Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 14 (2007) 220–247
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

Estimating the cross-sectional market response to an endogenous event: Naked vs. underwritten calls of convertible bonds

John T. Scruggs

Department of Banking and Finance, Terry College of Business, University of Georgia, Athens, GA, 30602-6253, USA

Accepted 6 February 2006
Available online 1 September 2006

Abstract

The literature on conditional event studies advocates the use of endogenous switching models to analyze cross-sectional variation in the stock market's response to corporate announcements of endogenous events. This paper proposes the use of a flexible Bayesian Markov chain Monte Carlo (MCMC) approach for estimating such models. The Bayesian MCMC approach offers several advantages over the Heckman–Lee two-stage estimation approach. In particular, analysis of the "treatment effect" (the difference between the observed and counter-factual outcomes) controls for the endogeneity of the firm's choice. As an application, the paper examines the market's response to naked and underwritten calls of convertible bonds. The paper reports evidence that the market's response to calls of convertible bonds is correlated with the private information partially revealed by the firm's choice of call type. However, although the average treatment effect associated with an underwritten call is negative, it is not significantly different from zero.

JEL classification: G14 (Information and Market Efficiency; Event Studies); C11 (Bayesian Analysis); C15 (Statistical Simulation Methods; Monte Carlo Methods)
Keywords: Convertible bond; Underwritten call; Conditional event study; Cross-sectional event study; Endogenous switching model; Self-selection model; Markov chain Monte Carlo; Gibbs sampling; Bayesian statistics

1. Introduction

Many empirical papers in corporate finance examine the market's response to a firm's announcement of a binary choice. In most instances, the choice is endogenous. Economically motivated managers have discretion over the type and timing of such announcements. The

E-mail address: jscruggs@terry.uga.edu.

0927-5398/$ - see front matter
doi:10.1016/j.jempfin.2006.02.002

manager's choice is conditioned on private or latent information relevant to valuation. The stock return associated with the announcement reflects the market's revised conditional expectation of the manager's private information. Traditional event study methodology (e.g., Fama et al., 1969) is well-suited for estimating *average* announcement effects. However, studies which attempt to explain *cross-sectional* variation in announcement effects are potentially misspecified. Acharya (1988, 1993), Eckbo, Maksimovic, and Williams (1990), Maddala (1996), and Prabhala (1997) advocate the use of "conditional" event study methods that explicitly model the latent information partially revealed by the announcement of an endogenous binary choice.

In this paper, I develop econometric methods for conditional event studies. I introduce an endogenous switching model to analyze the event window abnormal returns associated with corporate announcements. The paper's primary contribution is the use of flexible Bayesian Markov chain Monte Carlo (MCMC) methods, along the lines of Chib and Hamilton (2000), for estimating endogenous switching models. I focus on announcements which reveal an endogenous binary choice.[1] However, these methods could be used to estimate more general models in which the firm's choice reveals a polychotomous, multivariate binary or other complex choice. Bayesian MCMC methods have many attractive features. They provide exact small-sample posterior densities for the parameters. Since the model's choice and outcome equations are estimated simultaneously, posterior densities for parameters (and functions of interest) implicitly incorporate parameter uncertainty, and resulting inferences are free of errors-in-variables biases. Bayesian MCMC methods also offer a natural approach for examining "treatment effects". In a conditional event study, the "treatment effect" is the difference between the observed outcome (i.e., the event window abnormal return given the firm's actual choice) and the counter-factual outcome (i.e., the expected outcome under the alternative choice). Counter-factual outcomes are simulated conditional on the data, the model, and the other parameters. Analysis of average treatment effect controls for the endogeneity of the choice.

As an application, I examine the market's response to calls of convertible debt. The call announcement reveals whether the call is "naked" or underwritten by an investment bank. Thus, conditional on the announcement of the call, the call type (naked vs. underwritten) is an endogenous binary choice. Many papers have reported that the average abnormal return associated with calls of convertible bonds is negative and significant.[2] Singh et al. (1991) report that the average abnormal return for underwritten calls is negative and significant, while the average abnormal return for naked calls is negative, but not significantly different from zero. I investigate two potential explanations for the market's differential response to naked and underwritten calls of convertible debt. The *private information* hypothesis (see Cowan et al., 2000) predicts that an underwritten call is a more negative signal than a naked call. In terms of the model, it predicts that the private information partially revealed by the firm's choice of call type (i.e., naked vs. underwritten) is correlated with the event window abnormal return. The *announcement put option* hypothesis predicts that the market's response to calls of convertible bonds is related to the value of the put option given to convertible bondholders when the call is announced. These hypotheses are neither mutually exclusive nor exhaustive. I employ Bayesian MCMC methods to estimate a binary probit model of the firm's choice while simultaneously estimating a model of the market's

[1] As an application, Chib and Hamilton (2000) estimate an endogenous switching model to examine the efficacy of medical treatments when the treatment is not randomly assigned (i.e., the treatment is endogenous).

[2] See Mikkelson (1981), Ofer and Natarajan (1987), Campbell et al. (1991), Datta and Iskandar-Datta (1996), and Ederington and Goh (2001).

conditional response to the call announcement. The endogenous switching model explicitly tests both the *private information* and *announcement put option* hypotheses.

I find that the decision to have a convertible bond call underwritten is inversely related to the extent the conversion option is in-the-money (i.e., cushion), and positively related to pre-call stock price appreciation (i.e., run up). In the outcome model, I find some empirical support for the *private information* hypothesis. Conditional on the call event, the private information partially revealed by the firm's decision to have a convertible bond call underwritten is negatively related to event window abnormal returns.[3] However, the evidence is mixed. On average, the treatment effect associated with choosing an underwritten call is negative, but insignificant. In other words, I find no conclusive evidence that choosing the alternative call type (e.g., naked rather than underwritten or vice versa) would significantly change the outcome on average. I find no evidence to support the *announcement put option* hypothesis. Event window abnormal returns are negatively related (as predicted) to the imputed value of the announcement put option, but the relationship is not statistically reliable.

The remainder of the paper is organized as follows. Section 2 reviews the relevant theoretical and empirical literature on calls of convertible debt. Section 3 describes the endogenous switching model with a binary choice and two continuous potential outcomes. Section 4 discusses a flexible Bayesian approach for estimating the model. I discuss the use of Markov chain Monte Carlo (MCMC) methods to sample from and make inferences about the joint posterior density of the parameters of interest. Section 5 describes the data sample and discusses the empirical implications of the *private information* and *announcement put option* hypotheses. Section 6 reports empirical results from MCMC estimation of the endogenous switching model. For comparison, Section 7 discusses the Heckman–Lee two-stage estimation approach and reports maximum likelihood estimates of the Acharya (1988) latent information model. Section 8 concludes.

2. Naked vs. underwritten calls of convertible bonds

2.1. Calls of convertible bonds

Convertible call policy has long been a subject of interest in corporate finance. Ingersoll (1977a) and Brennan and Schwartz (1977) derive the optimal call policy for convertible bonds in a world where Modigliani and Miller (1958) capital structure irrelevance holds. They show that it is optimal for a firm to call a convertible bond as soon as its conversion value is equal to its call price. The empirical implications of these early theoretical papers were almost immediately refuted by Ingersoll (1977b) and Mikkelson (1981). Two empirical anomalies emerged. First, firms delay calls of convertible debt. Ingersoll (1977b) reports the conversion value exceeds call price by 43.9%, on average, at the time of call announcement. Second, the stock price of firms making conversion-forcing calls declines, on average, upon announcement. Mikkelson (1981) reports significant negative abnormal returns, on average, in response to call announcements.

Several papers attempt to explain, either theoretically or empirically, the early empirical anomalies. Harris and Raviv (1985) (hereafter HR) derive a sequential signaling model that attempts to simultaneously explain the late call and negative market response anomalies. In the HR model, managers observe private information about the firm's prospects. If the firm's

[3] I do not model the firm's decision to call a convertible bond, and have no evidence to offer regarding the information content of the call decision itself. Acharya (1988) examines the information content of convertible bond call announcements.

prospects are poor, managers will force conversion of in-the-money convertible bonds and thereby "share the woe" with new shareholders. Delaying a call of an in-the-money convertible bond is a positive signal, while calling a convertible bond is a negative signal. Ingersoll (1977b) suggests that the late call phenomenon is attributable to the existence of a call notice period. If firms are averse to raising the necessary cash to redeem the bonds, they will delay the call until the conversion value exceeds the call price by a sufficient margin to virtually assure conversion. Jaffee and Shleifer (1990) explicitly model the call notice period.

The call notice period is an important institutional feature. At the time of a call announcement, the firm effectively writes (or shorts) a put option on the firm's common stock to the convertible bondholders for the duration of the call notice period. I refer to this as the "announcement put option". After the announcement of a call, the bondholder's net position essentially consists of a long position in the underlying stock plus the announcement put option. If the conversion value of the bond (conversion ratio × stock price) exceeds the effective call price (face value + call premium + accrued interest) at the end of the call notice period, bondholders will exercise the bond's conversion option. This is equivalent to letting the announcement put option expire. If the conversion value of the bond is less than the effective call price of the bond at the end of the call notice period, bondholders will redeem their bond for cash rather than convert it into common stock. In other words, they will exercise the announcement put option.

2.2. Naked vs. underwritten calls

The call announcement reveals an endogenous binary choice by the firm's management: the firm chooses whether or not to engage the services of a standby underwriter. A standby underwriting agreement assures that all bonds will be converted into common stock rather than redeemed for cash. In an underwritten call, the standby underwriter agrees to: (1) purchase convertible bonds from bondholders at a small premium over the effective call price, (2) convert the bonds to common shares at the conversion price, and (3) resell the new shares in the market. If the convertible bond is in-the-money (i.e., conversion value exceeds effective call price at the end of the call notice period), then bondholders should convert their bonds into common shares rather than redeem them for cash. In this case, the underwriter pockets the standby fee and no further action is required. If the convertible bond "busts" (i.e., conversion value is less than effective call price at the end of the call notice period), then bondholders should tender to the standby underwriter. The standby underwriter is then obligated to convert the bonds into common shares. In this case, the underwriter's compensation consists of the standby fee plus the profit (or loss) on the sale of the new shares to the public. Standby underwriting agreements typically provide for a take-up fee to be paid to the underwriter for each bond actually purchased and converted by the underwriter. The take-up fee is essentially a discount that allows the underwriter to convert at a price lower than the conversion price.

From the firm's perspective, a standby underwriting agreement is analogous to an insurance policy that protects the firm from the undesired redemption of convertible bonds. In a naked call, the firm bears the risk of its short position in the announcement put option. In an underwritten call, the firm essentially buys a put option from the underwriter to hedge its short position in the put option. The underwriting agreement shifts the risk from the firm to the underwriter. The theoretical value of the standby underwriting agreement is the value of the put option. Whether the call is underwritten or naked, we would expect a wealth transfer from shareholders to convertible bondholders equal to the value of the announcement put option. This is the basis for the *announcement put option* hypothesis.

Cowan et al. (2000) extend the HR sequential signaling model to include the naked vs. underwritten call choice. They conclude that the decision to employ a standby underwriter conveys negative private information about the firm to the market and predict that the market response to convertible bond calls will be more negative for underwritten calls than for naked calls. This is the basis for the *private information* hypothesis. Singh et al. (1991) (hereafter SCN) empirically examine the firm's decision to have calls of convertible debt underwritten by investment bankers. They use the familiar Fama et al. (1969) method to compute two-day event window cumulative abnormal returns (CARs) for a sample of firms that announce conversion-forcing calls of convertible bonds. They report that the mean CAR for the subsample of firms making underwritten calls is negative (−2.00%) and significant, while the mean CAR for the subsample of firms making naked calls is negative (−0.84%) but insignificant.

SCN also attempts to explain cross-sectional variation in the market reaction to call announcements. Following the traditional cross-sectional event study approach, SCN regresses event window CARs on a dummy variable (set to one for underwritten calls) and other variables thought to influence the firm's call decision. SCN reports that event window CARs are negatively related to the underwritten call dummy variable, negatively related to the fraction of the original issue outstanding, and positively related to the extent that the bond's conversion option is in-the-money. SCN concludes that "the results are consistent with the idea that the market interprets an underwritten call as a stronger negative signal than a naked call".

