariance functions $\rho(j) = 0$ for all $j \neq 0$. A departure from unity is evidence against H_0. Under conditional homoskedasticity, Lo and MacKinlay (1988) show that the variance ratio has the following limiting distribution:

$$VR(p) = \frac{p\hat{\gamma}(0) + 2p\sum_{j=1}^{p}(1-j/p)\hat{\gamma}(j)}{p\hat{\gamma}(0)} \xrightarrow{d} N[0, 2(2p-1)(p-1)/3p].$$

However, volatilities of returns in financial markets often change over time. Under the null H_0 with conditional heteroskedasticity, the variance ratio has the following asymptotical distribution instead:

$$VR(p) = \sqrt{T/p}\sum_{j=1}^{p}(1-j/p)\hat{\gamma}(j) \Big/ \sqrt{\hat{\gamma}_2(j)} \xrightarrow{d} N[0, 2(2p-1)(p-1)/3p],$$

where $\hat{\gamma}(j)$ is the sample autocovariance function and

$$\hat{\gamma}_2(j) = T^{-1}\sum_{t=|j|+1}^{T}(X_t - \overline{X})^2(X_{t-j} - \overline{X})^2, \quad j = 0, \pm 1, \ldots, \pm(T-1).$$

When $p \to \infty$ as $T \to \infty$, the variance ratio test statistic is asymptotically equivalent to

$$VR(p) = \sqrt{T/p}\sum_{j=1}^{p}(1-j/p)\hat{\rho}(j) = \frac{\pi}{2}\sqrt{T/p}\left[\hat{f}(0) - \frac{1}{2\pi}\right],$$

where $\hat{f}(0)$ is a kernel-based normalized spectral density estimator at frequency zero with the Bartlett kernel $K(z) = (1 - |z|)1(|z| \leqslant 1)$, and $1(|z| \leqslant 1)$ is an indicator function that equals one if the random variable z is between $[-1,1]$ and zero, otherwise.

A.3. The generalized spectral test

This subsection outlines the generalized spectral test of Hong and Lee (2005) and the implementation procedure. Let a time series $\{X_t\}$ follow the process $X_t = g(I_{t-1}, \theta) + \varepsilon_t$, where I_{t-1} is the information set at time $t-1$, $g(I_{t-1}, \theta)$ is conditional mean $E(Y_t|I_{t-1})$, and $\theta \in \Theta$ is a parameter value in a limited dimensional parameter set Θ. In the present case, X_t can be the return on a stock or corporate bond index, and $g(I_{t-1}, \theta) = \mu$. Denote the demeaned return series as $\varepsilon_t \equiv X_t - \mu$. Under the null hypothesis H_0 of $E[\varepsilon_t|I_{t-1}] = 0$, this implies that $E\left[\varepsilon_t|I_{t-1}^{\varepsilon}\right] = 0$, $I_{t-1}^{\varepsilon} \equiv \{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\}$, and the return is unpredictable because the demeaned series $\{\varepsilon_t\}$ is a martingale difference sequence (MDS), where its conditional mean is serially independent.

Let $\{\varepsilon_t\}$ be a strictly stationary process with a marginal characteristic function $\varphi(u) = E(e^{iu\varepsilon_t})$ and a pairwise joint characteristic function $\varphi_j(u, v) \equiv E(e^{iu\varepsilon_t + iv\varepsilon_{t-|j|}})$, where $i = \sqrt{-1}$, $-\infty < u,\ v < \infty$, and $j = 0, \pm1, \pm2, \ldots$ The generalized spectral test uses the spectrum of the transformed series $\{e^{iu\varepsilon_t}\}$ in the MDS test. The generalized spectral density function can be written as

$$f(\omega, u, v) \equiv \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\sigma_j(u, v)e^{-ij\omega}, \quad \omega \in [-\pi, \pi], \ -\infty < u, \ v < \infty,$$

where ω is the frequency, and $\sigma_j(u, v) = \mathrm{cov}(e^{iuX_t}, e^{ivX_{t-|j|}}) = \varphi_j(u, v) - \varphi_t(u)\varphi_{t-|j|}(v)$ is the covariance function of the transformed series.

