Asset-Pricing Anomalies at the Firm Level

Scott Cederburg∗ Phil Davies† Michael O’Doherty‡

April 18, 2011§

Abstract

Portfolio-level tests linking CAPM alphas to a large number of firm characteristics suggest that the CAPM fails across multiple dimensions. There are, however, concerns that underlying firm-level associations may be distorted at the portfolio level. In this paper we use a hierarchical Bayes approach to model conditional firm-level alphas as a function of firm characteristics. Our empirical results indicate that much of the anomaly-based evidence against the conditional CAPM is overstated. Anomalies are primarily confined to small stocks, few characteristics are robustly associated with CAPM alphas out of sample, and the majority of firm characteristics do not contain unique information about abnormal returns.

JEL Classification: C11, G10, G12, G14

∗Tippie College of Business, University of Iowa, S217 Pappajohn Business Building, Iowa City, IA 52242. Email: scott-cederburg@uiowa.edu. Tel: (319) 541-5726.
†Rutgers Business School, 1 Washington Park, Newark, NJ 07102. Email: pdavies@andromeda.rutgers.edu.
‡Tippie College of Business, University of Iowa, S319 Pappajohn Business Building, Iowa City, IA 52242. Email: michael-odoherty@uiowa.edu. Tel: (319) 335-0974.
§We thank Doron Avramov, Ashish Tiwari, Paul Weller, Tong Yao, and seminar participants at Iowa State University, the University of Iowa, the University of Nebraska–Lincoln, the 2010 CRSP Forum, the 2011 Eastern Finance Association, and the 2011 Midwest Finance Association for helpful comments and suggestions.
1 Introduction

An anomaly is a pattern in average stock returns that is inconsistent with the predictions of the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965). Anomalies are commonly identified using a portfolio-based approach. The researcher sorts stocks on a firm characteristic and constructs a zero-cost hedge portfolio by taking long and short positions in the extreme groups. If the hedge portfolio earns abnormal returns relative to the CAPM, the sorting characteristic is classified as an anomaly. Over the past three decades, a large number of anomalies have been uncovered, suggesting the CAPM is unable to explain much of the cross-sectional variation in average stock returns.\footnote{Previous papers show a positive relation between average returns and book-to-market equity (Rosenberg, Reid, and Lanstein (1985), Chan, Hamao, and Lakonishok (1991), and Fama and French (1992)), stock return momentum (Jegadeesh (1990) and Jegadeesh and Titman (1993, 2001)), and profitability (Haugen and Baker (1996) and Cohen, Gompers, and Vuolteenaho (2002)). There is a negative relation between average returns and size (Banz (1981) and Fama and French (1992)), stock return reversal (DeBondt and Thaler (1985) and Chopra, Lakonishok, and Ritter (1992)), asset growth (Fairfield, Whisenant, and Yohn (2003), Titman, Wei, and Xie (2004), and Cooper, Gulen, and Schill (2008)), net stock issues (Loughran and Ritter (1995), Ikenberry, Lakonishok, and Vermaelen (1995), Daniel and Titman (2006a), and Pontiff and Woodgate (2008)), accruals (Sloan (1996), Collins and Hribar (2000), and Xie (2001)), and financial distress (Dichev (1998) and Campbell, Hilscher, and Szilagyi (2008)).}

There are, however, growing concerns in the literature about the use of portfolios to identify anomalies and, more generally, to test asset-pricing models. These arguments are centered around the idea that grouping firms into portfolios and aggregating returns wastes and potentially distorts valuable information about cross-sectional patterns in abnormal returns. For example, Roll (1977), Kandel and Stambaugh (1995), and Fama and French (2008) discuss how patterns in firm-level pricing errors can be distorted at the portfolio level. Lo and MacKinlay (1990) highlight the data-snooping biases inherent in portfolio-based asset-pricing tests, while Litzenberger and Ramaswamy (1979) and Ang, Liu, and Schwarz (2010) consider the loss in efficiency from using portfolios rather than individual firms in asset-pricing tests. More recently, Ahn, Conrad, and Dittmar (2009) and Lewellen, Nagel, and Shanken (2010) show that inferences in asset-pricing tests are sensitive to the choice of test portfolios.\footnote{For other issues, see Conrad, Cooper, and Kaul (2003), Kan (2004), and Daniel and Titman (2006b).}

One way to avoid the concerns with portfolio-level inferences is to use firm-level data. However, examining anomalies at the firm level is a challenging problem. The researcher has to relate firm characteristics to abnormal returns which are not directly observable. Therefore, the researcher must also model and estimate the evolution of market risk for each individual firm.

Two recent firm-level studies adopt contrasting approaches to control for market risk. Fama
and French (2008) argue that market risk should not be related to firm characteristics, removing
the need to examine abnormal returns, while Avramov and Chordia (2006) model market risk as an
exact linear function of firm size, book-to-market, and macroeconomic variables. Both approaches
are problematic. Examining the relation between firm characteristics and raw returns is likely
to overstate the CAPM’s failings if firm-level betas are associated with firm characteristics, while
specifying betas as an exact linear function of covariates is only valid if the researcher knows the
full set of variables associated with variation in betas. If betas are related to firm characteristics
not included in the model specification then betas will be systematically mismeasured, which in
turn may give rise to spurious relations between firm characteristics and mismeasured alphas.

Motivated by these concerns, we develop a new hierarchical Bayes approach to explore anomali-
ies at the firm level. Specifically, we simultaneously estimate (1) conditional CAPM parameters
for each firm using an approach similar to Lewellen and Nagel (2006) which specifies short time
periods and avoids the need for conditioning information, (2) the cross-sectional relation between
conditional alphas and firm characteristics in each time period, and (3) the systematic association
between alphas and firm characteristics across the entire sample period. Our approach has several
desirable features relative to the prior literature. The hierarchical Bayes approach eliminates a mea-
surement error problem encountered in traditional two-step approaches (e.g., Brennan, Chordia,
and Subrahmanyam (1998) and Avramov and Chordia (2006)). We also put little structure on the
dynamics of conditional betas, thereby minimizing potential model misspecification. In addition,
our approach implicitly controls for cross-sectional heteroskedasticity and cross-correlations among
stocks.

We use the hierarchical Bayes approach to examine nine anomalies over the period 1963 to 2008:
size, book-to-market, momentum, reversal, profitability, asset growth, net stock issues, accruals,
and financial distress. Studying each anomaly separately, we find that firm-level associations are
distorted at the portfolio level for four of the nine anomalies. For example, the traditional portfolio
approach suggests size and reversal are associated with abnormal returns, but using information
from the entire cross section of stocks there is no evidence of a robust relation between either of these
variables and firm alphas. Further analysis suggests the portfolio-level results for size and reversal
are driven by a small subset of stocks with extreme values for these characteristics. Nevertheless,
the initial firm-level evidence still paints a bleak picture for the conditional CAPM. Seven of the
nine characteristics are significantly associated with alphas, suggesting that the CAPM does indeed
fail across multiple dimensions. These results, however, may be misleading for three reasons.