Traditional cross-sectional event studies may be misspecified if the event is endogenous (see Acharya, 1988, 1993; Eckbo et al., 1990; Maddala, 1996; Prabhala, 1997). The misspecification is related to the familiar omitted variables bias in OLS regressions. The omitted variable is the market's expectation (conditional on the choice) of the manager's private information. This conditional expectation is a non-linear function of the firm-specific variables that enter into the manager's choice function. If the econometrician omits the conditional expectation from the model, but includes firm-specific variables that enter the manager's choice function, then the coefficients on the firm-specific variables will be biased.[4] Acharya (1988, 1993) and Eckbo et al. (1990) advocate the use of conditional event studies in which the market's response to an event is conditioned on the manager's choice function.

3. A switching model with endogenous self-selection

I examine a simultaneous equations model with an endogenous binary choice and two continuous potential outcomes. The model is well-known in econometrics and statistics. In labor economics, the model is known as an endogenous switching model. In biostatistics, the model is used to evaluate clinical trials where the treatment prescribed (e.g., experimental treatment vs. placebo) is not randomly assigned but correlated with the condition of the subject (i.e., the treatment is endogenous). Maddala (1983, Sections 8.3 and 9.2) provides detailed discussion of the endogenous switching model.

In finance, Acharya (1988, 1993) adapts the endogenous switching model to analyze the stock market's response to corporate announcements that reveal endogenous binary choices. Acharya's "latent variable model for endogenous events" assumes that outcomes for both *event* and *nonevent* firms are observed. Although similar in spirit, the switching model with endogenous self-selection analyzed in this paper is more general and flexible than the Acharya model. Acharya advocates the

[4] Prabhala (1997) argues that the Fama et al. (1969) event study methodology is well-specified as a test for the existence of average information effects. Prabhala further argues that although the traditional cross-sectional event study is misspecified, coefficient estimates are biased toward zero (attenuated) and conclusions are valid.

use of the two-stage estimation approach pioneered by Heckman (1976, 1979) and Lee (1979). I discuss the Heckman–Lee two-stage estimation approach in Section 7. Dunbar (1995) and Li and McNally (1998) employ variants of the endogenous switching model in which decisions or outcomes are multi-dimensional.

In this section, I discuss the endogenous switching model in general terms. In Section 4, I extend the model to allow for multivariate t errors and discuss the Bayesian MCMC estimation approach. I discuss details specific to the study of naked and underwritten calls of convertible bonds in Section 5.

Consider a firm i faced with a binary choice $s_i \in \{0, 1\}$. The choice is a function of a latent continuous variable, $s_i^* \in \mathbb{R}$, which is observed by the manager, but is not observed by the market. The manager chooses $s_i = 1$ if $s_i^* \geq 0$, and chooses $s_i = 0$ if $s_i^* < 0$. I assume that the latent variable is a linear function of publicly available information,

$$s_i^* = w_i' \gamma + \upsilon_i, \tag{1}$$

where w_i is a vector of observable variables and υ_i is the manager's private information. Prior to the announcement of the firm's choice, the market's conditional expectation of the manager's private information is zero (i.e., $E(\upsilon_i|w_i) = 0$). The announcement of the firm's choice reveals information about the manager's private information

$$s_i = I[s_i^* \geq 0] = \begin{cases} 0 \text{ if } \upsilon_i < -w_i'\gamma \\ 1 \text{ if } \upsilon_i \geq -w_i'\gamma \end{cases}.$$

The market learns that $\upsilon_i \geq -w_i'\gamma$ if $s_i = 1$, and that $\upsilon_i < -w_i'\gamma$ if $s_i = 0$.

The outcome in a conditional event study is the announcement effect associated with a choice. I model two potential outcomes, z_{i0} and z_{i1}, for each event. One of the potential outcomes is observed and the other potential outcome is "counter-factual". For example, if a firm chooses $s_i = 1$, then z_{i1} is the observed outcome and z_{i0} is the "counter-factual" outcome. Let $y_i \in \mathbb{R}$ denote the observed cumulative abnormal return (CAR) associated with an event. I assume that the observed outcome is generated by a switching model:

$$y_i = \begin{cases} z_{i0} = x_{i0}'\beta_0 + \epsilon_{i0} \text{ if } s_i = 0 \\ z_{i1} = x_{i1}'\beta_1 + \epsilon_{i1} \text{ if } s_1 = 1 \end{cases}, \tag{2}$$

where x_{i0} and x_{i1} are covariate vectors of publicly available information linearly related to the outcome. For each event ($i = 1, \ldots, n$), the econometrician observes the firm's choice s_i, the outcome y_i, and covariate vectors w_i, x_{i0} and x_{i1}. In the Bayesian MCMC approach, the latent variable s_i^* and the unobserved potential outcome are simulated conditional on the model (i.e., likelihood function), the data, and the parameters. This "data augmentation" is discussed in detail in the next section.

Stacking Eqs. (1) and (2) into a simultaneous system delivers the switching model with endogenous self-selection. I further assume that

$$\begin{pmatrix} s_i^* \\ z_{i0} \\ z_{i1} \end{pmatrix} \sim \mathcal{N}_3 \left(\begin{pmatrix} w_i'\gamma \\ x_{i0}'\beta_0 \\ x_{i1}'\beta_1 \end{pmatrix}, \begin{pmatrix} \sigma_\upsilon^2 & \sigma_{\upsilon 0} & \sigma_{\upsilon 1} \\ \sigma_{\upsilon 0} & \sigma_0^2 & \sigma_{01} \\ \sigma_{\upsilon 1} & \sigma_{01} & \sigma_1^2 \end{pmatrix} \right) \tag{3}$$

where w_i $(p \times 1)$, x_{i0} $(k_0 \times 1)$, and x_{i1} $(k_1 \times 1)$ are covariate vectors, γ, β_0, and β_1 are conforming parameter vectors, and $\mathcal{N}_3$ denotes a trivariate normal distribution. I can express Eq. (3) more compactly as

$$z_i \sim \mathcal{N}_3(X_i'\beta, \Sigma), \; i = 1, 2, \ldots, n, \tag{4}$$

where $z_i=(s_i^*\ z_{i0}\ z_{i1})'$ is a 3×1 vector, $\beta=(\gamma'\ \beta_0'\ \beta_1')'$ is a $k\times1$ ($k=p+k_0+k_1$) parameter vector, $X_i=\mathrm{diag}(w_i, x_{i0}, x_{i1})$ is a $k\times3$ matrix of covariates, and Σ is a 3×3 covariance vector. Let $\epsilon_i=(\upsilon_i\ \epsilon_{i0}\ \epsilon_{i1})'$ denote the 3×1 residual vector.

Economically, we are interested in the determinants of the firm's choice, and the determinants of the market's response to that choice. In the first line of Eq. (3), the parameter vector γ relates the latent information underlying the manager's choice (s_i^*) to a vector of publicly available covariates (w_i). The residual, υ_i, is the manager's private information. In the second and third lines of Eq. (3), the parameter vectors β_0 and β_1 relate the potential outcomes z_{i0} and z_{i1} to the covariate vectors x_{i0} and x_{i1}. Note that the covariate vectors x_{i0} and x_{i1} could, but do not need to contain the same elements. The parameter vector $\sigma=(\sigma_{\upsilon0}, \sigma_0^2, \sigma_{\upsilon1}, \sigma_1^2)$ consists of the four unique identifiable parameters in the covariance matrix Σ. σ_0^2 and σ_1^2 are the variances of the potential outcomes conditional on $s_i=0$ and $s_i=1$, respectively. The covariance terms $\sigma_{\upsilon0}$ and $\sigma_{\upsilon1}$ are of particular interest. These covariance terms permit a test of the hypothesis that the private information (υ_i) partially revealed by the firm's choice is correlated with the market's response. σ_υ^2 is set to one because the scale of s_i^* is indeterminate. σ_{01} is arbitrarily set to zero because it does not appear in the likelihood function.[5]

An extension of the model in Eq. (4) in which the error vector ϵ_i is distributed multivariate t is discussed in the following section.

4. Estimation by Bayesian MCMC methods

In this section, I describe a flexible Bayesian approach for estimating the endogenous switching model with multivariate t errors. My goal is to make inferences about the parameters of interest, $\beta=(\gamma, \beta_0, \beta_1)$ and $\sigma=(\sigma_{\upsilon0}, \sigma_0^2, \sigma_{\upsilon1}, \sigma_1^2)$, based on observations from n events. Bayesian analysis requires three elements: data, a likelihood function (i.e., sampling distribution), and prior densities on the parameters. Let $\pi(\beta, \sigma)=\pi(\beta)\pi(\sigma)$ be the prior density of the parameters and $f(y, s|\beta, \sigma)$ be the likelihood function for the model. By Bayes rule, the joint posterior density of the parameters is

$$\pi(\beta,\sigma|y,s)=\frac{\pi(\beta)\pi(\sigma)f(y,s|\beta,\sigma)}{\int\int\pi(\beta)\pi(\sigma)f(y,s|\beta,\sigma)\ d\beta\ d\sigma}\propto\pi(\beta)\pi(\sigma)f(y,s|\beta,\sigma). \tag{5}$$

$\pi(\beta, \sigma|y, s)$ cannot be summarized by analytical means due to the complexity of the likelihood function and restrictions on the parameters. This joint posterior density is, however, amenable to analysis by simulation-based methods, in particular those based on Markov chain Monte Carlo (MCMC) methods. $\pi(\beta, \sigma|y, s)$ is the *target distribution*. In MCMC methods, the econometrician specifies a Markov chain with a limiting invariant distribution identical to the target distribution (see Chib and Greenberg, 1996). This Markov chain is iterated a large number of times by Monte Carlo simulation. After a transient (or burn-in) phase, the output from the MCMC simulation represents a sample from the target distribution. The generated sample can then be used to make inferences about the posterior densities of the parameters of interest.

[5] The likelihood function for the endogenous switching model in Eq. (4) is

$$\mathcal{L}(\beta,\sigma)=\Pi\left[\int_{-\infty}^{-w_i'\gamma}f(z_{i0}-x_{i0}'\beta_0,\upsilon_i)\mathrm{d}\upsilon_i\right]^{1-s_i}\left[\int_{-w_i'\gamma}^{\infty}g(z_{i1}-x_{i1}'\beta_1,\upsilon_i)\mathrm{d}\upsilon_i\right]^{s_i}$$

where f and g are, respectively, the bivariate normal density functions of $(\upsilon_i,\epsilon_{i0})$ and $(\upsilon_i,\epsilon_{i1})$.

The endogenous switching model includes two latent (or unobserved) variables: the information underlying the manager's choice (s_i^*) and the counter-factual outcome. Including these latent variables in the model permits detailed analysis of the relation between the choice and the outcome. I employ a technique known as data augmentation to simulate these latent variables.[6] Let z_i^* denote the two latent elements of $z_i=(s_i^*\ z_{i0}\ z_{i1})'$. For example, if $s_i=1$, then the observed outcome would be $y_i=z_{i1}$, and z_i^* would be composed of the latent variable s_i^* and the counter-factual outcome z_{i0}. Let $\{z_i^*\}=(z_1^*, \ldots, z_n^*)$. In data augmentation, these latent variables are essentially added to the parameter space and simulated conditional on the model (i.e., likelihood function), the data, and the parameters. Inference regarding the relation between the choice and the outcome is accomplished by analyzing the posterior densities of the covariance terms $\sigma_{\upsilon 0}$ and $\sigma_{\upsilon 1}$ in Eq. (3), and of the "treatment effect" $z_{i1}-z_{i0}$.

4.1. Endogenous switching model with multivariate t errors

It is well-documented (e.g., Brown and Warner, 1985) that stock returns, especially daily stock returns, exhibit excess kurtosis. Thus, it is possible (perhaps likely) that the dependent variables z_{i0} and z_{i1} in Eq. (3) are fat-tailed. This is problematic since estimation of the model in Eq. (3) is known to be very sensitive to the Gaussian assumption. Chib and Hamilton (2000) extend the model such that $z_i=(s_i^*\ z_{i0}\ z_{i1})'$ follows a multivariate t distribution with location vector $X_i'\beta$, scale matrix Σ, and degrees of freedom v. I employ this endogenous switching model with multivariate t errors in my empirical work. Following Albert and Chib (1993), I represent the multivariate t distribution as a scale mixture of Gaussian distributions

$$z_i \sim \mathcal{N}_3(X_i'\beta, \lambda_i^{-1}\Sigma),\ i = 1, 2, \ldots, n, \tag{6}$$

where the mixing variables, $\{\lambda_i\}=(\lambda_1, \ldots, \lambda_n)$, are independent random variables from a gamma distribution $\mathcal{G}\left(\frac{v}{2}, \frac{v}{2}\right)$ with density function $\pi(\lambda_i) \propto \lambda_i^{\frac{v}{2}-1} e^{\frac{-v\lambda_i}{2}}$. To estimate the model with a multivariate t distribution, we further augment the parameter space with the mixing variables $\{\lambda_i\}$ that appear in Eq. (6). The likelihood function for the endogenous switching model in Eq. (6) is discussed in the Appendix.

