

Inference in Long-Horizon Event Studies: A Bayesian Approach with Application to Initial Public Offerings

ALON BRAV*

ABSTRACT

Statistical inference in long-horizon event studies has been hampered by the fact that abnormal returns are neither normally distributed nor independent. This study presents a new approach to inference that overcomes these difficulties and dominates other popular testing methods. I illustrate the use of the methodology by examining the long-horizon returns of initial public offerings (IPOs). I find that the Fama and French (1993) three-factor model is inconsistent with the observed long-horizon price performance of these IPOs, whereas a characteristic-based model cannot be rejected.

RECENT EMPIRICAL STUDIES IN FINANCE document systematic long-run abnormal price reactions subsequent to numerous corporate activities.¹ Since these results imply that stock prices react with a long delay to publicly available information, they appear to be at odds with the Efficient Markets Hypothesis (EMH).

Long-run event studies, however, are subject to serious statistical difficulties that weaken their usefulness as tests of the EMH. In particular, most studies maintain the standard assumptions that abnormal returns are independent and normally distributed although these assumptions fail to hold even approximately at long horizons. First, samples of long-horizon abnormal returns are not independently distributed because many of the sample

* Duke University, Fuqua School of Business. This paper has benefited from the comments of Brad Barber, Michael Barclay, Kobi Boudoukh, Michael Brandt, George Constantinides, Zvi Gilula, Paul Gompers, John Graham, David Hsieh, Shmuel Kandel, S. P. Kothari, Craig Mackinlay, Ernst Maug, Roni Michaely, Mark Mitchell, Nick Polson, Haim Reisman, Jay Ritter, Matthew Rothman, Jay Shanken, Erik Stafford, Robert Stambaugh, René Stulz (the editor), Richard Thaler, Sheridan Titman, Ingrid Tierens, Tuomo Vuolteenaho, Jerold Warner, Bob Whaley, Bob Winkler, two anonymous referees, and seminar participants at Boston College, Columbia, Cornell, Dartmouth, Duke, Harvard, London Business School, Ohio State, Rochester, Tel-Aviv, UCLA, Yale, and the 1999 AFA conference. I thank Jay Ritter for the IPO data set used in this study and Michael Bradley for access to the SDC database. Krishnamoorthy Narasimhan provided excellent research assistance. I owe special thanks to Eugene Fama, Campbell Harvey, and J. B. Heaton for their invaluable insights. All remaining errors are mine.

¹ See Ritter (1991) and Loughran and Ritter (1995) for initial public offerings; Ikenberry, Lakonishok, and Vermaelen (1995) for stock repurchases; Speiss and Affleck-Graves (1995) for seasoned equity offerings; Michaely, Thaler, and Womack (1995) for dividend initiations and omissions; and Womack (1996) for stock recommendations.

firms overlap in calendar time.² Second, abnormal returns are not normally distributed because long-horizon returns are skewed right by the compounding of single-period returns. The standard calculation of abnormal return—sample firm return minus the return on a well-diversified portfolio—results in a distribution of abnormal returns that is skewed right as well (see Barber and Lyon (1997) and Kothari and Warner (1997)). Both deviations from the standard assumptions imply that parametric inferences that rely on independence and normality are incorrect.³

In this paper I propose a methodology that confronts non-normality and cross-sectional dependence in abnormal returns. The methodology employs a Bayesian “predictive” approach, essentially a goodness-of-fit criterion, based on the idea that good models among those in consideration should make predictions close to what has been observed in the data.⁴ Given an asset-pricing model and a distribution for firm residuals, the model’s parameters are estimated for all the sample firms. Then, given the estimated parameters, long-horizon returns for all firms are simulated taking account of the estimated residual variations and covariations. These steps are repeated a large number of times, and the simulated averages are used to construct the null distribution for the sample mean. If the actual abnormal return is extreme relative to the range of predicted realizations, the entertained model and residual distribution are unlikely to have generated the observed sample return.

The paper proceeds as follows. Section I discusses recent attempts to address inference problems in long-horizon event studies. Section II presents the proposed methodology. Results are given in Section III. Comparison with an alternative approach is given in Section IV. Section V provides further details of implementation, and Section VI concludes.

I. Background

Recent papers by Kothari and Warner (1997) and Barber and Lyon (1997) also address biases in long-horizon event studies. Both document that for randomly chosen firms, the traditional *t*-test of abnormal performance is misspecified and indicates abnormal performance too frequently. Barber and Lyon (1997) argue that misspecification arises from three possible biases:

² Overlap in calendar time is associated with positive cross-sectional dependence because of unpriced industry factors in returns as this paper documents later. See also Collins and Dent (1984), Sefcik and Thompson (1986), and Bernard (1987).

³ Asymptotically, the normality of the sample mean is guaranteed using a Central Limit Theorem argument. The adequacy of this approximation is sample specific because it depends on the rate of convergence, which is negatively related to both the degree of cross-sectional dependency and non-normality. See also Cowan and Sergeant (1997).

⁴ See Box (1980); Rubin (1984); Gelfand, Dey, and Chang (1992); Ibrahim and Laud (1994); Laud and Ibrahim (1995); Gelman et al. (1995); and Gelman, Meng, and Stern (1996). For similar applications in the time-series literature see Tsay (1992) and Paparoditis (1996) and in the health-care research see Stangl and Huerta (1997).

the “new listing” bias, the “rebalancing” bias, and the “skewness” bias. The “new listing” bias arises because sample firms usually have a long pre-event return record, whereas the benchmark portfolio includes firms that have only recently begun trading and are known to have abnormally low returns (Ritter (1991)). The “rebalancing” bias arises because the compounded return on the benchmark portfolio implicitly assumes periodic rebalancing of the portfolio weights, whereas the sample firm returns are compounded without rebalancing. The “skewness” bias refers to the fact that with a skewed-right distribution of abnormal returns, the Student t -distribution is asymmetric with a mean smaller than the zero null.

Kothari and Warner (1997) present additional sources of misspecification. First, they argue that parameter shifts in the event-period can severely affect tests of abnormal performance. For example, the increase in variability of abnormal returns over the event-period needs to be incorporated when conducting inferences. Second, they stress the issue of firm survival and its effect both on the measured abnormal return and its variability.

The importance of these studies is in documenting the possibility of erroneous inferences in long-horizon event studies and thus the need for improved testing procedures that can potentially overcome these problems. Essentially, three methodologies have been proposed as a remedy.

The first, by Ikenberry, Lakonishok, and Vermaelen (1995), is a nonparametric bootstrap approach. In their approach, the researcher generates an empirical distribution of average long-horizon abnormal return and then infers if the observed performance is consistent with this distribution. To generate an empirical distribution when all that we observe is one realization of an average abnormal return, Ikenberry et al. suggest the following procedure. First, replace each firm from the original sample with another firm that has the same expected return and calculate the latter sample’s abnormal performance. Second, repeat this replacement a large number of times and then compare the observed average abnormal return to those generated by the new samples. The researcher can reject the null of no abnormal performance if it is unlikely that the realized average return came from the simulated distribution. The appeal of this approach is that it is easy to implement. Once the dimensions that determine expected returns have been specified, the replacement of the original sample is straightforward and the desired empirical density is easy to generate.

This method has two potential shortcomings, however. The replacement of original sample firms implies an assumption that the two samples are similar in *every* dimension, including, but not limited to, expected returns. This is unlikely for two reasons. First, if the two samples have systematically different residual variations then the resulting empirical distribution will be biased. Second, if the original sample’s abnormal returns are cross-sectionally correlated then the replacement with random samples, which by construction are uncorrelated, may lead to false inferences. Section IV provides direct evidence on both of these limitations and the magnitude of the biases that they lead to.

The second approach, by Lyon, Barber, and Tsai (1999), advocates the use of carefully constructed benchmark portfolios that are free of the “new listing” and “rebalancing” biases mentioned above. Moreover, to account for the “skewness” bias, they propose a skewness-adjusted t -statistic.⁵ They show that for randomly selected samples, their methods yield well-specified test statistics.

The fact that most of their analysis is conducted on *randomly* selected samples implies that their proposed methods are applicable to studies whose “events” occurred at random. While it may be the case that certain corporate events are uncorrelated across firms, observation suggests this is not true for initial public offerings, seasoned equity offerings, stock repurchases, and mergers, events that are frequently the subject of long-horizon event studies. Indeed, Lyon et al. recognize this and attempt to correct their methods for cross-sectional correlation. Regretfully, as they point out, the method proposed by Ikenberry et al. cannot be adjusted, and their skewness-adjusted t -statistic does not eliminate the misspecification in samples with overlapping returns.

Finally, a potential third approach, which was also examined by Lyon et al. (1999), is a calendar-time portfolio method, which has been advocated by Fama (1998) and applied by Jaffe (1974), Mandelker (1974), Loughran and Ritter (1995), Brav and Gompers (1997), and Mitchell and Stafford (1999). It is well known that the portfolio approach eliminates the problem of cross-sectional dependence among the sample firms, which is also the goal of the method suggested in this paper.

Lyon et al. (1999) find that the calendar-time approach is well specified in random samples but misspecified in nonrandom samples. Furthermore, as Mitchell and Stafford (1999) point out, the portfolio approach has several potential problems that arise from the changing composition of the portfolio through time. Specifically, the portfolio factor loadings, which are assumed constant, are likely to vary through time. Second, the change in the number of firms can potentially lead to heteroskedasticity. Loughran and Ritter (1999) raise another set of potential problems associated with the power of this approach. The methodology suggested in this paper does not suffer from the problems examined by Mitchell and Stafford (1999) and Loughran and Ritter (1999) because the unit of observation is an event firm rather than a calendar month. Furthermore, by avoiding the aggregation of firm returns, it allows the researcher the added flexibility to directly model the firm characteristics and their evolution to the extent that these affect inferences.

Next, I present the Bayesian approach in Section II by applying it directly to a small data set of IPOs.

II. Methodology

In this section I present a long-horizon event study methodology that addresses the statistical issues raised in Section I. The general approach to estimation undertaken in this paper is Bayesian. Within this framework,

⁵ Lyon et al. also support the use of the bootstrap approach due to Ikenberry et al. (1995).

the researcher begins with subjective prior beliefs regarding the model parameters. Posterior beliefs are then generated using sample information via the Bayes theorem. These beliefs summarize the researcher's knowledge about the parameters of interest conditional on both the model and the data.

An important feature of the Bayesian framework is that it enables the researcher to formally incorporate subjective prior beliefs regarding the parameters of the asset-pricing model. Consider, for example, a sample of firms drawn from the same industry. It may be reasonable to assume, *a priori*, that these firms have "similar" model parameters either because the systematic risk of their earnings may be driven by the same economy-wide factors or because of exposure to industry-specific supply and demand shocks (e.g., Pastor and Stambaugh (1999), Vasicek (1973)). When estimating an individual firm's parameters, the Bayesian approach can take account of these prior beliefs by exploiting the information contained in other firms' parameters explicitly. In practice, this is achieved by *shrinking* the mean of the posterior beliefs away from the firm's simple least squares estimate and toward the industry sample grand average (e.g., Lindley and Smith (1972), Blattberg and George (1991), Breslow (1990), Gelfand et al. (1990), and Stevens (1996)). Bayesian posterior beliefs also explicitly reflect the researcher's uncertainty about the firm's parameters known as estimation risk (e.g., Barry (1974), Jorion (1991), Klein and Bawa (1976)).

To ease exposition, I build the methodological approach in an application to a small data set of IPO firms (a one industry subset, chosen arbitrarily, from the full sample used in Section III). The goal is to construct a density for the sample average long-horizon abnormal return under the null hypothesis of no abnormal performance. Comparing the realized abnormal performance against this density provides a natural test for abnormal performance. The simulated density captures both the firm-specific residual standard variations that induce non-normality and the cross-sectional correlations as these are reflected in the researcher's posterior beliefs.