First, it is possible that the anomalous patterns are being driven primarily by small, illiquid stocks which represent only a tiny fraction of the total market capitalization. We investigate this possibility by allowing associations between alphas and firm characteristics to vary across micro, small, and big stocks. We find the associations are strongest in terms of statistical and economic magnitude for micro and small stocks. For big stocks, alphas are significantly associated with only three of the nine characteristics - asset growth, net stock issues, and accruals.

Second, while the existence of anomalies could indicate that the CAPM is fundamentally flawed, anomalies could also be the result of temporary market mispricing or data snooping by researchers. To distinguish between these competing explanations, we consider whether the relation between alphas and each firm characteristic attenuates or persists after the anomaly is established in the asset-pricing literature. Anomalies that persist post publication are more likely to reflect a fundamental failure of the CAPM, since investor mistakes are likely to be corrected once highlighted and results due to data snooping are unlikely to persist out of sample. Of the seven anomaly variables with sufficient post-publication sample periods, only two – book-to-market and accruals – are significantly related to abnormal returns after publication. Further analysis, however, shows that these relations are driven primarily by micro stocks for which transaction costs and liquidity concerns diminish investors’ ability to exploit anomalies and correct mispricings (e.g., Jensen (1978)). Among big stocks, no firm characteristic is significantly associated with abnormal returns post publication.

Third, firm characteristics could be correlated with each other and offer little unique information about abnormal returns. Asset-pricing tests that consider each firm characteristic in isolation are likely to suffer from an omitted variable bias that will result in the importance of an anomaly being overstated. Traditional portfolio approaches are unable to adequately address this omitted variable problem. Researchers typically rely on multi-dimensional sorts to isolate the effects of a particular characteristic, but controlling for more than one or two characteristics simultaneously is infeasible. In contrast, our approach is particularly well suited to assess which anomalies contain unique information; we simply specify conditional alphas as a function of multiple firm characteristics. Our results suggest that univariate tests do indeed suffer from a pronounced omitted variable bias.

Following Fama and French (2008), we classify stocks into three size groups - micro, small, and big. The breakpoints are based on the 20th and 50th percentiles of market capitalization for NYSE stocks at the end of June each year.
Considering all characteristics simultaneously we find that size, momentum, reversal, asset growth, and financial distress do not contain significant incremental information about abnormal returns, in contrast to the corresponding portfolio-level results.

Taken together, the results suggest that while the CAPM does not perfectly explain firm returns, much of the anomaly-based evidence against the CAPM is overstated. Relations between firm characteristics and conditional firm-level alphas are primarily focused among micro and small stocks and tend not to persist after the anomalies are first documented. Furthermore, few of the firm characteristics associated with alphas actually contain unique information.

The paper is organized as follows. Section 2 develops our econometric model for testing asset-pricing anomalies and discusses the advantages and disadvantages of the proposed approach. Section 3 describes the data. Section 4 presents the empirical results. Section 5 concludes.

2 Methodology

This section develops our firm-level approach for identifying anomalies relative to the CAPM. Section 2.1 outlines our empirical model. Section 2.2 provides details on the estimation procedure. Section 2.3 contrasts our framework with existing firm-level approaches.

2.1 Model Development

The Sharpe–Lintner version of the CAPM states that

$$E[r_{i,t}] = \beta_i E[r_{m,t}],$$

where $E[r_{i,t}]$ denotes the expected return on stock $i$ at time $t$ in excess of the risk-free rate, $E[r_{m,t}]$ is the market risk premium, and $\beta_i = \frac{\text{Cov}(r_{i,t}, r_{m,t})}{\text{Var}(r_{m,t})}$ captures stock $i$’s exposure to market risk. The Sharpe–Lintner CAPM is a static single-period model. In reality, as a firm grows and evolves, its exposure to market risk (and hence its expected return) will change. Similarly, the market risk premium is likely to vary depending on the state of the economy and the risk tolerance of investors. In the presence of time-varying risk exposures and risk premiums, a conditional version of the CAPM,

$$E_{t-1}[r_{i,t}] = \beta_{i,t} E_{t-1}[r_{m,t}],$$

where $E_{t-1}[r_{i,t}]$ denotes the expected return on stock $i$ at time $t$ in excess of the risk-free rate, $E_{t-1}[r_{m,t}]$ is the market risk premium, and $\beta_{i,t} = \frac{\text{Cov}(r_{i,t}, r_{m,t})}{\text{Var}(r_{m,t})}$ captures stock $i$’s exposure to market risk. The conditional CAPM is a dynamic single-period model. In reality, as a firm grows and evolves, its exposure to market risk (and hence its expected return) will change. Similarly, the market risk premium is likely to vary depending on the state of the economy and the risk tolerance of investors. In the presence of time-varying risk exposures and risk premiums, a conditional version of the CAPM,
may hold even if the unconditional CAPM does not (Jagannathan and Wang (1996)). The conditional CAPM implies that the expected conditional alpha, defined as

\[ E_{t-1}[\alpha_{i,t}] = E_{t-1}[r_{i,t}] - \beta_{i,t}E_{t-1}[r_{m,t}], \]  

(3)

should equal zero for all stocks.

A common way of testing this prediction is to examine whether alphas can be forecasted by firm characteristics. Many existing tests in the asset-pricing literature rely on portfolio-based approaches. However, grouping firms into portfolios and aggregating returns has adverse effects; valuable information is discarded by averaging across firms and cross-sectional patterns in firm returns can be distorted as a result of the portfolio formation procedure. Lewellen, Nagel, and Shanken (2010) argue that testing asset-pricing models using individual firms is a sensible alternative. The CAPM’s prediction that alphas are not forecastable can be tested at the firm level by examining a cross-sectional relation of the form,

\[ \alpha_{i,t} = \delta_0 + \delta_x x_{i,t} + \epsilon_{i,t}, \]  

(4)

where \( x_{i,t} \) is a firm characteristic that is observable at the beginning of period \( t \). The conditional CAPM implies that \( \delta_x = 0 \) in a cross-sectional regression based on equation (4). However, analysis of the cross-sectional regression in equation (4) is complicated by the fact that the dependent variable, \( \alpha_{i,t} \), is a latent variable. As such, a model for the latent alphas is necessary to examine the relation in equation (4).