After augmenting the parameter space with $\{z_i^*\}$ and $\{\lambda_i\}$, the augmented joint posterior density is

$$\pi(\beta, \sigma, \{z_i^*\}, \{\lambda_i\}|y, s) \propto \pi(\beta)\pi(\sigma)\pi(\{z_i^*\})\pi(\{\lambda_i\})f(y, s|\beta, \sigma, \{z_i^*\}, \{\lambda_i\}). \tag{7}$$

$\pi(\beta, \sigma, \{z_i^*\}, \{\lambda_i\}|y, s)$ is the target density for the MCMC simulations that follow.

4.2. Sampling strategy

I employ the Gibbs sampler algorithm to sample from the augmented joint posterior density in Eq. (7). The Gibbs sampler essentially cycles through the parameters (and augmented data), sampling from each full conditional density in turn (see Casella and George, 1992). The full conditional densities are detailed in the Appendix. The Gibbs sampler steps are summarized below:

(1) Initialize β, Σ, $\{\lambda_i\}$, and $\{z_i^*\}$. Set $\beta=\hat{\beta}_{\text{OLS}}$, $\Sigma=I_3$, $\lambda_i=1$ for all i, and $\{z_i^*\}$ to OLS fitted values.
(2) Sample β from $\pi(\beta|y, s, \Sigma, \{z_i^*\}, \{\lambda_i\})$.

[6] Data augmentation is introduced in Tanner and Wong (1987). Chib (1992) and Albert and Chib (1993) employ MCMC with data augmentation to estimate Tobit, probit and logit models with latent variables.

(3) Sample σ from $\pi(\sigma|y, s, \beta, \{z_i^*\}, \{\lambda_i\})$ using the Metropolis–Hastings algorithm.
(4) Sample λ_i independently from $\pi(\lambda_i|y, s, \beta, \Sigma, \{z_i^*\})$ for $i=1, \ldots, n$.
(5) Sample $\{z_i^*\}$ independently from $\pi(\{z_i^*\}|y, s, \beta, \Sigma, \{\lambda_i\})$.
(6) Repeat steps 2–5 using the most recent values of the conditioning variables.

Since the MCMC algorithm samples each parameter conditional on the data and the other parameters, the posterior densities obtained from the Gibbs sampler implicitly account for parameter uncertainty. Thus, sample statistics summarizing the posterior densities are free of errors-in-variables biases.

4.3. Model comparison

In the empirical work reported in Section 6, I estimate several variants of the model. I compare the relative fit of these models using Bayes factors. In Bayesian analysis, Bayes factors provide a unified way to compare the relative support that the data provide for competing model specifications (see Kass and Raftery, 1995). Let $m(y,s|\mathcal{M}_j)$ denote the marginal likelihood conditional on model $\mathcal{M}_j$. The Bayes factor comparing $\mathcal{M}_i$ to $\mathcal{M}_j$ is defined as

$$B_{ij} = \frac{m(y,s|\mathcal{M}_i)}{m(y,s|\mathcal{M}_j)}.$$

I estimate log marginal likelihoods following the method of Chib (1995) and Chib and Hamilton (2000). Following Kass and Raftery (1995), I evaluate the significance of a Bayes factor on a base 10 logarithmic scale:

$\log_{10}(B_{ij})$	Evidence against $\mathcal{M}_j$
0 to $\frac{1}{2}$	Not worth more than a bare mention
$\frac{1}{2}$ to 1	Substantial
1 to 2	Strong
>2	Decisive

5. Data

5.1. Sample construction

I used the Securities Data Corporation (SDC) New Issues Database to identify an initial list of 1155 convertible debt offerings between 1970 and 1998. The list includes both senior and subordinated bonds, notes and debentures. Hereafter, I use the generic term "convertible bond" to describe all of these issues. Each bond is callable, has a fixed coupon rate, and is convertible into the common stock of a U.S company listed on the NYSE, AMEX, or NASDAQ. Using this list of convertible bonds as a starting point, I searched for information on the disposition of each bond in Dow Jones News Retrieval (for Wall Street Journal, Dow Jones Newswire, and Business Wire articles), Moody's Bond Manual, and the SDC Underwritten Call Database. Bonds were not included in the sample if the nature of the firm's common stock changed due to merger or reorganization, if the bond was outstanding as of January 1999, or if the ultimate disposition of the bond could not be determined. Called bonds were dropped from the sample if the reason for the call was bankruptcy or merger, if there was a confounding news announcement regarding the company within ±2 days of the call announcement, or if data wasn't available from CRSP or Compustat. This process resulted in sample of 274 "uncontaminated" calls. The sample was further reduced by dropping calls that were not conversion-forcing (38 calls) or calls with missing data (7 calls). The final data sample contains 229 events. Each call is classified as either naked (144 calls) or underwritten (85 calls) using information available in news accounts or the SDC Underwritten Call Database.

For each convertible bond, the issue date, coupon rate, maturity date, CUSIP, initial conversion ratio, call protection code, and call price schedule (when available) were obtained from the SDC New Issues Database. Stock prices, returns, shares outstanding, and share adjustment factors for stock splits and stock dividends for the period from the issue date to the redemption date were obtained from the CRSP daily stock files. The effective conversion ratio at the time of the call announcement was computed by adjusting the initial conversion ratio to reflect stock splits and stock dividends. The daily conversion value for each bond is the product of the closing stock price and the effective conversion ratio. The effective call price for each event is the call price plus accrued interest to the redemption date. In most cases, the effective call price and effective conversion ratio are explicitly stated in the call announcement.

5.2. Outcome model variables

I define the observed event outcome, y_i, as the cumulative abnormal return (CAR) for the three-day window centered on the event date ($-1 \leq t \leq 1$). I choose a three-day event window because of the nature of the announcement sources. In some cases, the announcement date ($t=0$) is the date of a Wall Street Journal article. For these events, the announcement effect probably occurred on either event days $t=-1$ or $t=0$. In other cases, the announcement date is the date of a news wire article. In some of these cases, the time on the news wire article is after the close of trading on day $t=0$. For these events, the announcement effect may have occurred on either event day $t=0$ or $t=1$. Therefore, given the nature of the sample, I believe a three-day event window is more likely to capture the announcement effect. I define abnormal returns as the daily return on the stock less the daily return on the CRSP equally-weighted index of stocks trading on the NYSE, AMEX, and NAS-DAQ. I use market-adjusted returns rather than market model residuals for two reasons. First, Campbell et al. (1991) show that market model residuals are biased estimates of abnormal returns associated with calls of convertible debt. Since conversion-forcing calls of convertible bonds typically follow a period of abnormal stock price appreciation, market model parameters estimated from a pre-event estimation window are biased and the resulting event window market model residuals are likewise biased. My second reason for using market-adjusted returns is that there is little or nothing to be gained by using more complicated market-adjusted returns (see Brown and Warner, 1985).

Table 1 reports summary statistics for the sample. The mean CAR for the full sample is −1.581%, while the mean CARs for the underwritten and naked call subsamples are −2.415% and −1.085%, respectively. The mean CARs are all significantly negative. Furthermore, the mean CAR for underwritten calls is significantly less than the mean CAR for naked calls ($t=2.78$). The market's more negative response to underwritten calls is consistent with results reported in SCN. A Shapiro–Wilk test ($W=0.9879$, $p=0.0495$) rejects the hypothesis that the CARs for the full sample are drawn from a normal distribution.

The *announcement put option* hypothesis predicts that the market's response to calls of convertible bonds is related to the value of the announcement put option. I define the variable *Put Value* as the imputed value of the announcement put option scaled by the market value of equity. I

Table 1
Description of data

Variable	Sample	Mean	Median	SD	Min.	Max.
CAR (%)	All	−1.581	−1.754	3.775	−12.446	10.638
	Naked	−1.085	−1.225	3.980	−12.446	10.638
	Underwritten	−2.415	−2.655	3.254	−9.746	10.555
Cushion	All	0.410	0.304	0.417	0.013	3.974
	Naked	0.478	0.359	0.499	0.013	3.974
	Underwritten	0.293	0.251	0.164	0.052	0.846
Volatility (%)	All	2.258	2.137	0.940	0.871	10.023
	Naked	2.259	1.989	1.041	0.871	10.023
	Underwritten	2.258	2.192	0.744	0.889	4.372
Fraction Out	All	0.930	1.000	0.171	0.008	1.000
	Naked	0.918	1.000	0.196	0.008	1.000
	Underwritten	0.950	1.000	0.116	0.375	1.000
Cash Stress	All	0.206	0.163	1.656	−5.698	7.018
	Naked	0.110	0.075	1.593	−5.698	5.181
	Underwritten	0.369	0.245	1.755	−3.875	7.018
Put Value (%)	All	0.022	0.000	0.064	0.000	0.556
	Naked	0.023	0.000	0.068	0.000	0.556
	Underwritten	0.021	0.001	0.058	0.000	0.439
Run Up (%)	All	17.48	16.50	16.90	−26.32	74.82
	Naked	15.16	14.40	17.12	−26.32	68.89
	Underwritten	21.42	18.95	15.86	−10.65	74.82

CAR is the three-day event window market-adjusted cumulative abnormal return (CAR) in percent. *Cushion* is the ratio of conversion value less effective call price to effective call price on event date $t=-1$. *Volatility* is the standard deviation of daily stock returns for the period $-70 \le t \le -11$ in percent. *Fraction Out* is the ratio of the face value outstanding at the time of call to the initial face value issued. *Cash Stress* is the natural log of the ratio of the face value outstanding at the time of call to the firm's cash and marketable securities. *Put Value* is the theoretical put value divided by the market value of equity on event date $t=-1$ in percent. *Run Up* is the total log return for the period $-70 \le t \le -11$ in percent. Summary statistics are reported for the total sample (229 calls), the naked subsample (144 calls) and the underwritten subsample (85 calls).

compute the imputed put option value for each bond using the standard Black–Scholes pricing equation for a European put option on a non-dividend paying stock. I use the conversion value of the bond (stock price × conversion ratio) on event date $t=-1$ in place of stock price. I use the effective call price (face value + call premium + accrued interest) on event date $t=-1$ in place of strike price. For annual volatility, I multiply daily volatility (measured over the period $-70 \le t \le -11$) by $\sqrt{252}$. Time to expiration (in years) is the number of trading days from the call announcement to the expiration of the bondholder's conversion option divided by 252 trading days per year. In most cases, the bondholder's conversion option expires on the redemption date. However, the bondholder's conversion option expires prior to the redemption date in a significant number of cases. These dates are typically reported in the call announcement. For the risk-free interest rate, I use the three-month Treasury bill rate (secondary market) provided by the Federal Reserve Board of Governors. *Put Value* is computed by multiplying the approximate put value per bond by the number of outstanding bonds, and dividing that amount by the market value of equity on event date $t=-1$. *Put Value* is expressed in percent.

I include *Put Value* as a covariate in both potential outcome equations. The *announcement put option* hypothesis predicts that the coefficients on *Put Value* (β_{01} and β_{11}) will be negative. For the conversion-forcing calls of convertible debt that are the focus of this study, the announcement put options are typically well out of the money. Table 1 indicates that, on average, the imputed value of the announcement put option is only 0.022% of the market value of shareholder equity.

The means for naked and underwritten calls are 0.023% and 0.021%, respectively. Given that the mean three-day abnormal returns for these events are −1.085% and −2.415%, respectively, it is unlikely that the small wealth transfers associated with announcement put options will explain the market's negative response to calls of convertible debt.

5.3. Choice model variables

The covariate vector w_i for the choice model in Eq. (1) consists of a constant and six variables related to the firm's underwriting choice. The choice covariates are:

(1) ***Cushion***. *Cushion* is a measure of the extent that the bondholder's conversion option is in-the-money. I compute *Cushion* as the ratio of conversion value (effective conversion ratio × closing stock price on event date $t=-1$) to the effective call price (face value + call premium + accrued interest), less one. Given the sample selection criteria that all calls are conversion-forcing, this variable is bounded from below by zero. Since a large cushion is effectively a substitute for a standby underwriting agreement, I expect that the likelihood that a firm chooses to have a call underwritten is inversely related to *Cushion* (i.e., $\gamma_1 < 0$). The theoretical model of Cowan et al. (2000) formally makes this prediction. Mean *Cushion* is 0.410 for all calls. Mean *Cushion* is higher (0.478) for naked calls than for underwritten calls (0.293).