However, the generalized spectrum $f(\omega, u, v)$ itself is not suitable for testing H_0, since it involves all pairwise serial dependencies in $\{X_t\}$ at various lags. To cope with this problem, Hong and Lee (2005) propose the use of the partial derivative of the generalized spectral function in the MDS test, which focuses on the serial dependence in the conditional mean. The partial derivative function is not affected by serial dependence of higher-order moments. The partial derivative of the generalized spectral function is

$$f^{(0,1,0)}(\omega, u, v) \equiv \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\sigma_j^{(1,0)}(u, v)e^{-ij\omega}, \quad \omega \in [-\pi, \pi], \ -\infty < v < \infty,$$

where

$$\sigma_j^{(1,0)}(0, v) = \frac{\partial}{\partial u}\sigma_j(u, v)|_{u=0} = \mathrm{cov}(i\varepsilon_t, e^{iv\varepsilon_{t-|j|}}).$$

The covariance measure $\sigma_j^{(1,0)}(0, v)$ checks whether the autoregression function $E(\varepsilon_t|\varepsilon_{t-j})$is zero at lag j. An important property is that $\sigma_j^{(1,0)}(0, v) = 0$ for all $v \in R$ if and only if $E(\varepsilon_t|\varepsilon_{t-j}) = 0$ a.s. Therefore, it can detect any linear and nonlinear serial dependence in mean.

Under H_0, the partial derivative function $f^{(0,1,0)}(\omega, 0, v)$ becomes a "flat" spectrum:

$$f^{(0,1,0)}(\omega, 0, v) = f_0^{(0,1,0)}(\omega, 0, v) \equiv \frac{1}{2\pi}\sigma_0^{(1,0)}(0, v),$$
$$\omega \in [-\pi, \pi], \quad -\infty < v < \infty.$$

We can estimate $f^{(0,1,0)}(\omega, 0, v)$ and check whether it is a flat spectrum. If it is, then H_0 is true. Tests involve comparing consistent estimators for $f^{(0,1,0)}(\omega, 0, v)$ and $f_0^{(0,1,0)}(\omega, 0, v)$. Hong and Lee (2005) propose a test of whether $\hat{\varepsilon}_t = X_t - g(I_{t-1}, \hat{\theta})$ satisfies the condition that $E(\hat{\varepsilon}_t|I_{t-1}) = 0$. They show that $f^{(0,1,0)}(\omega, 0, v)$ can be consistently estimated by a smoothed kernel:

$$\hat{f}^{(0,1,0)}(\omega, 0, v) \equiv \frac{1}{2\pi}\sum_{j=1-T}^{T-1}(1 - |j|/T)^{1/2}k(j/p)\hat{\rho}_j^{(1,0)}(0, v)e^{-ij\omega},$$
$$\omega \in [-\pi, \pi], \quad -\infty < v < \infty,$$

where

$$\hat{\rho}_j^{(1,0)}(0, v) = (T - |j|)^{-1}\sum_{t=|j|+1}^{T} i\hat{\varepsilon}_t\hat{h}_{t-|j|}(v), \quad j = 0, \pm1, \ldots, \pm(T-1),$$

$$\hat{h}_{t-|j|}(v) = \hat{\varphi}_{t-|j|}(v) - \hat{G}_t'\hat{\beta}_{|j|}(v), \quad \hat{G}_t = \frac{\partial}{\partial\theta}g(\hat{I}_{t-1}, \hat{\theta}),$$

$$\hat{\beta}_j(v) = \left(\sum_{t=1}^{T}\hat{G}_t\hat{G}_t'\right)^{-1}\sum_{t=|j|+1}^{T}\hat{G}_t\hat{\varphi}_{t-|j|}(v),$$

$p \equiv p(T)$ is a bandwidth, and $k:(-\infty, \infty) \to [-1,1]$ is a symmetric kernel with $k(0) = 0$. An example of $k(\cdot)$ is the Bartlett kernel function, i.e., $k(z) = (1 - |z|)I(|z| \leqslant 1)$, where $I(\cdot)$ is an indicator function that equals to one when z is between $[-1,1]$ and zero, otherwise.

The flat spectrum can be estimated by

$$\hat{f}_0^{(0,1,0)}(\omega, 0, v) = \frac{1}{2\pi}\hat{\rho}_0^{(1,0)}(0, v), \quad \omega \in [-\pi, \pi], \ -\infty < v < \infty.$$