Motivated by the existing criticisms of portfolio-based methods, we develop a firm-level test of the CAPM’s implication that alphas are not predictable. Specifically, we propose a system of equations in which we simultaneously model conditional CAPM alphas and analyze the cross-sectional relation between firm-level alphas and firm characteristics. The model takes on the following structure:

\[ r_{i,t,y} = \alpha_{i,y} + \beta_{i,y} r_{m,t,y} + \epsilon_{i,t,y}, \quad \epsilon_{i,t,y} \sim N(0, \sigma_{i,y}^2), \]  

(5)

\[ \alpha_{i,y} = X_{i,y} \delta_y + \eta_{i,y}, \quad \eta_{i,y} \sim N(0, \sigma_{\alpha,y}^2), \]  

(6)

\[ \delta_y = \bar{\delta} + \nu_y, \quad \nu_y \sim MVN(0, \mathbf{V}), \]  

(7)
where \( r_{i,t,y} \) denotes the excess return on stock \( i \) in subperiod \( t \) of time period \( y \), \( r_{m,t,y} \) is the excess market return, and \( X_{i,y} \) is a matrix including a constant and firm characteristics observable at the beginning of period \( y \).

In the primary model specification, we use monthly subperiods \((t)\) and annual periods \((y)\). We therefore allow firm alphas and betas to change each year, building on the short-window regression approach of Lewellen and Nagel (2006) to test the conditional CAPM. In equation (6), \( \delta_y \) measures the year-by-year relations between alphas and firm characteristics. In a given year, however, conditional alphas may be related to characteristics purely by chance. To examine whether there is a systematic relation between firm characteristics and alphas throughout the entire sample period, we assume that the parameter vectors, \( \{\delta_y\}_{y=1}^{Y} \), are drawn from a multivariate normal distribution centered at \( \bar{\delta} \), as specified in equation (7). If an element of \( \bar{\delta} \) is focused away from zero, there is evidence of an anomaly that persists through time. In our empirical analysis, we analyze \( \bar{\delta} \) when assessing the importance of firm characteristics in forecasting alphas.

The model specified in equations (5) to (7) implicitly takes into account cross-sectional heteroskedasticity and cross-correlations among firms. These features of the return distribution will influence the precision of \( \delta_y \) in each period and will therefore be reflected in the posterior distributions for \( \bar{\delta} \) and \( V \) (Shanken and Zhou (2007)). Thus, a large number of test assets can be considered without requiring the estimation of a variance-covariance matrix.

### 2.2 Model Estimation

Estimating equations (5) to (7) simultaneously is a challenging problem. The model involves a high-dimensional parameter space since firm-specific parameters must be estimated for thousands of firms in each year. Estimation is further complicated by the fact that the latent variables \( \alpha_{i,y} \) and \( \delta_y \) appear in multiple equations within the system.

Fortunately, the problem can be greatly simplified by recognizing the hierarchical structure of the model. Equation (7) is a hierarchical prior for \( \delta_y \) in equation (6), while equation (6) is a hierarchical prior for \( \alpha_{i,y} \) in equation (5). Thus, we adopt a hierarchical Bayes approach to estimate the system of equations described by equations (5) to (7) simultaneously. In addition

---

\(^4\)Several papers have used Bayesian techniques to examine asset-pricing models. McCulloch and Rossi (1991) and Geweke and Zhou (1999) develop Bayesian analyses of the Arbitrage Pricing Theory (APT), while Shanken (1987), Harvey and Zhou (1990), Kandel, McCulloch, and Stambaugh (1995), and Cremers (2006) propose Bayesian tests for the mean-variance efficiency of a given portfolio. Ang and Chen (2007) use Bayesian methods to examine whether the conditional CAPM can explain the value premium.

\(^5\)See Rossi, Allenby, and McCulloch (2005, Ch. 5) for a discussion of hierarchical Bayes models.
to greatly reducing the computational burden relative to using maximum likelihood estimation or the generalized method of moments, the Bayesian approach provides a complete accounting of parameter uncertainty and exact finite sample inference.

The Bayesian approach does require the researcher to specify explicit priors and hyperparameters for all model parameters. We specify the prior for the parameter vector of interest, $\bar{\delta}$, to be

$$\bar{\delta} \sim MVN(0, 100I).$$  \hspace{1cm} (8)

The prior mean of zero implies that firm-level alphas are not associated with firm characteristics, which is not consistent with the considerable empirical evidence to the contrary. However, the informativeness of the prior depends on the prior variance. We specify a large prior variance indicating that we have little prior information about $\bar{\delta}$, so the prior mean has little effect on the posterior distribution of $\bar{\delta}$.\footnote{In unreported results, we considered non-zero prior means for each firm characteristic based on the evidence in the asset-pricing literature, but the impact on the posterior distributions was minimal due to the large prior variance.}

We specify the prior for firm-level betas as

$$\beta_{i,y} \sim N(1, 10).$$  \hspace{1cm} (9)

We use a prior mean equal to one because the average beta of firms in the market must equal one. We set the prior variance at 10, so the prior mean should have little impact on the posterior distribution of betas for most firms. For comparison, Vasicek (1973) recommends a prior variance of 0.25, which has a much stronger effect of shrinking firm betas toward one.\footnote{We also considered a hierarchical model structure for firm betas, similar to the model specified in equations (6) to (7) for firm-level alphas. However, we found that the posterior distributions for the parameter vector of interest, $\delta$, are almost identical using either the hierarchical prior or the prior specified above so we opt for the more parsimonious specification.}

It is also necessary to specify priors for $\{\sigma^2_{i,y}\}$, $\{\sigma^2_{\alpha,y}\}$, and $V$. We model $\{\sigma^2_{i,y}\}$ and $\{\sigma^2_{\alpha,y}\}$ using the Inverse Gamma distribution and $V$ with the Inverse Wishart distribution. The hyperparameters for these distributions are chosen to ensure that they have minimal influence on the posterior distributions. Our results are not sensitive to either doubling or halving the hyperparameter values.

We estimate the model using standard Markov chain Monte Carlo (MCMC) techniques. We are able to draw directly from the conditional posterior distributions for all model parameters using a Gibbs sampler. A detailed description of the estimation algorithm and the prior distributions and
associated hyperparameters is provided in Appendix A.1. We also conduct a series of simulation experiments to demonstrate the validity of the estimation approach as well as the robustness of inferences to various features of the cross section of firm returns. A summary of these results is provided in Appendix A.2.

2.3 Discussion

Our methodology has several desirable features relative to existing firm-level approaches in the literature such as Brennan, Chordia, and Subrahmanyam (1998), Avramov and Chordia (2006), and Fama and French (2008). The first advantage is that we make limited assumptions about the evolution of betas over time, assuming only that betas are relatively stable within each year. Fama and French (2006) note that minimizing the assumed structure on betas yields inferences that are less vulnerable to specification issues. Properly modeling betas is critical when testing for anomalies, since misspecifying betas can introduce spurious relations between alphas and firm characteristics.

In contrast to our parsimonious specification for firm betas, Avramov and Chordia (2006) allow betas to vary as an exact linear function of size, book-to-market, and macroeconomic variables. Such an approach requires the econometrician to know the “right” state variables (e.g., Harvey (1989), Shanken (1990), Jagannathan and Wang (1996), and Lettau and Ludvigson (2001)). If betas are related to other firm characteristics that are not included in the model, such as profitability or leverage, then firm alphas and betas will be systematically mismeasured. Thus, there is a joint hypothesis problem, as a statistical rejection of the CAPM may reflect either a failure of the model or a poor specification for firm betas.