(2) ***Volatility***. As a measure of stock volatility, I use the standard deviation (in percent) of daily stock returns from the period $-70 \leq t \leq -11$. The probability that a bond's conversion value will fall below the effective call price during the call notice period (resulting in redemption rather than conversion) increases with volatility. I expect that the likelihood of a firm having a call underwritten is positively related to *Volatility* (i.e., $\gamma_2 > 0$). Mean volatility reported in Table 1 is 2.258% per day for the whole sample. There is virtually no difference between the means of the naked and underwritten call subsamples, although the median volatility is slightly higher for the underwritten call subsample.

(3) ***Fraction Out***. *Fraction Out* is the ratio of the face value of outstanding bonds at the time of the call to the face value of the bonds at issue. Data on the face value of outstanding bonds at the time of call was collected from news accounts and from the SDC Underwritten Call database. When this information was unavailable, *Fraction Out* was set to one. Data on the face value of bonds at issue reported in the SDC New Issues database is collected from SEC filings. In many cases, the actual amount issued exceeds the amount reported by SDC due to exercise of overallotment options by the original underwriter. If face value at time of call exceeds face value at time of issue, *Fraction Out* was set to one. If *Fraction Out* is small, it is possible that the call is motivated by a desire to reduce the administrative costs of the bond. It is unlikely that the firm would have such a "cleanup" call underwritten. I expect that the likelihood of an underwritten call is positively related to *Fraction Out* (i.e., $\gamma_3 > 0$). The mean (median) *Fraction Out* for the whole sample is 0.930 (1.000). Mean *Fraction Out* is only slightly higher for underwritten calls (0.950) than for naked calls (0.918).

(4) ***Cash Stress***. Jaffee and Shleifer (1990) suggest that firms in or near financial distress are more likely to demand insurance against redemption (rather than conversion) of a called bond. This demand arises from the difficulty distressed firms are likely to face in raising the cash necessary to finance a redemption. *Cash Stress* is a variable designed to capture the stress that an unexpected redemption would place on a firm's cash reserves. *Cash Stress* is

the natural log of the ratio of the face value of outstanding bonds at the time of call to the firm's cash and marketable securities.[7] I expect that the likelihood of an underwritten call is positively related to *Cash Stress* (i.e., $\gamma_4 > 0$).

(5) ***Put Value***. I also include *Put Value*, defined above, in the manager's choice model. In a world with perfect markets, no information asymmetries, and risk-neutral managers and underwriters, the theoretical value of a standby underwriting agreement is the value of an equivalent put option.[8] If these assumptions are violated, the value of the put option represents a reasonable lower bound on the fee for a standby underwriting agreement. To the extent that managerial interests are aligned with stockholder interests, the likelihood of having a convertible call underwritten should be negatively related to *Put Value* (i.e., $\gamma_5 < 0$). It is important to note that *Put Value* is negatively related to *Cushion* and positively related to *Volatility* through the Black–Scholes option pricing model. For example, I predict above that the likelihood of having a call underwritten is positively related to *Volatility.* However, that positive relation should be somewhat mitigated by higher standby underwriting fees required when the underlying stock has high volatility. Including *Put Value* as an independent variable in the choice equation should control for this potentially confounding effect.

(6) ***Run Up***. *Run Up* is the total log return (in percent) for the period $-70 \le t \le -11$. Mean *Run Up* is 17.48% for the full sample. The underwritten subsample has a higher mean (21.42%) than the naked subsample (15.16%). The magnitude of pre-call stock returns is not surprising given that: (1) convertible bonds are issued at a substantial discount, and (2) the sample selection criteria requires that calls be conversion-forcing. A conversion-forcing call is more likely to be underwritten if managers believe that the firm is overvalued, and if a correction is foreseen during the call notice period. If *Run Up* is correlated with overvaluation by the market, then the likelihood of a conversion-forcing call being underwritten should be positively related to *Run Up* (i.e., $\gamma_6 > 0$).

6. Empirical results: Bayesian MCMC approach

I report empirical results for three variants of the endogenous switching model presented in Sections 3 and 4:

$\mathcal{M}_1$ Assumes that z_i is distributed multivariate t with $\nu = 6$ degrees of freedom, and imposes the restriction $\sigma_{\upsilon 0} = \sigma_{\upsilon 1}$.

$\mathcal{M}_2$ Same as $\mathcal{M}_1$, except that $\sigma_{\upsilon 0}$ and $\sigma_{\upsilon 1}$ are estimated separately.

$\mathcal{M}_3$ Same as $\mathcal{M}_1$, except that z_i is assumed to be distributed approximately multivariate normal.

Model comparisons based on Bayes factors indicate that $\mathcal{M}_1$ is decisively favored by the data. The Bayes factors comparing $\mathcal{M}_1$ to $\mathcal{M}_2$ and $\mathcal{M}_3$, respectively, are $\log_{10}(B_{12}) = 4.07$ and $\log_{10}(B_{13}) = 7.99$. Accordingly, I allocate most of my discussion to $\mathcal{M}_1$. Discussion of $\mathcal{M}_2$ and $\mathcal{M}_3$ will focus on how specification changes affect estimation of the model.

[7] Cash and marketable securities is Item 1 from the Compustat annual files. If the call announcement date is less than three months after the end of the fiscal year, I use annual data from the previous fiscal year.

[8] See Butler (1999) and Cowan et al. (1999) for more detailed analysis of the determinants of fees for underwritten calls of convertible bonds. These papers suggest that standby fees for underwriting a convertible call far exceed imputed put option values.

6.1. Priors

I employ relatively uninformative priors in the analysis that follows. The form of the priors is discussed in detail in the Appendix. My prior on $\beta=(\gamma' \ \beta_0' \ \beta_1')'$ is distributed normal with $k\times 1$ mean vector $b_0=0$ and covariance matrix $B_0=100\cdot I_k$, where I_k is a $k\times k$ identity matrix. This implies that the prior density for each element of β is normal with mean zero and a standard deviation of ten. My priors for the variance terms σ_0^2 and σ_1^2 are normal (truncated to insure non-negativity) with a mean of 9 and a standard deviation of 10. The priors for the covariance terms $\sigma_{\upsilon 0}$ and $\sigma_{\upsilon 1}$ are normal with a mean of 0 and a standard deviation of 10. Since my priors on β and σ are relatively uninformative (i.e., imprecise), the posterior densities estimated below should be primarily determined by the data.

6.2. Bayesian inference regarding model $\mathcal{M}_1$

Model $\mathcal{M}_1$ is the endogenous switching model in Eq. (6) with two further assumptions: (1) the vector z_i is distributed multivariate t with $\nu=6$ degrees of freedom, and (2) the restriction $\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$ is imposed. Of all specifications estimated, $\mathcal{M}_1$ is the model most favored by the data.[9] The Gibbs sampler was run for 10,000 iterations after an initial burn-in period of 1000 iterations. Table 2 summarizes the posterior densities of β and σ. The mean of the posterior density (i.e., the posterior mean) is analogous to the point estimate in frequentist methods. I evaluate the precision of the estimates in two ways. The standard deviation (SD) of the MCMC draws from the posterior density is analogous to the standard error of the estimate. The 5% and 95% quantile values are the endpoints of a 90% Bayesian confidence interval.

6.2.1. Choice model parameter estimates

Panel A of Table 2 reports summary statistics describing the posterior densities of parameters for the binary choice model embedded in Eq. (6). As predicted, the decision to underwrite calls of convertible debt is negatively related to *Cushion*. The posterior mean (median) of γ_1 is -2.050 (-2.021). The 90% Bayesian confidence interval is $[-3.066, -1.145]$, indicating that 95% of the parameter draws are less than -1.145. Since zero falls outside of the Bayesian confidence interval, I conclude that the negative relation between *Cushion* and the propensity to underwrite a call is statistically reliable. This result suggests that delaying the call of a convertible bond until conversion value substantially exceeds effective call price acts as a substitute for having the call underwritten by an investment bank.

The decision to underwrite calls is positively related to *Run Up*. The posterior mean (median) of γ_6 is 0.018 (0.018), and the 90% Bayesian confidence interval is [0.007, 0.030]. All else equal, firms with higher than average stock returns in the pre-call period are more likely to have calls underwritten. Posterior means of the coefficients for *Volatility*, *Fraction Out*, *Cash Stress* and *Put Value* all have the predicted signs. However, the Bayesian 90% Bayesian confidence intervals for γ_2, γ_3, γ_4, and γ_5 all bracket zero. Therefore, I can't claim that these relations are statistically reliable.

Are these results driven by the prior or by the sample data? Given a large enough sample, Bayesian estimates of parameters should be relatively insensitive to my relatively uninformative priors. This is confirmed by examining the prior and posterior densities of the parameters. Fig. 1

[9] The model was estimated using many values for ν. Based on the marginal likelihood criterion for model selection (see Chib, 1995), $\nu=6$ provides the best fit of the data. This is not surprising given the high degree of leptokurtosis exhibited by daily stock return data.

Table 2
Bayesian posterior densities of model parameters: model $\mathcal{M}_1$

Parameter (variable)	Mean	Median	SD	5%	95%	ρ_1
Panel A: $s_i^* = w_i'\gamma + \upsilon_i$ (choice equation)						
γ_0 (Constant)	−0.610	−0.590	0.685	−1.761	0.474	0.677
γ_1 (Cushion)	−2.050	−2.021	0.582	−3.066	−1.145	0.893
γ_2 (Volatility)	0.078	0.079	0.145	−0.159	0.316	0.638
γ_3 (Fraction Out)	0.561	0.538	0.661	−0.481	1.681	0.675
γ_4 (Cash Stress)	0.053	0.053	0.061	−0.048	0.152	0.574
γ_5 (Put Value)	−3.743	−3.560	2.525	−8.226	0.069	0.748
γ_6 (Run Up)	0.018	0.018	0.007	0.007	0.030	0.634
Panel B: $z_{i0} = x_{i0}'\beta_0 + \epsilon_{i0}$ (outcome for naked calls)						
β_{01} (Put Value)	−2.557	−2.604	4.207	−9.426	4.387	0.342
Panel C: $z_{i1} = x_{i1}'\beta_1 + \epsilon_{i1}$ (outcome for underwritten calls)						
β_{11} (Put Value)	−7.232	−7.376	5.570	−16.040	2.160	0.770
Panel D: free covariance matrix elements						
$\sigma_{\upsilon 0} = \sigma_{\upsilon 1}$	−0.607	−0.614	0.282	−1.071	−0.145	0.395
σ_0^2	12.395	12.244	1.801	9.717	15.578	0.492
σ_1^2	7.264	7.119	1.402	5.249	9.824	0.526

This table reports summary statistics describing the Bayesian posterior densities of parameters in the endogenous switching model assuming that residuals follow a multivariate t distribution with $v = 6$ degrees of freedom. The sample contains 229 calls of convertible bonds (85 underwritten and 144 naked). The model was estimated using Markov chain Monte Carlo (MCMC) techniques. The posterior mean, median, standard deviation, 5th and 95th percentile critical values, and 1st order autocorrelation for each parameter are based on 10,000 Gibbs sampler iterations from a suitably constructed Markov chain.

provides graphical summaries of the posterior densities for two coefficients from the binary choice model: γ_1 (*Cushion*) and γ_6 (*Run Up*). The top graphs plot 10,000 Gibbs sampler draws for each parameter. These plots indicate that the Gibbs sampler has converged and is behaving well after 1000 burn-in iterations. The middle graphs plot histograms describing the prior and posterior densities of each parameter. The posterior density histograms in Fig. 1 convey the same information as the summary statistics reported in Table 3: both γ_1 and γ_6 are reliably different than zero. For each parameter, zero falls well outside of posterior density's central mass. The middle graphs also plot slices of the prior densities over the range of the posterior densities. Note that the priors and posteriors are plotted on the same scale.[10] The priors for γ_1 and γ_6 are both normal with mean zero and a standard deviation of ten. Thus, a complete histogram for either prior would have a range of approximately [−30, 30] (i.e., ±three standard deviations). Since the priors are relatively uninformative, they appear nearly flat over the range of the posterior. For γ_6, the prior is so dispersed relative to the posterior that it is barely perceptible. It is evident from these histograms that the posterior densities are determined primarily by the data sample. The bottom graphs plot the 1st through 20th order autocorrelation coefficients for the Gibbs sampler draws of γ_1 and γ_6. Recall that Gibbs sampler draws are correlated, not independent. The correlogram is a diagnostic tool used to determine the degree of serial correlation between draws. If serial correlation is high, the Gibbs sampler may not be transiting the entire support of the density. The low degree of serial correlation indicated by the correlograms indicates that the

[10] The prior density histograms are constructed analytically by computing the cumulative distribution between the breakpoints for the posterior histogram. The resulting fractions are multiplied by 10,000 to simulate a Gibbs sampler run.