A. Data Description

The sample is from Ritter (1991), and it comprises 113 initial public offerings from the computer and data processing services industry conducted over the period from 1975 to 1984. Table I, Panel A, provides the mean and median market capitalization (size) and book-to-market ratios for these firms. Size is calculated using the first closing price on the CRSP daily tape, whereas pre-issue book values are from Ritter (1991). Panel B gives the allocation of these firms into size and book-to-market quintiles formed using size and book-to-market cutoffs of NYSE firms. Panel C provides the annual volume of issuance.

Panel A shows that the typical firm in this sample is small, with median market capitalization equal to \$29.7 million and median book-to-market ratio equal to 0.06. In fact, as shown in Panel B, most of the sample firms belong to the bottom quintile of size and book-to-market using NYSE firm breakpoints. From Panel C it is evident that, at least in this industry, equity issues were clustered in calendar time, mostly in 1981 and 1983.

Table I
Descriptive Statistics for Computer and
Data Processing Services IPOs, 1975–1984

The sample consists of 113 firms from the computer and data processing services industry (SIC = 737). The three-digit SIC code used to assign IPOs to this industry is from Ritter (1991). Panel A provides the mean and median market capitalization (size) and book-to-market ratios for these firms. Size is calculated using the first closing price on the CRSP daily tape, whereas pre-issue book values are from Ritter (1991). Book values were unavailable for 11 firms. Panel B gives the allocation of these firms into size and book-to-market quintiles that were formed using size and book-to-market cutoffs of NYSE firms. Panel C provides the annual volume of issuance.

Panel A: Size and Book-to-Market Statistics									
Size (mil\$)					Book-to-Market				
Mean		Median			Mean		Median		
60.2		29.7			0.09		0.06		
Panel B: Allocation into Size and Book-to-Market Quintiles									
		Low/Small		2	3	4	High/Large		Missing
Size		90		16	6	1	0		0
Book-to-market		100		0	1	0	1		11
Panel C: Annual Volume									
1976	1977	1978	1979	1980	1981	1982	1983	1984	Total
2	1	2	2	2	20	13	53	18	113

Table II provides descriptive statistics regarding these firms' five-year aftermarket return performance relative to the NYSE-AMEX value-weighted index. The appropriateness of this and other benchmarks will be discussed later in this section. Also given are the cross-sectional standard deviation of abnormal returns and the skewness of the abnormal return distribution.

The five-year returns on these firms are striking. Investors holding shares of IPOs in this industry earned 24.5 percent, on average, over a five-year period, and the market as a whole nearly doubled in value. Note also that the median firm lost 47.2 percent of its value over this period. Furthermore, the skewness and the standard deviation of the abnormal return distribution are extremely large. These estimates reflect the success of a few firms among the abysmal performance of most of the sample. For example, Legent Corporation (IPO in January 1984) earned 564 percent in excess of the market; Manufacturing Data Sys (IPO in February 1976) earned a 679 percent excess return; and Cullinet Software (IPO in August 1978) earned 866 percent excess return. On the other hand, B.P.I. Systems (IPO in June 1982), Intermetrics (IPO in June 1982), and Computone Sys (IPO in November

Table II
Five-Year Aftermarket Performance Relative to
the NYSE-AMEX Value-Weight Index

For each IPO buy-and-hold return is calculated starting in the month after the issuance through the end of its fifth seasoned year (60 calendar months). The benchmark return is the buy-and-hold return on the NYSE-AMEX value-weight index. If a firm delists prematurely the buy-and-hold return is calculated until the delisting month both for the sample firm and its benchmark.

Number of IPOs	IPO Average Five-Year Return	IPO Median Five-Year Return	NYSE-AMEX VW Five-Year Return	Five-Year Abnormal Return		
				Average	Standard Deviation	Skewness
113	24.5%	-47.2%	90.2%	-65.7%	185.0%	2.8

1981) underperformed the market by 305 percent, 249 percent, and 235 percent, respectively. Finally, normality of the abnormal return distribution is rejected for any traditional level of significance.⁶

B. Basic Setup

B.1. The Asset-Pricing Model

Fama (1970, 1998) emphasizes that all tests of market efficiency are joint tests of market efficiency and a model of expected returns. Since the goal in this part of the analysis is to apply the Bayesian approach to the estimation of the model parameters, I turn to the specification of the asset-pricing model.

I assume that asset returns are generated by a characteristic-based pricing model (Daniel and Titman (1997), and Daniel et al. (1997)).⁷ With this model, asset returns are determined by individual firm attributes such as size and book-to-market ratio. Given these attributes, each firm is matched to a benchmark portfolio of stocks with the same characteristics providing a measure of “normal” return. Specifically, for a firm return in period t , y_t , and attributes θ_t , we have

$$y_t = f(\theta_t) + \nu_t \quad (1)$$

where $f(\cdot)$ maps the firm characteristics to the return on the benchmark portfolio and ν_t is the firm idiosyncratic error.

The benchmark portfolios returns are formed as follows. Beginning in July 1974, I form size and book-to-market quintile breakpoints based on NYSE firms’ information. I allocate all NYSE, AMEX, and Nasdaq firms into the

⁶ The chi-square test that was used is described in Davidson and MacKinnon (1993, pp. 567–570).

⁷ The approach can be applied for any model. For example, a k -factor asset-pricing model is easily implementable. See Section V.

resulting 25 portfolios based on their known book values and market capitalizations (for additional details, see Fama and French (1993)). As in Mitchell and Stafford (1999) and Lyon et al. (1999), I include in the analysis only firms that have ordinary common share codes (CRSP share codes 10 and 11). To make sure that IPOs are not compared to themselves, I exclude all IPOs from the benchmark construction. I also exclude from the benchmarks firms that have conducted seasoned equity offerings (SEOs) within the previous five years because it has been documented that these firms tend to underperform up to five years after issuance (Loughran and Ritter (1995)).⁸ I repeat this procedure for every July through 1989, recording the resulting breakpoints and firm allocations.

Next, each sample firm is matched to one of the 25 portfolios based on its size and book-to-market attributes that were known at the month of the IPO. Then, I calculate benchmark buy and hold returns by equally weighting the buy and hold returns of all the firms in that relevant portfolio. I make sure that the length of the benchmark return horizon is either 60 months or shorter if the IPO delisted prematurely. Finally, if a benchmark firm delists prematurely, I reinvest the proceeds in an equally weighted portfolio of the remaining firms.

B.2. Regression Setup

Given a sample of N firms with T_i , $i = 1, \dots, N$ monthly observations, define y_i as the $(T_i \times 1)$ column vector of firm i 's returns and f_i as the $(T_i \times 1)$ vector of matched benchmark portfolio returns. The firm return is modeled as follows.

$$y_i = f_i + v_i \quad \forall i = 1, \dots, N. \quad (2)$$

The system of all N assets is written using a seemingly unrelated regressions (SUR) setup (see Zellner (1962) and Gibbons (1982)):

$$Y = F\beta + V \quad (3)$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \quad F = \begin{bmatrix} f_1 & 0 & 0 & 0 \\ 0 & f_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & f_N \end{bmatrix}, \quad V = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix}.$$

⁸ Specifically, the first five years of return history since going public were deleted for all IPOs over the period from 1975 to 1989. These data come from two sources. For the period from 1975 to 1984 I use Ritter's (1991) database, and for the period from 1985 to 1989 I use 1966 common equity IPOs identified from the SDC database. Also deleted from the benchmark construction are returns of firms that had conducted a seasoned equity offering within the previous five years. The seasoned equity offerings data is taken from Brav, Geczy, and Gompers (1999).

Y is a $(\sum_{i=1}^N T_i \times 1)$ stacked vector of firm returns. F is a $(\sum_{i=1}^N T_i \times N)$ block diagonal matrix of factor realizations, ι is a $(N \times 1)$ vector of ones and V is a $(\sum_{i=1}^N T_i \times 1)$ stacked vector of firm residuals.

I assume throughout the analysis a multivariate normal distribution for V with mean zero and a $(\sum_{i=1}^N T_i \times \sum_{i=1}^N T_i)$ variance-covariance matrix Σ . The residuals are assumed to be temporally independent and to share a common contemporaneous correlation. This correlation, denoted ρ , reflects joint co-variation in returns that is driven by unpriced factors in returns such as industry factors (Bernard (1987)).

B.3. Estimation Approach

I obtain posterior distributions for the model parameters using a Bayesian procedure. This procedure exploits prior beliefs that firms within a given industry tend to have similar residual variations, resulting in extreme estimates being “pooled” toward the sample average. Since different prior beliefs may yield different inferences, I report results consistent with varying degrees of strength regarding these prior beliefs.⁹

The likelihood function $l(V|\Sigma)$ is multivariate normal,

$$l(V|\Sigma) \propto |\Sigma|^{1/2} \exp\{-\frac{1}{2}(Y - F\iota)' \Sigma^{-1}(Y - F\iota)\}. \quad (4)$$

The prior for Σ is specified following Barnard, McCulloch, and Meng (1997) and is written in terms of two matrices: $\Sigma = SRS$, where S is a diagonal matrix with standard deviations on its diagonal and R is a correlation matrix, both of dimension $(\sum_{i=1}^N T_i \times \sum_{i=1}^N T_i)$. Prior beliefs are specified separately for S and R .

I specify a lognormal prior for the N distinct elements of S ,

$$\log(\sigma_i) \sim N(\bar{s}, \delta_\sigma) \quad \forall i = 1, \dots, N. \quad (5)$$

The specification of the parameters \bar{s} and δ_σ is complicated by the fact that we have no available data, prior to the IPO, to construct informative prior beliefs. Moreover, since it is unlikely that a seasoned firm’s residual variability is as high as that of an IPO, I choose not to make use of such information. Consequently, I take an Empirical Bayes (EB) approach as in Jorion (1986), McCulloch and Rossi (1991), and Pastor and Stambaugh (1999). For each firm, I calculate the time-series standard deviation of its abnormal return relative to the matched benchmark portfolio. Then, I calculate the grand mean and variance of these estimates. \bar{s} and δ_σ are specified such that my prior beliefs regarding the standard deviations are centered at the observed mean of the residual standard deviations. The advantage of this

⁹ See Kothari and Shanken (1997), Pastor and Stambaugh (1999), Stambaugh (1997), and Harvey and Zhou (1990) for discussions regarding the use of informative priors.

EB implementation is that it allows for shrinkage toward a central tendency that is completely determined by the data. Obviously, alternative informative beliefs that specify different central tendencies may prove useful in circumstances where the researcher knows a priori more about IPO residual variations.

Finally, shrinkage is induced by using either one-half or one-sixteenth of the observed variance of the standard deviations. In the analysis below, results will be reported for both levels of shrinkage (denoted by “mild” and “strong” correspondingly).^{10,11}

Prior specification for R requires a prior for the common correlation coefficient ρ . The only information that I am willing to impound in my inferences is that the residual covariance matrix is positive definite. I employ the following uniform prior.

$$\rho \sim Un(\lfloor \rho^* 1 \rfloor), \quad (6)$$

where ρ^* is given in the Appendix, Section C.

Using the Bayes Theorem, I combine the prior beliefs and likelihood function to obtain the joint posterior distribution for the parameters of the model,

$$p(\Sigma|Y, F) \propto l(V|\Sigma)p(\Sigma). \quad (7)$$

Although analytically intractable, the posterior distribution can be simulated using the Gibbs sampler.¹² To operationalize the sampler, the conditional distributions of the parameters are specified as follows:

The conditional distribution for ρ is proportional to

$$\rho|S \propto |R(\rho)|^{-1/2} \exp\{-\frac{1}{2}(Y - F\iota)'(SR(\rho)S)^{-1}(Y - F\iota)\}, \quad (8)$$

¹⁰ The levels of shrinkage were chosen, given the data, to reflect two extreme views regarding the dispersion of residual variance.

¹¹ To limit the effect that outliers may have on the estimation of these standard deviations, I winsorize extreme observations that lie beyond two standard deviations from a firm sample average abnormal return.