Fama and French (2008) avoid defining complex dynamics for betas by regressing raw returns on firm characteristics to examine anomalies, implicitly assuming that all stocks have betas of one. However, even in the absence of a relation between alphas and firm characteristics, this approach will erroneously identify anomalies if betas are correlated with characteristics. There is considerable theoretical and empirical evidence that betas are related to firm characteristics (e.g., Karolyi (1992), Gomes, Kogan, and Zhang (2003), and Avramov and Chordia (2006)), so properly adjusting for market risk is important while testing whether alphas are forecastable.

The second advantage of our approach is that we estimate all model parameters simultaneously.

\footnote{In results not reported, we also estimated a version of our model using weekly subperiods and quarterly periods to allow for more frequent changes in firm betas. Our inferences were unchanged relative to the reported results.}
In contrast, other firm-level tests rely on two-step methods to examine the relation between alphas and firm characteristics (e.g., Brennan, Chordia, and Subrahmanyam (1998) and Avramov and Chordia (2006)). The two-step approach entails estimating alphas in a first step before regressing the alpha estimates on firm characteristics. Since alphas are estimated with error in the first step, the variance of the estimated alphas will be greater than the variance of the true firm alphas. Therefore, the uncertainty (i.e., the standard error) of the relation between alphas and firm characteristics will be overstated in the second-step regression. This measurement error problem may cause alphas to appear to be unforecastable even if a significant relation exists in the data. By estimating all of the model parameters simultaneously, the hierarchical Bayes approach eliminates this measurement error problem.

3 Data

This section outlines the sample construction and data requirements for estimating the model described in equations (5) to (7). We obtain accounting data from the Compustat Fundamentals Annual files and stock return data from CRSP. The sample includes all NYSE, Amex, and NASDAQ ordinary common stocks with the data required to compute at least one of the following firm characteristics: size ($M$), book-to-market ($BM$), momentum ($MOM$), reversal ($REV$), profitability ($ROA$), asset growth ($AG$), net stock issues ($NS$), accruals ($ACC$), and financial distress ($OS$).

Following Fama and French (1992), year $y$ runs from July of calendar year $y$ through June of calendar year $y + 1$. The characteristics are measured at the end of June in each calendar year $y$. The variables are matched to monthly returns from July of calendar year $y$ to June of calendar year $y + 1$. We exclude financial firms (SIC codes between 6000 and 6999) and firms with negative book equity. Based on Fama and French (2008), we classify firms into micro, small, and big categories using the 20th and 50th percentiles of market capitalization for NYSE stocks at the end of June of calendar year $y$.

The model described in Section 2 requires alphas and betas to be estimated for each firm-year observation. For a firm to be included in the estimation sample in a given year, we require 12 months of return data during that year. The final sample includes 163,603 firm-years of data from July 1963 to June 2008. We use the CRSP value-weighted stock market index as the proxy for the unobserved market portfolio. Monthly excess returns on the CRSP value-weighted stock market
4 Results

In this section we examine cross-sectional anomalies at the firm level using the hierarchical Bayes model developed in Section 2. Section 4.1 presents the initial firm-level results from the estimation of the model described in equations (5) to (7) and contrasts these results with those from traditional portfolio-level tests. Section 4.2 takes a more detailed look at CAPM anomalies at the firm level.

We use a Gibbs sampler to draw directly from the conditional posterior distributions of interest. The MCMC algorithm converges quickly. For all models we run the algorithm for 5,000 iterations and discard the first 2,500 as a burn-in period. To check the convergence of the algorithm, initially we ran the algorithm for 20,000 iterations and found that the posterior distributions characterized using iterations 2,500 to 5,000 were nearly identical to those based on iterations 17,500 to 20,000.

4.1 Firm-Level Tests

In Table I we examine the relation between conditional alphas and each firm characteristic in isolation. Panel A summarizes the posterior distribution of $\delta$ in equation (7), which measures the systematic relation between alphas and firm characteristics over the entire sample period. We find that seven of the nine firm characteristics are significantly associated with conditional firm-level alphas. Alphas are positively associated with book-to-market, momentum, and profitability and negatively associated with asset growth, net stock issues, accruals, and financial distress. Alphas are unrelated to size and reversal. In terms of economic significance, a one-standard-deviation change in any of the seven characteristics associated with alphas results in a change in alpha ranging in magnitude from 16 basis points (bps) per month for momentum ($0.51 \times 0.32$) to in excess of 20 bps per month for book-to-market, profitability, asset growth, and net stock issues.

For comparison, in Panel B of Table I we report results based on the traditional portfolio approach that is commonly used to identify anomalies. For each firm characteristic, we sort stocks

---

9http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/. We thank Kenneth French for making this data available.
10Following Avramov and Chordia (2006) and Fama and French (2008) we assume a linear relation between conditional alphas and firm characteristics.
Panel A presents the results from the estimation of the model described in equations (5) to (7) examining the cross-sectional relation between firm alphas and each firm characteristic separately. We report the posterior mean and standard deviation for the aggregate-level parameters, $\delta$, which provide information about the relation between alphas and firm characteristics across the entire sample period. An $^*$ (***) indicates that the 95% (99%) credible interval of the posterior distribution does not include zero. Panel B reports average conditional alphas in percent per month for hedge portfolios that are long the highest decile of stocks and short the lowest decile for each variable. Following Lewellen and Nagel (2006), the conditional CAPM alphas are estimated annually using monthly data. Standard errors are in parentheses. An $^*$ (***) indicates significance at the 5% (1%) level using a two-tailed test. The firm characteristics are described in Appendix B.

<table>
<thead>
<tr>
<th></th>
<th>$M$</th>
<th>$BM$</th>
<th>$MOM$</th>
<th>$REV$</th>
<th>$ROA$</th>
<th>$AG$</th>
<th>$NS$</th>
<th>$ACC$</th>
<th>$OS$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Base Specification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior Mean for the Aggregate-Level Parameters, $\delta$</td>
<td>0.05</td>
<td>0.24**</td>
<td>0.51**</td>
<td>-0.06</td>
<td>1.79**</td>
<td>-0.45**</td>
<td>-1.36**</td>
<td>-1.27**</td>
<td>-1.20*</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.19)</td>
<td>(0.08)</td>
<td>(0.51)</td>
<td>(0.09)</td>
<td>(0.21)</td>
<td>(0.22)</td>
<td>(0.46)</td>
</tr>
<tr>
<td>Average Cross-Sectional Standard Deviation of Firm Characteristics</td>
<td>1.89</td>
<td>0.86</td>
<td>0.32</td>
<td>0.78</td>
<td>0.15</td>
<td>0.58</td>
<td>0.16</td>
<td>0.13</td>
<td>0.14</td>
</tr>
<tr>
<td><strong>Panel B: Performance of Hedge Portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Conditional CAPM Alpha, $\hat{\alpha}_{CAPM}$</td>
<td>-1.08**</td>
<td>1.37**</td>
<td>0.61**</td>
<td>-0.78**</td>
<td>0.17</td>
<td>-1.23**</td>
<td>-1.14**</td>
<td>-0.54**</td>
<td>-0.18</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(0.20)</td>
<td>(0.20)</td>
<td>(0.23)</td>
<td>(0.29)</td>
<td>(0.17)</td>
<td>(0.16)</td>
<td>(0.12)</td>
<td>(0.27)</td>
</tr>
</tbody>
</table>