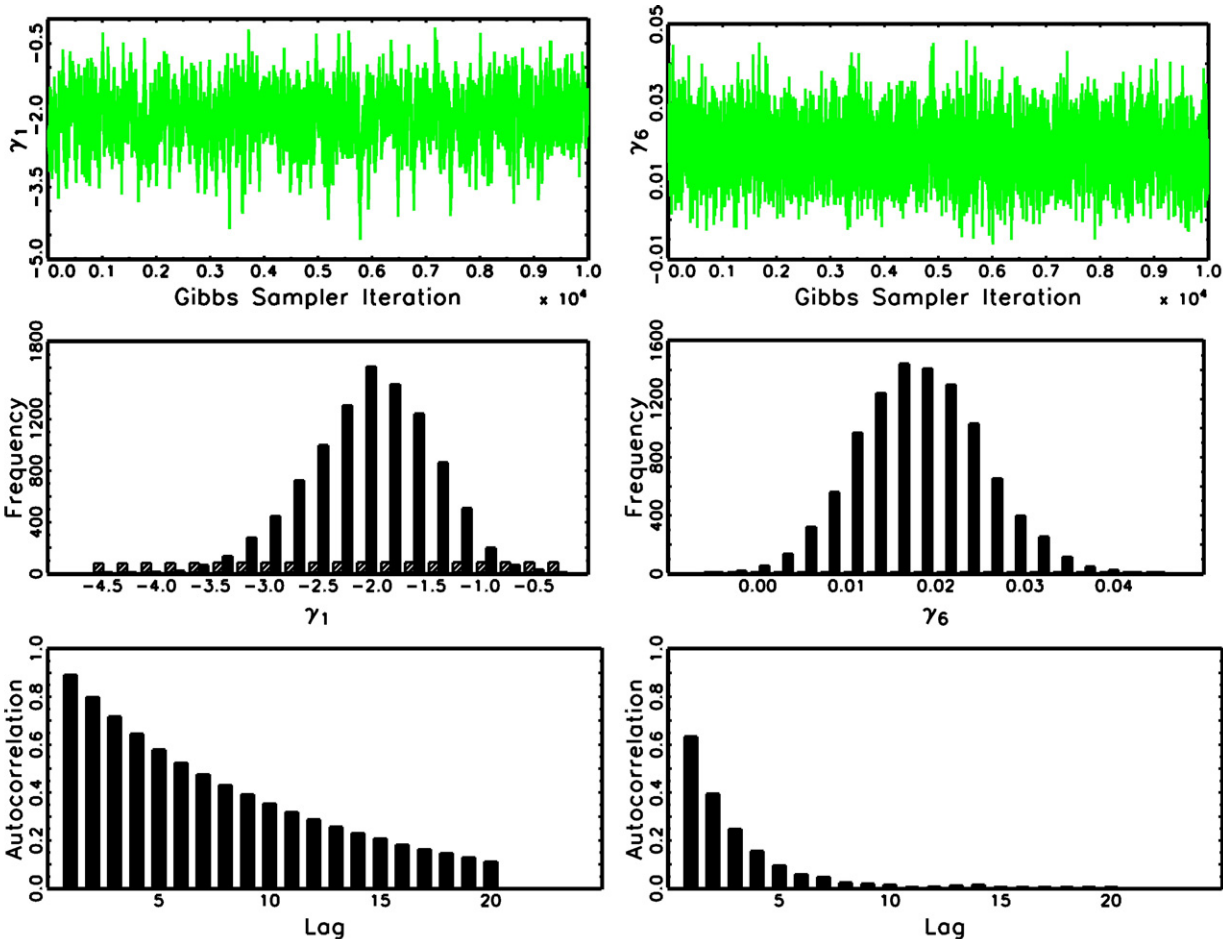

Fig. 1. Posterior densities of choice model parameters for model $\mathcal{M}_1$. This figure provides graphical summaries of the posterior densities of γ_1 (*Cushion*) and γ_6 (*Run Up*) for model $\mathcal{M}_1$ based on 10,000 iterations of the Gibbs sampler. Top: parameter draws from the Gibbs sampler for γ_1 (left) and γ_6 (right). Middle: prior and posterior Densities. Bottom: 1st through 20th order autocorrelations of parameter draws from the Gibbs sampler.

Gibbs sampler is behaving well (i.e., the sampler is regularly sampling from the entire support of the posterior densities). The correlograms are consistent with the plots at the top of Fig. 1.

6.2.2. Potential outcome model parameter estimates

The posterior densities of the coefficients from the potential outcome equations associated with naked and underwritten calls of convertible debt are summarized in Panels B and C, respectively, of Table 2. Since I am interested in explaining the cross-sectional response to announcements of call, I estimate the model with demeaned outcomes (i.e., I subtract the unconditional mean (−1.581%) from each event window CAR) and omit the intercept terms from the potential outcome equations.[11]

The *announcement put option* hypothesis predicts that the market's response to a call announcement is related to the value of the announcement put option, and predicts that β_{01} and β_{11} are negative. The evidence reported in Panels B and C of Table 2 does not support this hypothesis. For naked calls, the posterior mean (median) of β_{01} is −2.557 (−2.604), and the 90% Bayesian confidence interval is [−9.4264, 4.387]. For underwritten calls, the posterior mean (median) of β_{11} is −7.232 (−7.376), and the 90% Bayesian confidence interval is [−16.040, 2.160]. That the magnitudes of these coefficients exceed one or that $\beta_{11} > \beta_{01}$ is not surprising given the results of Butler (1999) and Cowan et al. (1999). Those papers report that standby underwriting fees typically far exceed their theoretical put option values. This is also true for the sample in this paper. Standby underwriting fees were available for 66 of the 85 underwritten calls in the sample. For these 66 calls, the mean fee was 0.137% of market capitalization. Recall from Table 1 that the mean theoretical fee (i.e., *Put Value*) was only 0.021% of market capitalization. However, since zero falls inside of the Bayesian confidence intervals for β_{01} and β_{11}, I must conclude that cross-sectional variation in announcement put option values is insufficient to explain cross-sectional variation in the market's response to call announcements.

6.2.3. Covariance matrix parameter estimates

The posterior densities of the residual variances, σ_0^2 and σ_1^2, are statistically summarized in Panel D of Table 2. In addition, Fig. 2 plots histograms depicting the prior and posterior densities of σ_0^2 and σ_1^2. The posterior mean (median) of σ_0^2 is 12.395 (12.244). The posterior mean (median) of σ_1^2 is 7.264 (7.119). These results indicate that the variance of event window abnormal returns is somewhat higher for naked calls than for underwritten calls. This is consistent with the summary statistics reported in Table 1.

Next, I address one of the central questions examined in this paper: Is the private information partially revealed by the manager's choice of call method (naked vs. underwritten) correlated with the market's response? If the *private information* hypothesis is true, positive values of υ_i should be associated with negative market responses. I test this hypothesis by examining the posterior

[11] In the course of estimating the model, I found that including intercept terms in the potential outcome equations led to severe problems identifying the covariance terms $\sigma_{\upsilon 0}$ and $\sigma_{\upsilon 1}$. I investigated this problem by estimating the model on simulated data sets. I created 400 data sets following the procedure described in Prabhala (1997). I estimated the model for each simulated sample and retained the posterior means for each parameter. Estimates of the choice equation parameters and the variance terms (σ_0^2 and σ_1^2) were unbiased and very precise. However, estimates of the outcome equation intercepts and covariance terms ($\sigma_{\upsilon 0}$ and $\sigma_{\upsilon 1}$) were poor. The outcome equation intercepts reflected the conditional means of the outcome and the covariance terms were biased toward zero. Intuitively, the (conditional) intercept terms were capturing much of the cross-sectional variation in the outcomes and left little for the covariance terms to explain. When I re-estimated the model using demeaned outcomes and no intercept terms, the covariance terms were unbiased and very precise. Using this procedure appears to eliminate the identification problem and significantly improves the performance of the Gibbs sampler.

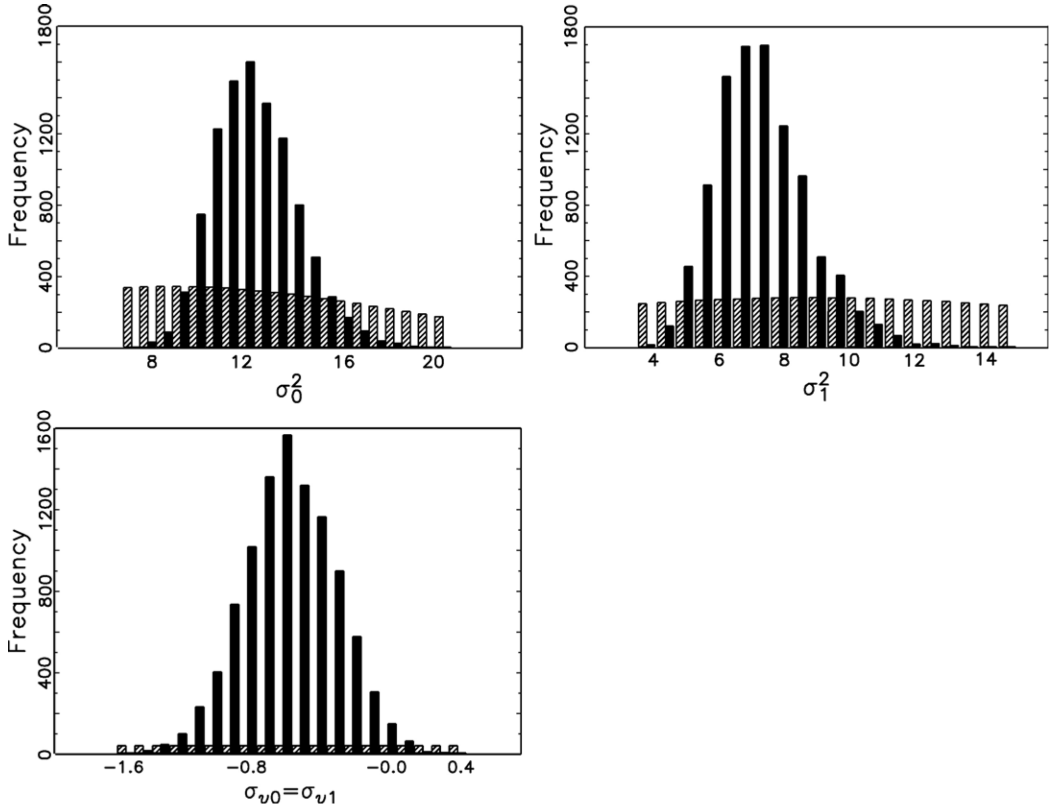

Fig. 2. Prior and posterior densities of covariance matrix parameters for model $\mathcal{M}_1$. Top: σ_0^2 and σ_1^2. Bottom: $\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$. Histograms are based on 10,000 iterations of the Gibbs sampler for the multivariate t model with $\nu=6$ degrees of freedom.

densities of $\sigma_{\upsilon 0}$ and $\sigma_{\upsilon 1}$. If the manager's choice of call method does not reveal private information correlated with the value of the firm, then I should observe $\sigma_{\upsilon 0}=\sigma_{\upsilon 1}=0$. If the decision to underwrite a call does reveal private information negatively correlated with stock price, then I should observe $\sigma_{\upsilon 0}<0$ and $\sigma_{\upsilon 1}<0$.

In Model $\mathcal{M}_1$, I impose the restriction $\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$. The posterior density of $\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$ is statistically summarized in Panel D of Table 2, and graphically summarized in Fig. 2. As predicted by the *private information* hypothesis, the market responds negatively to the private information partially revealed when a firm announces an underwritten conversion-forcing call. The posterior mean (median) of $\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$ is -0.607 (-0.614), and the 90% Bayesian confidence interval is $[-1.071, -0.145]$. These covariance terms are reliably negative. This suggests that the private information partially revealed by the firm's underwriting choice does help to explain the difference in event-window CARs between naked and underwritten calls.

6.2.4. Analysis of counter-factuals

What is the relation between the call type (naked vs. underwritten) and the event window CAR? Answering this question is problematic due to the endogeneity of the choice. Consistent with SCN, I find that the announcement effect is more negative, on average, for underwritten calls than for naked calls. Is the information content of an underwritten call more negative? Or, are firms that announce underwritten calls fundamentally different, on average, from firms that announce naked calls?