¹² The idea behind the Gibbs sampler is as follows. Suppose that we are interested in the posterior distribution of a vector of unknown random parameters $\Lambda = (\lambda_1, \lambda_2, \dots, \lambda_d)$. Denote by $\lambda_i|(\cdot)$ the conditional posterior distribution of λ_i . Then starting with an initial starting point $\Lambda = (\lambda_1^0, \lambda_2^0, \dots, \lambda_d^0)$, the algorithm iterates the following loop:

- a. Sample λ_1^{i+1} from $p(\lambda_1|\lambda_2^i, \dots, \lambda_d^i)$
- b. Sample λ_2^{i+1} from $p(\lambda_2|\lambda_1^{i+1}, \lambda_3^i, \dots, \lambda_d^i)$
- ⋮
- d. Sample λ_d^{i+1} from $p(\lambda_d|\lambda_1^{i+1}, \dots, \lambda_{d-1}^{i+1})$

The vectors $\lambda^0, \lambda^1, \dots, \lambda^t, \dots$ are a realization from a Markov chain. It can be shown (Geman and Geman (1984)) that the joint distribution of $(\lambda_1^i, \dots, \lambda_d^i)$ converges to $p(\lambda_1, \dots, \lambda_d|DATA)$, that is, the joint posterior distribution, as $i \rightarrow \infty$ under mild regularity conditions. For applications of the Gibbs sampler see Kandel, McCulloch, and Stambaugh (1995), Gelfand and Smith (1990), and Gelfand et al. (1990).

where $R(\rho)$ is the correlation matrix and the parentheses emphasize that it is a function of ρ . I draw from this conditional distribution using the Griddy-Gibbs approach (see Ritter and Tanner (1992) and Tanner (1996)). The details are given in the Appendix, Section D.

The conditional distributions for each $\sigma_i \forall i = 1, \dots, N$ are proportional to

$$\sigma_i | R(\rho), S_{-i} \propto \sigma_i^{-(T_i+1)} \times \exp \left\{ -\frac{1}{2} \left[\frac{(\log(\sigma_i) - \bar{s})^2}{\delta_\sigma} + (Y - F_i)' (S_{-i} R(\rho) S_{-i})^{-1} (Y - F_i) \right] \right\}, \quad (9)$$

where S_{-i} denotes the standard deviation matrix conditional on the other $N - 1$ standard deviation draws. As with the conditional distribution for ρ , I draw from this density using the Griddy-Gibbs approach.

I obtain the Gibbs sampler's initial values by first running firm by firm OLS regressions and then setting the initial values for the elements in S equal to the sample standard deviations. The initial value for ρ is set to zero. The sampler is iterated 600 times, and the first 100 draws are discarded.¹³

B.4. Model Estimation

Before discussing the regression results, it is worthwhile to explore the effect of different amounts of shrinkage on parameter estimation.

Figure 1 displays the dispersion of the residual standard deviation posterior means as a function of increasing shrinkage. I start from the left, where no shrinkage is used (denoted "OLS"); shrinkage is increased as we move to the right, resulting in stronger shrinkage toward the prior mean (16 percent in this case). With strong shrinkage, information is shared across different firms, and the posterior means are tightly clustered relative to the least squares estimates. Consequently, if the basic premise that firms within industries have similar residual variations is correct, then incorporating this information will benefit the precision of estimation.

Figure 2 presents both the prior and posterior distributions for ρ . The left panel plots the uniform prior described in equation (6) which puts mass only on those values of ρ that guarantee that the variance-covariance matrix is positive definite. The right plot gives the marginal posterior distribution given the data.

Table III gives the regression results. For each shrinkage scenario, I calculate the 113 residual standard deviation posterior means and report their grand average. The cross-sectional standard deviations of these means are given in parentheses. I also report the posterior mean and standard deviation of the correlation coefficient ρ .

¹³ Convergence of the Gibbs sampler in particular, and Markov chain Monte Carlo methods in general, has received considerable attention (see, e.g., Gilks, Richardson, and Spiegelhalter (1996)). Convergence was monitored by comparing the results of multiple chains started at different (random) initial values and also by the use of time-series plots of the parameter draws.

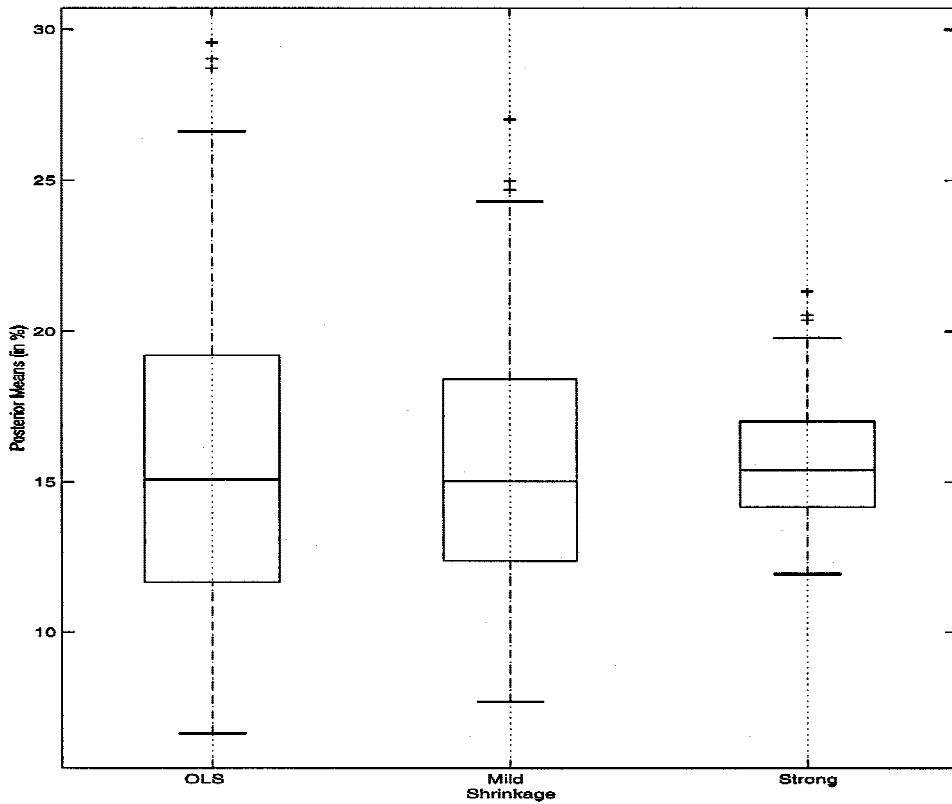


Figure 1. Shrinkage Estimation of Residual Standard Deviations. For the 113 computer and data processing IPOs the figure exhibits the dispersion of the residual standard deviations' posterior means as a function of the degree of their shrinkage. The figure shows the effect of increasing shrinkage (moving from left to right) on the dispersion of the estimates. Each boxplot gives lines at the lower quartile, median, and upper quartile of the distribution. The "whiskers" are lines extending from the boxes showing the extent of the rest of the data. The labels on the x-axis describe the degree of shrinkage used. Note that the leftmost boxplot corresponds to estimation via OLS (reflecting diffuse priors).

Consider the posterior beliefs for the monthly residual standard deviations. The reported average of 15.7 percent is large and nearly 50 percent larger than the residual standard deviation of an average firm traded on the NYSE or AMEX as reported in Kothari and Warner (1997).¹⁴ The dispersion of the residual standard deviation posterior means is as high as 5.6 percent, and it declines to 4.2 percent as shrinkage is increased.

¹⁴ See their Table 5. They report an average of 10.4 percent standard deviation of monthly abnormal returns over their test period.

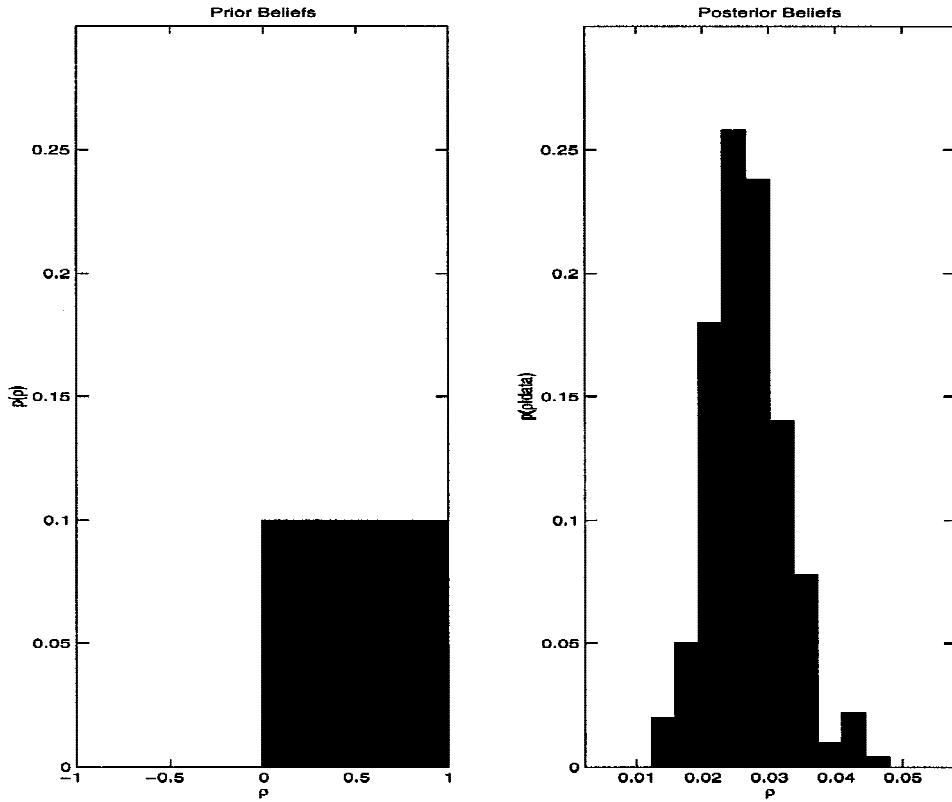


Figure 2. Shrinkage Estimation of the Common Correlation Coefficient. For the 113 computer and data processing IPOs the figure exhibits the prior and posterior beliefs regarding the common correlation coefficient ρ . The left plot shows the uniform prior described in equation (6). The right plot shows the marginal posterior beliefs given the data.

Finally, consider the estimated average correlation. The reported range of 2.5 percent to 2.7 percent is not large compared to prior studies of intraindustry correlation.¹⁵ In Section II.D.1 I show that these small correlations still affect the distribution of the sample abnormal mean return.

C. Predictive Distribution for Long-Horizon Returns

In this section I describe how to simulate buy-and-hold returns which are used later to construct the density of the sample mean abnormal return.

Let M denote the number of draws on Σ retained from the Gibbs sampler output and let Σ_j be the j th such draw. Then, conditional on the benchmark realizations, I draw K vectors of firm returns Y_j each of length $(\sum_{i=1}^N T_i \times 1)$

¹⁵ Bernard (1987) reports an average intra-industry correlation equal to 18 percent for market model residuals.

Table III
Descriptive Statistics for Posterior Distributions of
Correlations and Residual Standard Deviations

For each IPO, the time-series standard deviation of its abnormal return is calculated relative to a matched benchmark portfolio. “Mild” (“Strong”) shrinkage of σ_i is obtained by setting δ_σ equal to half (one-sixteenth) of the variance of these IPO residual standard deviations. The grand average of the resulting posterior means is reported below. The cross-sectional standard deviations of these means are given in parentheses. For example, with “Mild” shrinkage, the average of 113 residual standard deviation posterior means is 15.7 percent with a cross-sectional standard deviation equal to 5.6 percent. Also reported, for each shrinkage scenario, are the correlation coefficient’s mean and standard deviation.