The results indicate that underlying firm-level associations can be obscured at the portfolio level. Comparing the firm-level results in Panel A to the portfolio-based tests in Panel B, we find that inferences differ for four of the nine firm characteristics: size, reversal, profitability, and financial distress. The results in Table II provide evidence that underlying firm-level associations can be obscured at the portfolio level. Comparing the firm-level results in Panel A to the portfolio-based tests in Panel B, we find that inferences differ for four of the nine firm characteristics: size, reversal, profitability, and financial distress.

In Table II, we consider alternative model specifications to further characterize the discrepancies.

---

between Panels A and B in Table I. The biggest difference between the firm-level and portfolio-level tests is that the portfolio-level tests only consider firms in deciles one and ten, ignoring valuable information contained in the remaining 80% of stocks. The firm-level approach, on the other hand, utilizes information from the entire cross section. Thus, the portfolio-level analysis could be unduly influenced by a small number of outlier observations in the extreme deciles. To investigate this possibility, in Panel A of Table II we specify alphas as a function of a constant, the firm characteristic, and two dummy variables identifying whether a particular firm lies in the top or bottom decile for that characteristic. The results suggest that the portfolio-level associations for size and reversal are driven primarily by the extremes. For example, there is no linear relation between alphas and size, but firms in the smallest decile earn alphas that are nearly 0.4% per month higher than firms in the largest decile. For all other firm characteristics, inferences are not substantially altered by the introduction of dummy variables.


The table presents the results from the estimation of the model described in equations (5) to (7) examining the cross-sectional relation between firm alphas and each firm characteristic separately. We report the posterior mean and standard deviation for the aggregate-level parameters, $\delta$, which provide information about the relation between alphas and firm characteristics across the entire sample period. Panel A shows estimates from a nonlinear specification including a linear component and dummy variables for firms with characteristic values in the top or bottom decile. Panel B shows estimates using sum betas. An * (**) indicates that the 95% (99%) credible interval of the posterior distribution does not include zero.

<table>
<thead>
<tr>
<th></th>
<th>$M$</th>
<th>$BM$</th>
<th>$MOM$</th>
<th>$REV$</th>
<th>$ROA$</th>
<th>$AG$</th>
<th>$NS$</th>
<th>$ACC$</th>
<th>$OS$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Nonlinear Specification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{\delta}$, Linear</td>
<td>0.08</td>
<td>0.23**</td>
<td>0.50*</td>
<td>-0.01</td>
<td>1.74**</td>
<td>-0.42**</td>
<td>-0.91**</td>
<td>-1.44**</td>
<td>-0.79*</td>
</tr>
<tr>
<td>(0.07)</td>
<td>(0.09)</td>
<td>(0.24)</td>
<td>(0.09)</td>
<td>(0.51)</td>
<td>(0.10)</td>
<td>(0.23)</td>
<td>(0.28)</td>
<td>(0.33)</td>
<td></td>
</tr>
<tr>
<td>$\bar{\delta}$, Decile 1</td>
<td>0.19</td>
<td>-0.07</td>
<td>-0.42**</td>
<td>-0.15</td>
<td>-0.15</td>
<td>-0.20</td>
<td>0.17*</td>
<td>-0.46**</td>
<td>-0.01</td>
</tr>
<tr>
<td>(0.12)</td>
<td>(0.11)</td>
<td>(0.14)</td>
<td>(0.15)</td>
<td>(0.17)</td>
<td>(0.14)</td>
<td>(0.08)</td>
<td>(0.13)</td>
<td>(0.09)</td>
<td></td>
</tr>
<tr>
<td>$\bar{\delta}$, Decile 10</td>
<td>-0.18*</td>
<td>-0.02</td>
<td>-0.31*</td>
<td>-0.27*</td>
<td>-0.15</td>
<td>-0.08</td>
<td>-0.16</td>
<td>-0.21*</td>
<td>-0.33*</td>
</tr>
<tr>
<td>(0.08)</td>
<td>(0.11)</td>
<td>(0.13)</td>
<td>(0.11)</td>
<td>(0.10)</td>
<td>(0.12)</td>
<td>(0.11)</td>
<td>(0.10)</td>
<td>(0.16)</td>
<td></td>
</tr>
<tr>
<td>Panel B: Sum Betas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior Mean for the Aggregate-Level Parameters, $\delta$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{\delta}$</td>
<td>0.09</td>
<td>0.23**</td>
<td>0.48*</td>
<td>-0.07</td>
<td>2.03**</td>
<td>-0.47**</td>
<td>-1.49**</td>
<td>-1.42**</td>
<td>-1.48**</td>
</tr>
<tr>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.22)</td>
<td>(0.09)</td>
<td>(0.54)</td>
<td>(0.09)</td>
<td>(0.24)</td>
<td>(0.23)</td>
<td>(0.47)</td>
<td></td>
</tr>
</tbody>
</table>

When conducting firm-level tests of the CAPM it is also important to consider the potential impact of nonsynchronous returns. If trading is infrequent, the betas measured in equation (5) by relating firm returns to contemporaneous market returns will tend to underestimate exposure to

\[12\] In results not reported, we considered other nonlinear specifications, including the addition of squared and cubed terms for each characteristic, but our inferences were unchanged.
the market factor. This issue is particularly relevant for our analysis if the extent to which a firm has nonsynchronous returns is associated with a given firm characteristic. To control for nonsynchronicities we follow Dimson (1979) and include the lagged excess market return as an additional factor in equation (5) to correct for any downward bias in measured betas. Panel B of Table II shows that allowing for nonsynchronicities has little impact on the relations between alphas and firm characteristics.

Although there is evidence in Tables I and II that firm-level associations between alphas and firm characteristics are frequently distorted at the portfolio level, the firm-level analysis nonetheless finds substantial evidence against the conditional CAPM. Seven of the nine firm characteristics are significantly associated with conditional CAPM alphas even after allowing for the possibility of nonlinearities and nonsynchronous returns. In the next section we take a more detailed look at the empirical shortcomings of the conditional CAPM.