In order to answer these questions, the researcher must fully control for the endogeneity of the choice. Ideally, we would like to know how the market would have responded if a firm had made the alternative or counter-factual choice. For example, what would the CAR have been if a particular firm had announced a naked call instead of an underwritten call? The effect of the call type would be the difference between the observed outcome and the counter-factual outcome. The Bayesian MCMC approach provides a natural framework for this type of analysis.

The model considers two potential outcomes for each event: the CAR for a naked call (z_{i0}) and the CAR for an underwritten call (z_{i1}). Only one potential outcome is observable. For firms that announce naked calls (i.e., $s_i=0$), the observed outcome is $z_{i0}=y_i$ and the counter-factual outcome is z_{i1}. For firms that announce underwritten calls (i.e., $s_i=1$), the observed outcome is $z_{i1}=y_i$ and the counter-factual outcome is z_{i0}. In the Bayesian MCMC approach, the unobserved counter-factual outcome is simulated conditional on the model, the data, and the other parameters using a technique known as data augmentation.

Let T_i denote the "treatment effect" for the ith event. The treatment effect for a given event is the difference between the observed outcome and the counter-factual outcome, and is determined by the following switching model:

$$T_i = \begin{cases} z_{i1}-y_i & \text{if } s_i = 0 \\ y_i-z_{i0} & \text{if } s_i = 1 \end{cases}. \tag{8}$$

Samples from the posterior density of T_i are easily constructed from the MCMC output for each event.[12] These densities can be used to make inferences for a given event. I construct the

[12] Maddala (1983, Ch. 9) discusses frequentist estimation of treatment effects, but does not provide guidance regarding their distribution. Unable to identify an appropriate distribution for testing average treatment effects, Dunbar (1995) resorts to non-parametric binomial tests. In contrast, the Bayesian approach provides exact small-sample posterior densities for making inferences regarding treatment effects.

posterior density of the expected treatment effect (i.e., $E[z_{i1}-z_{i0}]$) by computing the cross-sectional mean of T_i for each iteration of the Gibbs sampler. I use this posterior density to make inferences regarding the expected treatment effect of an underwritten call.

On average, the expected treatment effect associated with the announcement of an underwritten call is negative, but not significant. The posterior mean (median) of $E[z_{i1}-z_{i0}]$ is −0.213 (−0.218), and the 90% Bayesian confidence interval is [−0.810, 0.405]. This means that, on average, the market is expected to respond more negatively to an underwritten conversion-forcing call than to a naked call by the same firm. However, the expected treatment effect is not reliably different from zero. I also analyze the average treatment effects conditional on call type. For underwritten calls, the posterior mean (median) of $E[z_{i1}-z_{i0}|s_i=1]$ is −0.119 (−0.122), and the 90% Bayesian confidence interval is [−1.067, 0.841]. Had these firms chosen naked rather than underwritten calls, the market response would have been less negative, but not significantly so. For naked calls, the posterior mean (median) of $E[z_{i1}-z_{i0}|s_i=0]$ is −0.268 (−0.273), and the 90% Bayesian confidence interval is [−0.861, 0.357]. Again, the conditional treatment effect is negative, but not significant.

After carefully controlling for the endogeneity of the naked vs. underwritten call choice, the data does not support the *private information* hypothesis. As noted previously, cross-sectional variation in event window abnormal returns is large. This contributes to the relatively imprecise estimates of average treatment effects, and to the finding that average treatment effects are not reliably different from zero.

Table 3
Bayesian posterior densities of model parameters: model $\mathcal{M}_2$

Parameter (variable)	Mean	Median	SD	5%	95%	ρ_1
Panel A: $s_i^* = w_i'\gamma + \upsilon_i$ (choice equation)						
γ_0 (Constant)	−0.618	−0.603	0.682	−1.769	0.472	0.686
γ_1 (Cushion)	−2.092	−2.059	0.615	−3.164	−1.134	0.906
γ_2 (Volatility)	0.083	0.084	0.148	−0.159	0.325	0.655
γ_3 (Fraction Out)	0.569	0.554	0.652	−0.465	1.670	0.682
γ_4 (Cash Stress)	0.058	0.058	0.061	−0.042	0.161	0.605
γ_5 (Put Value)	−3.909	−3.679	2.577	−8.566	−0.107	0.756
γ_6 (Run Up)	0.018	0.018	0.007	0.007	0.030	0.652
Panel B: $z_{i0} = x_{i0}'\beta_0 + \epsilon_{i0}$ (outcome for naked calls)						
β_{01} (Put Value)	−1.670	−1.703	4.356	−8.673	5.587	0.365
Panel C: $z_{i1} = x_{i1}'\beta_1 + \epsilon_{i1}$ (outcome for underwritten calls)						
β_{11} (Put Value)	−6.164	−6.335	5.679	−15.182	3.595	0.772
Panel D: free covariance matrix elements						
$\sigma_{\upsilon 0}$	−0.184	−0.174	0.529	−1.063	0.681	0.424
σ_0^2	12.297	12.194	1.812	9.550	15.455	0.515
$\sigma_{\upsilon 1}$	−0.792	−0.784	0.344	−1.367	−0.239	0.454
σ_1^2	7.415	7.270	1.417	5.370	10.010	0.548
$\sigma_{\upsilon 1}-\sigma_{\upsilon 0}$	−0.608	−0.616	0.629	−1.632	0.445	0.439

This table reports summary statistics describing the Bayesian posterior densities of parameters in the endogenous switching model assuming that residuals follow a multivariate *t* distribution with $\nu=6$ degrees of freedom. The sample contains 229 calls of convertible bonds (85 underwritten and 144 naked). The model was estimated using Markov chain Monte Carlo (MCMC) techniques. The posterior mean, median, standard deviation, 5th and 95th percentile critical values, and 1st order autocorrelation for each parameter are based on 10,000 Gibbs sampler iterations from a suitably constructed Markov chain.

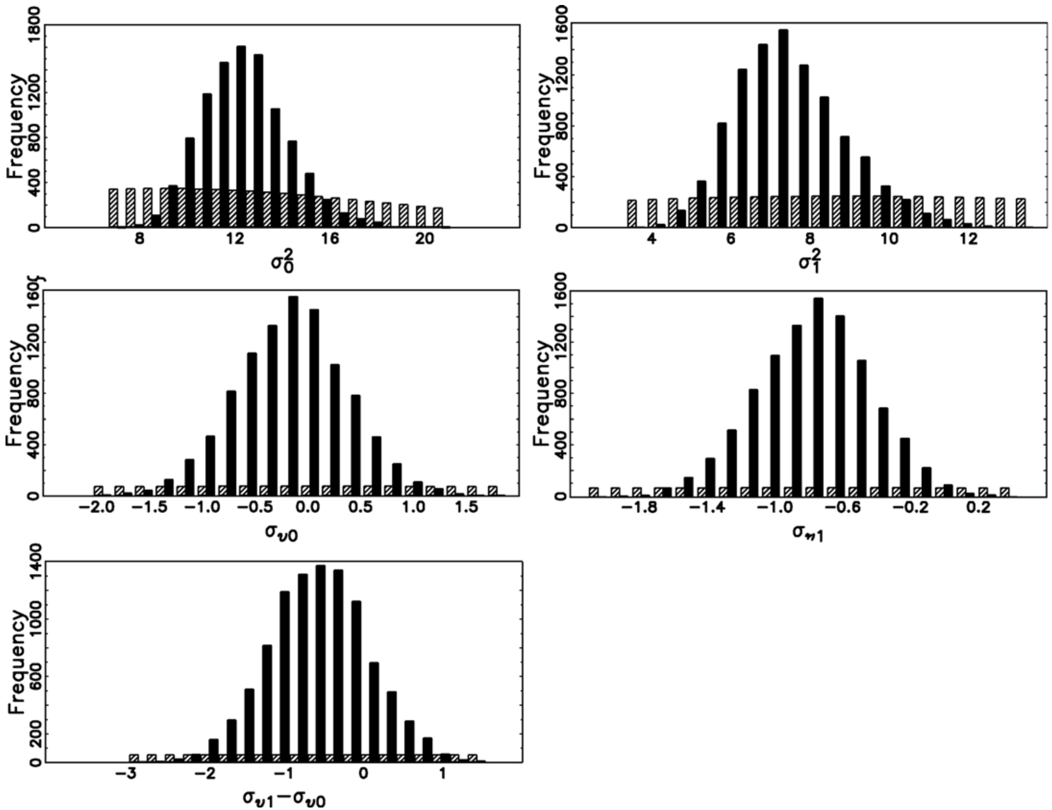

Fig. 3. Prior and posterior densities of covariance matrix parameters for model $\mathcal{M}_2$. Top: σ_0^2 and σ_1^2. Middle: $\sigma_{\upsilon 0}$ and $\sigma_{\upsilon 1}$. Bottom: $\sigma_{\upsilon 1}-\sigma_{\upsilon 0}$. Histograms are based on 10,000 iterations of the Gibbs sampler for the multivariate t model with $v=6$ degrees of freedom.

6.3. Bayesian inference regarding models $\mathcal{M}_2$ and $\mathcal{M}_3$

In $\mathcal{M}_2$, I relax the restriction that $\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$. $\mathcal{M}_2$ is otherwise identical to $\mathcal{M}_1$. Table 3 reports sample statistics summarizing the posterior densities of the model's parameters. Most of the posterior densities are quite similar to those for $\mathcal{M}_1$ reported in Table 2. In Panel A, the coefficient on *Put Value* is reliably negative. The posterior mean (median) of γ_5 is −3.909 (−3.679) and the 90% confidence interval is [−8.566, −0.107]. All other inferences regarding the binary choice model are the same as $\mathcal{M}_1$. The most noteworthy differences between $\mathcal{M}_1$ and $\mathcal{M}_2$ concern the covariance matrix. Panel D of Table 3 reports summary statistics describing the posterior densities of the covariance matrix elements. With no restriction imposed, the posterior densities of $\sigma_{\upsilon 0}$ and $\sigma_{\upsilon 1}$ diverge somewhat. $\sigma_{\upsilon 1}$ is reliably negative. The posterior mean (median) of $\sigma_{\upsilon 1}$ is −0.792 (−0.784) and the 90% confidence interval is [−1.367, −0.239]. $\sigma_{\upsilon 0}$ is negative, but not reliably different from zero. The posterior mean (median) of $\sigma_{\upsilon 0}$ is −0.184 (−0.174) and the 90% confidence interval is [−1.063, 0.681]. The histograms in Fig. 3 provide graphical summaries of the prior and posterior densities of the covariance matrix terms for model $\mathcal{M}_2$.

I test the hypothesis that $\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$ in $\mathcal{M}_2$ by examining the posterior density of $\sigma_{\upsilon 1}-\sigma_{\upsilon 0}$. Gibbs sampler draws of $\sigma_{\upsilon 1}-\sigma_{\upsilon 0}$ are easily constructed from existing draws of $\sigma_{\upsilon 0}$ and $\sigma_{\upsilon 1}$. Summary statistics for this posterior density are reported in Panel D of Table 3. Fig. 3 presents graphical summaries of the prior and posterior densities of $\sigma_{\upsilon 1}-\sigma_{\upsilon 0}$. The mean (median) of the posterior density is −0.608 (−0.616). The 90% Bayesian confidence interval is [−1.632, 0.445]. Since zero clearly falls within the central mass of this posterior density, I can't reject the hypothesis that $\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$. This helps explain why the data favors $\mathcal{M}_1$ over $\mathcal{M}_2$.