Shrinkage of σ_i	Residual Std. (%)	ρ (%)	Shrinkage of σ_i	Residual Std. (%)	ρ (%)
Mild	15.7 (5.6)	2.7 (0.58)	Strong	15.7 (4.2)	2.5 (0.58)

from the likelihood function in equation (4). By repeating this procedure M times I obtain a set $\{Y_1, Y_2, \dots, Y_{KM}\}$ of draws from an “averaged” likelihood function that incorporates the additional parameter uncertainty. This density is called the “predictive” distribution in the Bayesian literature because, given the modeling assumptions, it generates all possible realizations of the vector Y .¹⁶

I calculate long-horizon firm returns by compounding the single-period returns for each firm. Consequently, these $K \times M$ draws yield a set of $K \times M$ abnormal mean returns that is used to construct the predictive distribution of the sample mean. This distribution is centered at zero by construction and therefore is free from the “new listing” and “rebalancing” biases identified by Barber and Lyon (1997). In the analysis below, K is set equal to four (rather than one) to reduce variation due to simulation error. Because M is equal to 500, the predictive density is based on 2,000 draws.

D. Statistical Inferences

The predictive density constructed in the previous section provides the basis for statistical inferences. The null of no abnormal performance will be “called into question” at the α percent level if the average abnormal return obtained from the original IPO sample is greater [smaller] than the $(1 - \alpha)$ [α] percentile abnormal return observed in the constructed distribution.

It is important to emphasize that the current implementation of the approach is to assess the validity of a *single* model relative to observed data. Concentrating on a specific asset pricing model is in the spirit of Box (1980), who argued:

¹⁶ See Tanner (1996) and Box (1980).

Table IV
Predictive Densities of Five-Year Average Buy-and-Hold
Returns under Two Shrinkage Scenarios

Predictive densities are generated for each of the two shrinkage scenarios based on 2,000 simulated average buy-and-hold abnormal returns (see Table III and Section II.B for additional details). Each row provides the 1st, 5th, 50th, and 95th percentiles in addition to the means of these densities. The rightmost column gives the sample abnormal performance calculated using the firms' matched benchmark portfolio returns.

Shrinkage of σ_i	Simulated Distribution					Abnormal Return (%)
	1%	5%	50%	95%	Mean	
Mild	-49.4	-36.8	-4.7	50.5	0.0	-5.8
Strong	-49.1	-38.5	-4.9	48.4	0.0	-5.8

In making a predictive check it [is] not necessary to be specific about an alternative model. This issue is of some importance for it seems a matter of ordinary human experience that an appreciation that a situation is unusual does not necessarily depend on the immediate availability of an alternative. (p. 387).

The predictive approach can be easily extended to select among *competing* asset pricing models using Bayesian posterior odds ratios (Harvey and Zhou (1990), McCulloch and Rossi (1991)). Yet, as Box (1980) argues, the difficulty with posterior odds ratios is that the researcher is required to specify *a priori all* the possible models that may have generated the observed sample. Otherwise, the interpretation of these odds is unclear. Odds ratios that declare the Fama and French three-factor model 1,000 times more probable to have generated the data than the CAPM tell the reader nothing about the validity of the three-factor model in describing IPO returns. The view taken in this paper is that we take the "best" model that we have and pit it against the data in search of refinements to our understanding of the underlying process of interest. Alternative models usually emerge when the existing paradigm is discredited by the data.

Table IV reports descriptive statistics for the predictive densities under different shrinkage scenarios.

In each row I present the properties of these densities for two different levels of residual standard deviation shrinkage. I report the first, fifth, 50th, and 95th percentiles as well as the mean. The rightmost column gives the sample abnormal performance calculated using the firms' matched benchmark returns.

Four interesting results emerge. First, the range of these distributions is very large. For example, the results in the first row indicate that, at the five percent level, we cannot reject the characteristic-based model with a realized abnormal performance as low as -37 percent or as high as 50 percent. Second, the median firm under both shrinkage scenarios earns a neg-

Table V
Predictive Densities of Five-Year Average Buy-and-Hold
Return under Two Shrinkage Scenarios Assuming Independence

Predictive densities are generated for each of the two shrinkage scenarios based on 2,000 simulated average buy-and-hold abnormal returns (see Table III and Section II.B for additional details). These simulations are identical to those reported in Table IV except for the restriction that the variance–covariance matrix be diagonal ($\rho = 0$). Each row provides the first, fifth, 50th, and 95th percentiles in addition to the means of the resulting predictive densities.

Shrinkage of σ_i	Simulated Distribution				
	1%	5%	50%	95%	Mean
Mild	−39.1	−29.7	−2.0	37.9	0.0
Strong	−36.5	−29.0	−3.6	39.2	0.0

ative abnormal return. Recall that these densities were generated under the null of no abnormal performance, which means that for skewed-right distributions we should *expect* to observe that the median firm underperforms. Third, the resulting shape of the predictive density is insensitive to the different prior specifications employed in this paper. Fourth, and most importantly, under all shrinkage scenarios and at the five percent significance level, the observed abnormal performance fails to reject this model.

D.1. Do the Residual Covariations Matter?

This section explores whether estimated covariations affect inferences. Table V gives the predictive distributions generated by constraining the residual covariance matrix estimated earlier to be diagonal, that is, imposing independence.

Comparison of the results in Tables IV and V reveals that firm cross-sectional correlations (reported in Table III) have a large effect on inferences. Imposing diagonal covariance matrices resulted in a reduction in uncertainty regarding the sample mean abnormal return. Specifically, the first percentile has declined by 10 percent to 13 percent across the different shrinkage scenarios, whereas the fifth percentile has declined by eight percent to nine percent across different shrinkage scenarios. Similarly, the 95th percentile has decreased by as much as 12 percent.

D.2. A Check on Simulation Error

This section explores how sensitive these results are to simulation error. The extent of simulation error is examined as follows. For each shrinkage scenario I use the simulated means and resample 100 times, with replacement, samples of abnormal mean returns each containing 2,000 observations. For each such sample I calculate the first, fifth, 50th, and 95th percentiles. Summary statistics regarding the variation of these statistics across the different simulations are presented in Table VI.

Table VI
Sensitivity to Simulation Error

Each of the two possible shrinkage scenarios described in Section II.B yields a set of 2,000 simulated means. One hundred samples of abnormal mean returns, each containing 2,000 observations, are then drawn, with replacement. For each such sample the first, fifth, 50th, and 95th percentiles are calculated. Summary statistics for the variation of these percentiles across the different simulations are presented below. For example, in the first row, the first percentile was on average -49.4 percent with a 1.2 percent standard deviation across the 100 bootstraps.

Shrinkage of σ_i	1%		5%		50%		95%	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Mild	-49.4	1.2	-37.1	1.2	-4.7	0.7	51.1	2.0
Strong	-48.8	1.4	-38.1	0.8	-4.9	0.6	48.5	1.9

The sensitivity results reveal that all percentiles are measured accurately. The accuracy increases for quantiles that are closer to the mode of the distribution. The first and fifth percentiles and also the median are measured more accurately than the 95th percentile, which is more sensitive to extreme observations. As I show in the next section, accuracy is improved further as the sample size is increased.

III. Full Sample Analysis

In this section I conduct inferences regarding the long-term returns to a sample of 1,521 IPOs issued over the period from 1975 to 1984. The sample is the one used by Ritter (1991), who finds -27.4 percent size- and industry-adjusted three-year abnormal returns for these firms.¹⁷ Below, this interesting result is revisited using the characteristic-based model. Given the evidence in Loughran and Ritter (1995) that abnormal returns persist for five years after the event, I examine a five-year horizon as well.

A. Sample Description

Table VII gives the distribution of the sample firms by size and book-to-market. For each IPO I used the pre-issue book value reported by Ritter (1991). Size was determined using the first closing price available from CRSP. The 5×5 size and book-to-market cutoffs were determined using NYSE firm breakpoints. Each IPO was first allocated to a size quintile and then allocated to a book-to-market quintile. Panel A reports the number of observations in each cell. The last row in this panel gives the number of missing book-to-market observations. Panel B reports the mean market capitalization within each size quintile. Panel C reports the mean book-to-market for each cell.

¹⁷ The data are available at <http://www.cba.ufl.edu/fire/faculty/ritter.htm>. Three IPOs were deleted because I could not find monthly data for those firms on CRSP. Two additional IPOs were deleted since I had fewer than four monthly observations to use in the regressions below.

Table VII
Descriptive Statistics for Whole IPO Sample, 1975–1984

The sample consists of 1,521 firms. For each IPO the pre-issue book value is obtained from Ritter (1991), and market capitalization (size) is determined using the first closing price available from CRSP. The 5×5 size and book-to-market cutoffs were created using NYSE firm breakpoints. Each IPO is first allocated to a size quintile and then into a book-to-market quintile. Panel A reports the number of observations in each cell. Panel B reports the mean market capitalization within each size quintile. Panel C reports the mean book-to-market for each cell.

Panel A: Number of IPOs					
Book-to-Market	Size				
	Smallest	2	3	4	Largest
Low	990	183	65	21	6
2	41	3	4	3	0
3	39	1	4	0	0
4	22	1	0	0	0
High	26	0	0	0	0
Missing Book-to-Market Ratio	106	4	1	0	0

Panel B: Mean Size (mil\$)				
Smallest	Size Quintiles			
	2	3	4	Largest
24.6	104.7	223.1	451.8	1099.5

Panel C: Mean Book-to-Market					
Book-to-Market	Size				
	Smallest	2	3	4	Largest
Low	0.11	0.12	0.12	0.07	0.11
2	0.73	0.70	0.70	0.63	—
3	1.04	0.93	0.98	—	—
4	1.37	1.33	—	—	—
High	1.91	—	—	—	—

The evidence in Table VII indicates that the majority of the sample firms are concentrated in the smallest size and book-to-market quintiles, with approximately 80 percent of the sample belonging to the smallest size quintile.

To set the stage for the inferences in the next section, five-year returns to these IPOs are calculated versus the five-year return on the NYSE-AMEX value-weighted index. The purpose of this comparison is to provide further information regarding the average long-horizon return of this sample. The benchmark return that accounts for firm characteristics is calculated later in this section. Table VIII reports the sample average and median return and also the average market return. The last three columns give the average

Table VIII
IPO Sample Five-Year Aftermarket Performance Relative
to the NYSE-AMEX Value-Weight Index, 1975–1984

Buy-and-hold return is calculated for each IPO starting in the month after the issuance through the end of its fifth seasoned year (60 calendar months). The benchmark return is the buy-and-hold return on the NYSE-AMEX value-weight index. If a firm delists prematurely the buy-and-hold return is calculated until the delisting month both for the sample firm and its benchmark.

Number of IPOs	IPO Average Five-Year Return (%)	IPO Median Five-Year Return (%)	NYSE-AMEX VW Five-Year Return (%)	Five-Year Abnormal Return (%)		
				Average	Standard Deviation	Skewness
1,521	27.2	–37.1	92.9	–65.7	245.9	11.9

excess return in addition to its cross-sectional standard deviation and skewness (disaggregated information regarding industry performance is given in the Appendix, Section B).

From Table VIII we see that the five-year underperformance relative to this market index is –65.7 percent. Further inspection of the underperformance by industry (see the Appendix, Section B) reveals that 15 out of 17 industries underperformed. In fact, in only three industries, financial institutions, insurance, and drug and genetic engineering, did the median firm earn a positive *raw return* over this five-year period. The last two columns show that the standard deviation and skewness of the abnormal returns distribution are both extremely large. The latter statistic confirms that long-horizon excess returns are not normally distributed.

B. Statistical Inferences

The first step is to decompose the sample into industries, conduct inferences within industries and then aggregate the results. I form 17 industry classifications based on Ritter (1991) and Spiess and Affleck-Graves (1995). SIC codes for the IPO sample are from Compustat and Ritter (1991). The list of the original Ritter three-digit industry classifications and the additions made is given in the Appendix, Section A.

For each industry, the methodology outlined in Section II is used to estimate firm residual variations and cross-correlations. Then, using the parameters' posterior distributions and the procedure outlined in Section II.C, I simulate 2,000 long-horizon average returns for each industry.¹⁸

¹⁸ Note that the last industry definition “Other” contains IPOs that were not associated with any of the previous 16 industries. Because it was assumed that the source of residual cross-correlation was due to industry factors, the residual correlation for this industry is set equal to zero.