4.2 A Closer Look at CAPM Anomalies

Given the results in Tables I and II, it is tempting to conclude that the CAPM provides a poor characterization of stock returns. However, in order to rigorously evaluate the empirical performance of the CAPM, we further assess the model across three dimensions. First, from an economic perspective, it is important to know whether anomalous patterns in returns are market-wide or limited to illiquid stocks that represent a small fraction of the total market capitalization. Second, it is important to distinguish between anomalies that reflect a fundamental failure of the CAPM and those that arise due to temporary mispricing or data snooping by researchers. Third, it is important to examine the extent to which firm characteristics identified as anomalies contain unique information about conditional alphas.

To examine whether anomalies are pervasive across size groups, we repeat the firm-level analysis from Table I but allow $\delta$ to vary across micro, small, and big stocks. The posterior distributions are presented in Figure 1 for each firm characteristic. Of the nine characteristics considered, seven are significantly related to the conditional alphas of micro stocks based on 95% credible intervals. In contrast, only three anomaly variables – asset growth, net stock issues, and accruals

---

13The robustness of anomalies across size subgroups is an active area of interest. For example, Loughran (1997) argues that the value effect is restricted to small stocks, while Fama and French (2006) show Loughran’s (1997) results are specific to the value/growth indicator, the sample period, and US stocks. Several other papers documenting individual anomalies conduct double sorts on size and a particular anomaly variable, with mixed results. Fama and French (2008) take a more comprehensive approach by analyzing the relations between returns and several firm characteristics within size subgroups.
Figure 1: Firm Characteristics and CAPM Alphas by Size Group
The figure presents the results from the estimation of the model described in equations (5) to (7) examining the cross-sectional relation between firm alphas and each firm characteristic separately. We report the posterior distributions for the aggregate-level parameters, $\delta$, which provide information about the relation between alphas and firm characteristics across the entire sample period. We estimate a model for each anomaly in which the aggregate-level parameters ($\delta$) vary across micro (dotted), small (dashed), and big (line) stocks.

- are significantly associated with the abnormal returns of big stocks. Moreover, the economic magnitudes of the relations are greatly reduced among big stocks relative to micro stocks. For example, a one-standard-deviation shock in asset growth has a 32 bps per month impact on micro stocks compared to just 12 bps for big stocks. The results in Figure 1 suggest that the CAPM provides a much more effective characterization of the returns of big stocks, which constitute over 90% of the total market capitalization.

To distinguish between anomalies that arise due to fundamental flaws in the CAPM and anomalies that arise due to temporary mispricing or data snooping by researchers, we examine the extent to which relations between firm characteristics and alphas persist after each characteristic is first documented as an anomaly. Anomalies that arise due to fundamental flaws in the CAPM are likely to persist over time, while anomalies that arise due to temporary mispricing or data snooping are likely to disappear after they are first documented.\textsuperscript{14}

In Table \textbf{III}, we re-estimate the model in equations (5) to (7), but unlike Panel A of Table \textbf{I}, in which $\delta$ is constant across the whole sample period, we allow $\delta$ to vary across the pre- and post-publication periods. We use publication dates based on the following papers: Banz (1981) (size), Schwert (2003) finds that the size and book-to-market effects appear to have attenuated after the anomalies were documented, while the momentum anomaly has persisted. Jegadeesh and Titman (2001) also find that the momentum anomaly appears to have persisted throughout the 1990s.

\textsuperscript{14}In prior research regarding the persistence of anomalies, Schwert (2003) finds that the size and book-to-market effects appear to have attenuated after the anomalies were documented, while the momentum anomaly has persisted. Jegadeesh and Titman (2001) also find that the momentum anomaly appears to have persisted throughout the 1990s.

In pre-publication periods the results in Table III show that alphas are positively related to book-to-market, momentum, and profitability, and negatively related to accruals and financial distress. In the post-publication periods, only book-to-market and accruals remain significantly associated with firm alphas. Moreover, the results for book-to-market and accruals are driven by micro stocks. Among big stocks, there is no evidence of any robust relations between firm characteristics and conditional alphas post publication.

The table presents the results from the estimation of the model described in equations (5) to (7) examining the cross-sectional relation between firm alphas and each firm characteristic separately. We report the posterior mean and standard deviation for the aggregate level parameters, $\delta$, which provide information about the relation between alphas and firm characteristics across time. We allow for different aggregate-level parameters in the pre- and post-publication periods. We estimate two models for each anomaly, one in which $\delta$ is restricted to be the same across all firms (All) and one in which $\delta$ varies across micro, small, and big stocks. An * (***) indicates that the 95% (99%) credible interval of the posterior distribution does not include zero.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Publication – Posterior Means for the Aggregate-Level Parameters, $\delta$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M</strong></td>
<td><strong>BM</strong></td>
<td><strong>MOM</strong></td>
<td><strong>REV</strong></td>
<td><strong>ROA</strong></td>
<td><strong>ACC</strong></td>
<td><strong>OS</strong></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>-0.00</td>
<td>0.24*</td>
<td>0.64**</td>
<td>-0.19</td>
<td>1.91**</td>
<td>-1.32**</td>
<td>-1.18*</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.11)</td>
<td>(0.24)</td>
<td>(0.11)</td>
<td>(0.58)</td>
<td>(0.25)</td>
<td>(0.52)</td>
</tr>
<tr>
<td>Micro</td>
<td>-0.03</td>
<td>0.31*</td>
<td>0.52*</td>
<td>-0.17</td>
<td>1.48**</td>
<td>-1.31**</td>
<td>-1.18**</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.13)</td>
<td>(0.20)</td>
<td>(0.13)</td>
<td>(0.46)</td>
<td>(0.22)</td>
<td>(0.40)</td>
</tr>
<tr>
<td>Small</td>
<td>0.12</td>
<td>0.24</td>
<td>0.53</td>
<td>-0.19</td>
<td>2.53**</td>
<td>-1.35**</td>
<td>-2.25**</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.15)</td>
<td>(0.29)</td>
<td>(0.14)</td>
<td>(0.60)</td>
<td>(0.38)</td>
<td>(0.60)</td>
</tr>
<tr>
<td>Big</td>
<td>-0.08</td>
<td>0.09</td>
<td>0.57</td>
<td>-0.12</td>
<td>1.16</td>
<td>-1.31**</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.14)</td>
<td>(0.37)</td>
<td>(0.15)</td>
<td>(0.93)</td>
<td>(0.39)</td>
<td>(0.84)</td>
</tr>
<tr>
<td>Post-Publication – Posterior Means for the Aggregate-Level Parameters, $\delta$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>0.08</td>
<td>0.23*</td>
<td>0.33</td>
<td>0.07</td>
<td>1.29</td>
<td>-0.97*</td>
<td>-1.19</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.11)</td>
<td>(0.33)</td>
<td>(0.11)</td>
<td>(0.98)</td>
<td>(0.43)</td>
<td>(0.94)</td>
</tr>
<tr>
<td>Micro</td>
<td>0.01</td>
<td>0.41**</td>
<td>0.22</td>
<td>-0.02</td>
<td>0.99</td>
<td>-1.35**</td>
<td>-0.92</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.13)</td>
<td>(0.26)</td>
<td>(0.13)</td>
<td>(0.75)</td>
<td>(0.40)</td>
<td>(0.72)</td>
</tr>
<tr>
<td>Small</td>
<td>0.11</td>
<td>0.24</td>
<td>0.39</td>
<td>0.04</td>
<td>1.67</td>
<td>-0.50</td>
<td>-1.91**</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.15)</td>
<td>(0.37)</td>
<td>(0.14)</td>
<td>(0.95)</td>
<td>(0.64)</td>
<td>(0.74)</td>
</tr>
<tr>
<td>Big</td>
<td>-0.01</td>
<td>0.12</td>
<td>0.36</td>
<td>-0.01</td>
<td>0.49</td>
<td>0.11</td>
<td>-1.80</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.14)</td>
<td>(0.47)</td>
<td>(0.16)</td>
<td>(1.58)</td>
<td>(0.70)</td>
<td>(1.05)</td>
</tr>
</tbody>
</table>
The figure presents the results from the estimation of the model described in equations (5) to (7) examining the cross-sectional relation between firm alphas and accruals. We report the posterior distributions for the aggregate-level parameters, $\delta$, which provide information about the relation between alphas and firm characteristics across time. We allow for different aggregate-level parameters in the pre- and post-publication periods and also allow the aggregate-level parameters to vary across micro (dotted), small (dashed), and big (line) stocks.