Table 4
Bayesian posterior densities of model parameters: model $\mathcal{M}_3$

Parameter (variable)	Mean	Median	SD	5%	95%	ρ_1
Panel A: $s_i^*=w_i'\gamma+\upsilon_i$ (choice equation)						
γ_0 (Constant)	−0.564	−0.556	0.634	−1.636	0.443	0.624
γ_1 (Cushion)	−1.805	−1.789	0.502	−2.652	−1.004	0.884
γ_2 (Volatility)	0.040	0.041	0.126	−0.167	0.245	0.520
γ_3 (Fraction Out)	0.543	0.519	0.610	−0.444	1.575	0.614
γ_4 (Cash Stress)	0.042	0.041	0.057	−0.051	0.136	0.527
γ_5 (Put Value)	−2.730	−2.722	1.769	−5.690	0.151	0.556
γ_6 (Run Up)	0.017	0.017	0.006	0.006	0.028	0.571
Panel B: $z_{i0}=x_{i0}'\beta_0+\epsilon_{i0}$ (outcome for naked calls)						
β_{01} (Put Value)	−2.500	−2.460	4.331	−9.600	4.610	0.265
Panel C: $z_{i1}=x_{i1}'\beta_1+\epsilon_{i1}$ (outcome for underwritten calls)						
β_{11} (Put Value)	−6.720	−6.682	5.305	−15.353	1.952	0.666
Panel D: free covariance matrix elements						
$\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$	−0.687	−0.686	0.319	−1.206	−0.161	0.298
σ_0^2	16.403	16.211	2.009	13.396	20.062	0.239
σ_1^2	10.576	10.390	1.679	8.103	13.632	0.477

This table reports summary statistics describing the Bayesian posterior densities of parameters in the endogenous switching model assuming that residuals follow a multivariate *t* distribution with $\nu=100$ degrees of freedom (i.e., multivariate normal). The sample contains 229 calls of convertible bonds (85 underwritten and 144 naked). The model was estimated using Markov chain Monte Carlo (MCMC) techniques. The posterior mean, median, standard deviation, 5th and 95th percentile critical values, and 1st order autocorrelation for each parameter are based on 10,000 Gibbs sampler iterations from a suitably constructed Markov chain.

One contribution of this paper is the use of the multivariate t distribution rather than the multivariate normal distribution for the residuals in Eq. (3). In order to understand how this generalization affects my results, I estimate a third model ($\mathcal{M}_3$) with degrees of freedom $\nu=100$ to approximate a multivariate normal distribution. Model $\mathcal{M}_3$ is otherwise identical to $\mathcal{M}_1$, and imposes the restriction $\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$. Of the models estimated in this paper, $\mathcal{M}_3$ is the closest in spirit to the latent information model of Acharya (1988). The posterior densities of the parameters are summarized in Table 4. The posterior densities of parameters in the choice and potential outcome equations (Panels A, B and C) are similar in sign and magnitude to those reported in Table 2 for $\mathcal{M}_1$. For each parameter, the posterior mean shrinks toward zero and the posterior standard deviation is slightly smaller (i.e., the posterior density is more precise). However, none of the inferences from $\mathcal{M}_3$ are different from $\mathcal{M}_1$. The parameters in the covariance matrix (Panel D) change more substantially. The posterior means of σ_0^2, σ_1^2, and $\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$ shift away from zero, and the posterior standard deviations become larger. The explanation is simple. In $\mathcal{M}_3$, the errors are distributed multivariate t with $\nu=100$ rather than $\nu=6$. Intuitively, estimates of the variance and covariance terms in $\mathcal{M}_3$ must increase in magnitude to fit the dispersion of residuals formerly captured by fatter tails in $\mathcal{M}_1$.

7. Empirical results: Heckman–Lee two-stage approach

Estimation of the endogenous switching model is cumbersome, but feasible, using frequentist (or classical) econometric methods. Heckman (1976, 1979) and Lee (1979) discuss a two-stage approach for estimating the model for the case where $\epsilon_i=(\upsilon_i\ \epsilon_{i0}\ \epsilon_{i1})'$ has a trivariate normal distribution.[13] In the first stage, the maximum likelihood method is used to estimate a simple probit model of the firm's choice:

$$s_i = \Phi(w_i'\gamma) + \upsilon_i. \tag{9}$$

The difficulty in estimating the outcome equations arises from a truncated residuals problem. The expected residuals from Eq. (2), conditional on the binary endogenous choice s_i, are

$$E(\epsilon_{i0}|s_i = 0, x_{i0}) = E(\epsilon_{i0}|\upsilon_i < -w_i'\gamma, x_{i0}) = -\sigma_{\upsilon 0}\frac{\phi(-w_i'\gamma)}{\Phi(-w_i'\gamma)} = -\sigma_{\upsilon 0}\frac{\phi(w_i'\gamma)}{1-\Phi(w_i'\gamma)} \tag{10}$$

and

$$E(\epsilon_{i1}|s_i = 1, x_{i1}) = E(\epsilon_{i1}|\upsilon_i \geq -w_i'\gamma, x_{i1}) = \sigma_{\upsilon 1}\frac{\phi(-w_i'\gamma)}{1-\Phi(-w_i'\gamma)} = \sigma_{\upsilon 1}\frac{\phi(w_i'\gamma)}{\Phi(w_i'\gamma)}, \tag{11}$$

where $\phi(w_i'\gamma)$ and $\Phi(w_i'\gamma)$ are, respectively, the standard normal density function and standard normal cumulative distribution function evaluated at $w_i'\gamma$. Note that these expected residuals are conditionally non-zero if $\sigma_{\upsilon 0}\neq 0$ or $\sigma_{\upsilon 1}\neq 0$. Non-zero residuals are a violation of the standard ordinary least squares (OLS) assumptions.

The first-stage probit regression provides a consistent estimate of γ. This estimate, $\hat{\gamma}$, can be used to obtain estimates of the correction terms (also known as inverse Mill's ratios) defined in Eqs. (10) and (11) for each event in the sample.

[13] Section 8.3 of Maddala (1983) provides a textbook treatment of the two-stage method for estimating a switching regression model with endogenous switching. See Acharya (1988, 1993) for a more detailed discussion of the estimation methodology and an application to conditional event studies.

Conditional on the choice s_i, the second-stage regressions are

$$y_i = x_{i0}'\beta_0 - \sigma_{\upsilon 0}\frac{\phi(w_i'\hat{\gamma})}{1-\Phi(w_i'\hat{\gamma})} + u_{i0} \quad \text{for } s_i = 0 \tag{12}$$

$$y_i = x_{i1}'\beta_1 + \sigma_{\upsilon 1}\frac{\phi(w_i'\hat{\gamma})}{\Phi(w_i'\hat{\gamma})} + u_{i1} \quad \text{for } s_i = 1. \tag{13}$$

The residuals from Eqs. (12) and (13) can be written

$$u_{i0} = \epsilon_{i0} + \sigma_{\upsilon 0}\frac{\phi(w_i'\gamma)}{1-\Phi(w_i'\gamma)}$$
$$u_{i1} = \epsilon_{i1} - \sigma_{\upsilon 1}\frac{\phi(w_i'\gamma)}{\Phi(w_i'\gamma)}.$$

Note that these residuals are conditionally mean zero (i.e., $E(u_{i0}|s_i=0, x_{i0}) = E(u_{i1}|s_i=1, x_{i1}) = 0$), so estimation by OLS is appropriate. Although OLS parameter estimates are consistent, OLS standard errors are biased downward because they ignore the sampling variation in the first-stage estimates of γ. Maddala (1983, Chapter 8) derives an asymptotic covariance matrix for the second-stage estimates that accounts for this errors-in-variables problem.

Note that Eqs. (12) and (13) differ from second-stage regression in Acharya (1988). Acharya (1988) employs a variant of the Heckman–Lee procedure in which Eqs. (12) and (13) are replaced by the single equation

$$y_i = x_i'\beta - \pi(1-s_i)\frac{\phi(w_i'\hat{\gamma})}{1-\Phi(w_i'\hat{\gamma})} + \pi s_i\frac{\phi(w_i'\hat{\gamma})}{\Phi(w_i'\hat{\gamma})} + u_i \tag{14}$$

in the second stage. In this model, the binary choice variable s_i serves as a switching variable that activates the appropriate correction term. This approach estimates a single covariance term

Table 5
Two-stage parameter estimates for the latent information model

Parameter (variable)	Estimate	SE	*t*-stat
Panel A: $s_i = \Phi(w_i'\gamma) + \upsilon_i$			
γ_0 (Constant)	−0.543	0.639	−0.85
γ_1 (Cushion)	−1.819	0.518	−3.51
γ_2 (Volatility)	0.049	0.126	0.39
γ_3 (Fraction Out)	0.516	0.621	0.83
γ_4 (Cash Stress)	0.030	0.056	0.53
γ_5 (Put Value)	−2.619	1.763	−1.49
γ_6 (Run Up)	0.017	0.007	2.58
Panel B: $y_i = x_i'\beta - \pi(1-s_i)\frac{\phi(w_i'\hat{\gamma})}{1-\Phi(w_i'\hat{\gamma})} + \pi s_i\frac{\phi(w_i'\hat{\gamma})}{\Phi(w_i'\hat{\gamma})} + u_i$			
β_0 (Constant)	−1.461	0.261	−5.59
β_1 (Put Value)	−5.428	3.846	−1.41
π (Private Information)	−0.701	0.331	−2.11

Panel A reports maximum likelihood estimates of the first-stage probit regression where $s_i=1$ ($s_i=0$) for underwritten (naked) calls, and w_i is a vector of firm-specific covariates. Panel B reports OLS estimates of the second-stage cross-sectional regression where y_i is the three-day event window cumulative abnormal return (CAR), x_i is a vector of firm-specific covariates, $\phi(\cdot)$ is the standard normal density function, and $\Phi(\cdot)$ is the standard normal cumulative distribution function. The sample contains 229 calls of convertible bonds (85 underwritten and 144 naked).

($\pi=\sigma_{\upsilon 0}=\sigma_{\upsilon 1}$) in lieu of two covariance terms ($\sigma_{\upsilon 0}$ and $\sigma_{\upsilon 1}$). It also requires that the same covariate vector (x_i) be used for both choices.

Table 5 reports maximum likelihood estimates of the model using the two-stage method of Acharya (1988). Panel A reports parameter estimates for the first-stage probit equation. The signs and relative magnitudes of the coefficient estimates and standard errors are very similar to the MCMC estimates of model $\mathcal{M}_3$ (the most closely-related specification) reported in Table 4. The decision to underwrite a call is inversely related to *Cushion*, and positively related to *Run Up*. The point estimate of γ_1 is -1.819 ($t=-3.51$). The point estimate of γ_6 is 0.017 ($t=2.58$). None of the other coefficients in the first-stage probit equation are significantly different from zero.

Panel B of Table 5 reports OLS estimates for the second-stage outcome equation. The covariance term π is negative and significant ($\pi=-0.701$, $t=-2.11$), which indicates that the private information partially revealed by the firm's decision to have a call underwritten is negatively correlated with the market's response. This evidence supports the *private information* hypothesis and is consistent with the MCMC results reported in the previous section. The coefficient β_1 is negative, but not significant ($\beta_1=-5.428$, $t=-1.41$). This suggests that the value of announcement put options does not explain cross-sectional variation in event window abnormal returns. The intercept is negative and significant ($\beta_0=-1.461$, $t=-5.59$), and close to the unconditional mean CAR for the full sample (-1.58%). Since the second-stage regression is not estimated conditional on the firm's choice (s_i), there is no need to work with demeaned CARs or omit the intercept. The adjusted R^2 for the second-stage regression is 0.0198.

8. Conclusion

This paper employs flexible Bayesian MCMC methods to estimate switching models with endogenous self-selection. The literature on conditional event studies advocates the use of such models to control for the endogeneity of the firm's choice. The MCMC approach offers several significant advantages over the Heckman–Lee two-stage estimation approach. Since it is not necessary to maximize the likelihood function, complex models are not an insurmountable obstacle. Inferences are based on sample distributions drawn from the exact posterior densities of parameters of interest. The resulting exact small sample results are attractive given the small sample sizes typical of corporate event studies. The Gibbs sampler algorithm implicitly accounts for parameter uncertainty and eliminates the errors-in-variables problem associated with two-stage procedures. And, Bayesian MCMC methods provide a natural approach for estimating the treatment effects that arise in models with endogenous self-selection.

As an application, I analyze cross-sectional variation in the market's response to naked and underwritten calls of convertible bonds. The call announcement reveals an endogenous binary choice by the firm's management: the call is either naked or underwritten by an investment bank. The firm's choice is a function of observable firm-specific attributes and the manager's private information. I investigate two potential explanations for the market's more negative response to underwritten calls of convertible bonds. I find some evidence to support the *private information* hypothesis. Private information partially revealed by announcement of an underwritten call is negatively correlated with the market's response to the announcement. This evidence suggests that the decision to have a call underwritten by an investment bank is interpreted as "bad news" by the market. However, the evidence is mixed. On average, the treatment effect associated with choosing an underwritten call is negative, but insignificant. The treatment effect is the difference between the observed and counter-factual outcomes. Analysis of the treatment effect is important and appropriate because it controls for the endogeneity of the firm's choice. Such analysis is possible because the unobserved counter-factual outcomes can be simulated using a Bayesian MCMC technique known as data augmentation. I find no evidence to support the *announcement put option* hypothesis. Announcement put option values are relatively small and not significantly related to announcement window abnormal returns.