Under H_0, $f^{(0,1,0)}(\omega, 0, v) = f_0^{(0,1,0)}(\omega, 0, v)$. When $\{\hat{\varepsilon}_t\}$ is an MDS, the estimated function $\hat{f}^{(0,1,0)}(\omega, 0, v)$ will be very close to $\hat{f}_0^{(0,1,0)}(\omega, 0, v)$. If $\hat{f}^{(0,1,0)}(\omega, 0, v)$ turns out to be quite different from $\hat{f}_0^{(0,1,0)}(\omega, 0, v)$, $\{\hat{\varepsilon}_t\}$ is not an MDS. Based on this relationship, Hong and Lee (2005) construct the following test statistic:

$$M(p) = \left[\sum_{j=1}^{T-1}k^2(j/p)(T-j)\int\left|\hat{\rho}_j^{(1,0)}(0, v)\right|^2 dW(v) - \widehat{C}_1(p)\right]\Big/\sqrt{\widehat{D}_1(p)},$$

where

$$\widehat{C}_1(p) = \widehat{S}^2\int\left|\frac{1}{T}\sum_{t=1}^{T}\hat{h}_t(v)\right|^2 dW(v)\sum_{j=1}^{T-1}k^2(j/p),$$

$$\widehat{D}_1(p) = 2\widehat{S}^2\int\int\left|\frac{1}{T}\sum_{t=1}^{T}\hat{h}_t(u)\hat{h}_t(v)\right|^2 dW(u)dW(v)\sum_{j=1}^{T-2}k^4(j/p),$$

$\widehat{S}^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t^2$, and $W(\cdot):R \to R^+$ is a nondecreasing function that sets weights symmetrically around zero. An example of $W(\cdot)$ is the $N(0,1)$ cumulative density function commonly used in the characteristic function literature. It can be shown that $M(p) \xrightarrow{d} N(0,1)$ under H_0. Otherwise, $M(p)$ goes to positive infinity as the sample size increases. Thus, if $M(p)$ is greater than the critical value of $N(0,1)$ at a significant level, the difference between $\hat{f}^{(0,1,0)}(\omega, 0, v)$ and $\hat{f}_0^{(0,1,0)}(\omega, 0, v)$ is significant and the MDS hypothesis is rejected.

The steps for conducting the generalized spectral test of the MDS hypothesis are as follows:

(1) Estimate $\hat{\mu}$ and compute the error $\{\hat{\varepsilon}_t\} \equiv y_t - \hat{\mu}$.
(2) Calculate the derivative $\widehat{G}_t = \frac{\partial}{\partial \theta} g(I_{t-1}, \hat{\theta}) = 1$.
(3) Regress $\hat{\varphi}_{t-|j|}(v)$ on $\widehat{G}_t$ and compute the residual $\hat{h}_{t-|j|}(v)$. We separately regress the real and imaginary components of $\hat{\varphi}_{t-|j|}(v)$ on $\widehat{G}_t$.
(4) Compute the test statistic $M(p)$.
(5) Compare $M(p)$ with the critical value of standard normal distribution $N(0,1)$ at a significance level. For example, the critical value at the 1% level is 2.33. If the test statistic is greater than this critical value, the hypothesis H_0 is rejected at the 1% level, and we conclude that the series is not an MDS.

A.4. Nonlinear Granger causality tests

A.4.1. The causality model under heteroskedasticity

Standard Granger causality tests are not robust to heteroskedasticity. To overcome this problem, we modify the model to accommodate conditional heteroskedasticity in asset returns:

$$Z_t = \delta + \Phi_1 Z_{t-1} + \cdots + \Phi_p Z_{t-p} + \varepsilon_t,$$

where $\varepsilon_t | I_{t-1} \sim N(0, H_t)$, and $H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{bmatrix}$ is the conditional variance–covariance matrix. H_t can have a dynamic conditional correlation (DCC) representation (Engle, 2002), a BEKK representation (Baba et al., 1990), or a BEW representation (Bollerslev et al., 1988). In empirical tests, we adopt the BEKK representation in bivariate GARCH (1,1):

$$H_t = c^T c + a^T \varepsilon_{t-1} \varepsilon_{t-1}^T a + g^T H_{t-1} g,$$

where

$$c = \begin{bmatrix} c_{11} & c_{12} \\ 0 & c_{21} \end{bmatrix}, \quad a = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad g = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix},$$

$$\varepsilon_t = [\, \varepsilon_{1t-12t-1} \,].$$

A.4.2. The GARCH-M model

Both linear and nonlinear Granger causality tests deal with the cross dependence in returns without considering the effects of higher-order moments on the mean return. It is possible that some higher-order moments of returns may affect the mean return of an asset. A classic example is the GARCH-M effect where the second moment (conditional variance) affects conditional mean returns.