To highlight the patterns in the differences between pre- and post-publication relations between alphas and firm characteristics, in Figure 2 we plot the relation between accruals and firm alphas before and after the initial publication by Sloan in 1996. Pre publication there is a robust negative relation between accruals and alphas across stocks of all sizes. Post publication the negative relation persists among micro stocks, diminishes for small stocks, and disappears in big stocks. This pattern is consistent with market participants attempting to exploit the anomaly to earn abnormal returns. Among big stocks, where transaction costs are lowest and there are few, if any, short selling constraints, deviations from CAPM pricing are quickly eliminated. In contrast, investors appear to be unable to trade away the anomaly among micro stocks, where transaction costs are high, liquidity is low, and short selling is often difficult to implement (e.g., Jensen (1978)).

The pre-post analysis in Table III and Figure 2 provides little evidence that anomalies persist after they are first documented, especially among big firms. As such, our evidence is more consistent with the hypothesis that anomalies arise in the data due to either temporary market mispricing or data snooping by researchers.

Thus far our analysis has focused on the relation between conditional alphas and each firm characteristic in isolation. If firm characteristics are correlated with each other and offer little
Figure 3: Individual and Multiple Anomaly Variables

The figure presents the results from the estimation of the model described in equations (5) to (7) examining the cross-sectional relation between firm alphas and multiple firm characteristics simultaneously. We report the posterior distributions (line) for the aggregate-level parameters, $\delta$, which provide information about the relation between alphas and firm characteristics across the entire sample period. For comparison, for each anomaly variable we also present the posterior distribution (dashed) of $\delta$ from estimation of the model described in equations (5) to (7) for each characteristic in isolation using the same data sample.

unique information about alphas then studying each characteristic in isolation will overstate the failings of the conditional CAPM. The traditional portfolio approach is unable to adequately address this omitted variable problem. Researchers typically rely on two- or possibly three-way sorts to isolate the effects of a particular characteristic. Controlling for more than one or two characteristics simultaneously, however, is infeasible, and inferences are sensitive to both the sorting technique and the sorting sequence (e.g., Conrad, Cooper, and Kaul (2003)). In contrast, our approach is particularly well suited to assess which anomalies contain unique information; we simply specify firm-year alphas in equation (6) as a function of all nine firm characteristics from Table I.

In Figure 3 we compare the posterior distributions from two analyses – one in which each firm characteristic is considered in isolation and one in which all characteristics are considered simultaneously. Momentum, asset growth, and financial distress are significantly associated with CAPM alphas when considered in isolation, but as Figure 3 highlights, none of these characteristics contain significant incremental information when all firm characteristics are considered simultaneously. The only firm characteristics that are significantly related to firm-level alphas when multiple characteristics are considered simultaneously are book-to-market, profitability, net stock issues,
and accruals. Our analysis therefore suggests that univariate tests provide a low hurdle for firm characteristics to be classified as anomalies.

5 Conclusion

In this paper, we develop a hierarchical Bayes model to examine asset-pricing anomalies, modeling firm-year alphas as a function of one or more firm characteristics. We investigate nine anomalies – size, book-to-market, momentum, reversal, profitability, asset growth, net stock issues, accruals, and financial distress – over the period 1963 to 2008. Studying each anomaly separately we find robust evidence that conditional CAPM alphas are positively associated with book-to-market, momentum, and profitability, and negatively associated with asset growth, net stock issues, accruals, and financial distress.

These initial results imply the failings of the CAPM are widespread. A deeper investigation of anomalies, however, suggests that while the CAPM may not perfectly explain firm returns, the anomaly-based evidence against the CAPM is greatly overstated. Relations between firm characteristics and conditional firm-level alphas are primarily confined to micro and small stocks and tend not to persist after the anomalies are first documented. Furthermore, few of the firm characteristics associated with alphas actually contain unique information.

\[^{15}\text{In results not reported we also considered a model specification in which conditional alphas were modeled as a function of multiple firm characteristics and the relations were allowed to vary across micro, small, and big stocks. As in Figure 1 the relations between characteristics and alphas are generally driven by micro and small stocks.}\]
A Model Appendix

Section A.1 provides details about the MCMC estimation algorithm, and Section A.2 presents a simulation study that demonstrates the ability of the algorithm to accurately recover parameters.

A.1 Estimation Methodology

The model outlined in equations (5) to (7) can be estimated by repeatedly cycling through steps 1 to 6 below. As discussed in the text, we place a hierarchical structure on alphas, but not on betas. Instead we impose a proper, but diffuse, prior directly on betas in the base specification, $\beta_{i,y} \sim N(1, 10)$. Let $r_{i,t,y}$ denote the excess return on stock $i$ in month $t$ of year $y$ and $r_{m,t,y}$ the excess return on the market portfolio. Further, let $Z_y$ denote a matrix in which the first column is a vector of ones and the second column is the excess returns on the market portfolio, and let $X_y$ denote a matrix of a constant and firm-year characteristics associated with anomalies.