The Bayesian MCMC approach introduced in this paper can be extended to analyze the effects of corporate announcements of more complex choices. Recent advances in Bayesian statistics, particularly MCMC methods, make more complex models with self-selection amenable to analysis.

Acknowledgments

I am grateful to Siddhartha Chib for many helpful discussions. I appreciate the comments of Kerry Back, Phil Dybvig, Bart Hamilton, Paul Irvine, Avner Kalay, Kai Li, Mark Loewenstein, Tom Miller, Federico Nardari, Annette Poulsen, Anjolein Schmeits, Raj Singh, Guofu Zhou, two anonymous referees and seminar participants at the University of Georgia, University of Minnesota, University of Missouri, University of Utah, Washington University in St. Louis and the 2000 meetings of the European Finance Association in London. Paskalis Glabadanidis, Qiang Li and Ryan McKeon provided valuable research assistance. The usual disclaimer applies.

Appendix A. Markov chain Monte Carlo sampling methodology

A.1. Prior densities

My prior on β is multivariate Gaussian and denoted as $\mathcal{N}_k(b_0, B_0)$. My prior on σ, the vector of unique free elements in Σ, is a little trickier. Let q denote the number of free parameters in σ. I need to truncate the distribution of σ to the region $S \subset \mathbb{R}^q$ where Σ is positive definite. Following Chib and Greenberg (1998), I let the prior distribution be truncated multivariate Gaussian, $\sigma \propto \mathcal{N}_q(g_0, G_0) I_S(\sigma \in S)$, where g_0 and G_0 denote hyperparameters, and I_S is an indicator function that takes value one if $\sigma \in S$ and takes value zero otherwise.

A.2. Likelihood function

The likelihood function for the model in Eq. (6), where the choice is binary and the potential outcomes are distributed multivariate t, is

$$\begin{aligned} f(y, s|\beta, \sigma, \{\lambda_i\}) &\propto f(y|\beta, \Sigma, \{\lambda_i\}) \Pr(s|y, \beta, \Sigma, \{\lambda_i\}) \\ &\propto \prod_{i:s_i=0} \phi(z_{i0}|x_{i0}'\beta_0, \lambda_i^{-1}\sigma_0^2)\ \Phi\left(\frac{-w_i'\gamma - (\sigma_{\upsilon 0}/\sigma_0^2)(z_{i0} - x_{i0}'\beta_0)}{\lambda_i^{-\frac{1}{2}}(1 - \sigma_{\upsilon 0}^2/\sigma_0^2)^{\frac{1}{2}}}\right) \\ &\times \prod_{i:s_i=1} \phi(z_{i1}|x_{i1}'\beta_1, \lambda_i^{-1}\sigma_1^2)\ \Phi\left(\frac{w_i'\gamma + (\sigma_{\upsilon 1}/\sigma_1^2)(z_{i1} - x_{i1}'\beta_1)}{\lambda_i^{-\frac{1}{2}}(1 - \sigma_{\upsilon 1}^2/\sigma_1^2)^{\frac{1}{2}}}\right) \end{aligned} \tag{A.1}$$

where $\beta = (\gamma, \beta_0, \beta_1) \in \mathbb{R}^k$, $\sigma = (\sigma_{\upsilon 0}, \sigma_0^2, \sigma_{\upsilon 1}, \sigma_1^2) \in S$, and $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the standard normal density function and standard normal cumulative distribution function.

A.3. Full conditional distributions

The Gibbs sampler algorithm employed in this paper is summarized in Subsection 4.2. After initializing the values of β, σ, $\{\lambda_i\}$, and $\{z_i^*\}$, the Gibbs sampler samples recursively from the full conditional distributions of the parameters and augmented data. The full conditional distributions are detailed below.

$\beta=(\gamma, \beta_0, \beta_1)$ is the coefficient vector from Eq. (3). The full conditional distribution of β is $\pi(\beta|y,s,\Sigma,\{z_i^*\},\{\lambda_i\})\propto\mathcal{N}_k(b,B)$, where $b=B(b_0B_0^{-1}+\sum_{i=1}^{n}X_i'\lambda_i\Sigma^{-1}z_i)$ and $B=(B_0^{-1}+\sum_{i=1}^{n}X_i'\lambda_i\Sigma^{-1}X_i)^{-1}$.

$\{z_i^*\}=(z_1,\ldots,z_n)$ is the set of latent variables and potential outcomes for events $i=1,\ldots,n$. $\{z_i^*\}$ is sampled from the full conditional distribution $\pi(\{z_i^*\}|y, s, \beta, \Sigma, \lambda_i)$ independently for $i=1,\ldots,n$. Recall that conditional on λ_i the full conditional distribution for z_i is Gaussian with mean $X_i'\beta$ and covariance matrix $\lambda_i^{-1}\Sigma$. For underwritten calls ($s_i=1$), the latent variable s_i^* is sampled from $\pi(s_i^*|y_i, s_i, z_{i0}, \beta, \Sigma, \lambda_i)$, a Gaussian distribution truncated on the interval $(0, \infty)$. The corresponding potential outcome z_{i0} is then sampled from $\pi(z_{i0}|y_i, s_i, s_i^*, \beta, \Sigma, \lambda_i)$, an unrestricted Gaussian distribution. For naked calls ($s_i=0$), s_i^* is sampled from $\pi(s_i^*|y_i, s_i, z_{i1}, \beta, \Sigma, \lambda_i)$, a Gaussian distribution truncated on the interval $(-\infty, 0)$. The corresponding potential outcome z_{i1} is sampled from $\pi(z_{i1}|y_i, s_i, s_i^*, \beta, \Sigma, \lambda_i)$, an unrestricted Gaussian distribution.

Due to the positive definite constraint on Σ, solving analytically for the full conditional distribution of σ is not feasible. Here I resort to simulation and employ the Metropolis–Hastings algorithm to sample from the conditional distribution.[14] Let $\pi(\sigma|y,s,\beta)\propto g(\sigma)=\mathcal{N}_4(g_0,G_0)I_S(\sigma)$. Let $q(\sigma|\mu, V)$ denote a multivariate t density with location parameter μ and scale parameter V defined as the mode and inverse of the negative Hessian, respectively, of log $g(\sigma)$. The Metropolis–Hastings algorithm:

(1) Samples a proposal value σ' from the density $q(\sigma|\mu, V)$.
(2) Moves to σ' from the current point σ with probability [see Chib and Greenberg, 1995]

$$\alpha=\min\left[\frac{f_N(\sigma'|g_0,G_0)I_S(\sigma')f(y,s|\beta,\Sigma')}{f_N(\sigma|g_0,G_0)I_S(\sigma)f(y,s|\beta,\Sigma)}\frac{q(\sigma|\mu,V)}{q(\sigma'|\mu,V)},1\right], \tag{A.2}$$

otherwise, stays at σ.

λ_i is the set of mixing variables used to represent the multivariate t distribution. The full conditional distribution for λ_i is

$$\pi(\lambda_i|y,s,\beta,\Sigma,\{z_i^*\})\sim\mathcal{G}\left(\frac{v+3}{2},\frac{v+(z_i-X_i'\beta)'\Sigma^{-1}(z_i-X_i'\beta)}{2}\right), \tag{A.3}$$

where v is degrees of freedom.

References

Acharya, S., 1988. A generalized econometric model and tests of a signalling hypothesis with two discrete signals. Journal of Finance 43, 413–429.

Acharya, S., 1993. Value of latent information: alternative event study methods. Journal of Finance 48, 363–385.

Albert, J.H., Chib, S., 1993. Bayesian analysis of binary and polychotomous response data. Journal of the American Statistical Association 88, 669–679.

Brennan, M.J., Schwartz, E.S., 1977. Convertible bonds: valuation and optimal strategies for call conversion. Journal of Finance 32, 1699–1715.

Brown, S.J., Warner, J.B., 1985. Using daily stock returns: the case of event studies. Journal of Financial Economics 14, 3–31.

[14] See Chib and Greenberg (1995) for an explanation of the Metropolis–Hastings algorithm.

Butler, A.W., 1999. Why do firms hire underwriters? An examination of underwritten and naked calls of convertible bonds, working paper. Louisiana State University.

Campbell, C.J., Ederington, L.H., Vankudre, P., 1991. Tax shields, sample-selection bias, and the information content of conversion-forcing bond calls. Journal of Finance 46, 1291–1324.

Casella, G., George, E.I., 1992. Explaining the Gibbs sampler. The American Statistician 46, 167–174.

Chib, S., 1992. Bayes inference in the Tobit censored regression model. Journal of Econometrics 51, 79–99.

Chib, S., 1995. Marginal likelihood from the Gibbs output. Journal of the American Statistical Association 90, 1313–1321.

Chib, S., Greenberg, E., 1995. Understanding the Metropolis–Hastings algorithm. The American Statistician 49, 327–335.

Chib, S., Greenberg, E., 1996. Markov chain Monte Carlo simulation methods in econometrics. Econometric Theory 12, 409–431.

Chib, S., Greenberg, E., 1998. Analysis of multivariate probit models. Biometrika 85, 347–361.

Chib, S., Hamilton, B., 2000. Bayesian analysis of cross-section and clustered data treatment models. Journal of Econometrics 97, 25–50.

Cowan, A.R., Nayar, N., Singh, A.K., 1999. How high are investment banking fees? The case of standby underwritten convertible calls, working paper. Iowa State University.

Cowan, A.R., Nayar, N., Singh, A.K., 2000. Underwriting and calls of convertible bonds. Decision Sciences 31, 57–77.

Datta, S., Iskandar-Datta, M., 1996. New evidence on the valuation effects of convertible bond calls. Journal of Financial and Quantitative Analysis 31, 295–307.

Dunbar, C.G., 1995. The use of warrants as underwriter compensation in initial public offerings. Journal of Financial Economics 38, 59–78.

Eckbo, B.E., Maksimovic, V., Williams, J., 1990. Consistent estimation of cross-sectional models in event studies. Review of Financial Studies 3, 343–365.

Ederington, L.H., Goh, J.C., 2001. Is a convertible bond call really bad news? Journal of Business 74, 459–476.

Fama, E.F., Fisher, L., Jensen, M.C., Roll, R., 1969. The adjustment of stock prices to new information. International Economic Review 10, 1–21.

Harris, M., Raviv, A., 1985. A sequential signalling model of convertible debt call policy. Journal of Finance 40, 1263–1281.

Heckman, J.J., 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables in a simultaneous equation system. Annals of Economic and Social Measurement 5, 475–492.

Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47, 153–161.

Ingersoll, J.E., 1977a. A contingent-claims valuation of convertible securities. Journal of Financial Economics 4, 289–322.

Ingersoll, J.E., 1977b. An examination of corporate call policies on convertible securities. Journal of Finance 32, 463–478.

Jaffee, D., Shleifer, A., 1990. Costs of financial distress, delayed calls of convertible bonds, and the role of investment banks. Journal of Business 63, S107–S123.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. Journal of the American Statistical Association 90, 773–795.

Lee, L.-F., 1979. Identification and estimation in binary choice models with limited (censored) dependent variables. Econometrica 47, 977–996.

Li, K., McNally, W., 1998. Open market versus tender offer share repurchases: A conditional event study, working paper. University of British Columbia.

Maddala, G.S., 1983. Limited-dependent and qualitative variables in econometrics. Econometric Society Monographs. Cambridge University Press, Cambridge, England.

Maddala, G.S., 1996. Applications of limited dependent variable models in finance. In: Maddala, G.S., Rao, C.R. (Eds.), Handbook of Statistics, vol. 14. Elsevier Science B. V., Amsterdam, pp. 553–566.

Mikkelson, W.H., 1981. Convertible bonds and security returns. Journal of Financial Economics 9, 237–264.

Modigliani, F., Miller, M.H., 1958. The cost of capital, corporation finance, and the theory of investment. American Economic Review 48, 261–297.

Ofer, A.R., Natarajan, A., 1987. Convertible call policies: an empirical analysis of an information-signalling hypothesis. Journal of Financial Economics 19, 91–108.

Prabhala, N.R., 1997. Conditional methods in event studies and an equilibrium justification for standard event-study procedures. Review of Financial Studies 10, 1–38.

Singh, A.K., Cowan, A.R., Nayar, N., 1991. Underwritten calls of convertible bonds. Journal of Financial Economics 29, 173–196.

Tanner, M.A., Wong, W.H., 1987. The calculation of posterior distributions by data augmentation. Journal of the American Statistical Association 82, 528–540.