Table IX
Predictive Densities of Five-Year Average Buy-and-Hold
Return under Various Shrinkage Scenarios

Predictive densities are generated for each of the two shrinkage scenarios based on 2,000 simulated average buy-and-hold abnormal returns (see Table III and Section II.B for additional details). Each row provides the first, fifth, 50th, and 95th percentiles in addition to the means of these densities. The rightmost column gives the sample abnormal performance calculated using the firms' matched benchmark portfolio returns.

Shrinkage of σ_i	Simulated Distribution					Abnormal Return (%)
	1%	5%	50%	95%	Mean	
Mild	-18.4	-14.0	-0.9	16.7	0.1	-4.9
Strong	-18.0	-13.3	-1.1	15.3	-0.2	-4.9

Table IX presents the summary statistics regarding the distribution of the sample mean aggregated across all 17 industries. The table reports the first, fifth, 50th, and 95th percentiles of the distribution as well as the mean. The realized abnormal return corrected for each firm's characteristics is given in the last column. I present results corresponding to the two different shrinkage scenarios.

Under both shrinkage scenarios the observed IPO returns are consistent with the characteristic-based model. This result confirms the finding in Brav and Gompers (1997), who argue that the five-year returns to recent IPOs do not differ from the average return on benchmarks constructed based on size and book-to-market ratios.

IV. Comparison with an Alternative Approach

In this section, I compare the proposed methodology to an alternative, nonparametric, bootstrap approach advanced by Ikenberry et al. (1995). This section is organized as follows. Section A provides the construction of the bootstrapped density that is then used to conduct inferences regarding the abnormal performance of the IPO sample. Section B discusses the differences in the results between the suggested methodology and the bootstrap approach.¹⁹

A. Inferences Based on the Bootstrapped Density

Given that each IPO has been assigned to a size and book-to-market portfolio allocation, I randomly select from that allocation a *replacement* firm with the same return horizon as the original firm. This replacement is repeated for all IPOs in the sample, resulting in a new "pseudo" sample. I

¹⁹ I thank Mark Mitchell and Erik Stafford for the data used in this section.

Table X
Bootstrapped Distribution of Five-Year Average
Buy-and-Hold Abnormal Return

The bootstrapped distribution is constructed by first allocating each IPO to one of 25 size and book-to-market portfolios. Then, for each of these IPOs, a replacement firm is randomly selected from their respective portfolio allocations with the same return horizon as the original firm. If a replacement firm delists prematurely, the proceeds from the delisting firm are re-invested in another randomly selected firm from the same portfolio for the remaining period. This replacement results in a new “pseudo” sample containing 1,521 replacement firms, and this sample’s average abnormal performance relative to the original size and book-to-market portfolios is recorded. Repeating this replacement 2,000 times results in 2,000 average abnormal returns, which are used to construct the bootstrapped density for the sample mean. The table reports the first, fifth, 50th, and 95th percentiles of the distribution in addition to its standard deviation, mean, and skewness coefficient. The IPO abnormal return adjusted by size and book-to-market is given in the last column.

Simulated Distribution (2,000 bootstraps)							Abnormal Return
1%	5%	50%	95%	Mean	Std	Skewness	(%)
-11.5	-8.4	-0.3	9.0	-0.1	5.2	0.2	-4.9

proceed to calculate the latter sample abnormal performance relative to the original size and book-to-market portfolios. Repeating this replacement 2,000 times results in 2,000 average abnormal returns that are used to construct the bootstrapped density for the sample mean. The null of no abnormal performance is rejected at the α percent level if the average abnormal return obtained from the original sample is greater [smaller] than the $(1 - \alpha)$ [α] percentile abnormal return observed in the bootstrapped distribution.

Table X presents the bootstrapped distribution. The table reports the first, fifth, 50th, and 95th percentiles in addition to its standard deviation, mean, and skewness coefficient. The abnormal return adjusted by size and book-to-market is given in the last column.

Not surprisingly, the negligible average abnormal return is consistent with the attribute-based model as well. Yet, comparison of the fractiles in Table IX and Table X reveals that the densities differ substantially. The next section explains the reason for these differences.

B. Differences between the Two Methodologies

Comparison of Table IX and Table X reveals that the range and shape of the bootstrapped density are different from those derived using the proposed approach. The bootstrapped density is less skewed, and the range between its fifth and 95th percentiles is shorter by approximately 30 percent relative to the range reported in Table IX. The differences between the bootstrapped distribution and the predictive density are not surprising. As argued in Section I, the replacement of sample firms with nonissuing firms neglects two

Table XI
Predictive Densities of Five-Year Average Buy-and-Hold
Return under Two Shrinkage Scenarios

Each row gives the properties of the simulated densities for two levels of shrinkage under the restriction that IPO returns be independent ($\rho = 0$). Each row provides the 1st, 5th, 50th, and 95th percentiles in addition to the means of these densities. The rightmost column gives the sample abnormal performance calculated using the attribute matched portfolio returns.

Shrinkage of σ_i	Simulated Distribution					Abnormal Return (%)
	1%	5%	50%	95%	Mean	
Mild	-14.5	-10.6	-1.0	14.0	0.4	-4.9
Strong	-14.4	-10.8	-0.6	11.4	-0.1	-4.9

important features of the data. First, cross-sectional correlation in IPO abnormal returns is not taken into account by the bootstrap methodology. Second, IPO residual standard deviations might be larger than those of the replacing firms.

Consider first the effect that residual cross-correlation has on inferences. Table XI presents the properties of the simulated density generated using the proposed methodology, imposing independence across firm abnormal returns.

Comparison of Tables IX and XI provides direct support for the importance of IPO cross-correlations. Imposing lack of correlation leads to a substantial reduction in uncertainty regarding the sample average. Indeed, about half of the observed difference between the bootstrapped and predictive densities is driven by imposing independence.

To verify the conjecture that differences in residual standard deviations between IPOs and their replacement firms drive the rest of the discrepancy between the bootstrap and predictive densities, the following analysis was conducted. For all IPOs and replacing firms, I calculated the residual standard deviations relative to their respective size and book-to-market benchmarks. Then, for each IPO, I determined the percentile of its standard deviation relative to the standard deviations of its potential replacing firms, yielding 1,521 percentiles. If IPO standard deviations do not differ systematically from replacing firms' standard deviations, we should expect these percentiles to be equally distributed between one percent and 100 percent. Accordingly, in Figure 3, I plot the histogram of these percentiles and the expected count of these percentiles in each bin (the dashed horizontal bar).

The histogram of percentiles indicates that there are disproportionately more IPOs with high residual standard deviations than predicted by the matching firm approach. The inability of the bootstrap approach to correctly account for these high residual standard deviations results in an empirical distribution that understates the uncertainty regarding the IPO average abnormal return.

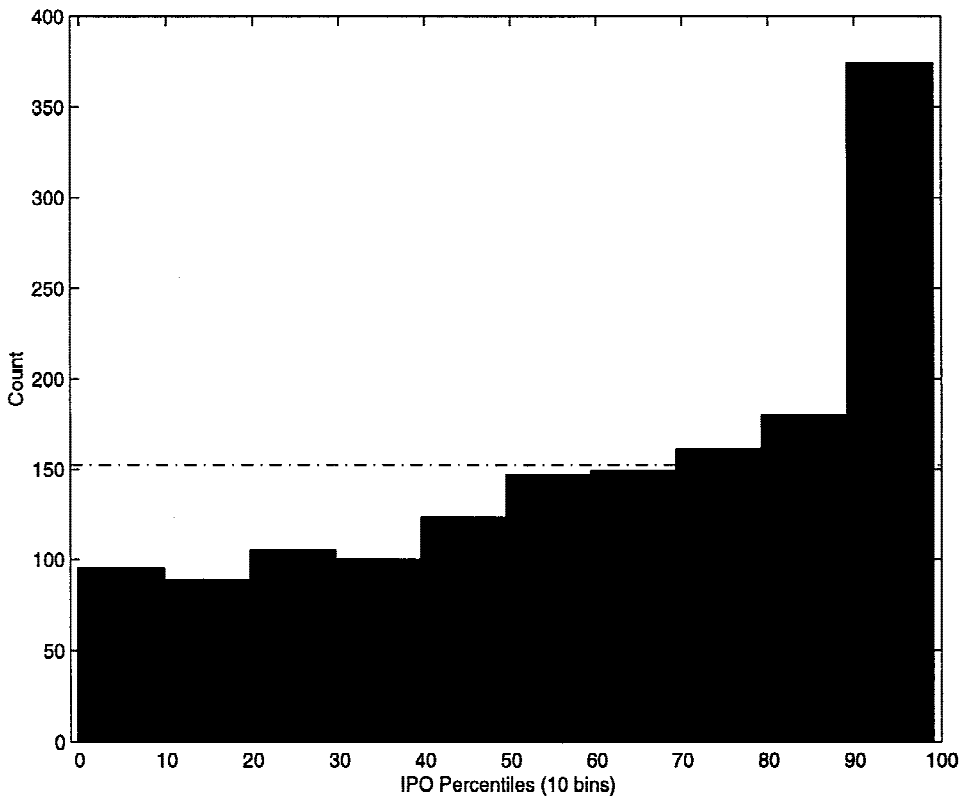


Figure 3. Comparison of IPO and Replacing Firms Residual Standard Deviations. The sample includes 1,521 IPOs. For each IPO and all possible replacing firms in the relevant size and book-to-market allocation (see Section IV), I calculate the residual standard deviations relative to the size and book-to-market benchmark return. Then, for each IPO, I determine the percentile of its standard deviation relative to the standard deviations of its potential replacing firms. By repeating this calculation I obtain 1,521 percentiles. The figure provides a histogram of these 1,521 percentiles. If IPO standard deviations do not differ systematically from replacing firm standard deviations, we should expect these percentiles to be equally distributed between 1 percent and 100 percent. The (dashed) horizontal bar gives the expected count of percentiles in each bin (approximately 152).

The results in this section highlight the potential dangers of applying the bootstrap methodology to samples of firms whose abnormal performance is cross-sectionally correlated and whose post-event residual variations may be systematically different than their attribute-matched control firms.

V. Alternative Specification

The methodology presented in this paper is not limited to characteristic-based models. In this section I show how to extend the approach under the assumption that the asset pricing model is a k -factor model, entertaining the Fama and French (1993) three-factor model in particular.

This model consists of three factors. The first, RMRF, is the value-weighted monthly return on the market portfolio less the one-month T-bill rate. The second, HML, is the difference in returns between high and low book-to-market firms, and the third, SMB, is the difference between the returns of small and large firms. Thus, for any asset excess return in period t , $y_t - r_{f,t}$, we have²⁰

$$y_t - r_{f,t} = \alpha + \beta_1 \text{RMRF}_t + \beta_2 \text{HML}_t + \beta_3 \text{SMB}_t + v_t. \quad (10)$$

The estimation of the model parameters is conducted under the null that this model generated the observed sample returns. The intercept in this regression, α , is a measure of monthly abnormal performance that controls for size, book-to-market, and market effects in average returns. A nonzero intercept would indicate that this model does not price the sample firms, which would be a contradiction of the null hypothesis. Consequently, regression (10) is estimated *without the intercept term*.

A. Regression Setup

Keeping the same notation as in Section II.B.2, let N be the number of sample firms with T_i $i = 1, \dots, N$ monthly observations. Define y_i as the $(T_i \times 1)$ column vector of firm i 's returns in excess of the risk-free rate, f_i as the $(T_i \times 3)$ matrix of factor mimicking portfolios' returns, and β_i as the (3×1) vector of factor loadings. The firm's excess returns are modeled as follows:

$$y_i = f_i \beta_i + v_i \quad \forall i = 1, \dots, N. \quad (11)$$

The system of all N assets is written using the SUR setup:

$$Y = FB + V, \quad (12)$$

where Y is a $(\sum_{i=1}^N T_i \times 1)$ stacked vector of firm returns. F is a $(\sum_{i=1}^N T_i \times 3N)$ block diagonal matrix of factor realizations, and B is a $(3N \times 1)$ vector of stacked factor loadings β_1, \dots, β_N . V is a $(\sum_{i=1}^N T_i \times 1)$ stacked vector of firm residuals.