Considering the GARCH-M effect in a bivariate VAR-GARCH (1,1) setting, we have

$$Z_t = \delta + \Phi_1 Z_{t-1} + \cdots + \Phi_p Z_{t-p} + \Psi_1 \tilde{H}_t + \varepsilon_t, \quad Z_t = [Z_{1t}, Z_{2t}]^T,$$

where $\Psi_1 = \begin{bmatrix} \Psi_{11} & \Psi_{12} & \Psi_{13} \\ \Psi_{21} & \Psi_{22} & \Psi_{23} \end{bmatrix}$, $\tilde{H}_t = [h_{11,t}, h_{12,t}, h_{22,t}]^T$, and $\varepsilon_t | I_{t-1} \sim N(0, H_t)$. The matrix $H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{bmatrix}$ again can be formulated by a BEKK representation.

References

Abhyankar, A., Gonzales, A., 2009. News and the cross-section of expected corporate bond returns. Journal of Banking and Finance 33, 996–1004.
Ang, A., Bekaert, G., 2007. Stock return predictability: is it there? Review of Financial Studies 20, 651–707.
Baba, Y., Engle, R.F., Kraft, D.F., Kroner, K.F., 1990. Multivariate Simultaneous Generalized ARCH. Mimeo.
Bollerslev, T., Engle, R.F., Wooldridge, J.M., 1988. A capital asset pricing model with time-varying covariances. Journal of Political Economy 96, 116–131.
Campbell, J.Y., Thompson, S.B., 2008. Predicting the equity premium out of sample: can anything bear the historical average? Review of Financial Studies 21, 1509–1531.
Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. Econometrics of Financial Markets. Princeton University Press, Princeton, NJ.
Chen, X., Maringer, D., 2011. Detecting time-variation in corporate bond index returns: a smooth transition regression model. Journal of Banking and Finance 35, 95–103.
Domínguez, M., Lobato, I., 2003. Testing the martingale difference hypothesis. Econometric Reviews 22, 351–377.
Downing, C., Underwood, S., Xing, Y., 2009. The relative information efficiency of stocks and bonds: an intraday analysis. Journal of Financial and Quantitative Analysis 44, 1081–1102.
Engle, R.F., 2002. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. Journal of Business and Economic Statistics 20, 339–350.
Fan, L., Tian, S., Zhang, C., 2012. Why are excess returns on China's treasury bonds so predictable? The role of the monetary system. Journal of Banking and Finance 36, 239–248.
Friewald, N., Jankowitsch, R., Subrahmanyam, M., 2012. Illiquidity or credit deterioration: a study of liquidity in the US corporate bond market during financial crises. Journal of Financial Economics, forthcoming.
Gebhardt, W.R., Hvidkjaer, S., Swaminathan, B., 2005. Stock and bond market interaction: does momentum spill over? Journal of Financial Economics 75, 651–690.
Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424–438.
Hamao, Y., Masulis, R.W., Ng, V., 1990. Correlations in price changes and volatility across international stock markets. Review of Financial Studies 3, 281–307.
Hong, Y.M., 1996. Consistent testing for serial correlation of unknown form. Econometrica 64, 837–864.
Hong, Y.M., Lee, Y.J., 2005. Generalized spectral tests for conditional mean models in time series with conditional heteroskedasticity of unknown form. Review of Economic Studies 73, 499–541.
Hotchkiss, E.S., Ronen, T., 2002. The informational efficiency of the corporate bond markets: an intraday analysis. Review of Financial Studies 15, 1325–1354.
Keim, D.B., Stambaugh, R.F., 1986. Predicting returns in the stock and bond markets. Journal of Financial Economics 17, 357–390.
Kessler, S., Scherer, B., 2009. Varying risk premia in international bond markets. Journal of Banking and Finance 33, 1361–1375.
Kwan, S.H., 1996. Firm-specific information and the correlation between individual stocks and bonds. Journal of Financial Economics 40, 63–80.
Lo, A., MacKinlay, A.C., 1988. Stock market prices do not follow random walks: evidence from a simple specification test. Review of Financial Studies 1, 41–66.
Rapach, E.D., Strauss, J.K., Zhou, G., 2010. Out-of sample equity premium prediction: combination forecasts and links to the real economy. Review of Financial Studies 23, 821–862.
Sekkel, R., 2011. International evidence on bond risk premia. Journal of Banking and Finance 35, 174–181.
Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies 21, 1455–1508.
Wu, C., Li, J., Zhang, W., 2005. Intraday return and volatility spillovers between international stock index futures markets. Journal of Futures Markets 25, 553–585.