As in Section II.B, I assume a multivariate normal distribution for V with mean zero and a $(\sum_{i=1}^N T_i \times \sum_{i=1}^N T_i)$ variance-covariance matrix Σ . The residuals are assumed to be temporally independent and to share a common contemporaneous correlation, denoted ρ , that reflects joint covariation in returns driven by unpriced factors in returns.

²⁰ The implicit assumption here is that the short term rate is independent of the factor realizations.

B. Estimation Approach

The likelihood function $l(V|\beta, \Sigma)$ is multivariate normal,

$$l(V|\beta, \Sigma) \propto |\Sigma|^{-1/2} \exp\{-\frac{1}{2}(Y - FB)' \Sigma^{-1} (Y - FB)\}. \quad (13)$$

The prior for B is formed using a hierarchical multivariate normal setup. Each β_i is assumed to be an independent and identical draw from the following multivariate normal distribution:

$$\beta_i \sim N(\bar{\beta}, \Delta_\beta) \quad \forall i = 1, \dots, N, \quad (14a)$$

where $\bar{\beta}$ is a (3×1) mean vector and Δ_β is a (3×3) diagonal matrix with elements $(\delta_1, \dots, \delta_3)$. I add an additional layer of uncertainty by modeling the precision of the prior beliefs regarding $\bar{\beta}$ (i.e., Δ_β^{-1}) as a random draw from the following Wishart prior,²¹

$$\Delta_\beta^{-1} \sim W(\nu_\beta, \Psi^{-1}). \quad (14b)$$

As Lindley and Smith (1972) note, the specification of ν_β and Ψ^{-1} determines the amount of shrinkage used. Specifically, Ψ^{-1} determines the location of the prior distribution while ν_β , the degrees of freedom, determines its dispersion. Furthermore, one can interpret the above prior as a posterior distribution obtained after observing an imaginary sample of size ν_β and mean centered at Ψ^{-1} .

Since Δ_β^{-1} determines the extent of the shrinkage, setting ν_β large along with a large location value in Ψ^{-1} results in high degree of shrinkage. Conversely, small values for ν_β and small diagonal elements in Ψ^{-1} result in low amount of shrinkage and large variation across the factor loadings.

The assessment of the diagonal elements in Ψ^{-1} is complicated by the fact that it is impossible to observe the dispersion of IPO factor loadings before the IPO date. I take the Empirical Bayes approach here as well. I estimate N individual factor loadings from separate OLS regressions and calculate the sum of squares for each of the three sets of loadings about their grand average. Shrinkage is induced by employing either half or a quarter of these three sums of squares and then using the reciprocals as the diagonal entries in Ψ^{-1} . These levels of shrinkage will be referred to later as “mild” and “strong,” respectively. Finally, the degrees of freedom ν_β is set equal to N .

$\bar{\beta}$ is modeled as a draw from the following multivariate normal distribution:

$$\bar{\beta} \sim N(\bar{\bar{\beta}}, \varphi_{\bar{\beta}}) \quad (14c)$$

²¹ See Zellner (1971, p. 389) for the properties of this distribution.

The vector $\bar{\beta}$ specifies my beliefs about the central tendency of the factor loadings and $\varphi_{\bar{\beta}}$ specifies the strength of this prior information. Because I have no prior information regarding $\bar{\beta}$, I let the data determine the central tendencies. Hence, the elements in $\varphi_{\bar{\beta}}^{-1}$ are set to zero.

The specification of the prior for Σ and the levels of shrinkage of the residual standard deviations are identical to the derivation in Section II.B for the characteristic-based model.

Using the Bayes Theorem, I combine the prior beliefs and likelihood function to obtain the joint posterior distribution for the parameters and hyperparameters of the model,

$$p(B, \Sigma, \bar{\beta}, \Delta_{\beta} | Y, F) \propto l(V | B, \Sigma) p(\Sigma) p(B | \bar{\beta}, \Delta_{\beta}) p(\bar{\beta}) p(\Delta_{\beta}). \quad (15)$$

As in Section II.B, I now specify the conditional distributions of the parameters and hyperparameters.

The multivariate normal distribution for B is

$$B | \bar{\beta}, \Delta_{\beta}, \Sigma, Y, F \sim N(b^*, (F' \Sigma^{-1} F + I_N \otimes \Delta_{\beta}^{-1})^{-1}), \quad (16a)$$

where

$$b^* = (F' \Sigma^{-1} F + I_N \otimes \Delta_{\beta}^{-1})^{-1} ((F' \Sigma^{-1} F) \hat{b}_{gls} + (I_N \otimes \Delta_{\beta}^{-1}) (\iota_N \otimes \bar{\beta})), \quad (16b)$$

ι_N is a $(N \times 1)$ vector of ones, I_N is an $(N \times N)$ identity matrix and \hat{b}_{gls} is a vector of GLS regression coefficients, namely, $\hat{b}_{gls} = (F' \Sigma^{-1} F)^{-1} F' \Sigma^{-1} Y$.

The multivariate normal distribution for the hyperparameter vector $\bar{\beta}$ is

$$\bar{\beta} | B, \Delta_{\beta} \sim N\left(\frac{1}{N} (\iota_N \otimes I_3)' B, \frac{1}{N} \Delta_{\beta}\right), \quad (16c)$$

where I_3 is a (3×3) identity matrix.

The Wishart distribution for the precision matrix Δ_{β}^{-1} is

$$\Delta_{\beta}^{-1} | B, \bar{\beta} \sim W(\nu_{\beta} + N, ((D - \bar{\beta} \otimes \iota'_N)(D - \bar{\beta} \otimes \iota'_N)' + \Psi)^{-1}), \quad (16d)$$

where D is a $(3 \times N)$ matrix whose N columns, each of length three, are taken sequentially from the vector B .

The conditional distribution for ρ is proportional to

$$\rho | B, S \propto |R(\rho)|^{-1/2} \exp\{-\frac{1}{2}(Y - FB)'(SR(\rho)S)^{-1}(Y - FB)\}, \quad (16e)$$

where $R(\rho)$ is the correlation matrix and the parentheses emphasize that it is a function of ρ . I draw from this conditional distribution using the Griddy-Gibbs approach. The details are given in the Appendix, Section D.

The conditional distributions for each $\sigma_i \forall i = 1, \dots, N$ are proportional to

$$\sigma_i | R(\rho), S_{-i}, B \propto \sigma_i^{-(T_i+1)} \times \exp \left\{ -\frac{1}{2} \left[\frac{(\log(\sigma_i) - \bar{s})^2}{\delta_\sigma} + (Y - FB)'(S_{-i}R(\rho)S_{-i})^{-1}(Y - FB) \right] \right\}, \quad (16f)$$

where S_{-i} denotes the standard deviation matrix conditional on the other $N - 1$ standard deviation draws. As with the conditional distribution for ρ , I draw from this density using the Griddy-Gibbs approach.

I obtain the Gibbs sampler's initial values by first running OLS univariate regressions and then setting the initial values for $\bar{\beta}$ equal to the average of the OLS parameter estimates and the elements in S equal to the sample standard deviations. The initial value for ρ is set to zero. The sampler is iterated 600 times, and the first 100 draws are discarded.

C. Model Estimation and Statistical Inferences

I estimate the model parameters, as in Section III separately for each of the 17 industries presented earlier. Then, using the parameters' posterior distributions and the procedure outlined in Section II.C, I simulate 2,000 long-horizon average abnormal returns for each industry.

Table XII presents the summary statistics regarding the distribution of the sample mean aggregated across all 17 industries. The table reports the first, fifth, 50th, and 95th percentiles of the distribution in addition to the mean. The realized abnormal return corrected for each firm's factor loadings is given in the last column. I present results corresponding to the four different shrinkage scenarios.

Under all shrinkage scenarios, the observed IPO returns are inconsistent with the three-factor model. The measured abnormal performance is approximately 40 percent lower than the one reported in Table IX. This large difference in abnormal returns is driven by the fact that the IPO factor loadings dictate much higher average return than is actually observed. The characteristic-based model results, on the other hand, indicate that the IPO firm realized returns are consistent with their attributes.

This striking result highlights the sensitivity of long-horizon event studies to the choice of the pricing model as discussed in Fama (1998) and Lyon et al. (1999).

VI. Conclusion

This paper proposes a new approach to inference in long-horizon event studies that overcomes two statistical difficulties plaguing traditional testing methods—non-normality and cross-sectional correlation of long-horizon abnormal returns. The methodology takes as given an asset-pricing model and a distribution for firm residual variation and uses these to simulate the predictive distribution of the long-horizon average abnormal return.

Table XII
Predictive Densities of Five-Year Average Buy-and-Hold
Return under Various Shrinkage Scenarios

For each of the four possible shrinkage scenarios 2,000 average buy-and-hold abnormal returns are simulated. Panels A and B give the properties of these densities for two different levels of shrinkage of the factor loadings. The effect of residual variance shrinkage is reported within each panel. Each row provides the 1st, 5th, 50th, and 95th percentiles as well as the means of these densities. The rightmost column gives the sample abnormal performance calculated using the firms' factor loadings.

Panel A: "Mild" Shrinkage of B						
Shrinkage of σ_i	Simulated Distribution					Abnormal Return (%)
	1%	5%	50%	95%	Mean	
Mild	-25.5	-18.5	-1.4	24.1	0.7	-47.9
Strong	-23.7	-17.9	-0.9	22.1	0.4	-46.9

Panel B: "Strong" Shrinkage of B						
Shrinkage of σ_i	Simulated Distribution					Abnormal Return (%)
	1%	5%	50%	95%	Mean	
Mild	-24.8	-18.6	-0.6	27.0	1.5	-45.0
Strong	-23.0	-17.3	-1.1	22.0	0.3	-45.3

The methodology is applied to a small data set of IPOs, demonstrating in detail how to implement the methodology and make inferences regarding long-horizon abnormal performance. The effects of both non-normality and residual cross-correlation on inference are shown. Next, the methodology is applied to a sample of 1,521 IPOs conducted over the period from 1975 to 1984 (Ritter (1991)). Even for this large sample, the distribution of average abnormal return is non-normal. Furthermore, residual variability and cross-correlation have a large impact on inferences, implying that methods that do not explicitly control for these statistical characteristics may yield erroneous results. Finally, I find that IPO returns are consistent with a characteristic-based pricing model, whereas the Fama and French (1993) three-factor model is inconsistent with the observed long-horizon price performance of these firms.

The methodology proposed in this paper has a number of applications. First, in light of this paper's results, it would be interesting to revisit long-horizon abnormal performance subsequent to other corporate events.²² Second, the approach can be extended to allow for time variation in factor loadings (e.g., Shanken (1990)) and also for various forms of heteroskedasticity and time variation in the common correlation of the assets under study.

²² For example, Brav (1998) applies the proposed methodology to a stock repurchase sample and finds that, contrary to previous results, the Fama and French three-factor model is not rejected once residual cross-correlation is taken into account.

Table AI
Industry Allocation for the IPO Sample

The main sources for the industry definitions are Ritter (1991) and Spiess and Affleck-Graves (1995). The column “Added” lists the additional SIC codes added to some of these industries.

Industry	SIC Codes		Number of IPOs
	Ritter	Added	
1. Electronic equipment	366, 367	369	146
2. Computer manufacturing	357		144
3. Financial institutions	602–603, 612, 671	620–628	139
4. Oil and gas	131, 138, 291, 679	492	129
5. Computer and data processing services	737		113
6. Optical, medical, and scientific equipment	381–384		111
7. Retailers	520–573, 591–599		70
8. Wholesalers	501–519		63
9. Health care and HMOs	805–809	800–804	57
10. Restaurant chains	581		54
11. Drug and genetic engineering	283		44
12. Business services	—	739	42
13. Airlines	451	452, 458, 372	32
14. Communications	—	480–489	29
15. Metal and metal products	351–356, 358–359		24
16. Insurance	—	631–641	17
17. Other	—		307
Total	—		1,521

Appendix

A. Industry SIC Codes

Table AI presents the industry allocation for the IPO sample.

B. Excess Return Relative to the NYSE-AMEX Value-Weight Index

Table AII presents the IPO sample five-year aftermarket performance.

C. Positive-Definiteness of R

In this section I prove that positive definiteness of R requires that we restrict the range of the common correlation coefficient ρ . The proof has two steps. I begin by formulating an alternative regression setup and derive the required condition for the positive definiteness of the correlation matrix in this case. Then I show that the regression setup used throughout this paper is just a transformation of the alternative formulation, which in turn implies the condition for the positive definiteness of R .

For the period January 1975 through December 1989, define t_{min} as the first calendar month for which we have at least one valid monthly observation and t_{max} as the last calendar month for which we have a valid obser-

Table AII
IPO Sample Five-Year Aftermarket Performance

Based on the industry classification given in Table AI, the 1,521 IPOs are allocated into 17 industries. Then, five-year abnormal return is calculated relative to the NYSE-AMEX value-weight index. Reported, for each industry, are the number of firms, the average and median industry return, and the corresponding average market return. The last three columns give the average abnormal return in addition to the cross-sectional standard deviation and skewness of excess returns. The last row provides these statistics for the full sample.

Industry	Number of IPOs	IPO Return (%)		NYSE-AMEX VW Avg. Return (%)	Abnormal Return (%)		
		Avg.	Median		Avg.	Std.	Skew
1. Electronic equipment	146	3.9	-48.3	97.6	-93.7	152.2	3.8
2. Computer manufacturing	144	19.3	-47.7	99.1	-79.8	201.1	3.8
3. Financial institutions	139	90.6	52.6	93.7	-3.1	150.3	0.8
4. Oil and gas	129	-50.7	-86.1	93.5	-144.2	152.9	7.9
5. Computer and data processing services	113	24.5	-47.2	90.2	-65.7	185.0	2.8
6. Optical, medical, and scientific equipment	111	-2.3	-53.5	96.2	-98.5	150.7	3.4
7. Retailers	70	35.2	-16.4	91.9	-56.7	154.4	2.2
8. Wholesalers	63	-10.9	-48.5	86.8	-97.7	93.6	1.2
9. Health care and HMO	57	19.9	-27.3	84.0	-64.1	146.9	1.4
10. Restaurant chains	54	166.4	-71.8	89.1	77.3	971.8	5.0
11. Drug and genetic engineering	44	72.7	41.8	100.7	-28.0	142.3	1.8
12. Business services	42	0.0	-35.3	90.3	-90.3	134.4	2.1
13. Airlines	32	61.1	-11.9	82.1	-21.0	180.1	2.0
14. Communications	29	-15.8	-55.6	79.3	-95.1	83.1	0.3
15. Metal and metal products	24	5.9	-50.6	90.5	-84.6	136.8	1.7
16. Insurance	17	101.3	84.9	96.4	4.9	159.4	0.8
17. Other	307	36.1	-26.1	91.2	-55.1	207.7	5.9
Total	1,521	27.2	-37.1	92.9	-65.7	245.9	11.9

vation. Consequently, month t ($t_{\min} \leq t \leq t_{\max}$) labels any calendar month over the period of investigation with valid monthly returns. Let n_t denote the number of IPOs with valid returns in month t , T^* the number of calendar months with valid returns, and N the number of IPOs in the sample.

Define y_t^* as the $(n_t \times 1)$ column vector of firm excess returns in month t , f_t^* an $(n_t \times 1)$ vector of month t characteristic-matched portfolio returns, and v_t^* , the $(n_t \times 1)$ column vector of residuals. The month t returns are modeled as follows:

$$y_t^* = f_t^* + v_t^* \quad \forall t = t_{\min}, \dots, t_{\max}. \quad (\text{A1})$$

Next, write the system of all T^* monthly regressions as follows,

$$Y^* = F^* \iota + V^*,$$

$$Y^* = \begin{bmatrix} y_{t_{\min}}^* \\ \vdots \\ y_{t_{\max}}^* \end{bmatrix}, \quad F^* = \begin{bmatrix} f_{t_{\min}}^* & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & f_{t_{\max}}^* \end{bmatrix}, \quad V^* = \begin{bmatrix} v_{t_{\min}}^* \\ \vdots \\ v_{t_{\max}}^* \end{bmatrix}, \quad (\text{A2})$$

where Y^* is a $(\sum_{t=t_{\min}}^{t_{\max}} n_t \times 1)$ stacked vector of firm excess returns, F^* is a $(\sum_{t=t_{\min}}^{t_{\max}} n_t \times T^*)$ matrix of factor realizations, ι is a $(T^* \times 1)$ vector of ones, and V^* is a $(\sum_{t=t_{\min}}^{t_{\max}} n_t \times 1)$ stacked vector of firm residuals. Finally, define R^* as the correlation matrix of the residuals in V^* .

I make the following two assumptions.

ASSUMPTION 1:

The month t ($n_t \times n_t$) residual covariance matrix is positive definite.

ASSUMPTION 2:

All IPOs within a given industry share the same residual cross-correlation ρ . ρ does not vary over time.

From Assumption 1 it follows (see Priest (1968) and Olkin (1981)) that the correlation coefficient for any month t , ρ_t , is bounded from below by ρ_t^* , where

$$\rho_t > \rho_t^* = -\frac{1}{n_t - 1}. \quad (\text{A3})$$

Given Assumption 2, the tightest lower bound on ρ , denoted by ρ^* , is equal to $\max_t \rho_t^* = -1/\max(n_t) - 1$. Because R^* is a correlation matrix containing the calendar month t correlation matrices on its diagonal, its positive definiteness is guaranteed by the above inequality.

The last step in the proof follows from the fact that the regression system in equation (A2) is just a transformation of the system studied in this paper (see equation (3)). Specifically, the vector of returns R^* contains the same elements as in R but the ordering has changed. Hence, there exists a permutation matrix²³ P such that $PY^* = Y$, $PF^* = F$, and $PV^* = V$.

As a result, the correlation matrix R of the residuals in V is equal to PR^*P' , which is positive definite given the positive definiteness of R^* and the nonsingularity of P . Q.E.D.

D. Sampling from the Conditional Distributions for ρ and σ_i

I sample from the conditional distributions of ρ and σ_i following Tanner (1996, pp. 164–165).

The conditional distribution for ρ was given in equation (8):

$$\rho|S \propto |R(\rho)|^{-1/2} \exp\{-\frac{1}{2}(Y - F\iota)'(SR(\rho)S)^{-1}(Y - F\iota)\}. \quad (\text{A4})$$

I sample from this distribution using the following algorithm.

1. Evaluate the density $\rho|b, S$ on a grid of M points $\{\rho_1, \rho_2, \dots, \rho_M\}$. This yields a vector $\{p_1, p_2, \dots, p_M\}$.
2. Use the vector $\{p_1, p_2, \dots, p_M\}$ to obtain an approximation of the inverse cdf of $\rho|b, S$.
3. Sample from a uniform distribution $\text{Un}([0, 1])$ and transform the observation via the approximation in step 2.

I begin with an initial grid centered at zero. Following Tanner's (1996) suggestion, I specify the grid such that more points are in the neighborhood of this "center" of the distribution. The grid is modified every 100 iterations by locating the mode of the previous draws and constructing the grid such that it is more dense around this mode. Finally, given the restriction on positive-definiteness of R , the grid is always bounded from below by ρ^* (see the discussion in Section C).

The conditional distributions for $\sigma_i \forall i = 1, \dots, N$ were given in equation (9).

$$\sigma_i|R(\rho), S_{-i} \propto \sigma_i^{-(T_i+1)} \times \exp\left\{-\frac{1}{2}\left[\frac{(\log(\sigma_i) - \bar{s})^2}{\delta_\sigma} + (Y - F\iota)'(S_{-i}R(\rho)S_{-i})^{-1}(Y - F\iota)\right]\right\}. \quad (\text{A5})$$

²³ A square matrix P is said to be a permutation matrix if each row and each column of P contain a single element 1 and the remaining elements are zero. Any permutation matrix is nonsingular and orthogonal, that is, $P^{-1} = P'$.

For each such conditional distribution I specify an initial grid centered at the least squares estimate for that σ_i . I sample from these conditional distributions as in steps 2 and 3. The grid points are modified every 100 iterations by recentering the grid at the mode of the retained draws. All grids are equally spaced.

E. Specification in Random Samples

In this section I assess the specification of the proposed methodology in random samples. I employ all NYSE/AMEX/Nasdaq firms with available data from CRSP. The period of analysis is from July 1973 through December 1995, and as in Lyon et al. (1999), I include in the analysis only firms that have ordinary common share codes (CRSP share codes 10 and 11). Following Kothari and Warner (1997), the specification is assessed based on 250 random samples of 200 event months. Two possible holding periods τ are considered: 36 or 60 months. Thus, for each randomly selected sample, the selected securities are held at most for τ months. If a security delists beforehand, it is held until the delisting month. Finally, the analysis is applied to a characteristic-based model as well as to a factor model.

If the proposed methodology is well specified, then at $\alpha = 5$ percent and 10 percent levels of significance, we should expect to find that 250 α of these simulations yield rejections of the null model under consideration. For a given significance level α , I report separately rejections of the null model in favor of either positive or negative abnormal performance. These one-sided tests should reject the null model in 250 $\alpha/2$ such samples.

E.1. Characteristic-Based Model

The derivation of the predictive density for each random sample is simplified considerably once it is recognized that the constituent firms' abnormal performance are independent by construction. Hence, the unknown parameters are the firm residual standard deviations.

Consistent with the motivation for the shrinkage of firm parameters outlined in Section II, I use an informative prior on the firm residual standard deviations. To simplify the computational burden I assume an inverse-gamma prior for the firm residual variances,²⁴

$$\sigma_i^2 \sim IG(\nu, \sigma_0^2) \quad \forall i = 1, \dots, 200, \quad (\text{A6})$$

where we can interpret the above prior as a posterior distribution obtained after observing an imaginary sample of size ν and mean centered at σ_0^2 . I set ν equal to the average number of time-series observations in sample s and σ_0^2 equal to the average of the sample residual variances.

²⁴ See Zellner (1971, pp. 70–72) and Gelman et al. (1995, pp. 46–47).

Table AIII

Specification of the Characteristic-Based Asset-Pricing Model

The specification of the proposed methodology is based on 250 random samples of 200 event months. Two possible holding periods τ are considered: 36 or 60 months. If a security delists prematurely, it is held until the delisting month. The sample is restricted to firms that have at least 12 monthly returns in order to be able to estimate the characteristic-based model below. Rejection rates are presented separately in favor of either negative or positive abnormal performance at the $\alpha = 5$ percent or 10 percent significance level.

Three-Year Returns				Five-Year Returns			
5%		10%		5%		10%	
2.5	97.5	5	95	2.5	97.5	5	95
0.8	1.6	2.8	2.0	0.8	1.6	2.8	4.4

The resulting posterior beliefs, for each firm residual variance, are also in the inverse-gamma form:

$$\sigma_i^2 \sim IG(\nu + T_i, \sigma_0^2 + \hat{\sigma}_i^2) \quad \forall i = 1, \dots, 200, \quad (\text{A7})$$

where T_i is the number of monthly observations for firm i and $\hat{\sigma}_i^2$ is the OLS estimate of the firm residual variance.

Given the regression setup in equation (2), and the posterior beliefs for the firms in sample s ($s = 1, \dots, 250$), I simulate 1,000 average buy-and-hold abnormal returns as described in Section II.C. These simulated averages are used to construct the predictive density for sample s . Last, long-horizon abnormal returns are computed using the size and book-to-market portfolios described in Section II.B. Table AIII reports the rejection frequencies for the characteristic-based model.

None of the rejection frequencies are significant, at the five percent level, using a one-sided binomial test statistic.

E.2. Factor Model

The regression setup in this case was given in Section V,

$$y_i = f_i \beta_i + v_i \quad \forall i = 1, \dots, 200. \quad (\text{A8})$$

As in the previous section, the derivation of the predictive density is simplified because the firms' abnormal performance are independent by construction. Hence, the unknown parameters are the firm factor loadings and residual standard deviations.

I employ an informative prior for all the model parameters. The prior for the firm residual variance is in the inverse gamma form (see equation (A6)), whereas the prior for the firm factor loadings is multivariate normal,

$$\beta_i \sim N(b, A) \quad \forall i = 1, \dots, 200 \quad (\text{A9})$$

Table AIV

Specification of the Factor-Based Asset-Pricing Model

The specification of the proposed methodology is based on 250 random samples of 200 event months. Two possible holding periods τ are considered: 36 or 60 months. If a security delists prematurely, it is held until the delisting month. The sample is restricted to firms that have at least 12 monthly returns to be able to estimate the factor models below. Rejection rates are presented separately in favor of either negative or positive abnormal performance at the $\alpha = 5$ percent or 10 percent significance level.

Three-Year Returns				Five-Year Returns			
5%		10%		5%		10%	
2.5	97.5	5	95	2.5	97.5	5	95
Panel A: Three-Factor Model							
3.2	6.0	2.0	0.4	2.0	5.2	0.4	0.0
Panel B: Four-Factor Model							
2.0	3.2	3.2	2.4	0.8	2.0	0.4	0.4

where b is set equal to the grand average of the 200 sample factor loadings obtained from separate OLS regressions. A is the prior covariance matrix, and I set its diagonal elements equal to half of the observed cross-sectional variance of these OLS estimates, leading to mild shrinkage toward the sample grand mean (see Section II.B). The resulting posterior distribution is normal conditional on the firm residual variance,

$$\beta_i | \sigma_i^2 \sim N(\bar{b}, \bar{A}) \quad \forall i = 1, \dots, 200, \quad (\text{A10})$$

where $\bar{b} = (\sigma_i^{-2} f' f + A^{-1})^{-1} (\sigma_i^{-2} f' f \beta_{ols} + A^{-1} \bar{b})$, $\bar{A} = (\sigma_i^{-2} f' f + A^{-1})^{-1}$, and β_{ols} is the firm OLS parameter estimates.

Given the posterior beliefs for the firms in sample s ($s = 1, \dots, 250$), I simulate 1,000 average buy-and-hold abnormal returns as described in Section II.C. These simulated averages are used to construct the predictive density for sample s .

Table AIV reports the rejection frequencies for two specifications of the factor model. Panel A presents results based on the Fama and French (1993) three-factor model, whereas Panel B provides rejection frequencies for a four-factor model that extends the Fama and French model by adding a momentum factor in returns (Jegadeesh and Titman (1993), Carhart (1997)).

For both models none of the rejection frequencies are significant, at the 5 percent level, using a one-sided binomial test statistic.

REFERENCES

Barber, Brad M., and John D. Lyon, 1997, Detecting long-run abnormal stock returns: The empirical power and specification of test-statistics, *Journal of Financial Economics* 43, 341–372.

- Barnard, John, Robert McCulloch, and Xiao-Li Meng, 1997, Modeling covariance matrices with application to shrinkage, *Statistica Sinica*, forthcoming.
- Barry, Christopher B., 1974, Portfolio analysis under uncertain means, variances and covariances, *Journal of Finance* 29, 515–522.
- Bernard, Victor L., 1987, Cross-sectional dependence and problems in inference in market-based accounting research, *Journal of Accounting Research* 25, 1–48.
- Blattberg, Robert C., and Edward I. George, 1991, Shrinkage estimation of price and promotional elasticities: Seemingly unrelated equations, *Journal of the American Statistical Association* 86, 304–315.
- Box, George E. P., 1980, Sampling and Bayes inference in scientific modelling and robustness (with discussion), *Journal of the Royal Statistical Society A* 143, 383–430.
- Brav, Alon, 1998, Inference in long horizon event studies: A re-evaluation of the evidence, Working paper, Duke University, Fuqua School of Business.
- Brav, Alon, Chris Geczy, and Paul Gompers, 2000, Is the abnormal return following equity issuances anomalous? *Journal of Financial Economics* 56, 209–249.
- Brav, Alon, and Paul Gompers, 1997, Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies, *Journal of Finance* 52, 1791–1821.
- Breslow, Norman, 1990, Biostatistics and Bayes, *Statistical Science* 5, 269–298.
- Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.
- Collins, Daniel W., and Warren T. Dent, 1984, A comparison of alternative testing methodologies used in capital market research, *Journal of Accounting Research* 22, 48–84.
- Cowan, Arnold R., and Anne M. A. Sergeant, 1997, Interacting biases, non-normal return distributions and the performance of parametric and bootstrap tests for long-horizon event studies, Working paper, Iowa State University.
- Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035–1058.
- Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.
- Davidson, Russell, and James G. MacKinnon, 1993, *Estimation and Inference in Econometrics* (Oxford University Press, New York).
- Fama, Eugene F., 1970, Efficient capital markets: A review of theory and empirical work, *Journal of Finance* 25, 383–417.
- Fama, Eugene F., 1998, Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics* 49, 283–306.
- Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.
- Gelfand, Alan E., Dipak K. Dey, and Hong Chang, 1992, Model determination using predictive distributions with implementation via sampling-based methods; in Jose M. Bernardo, James O. Berger, Philip A. Dawid, and Adrian F. M. Smith, eds.: *Bayesian Statistics*, vol. 4 (Oxford University Press, Oxford).
- Gelfand, Alan E., Susan E. Hills, Amy Racine-Poon, and Adrian F. M. Smith, 1990, Illustration of Bayesian inference in normal data models using Gibbs sampling, *Journal of the American Statistical Association* 85, 972–985.
- Gelfand, Alan E., and Adrian F. M. Smith, 1990, Sampling based approaches to calculating marginal densities, *Journal of the American Statistical Association* 85, 398–409.
- Gelman, Andrew, John B. Carlin, Hal H. Stern, and Donald B. Rubin, 1995, *Bayesian Data Analysis* (Chapman and Hall, London.).
- Gelman, Andrew, Xiao-Li Meng, and Hal Stern, 1996, Posterior predictive assessment of model fitness via realized discrepancies, *Statistica Sinica* 6, 733–807.
- Geman, Stuart, and Donald Geman, 1984, Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 721–741.

- Gibbons, Michael, 1982, Multivariate tests of financial models: A new approach, *Journal of Financial Economics* 10, 3–27.
- Gilks, Walter R., Sylvia Richardson, and David J. Spiegelhalter, 1996, *Markov Chain Monte Carlo in Practice* (Chapman and Hall, London).
- Harvey, Campbell, and Guofu Zhou, 1990, Bayesian inference in asset pricing tests, *Journal of Financial Economics* 26, 221–254.
- Ibrahim Joseph G., and Purushottam W. Laud, 1994, A predictive approach to the analysis of designed experiments, *Journal of the American Statistical Association* 89, 309–319.
- Ikenberry, David, Josef Lakonishok, and Theo Vermaelen, 1995, Market underreaction to open market share repurchases, *Journal of Financial Economics* 39, 181–208.
- Jaffe, Jeffrey, 1974, Special information and insider trading, *Journal of Business* 47, 411–428.
- Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 881–898.
- Jorion, Philippe, 1986, Bayes-Stein estimation for portfolio analysis, *Journal of Financial and Quantitative Analysis* 21, 279–292.
- Jorion, Philippe, 1991, Bayesian and CAPM estimators of the means: Implications for portfolio selection, *Journal of Banking and Finance* 15, 717–727.
- Kandel, Shmuel, Robert McCulloch, and Robert Stambaugh, 1995, Bayesian inference and portfolio efficiency, *Review of Financial Studies* 8, 1–53.
- Klein, Roger W., and Vijay S. Bawa, 1976, The effect of estimation risk on optimal portfolio choice, *Journal of Financial Economics* 3, 215–231.
- Kothari, S. P., and Jay Shanken, 1997, Book-to-market, dividend yield, and expected market returns: A time-series analysis, *Journal of Financial Economics* 44, 169–204.
- Kothari, S. P., and Jerold B. Warner, 1997, Measuring long-horizon security performance, *Journal of Financial Economics* 43, 301–339.
- Laud, Purushottam W., and Joseph G. Ibrahim, 1995, Predictive model selection, *Journal of the Royal Statistical Society B* 57, 247–262.
- Lindley, Dennis V., and Adrian F. M. Smith, 1972, Bayes estimates for the linear model, *Journal of the Royal Statistical Society B* 34, 1–41.
- Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50, 23–51.
- Loughran, Tim, and Jay R. Ritter, 1999, Uniformly least powerful tests of market efficiency, *Journal of Financial Economics*, forthcoming.
- Lyon, John D., Brad M. Barber, and Chih-Ling Tsai, 1999, Improved methods for tests of long-run abnormal stock returns, *Journal of Finance* 54, 165–201.
- Mandelker, Gershon, 1974, Risk and return: The case of merging firms, *Journal of Financial Economics* 1, 303–335.
- McCulloch, Robert, and Peter Rossi, 1991, A Bayesian approach to testing the arbitrage pricing theory, *Journal of Econometrics* 49, 141–168.
- Michael, Roni, Richard H. Thaler, and Kent L. Womack, 1995, Price reactions to dividend initiations and omissions: Overreaction or drift? *Journal of Finance* 50, 573–608.
- Mitchell, Mark, and Erik Stafford, 1999, Managerial decisions and long-term stock price performance, Working paper, Harvard University.
- Olkin, Ingram, 1981, Range restrictions for product-moment correlation matrices, *Psychometrika* 46, 469–472.
- Papadimitis, Efsthios, 1996, A frequency domain bootstrap-based method for checking the fit of a transfer function model, *Journal of the American Statistical Association* 91, 1535–1550.
- Pastor, Lubos, and Robert F. Stambaugh, 1999, Costs of equity capital and model mispricing, *Journal of Finance* 54, 67–121.
- Priest, Hugh F., 1968, Range of correlation matrices, *Psychological Reports* 22, 168–170.
- Ritter, Christian, and Martin A. Tanner, 1992, Facilitating the Gibbs sampler: The Gibbs stopper and the griddy Gibbs sampler, *Journal of the American Statistical Association* 87, 861–868.
- Ritter, Jay R., 1991, The long-run performance of initial public offerings, *Journal of Finance* 46, 3–28.
- Rubin, Donald B., 1984, Bayesianly justifiable and relevant frequency calculations for the applied statistician, *Annals of Statistics* 12, 1151–1172.

- Sefcik, Stephan E., and Rex Thompson, 1986, An approach to statistical inference in cross-sectional models with security abnormal returns as dependent variable, *Journal of Accounting Research* 24, 316–335.
- Shanken, Jay, 1990, Intertemporal asset pricing: An empirical investigation, *Journal of Econometrics* 45, 99–120.
- Spiess, Katherine D., and John Affleck-Graves, 1995, Underperformance in the long-run stock returns following seasoned equity offerings, *Journal of Financial Economics* 38, 243–267.
- Stambaugh, Robert F., 1997, Analyzing investments whose histories differ in length, *Journal of Financial Economics* 45, 285–331.
- Stangl, Dalene, and Gabriel Huerta, 1997, Using Bayesian hierarchical models to assess the impact of managed-care strategies, Working paper, Institute of Statistics and Decision Sciences, Duke University.
- Stevens, Ross L., 1996, New methods in asset pricing model estimation, Working paper, University of Chicago.
- Tanner, Martin A., 1996, *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions* (Springer-Verlag, New York).
- Tsay, Ruey S., 1992, Model checking via parametric bootstraps in time series analysis, *Applied Statistics* 41, 1–15.
- Vasicek, Oldrich A., 1973, A note on using cross-sectional information in Bayesian estimation of security betas, *Journal of Finance* 28, 1233–1239.
- Womack, L. Kent, 1996, Do brokerage analysts' recommendations have investment value? *Journal of Finance* 51, 137–167.
- Zellner, Arnold, 1962, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *Journal of the American Statistical Association* 57, 348–368.
- Zellner, Arnold, 1971, *An Introduction to Bayesian Inference in Econometrics* (John Wiley and Sons, New York).