1. Draw $\alpha_{i,y}, \beta_{i,y} | \sigma_{i,y}^2, \delta_y, \sigma_{\alpha,y}^2$ for each stock $i = 1, \ldots, N$, in each year $y = 1, \ldots, Y$. We obtain a draw from the marginal posterior distribution of $\alpha_{i,y}$ and $\beta_{i,y}$ as follows:

$$
\begin{bmatrix}
\alpha_{i,y} \\
\beta_{i,y}
\end{bmatrix} \sim N\left(\bar{\lambda}_i, \left(\sigma_{i,y}^{-2} Z'_y Z_y + V^{-1}_\lambda\right)^{-1}\right),
$$

(A.1)

where

$$
\bar{\lambda}_i = \left(\sigma_{i,y}^{-2} Z'_y Z_y + V^{-1}_\lambda\right)^{-1}\left(\sigma_{i,y}^{-2} Z'_y Z_y \hat{\lambda}_i + V^{-1}_\lambda \lambda_{i,y}\right),
$$

(A.2)

$$
\hat{\lambda}_i = \left(Z'_y Z_y\right)^{-1} Z'_y r_{i,y},
$$

(A.3)

$$
\bar{\lambda}_{i,y} = \begin{bmatrix} X_{i,y} \delta_y \\ 1 \end{bmatrix},
$$

(A.4)

and

$$
V_\lambda = \begin{bmatrix} \sigma_{\alpha,y}^2 & 0 \\ 0 & 10 \end{bmatrix}.
$$

(A.5)

2. Draw $\sigma_{i,y}^2 | \alpha_{i,y}, \beta_{i,y}$ for each stock $i = 1, \ldots, N$, in each year $y = 1, \ldots, Y$. We obtain a draw from the marginal posterior distribution of $\sigma_{i,y}^2$ as follows:

$$
\sigma_{i,y}^2 \sim \text{Inverse Gamma} \left(\frac{v_1 s_1^2}{2}, \frac{v_1}{2}\right),
$$

(A.6)
\[ v_1 = v_0 + M, \]  
(A.7)

and

\[ s_1^2 = \frac{v_0 s_0^2 + s^2}{v_0 + M}, \]  
(A.8)

where \( s^2 \) is the sample sum of squared errors and \( M \) denotes the number of observations. The priors, \( v_0 \) and \( s_0^2 \), are determined by the researcher. We set \( v_0 \) equal to 3 and \( s_0^2 \) equal to the variance of the monthly returns for stock \( i \) in year \( y \).

3. Draw \( \delta_y \mid \{ \alpha_{i,y} \}, \sigma_{\alpha,y}^2, \bar{\delta}, V \) for each year \( y = 1, ..., Y \). Let \( \alpha_y \) denote a column vector composed of draws of \( \alpha_{i,y} \) for all firms \( i \) in the dataset in year \( y \). We obtain a draw from the marginal posterior distribution of \( \delta_y \) as follows:

\[ \delta_y \sim N \left( \overline{\delta_y}, (\sigma_{\alpha,y}^{-2} X_y' X_y + V^{-1})^{-1} \right), \]  
(A.9)

where

\[ \overline{\delta_y} = (\sigma_{\alpha,y}^{-2} X_y' X_y + V^{-1})^{-1}(\sigma_{\alpha,y}^{-2} X_y' X_y \hat{\delta}_y + V^{-1} \bar{\delta}), \]  
(A.10)

and

\[ \hat{\delta}_y = (X_y' X_y)^{-1} X_y' \alpha_y. \]  
(A.11)

4. Draw \( \sigma_{\alpha,y}^2 \mid \{ \alpha_{i,y} \}, \delta_y \) for each year \( y = 1, ..., Y \). We obtain a draw from the marginal posterior distribution of \( \sigma_{\alpha,y}^2 \) as follows:

\[ \sigma_{\alpha,y}^2 \sim \text{Inverse Gamma} \left( \frac{v_1 s_1^2}{2}, \frac{v_1}{2} \right), \]  
(A.12)

\[ v_1 = v_0 + M, \]  
(A.13)

and

\[ s_1^2 = \frac{v_0 s_0^2 + s^2}{v_0 + M}, \]  
(A.14)

where \( s^2 \) is the sample sum of squared errors and \( M \) denotes the number of observations. The priors, \( v_0 \) and \( s_0^2 \), are determined by the researcher. We set \( v_0 \) equal to 3. We elicit priors for \( s_0^2 \) in the following manner. For each stock in year \( y \) we estimate equation (5) using OLS and store \( \hat{\alpha} \). We set \( s_0^2 \) equal to the variance of \( \hat{\alpha} \) across all firms in year \( y \).
Having drawn the firm- and year-level coefficients we proceed to draw the aggregate-level parameters. Let $P$ denote a $Y \times nvar$ matrix comprised of a draw of $\{\delta_y\}_{y=1}^Y$, where $nvar$ denotes the number of columns in $X$, and let $H$ be a matrix of covariates the researcher believes to be associated with the evolution of the parameter vector $\delta_y$ over time. In our specification, $H$ is a column vector of ones, but could easily be extended, for example, to include macroeconomic variables.

5. Draw $V | \{\delta_y\}$. We obtain a draw from the marginal posterior distribution of $V$ as follows:

$$V \sim \text{Inverse Wishart} \left( nvar + Nu + Y, V_0 + S \right), \quad (A.15)$$

where

$$S = \left( P - H\tilde{\Gamma} \right) \left( P - H\tilde{\Gamma} \right)' + \left( \tilde{\Gamma} - \Gamma \right)' A \left( \tilde{\Gamma} - \Gamma \right), \quad (A.16)$$
$$\tilde{\Gamma} = \left( \left( H'H + A \right)^{-1} \left( H'H\tilde{\Gamma} + A\Gamma \right) \right), \quad (A.17)$$

and

$$\tilde{\Gamma} = \left( H'H \right)^{-1} \left( H'P \right). \quad (A.18)$$

$A$, $\Gamma$, $Nu$ and $V_0$ are priors specified by the researcher. We set $A^{-1} = 100I$ and define $\Gamma$ to be an $n_H \times nvar$ matrix of zeros, where $n_H$ denotes the number of columns in $H$. $Nu$ is set to $nvar + 3$, and $V_0 = NuI$. $I$ denotes an appropriately dimensioned identity matrix.

6. Draw $\gamma | \{\delta_y\}, V$. We obtain a draw from the marginal posterior distribution of $\gamma$ as follows:

$$\gamma \sim N \left( \tilde{\gamma}, V \otimes \left( H'H + A \right)^{-1} \right), \quad (A.19)$$

where $\tilde{\gamma} = \text{vec} \left( \tilde{\Gamma} \right)$. Given that $H$ is a vector of ones, $\tilde{\delta} = \gamma$.

A.2 Model Simulation

In this section, we conduct a simulation exercise and show our estimation algorithm successfully recovers the parameters of interest. Data are created for 1,000 firms over a 45-year period. The length of each time period, $y$, is set to 12 months. We assume there are two firm characteristics associated with firm-level alphas, $x_1$ and $x_2$, which are both uniformly distributed over the range -0.5 to +0.5. The parameters in the simulation are set to ensure that the simulated firm-level returns, alphas, betas, and market returns are consistent with the actual values observed using the
CRSP return data.

1. Draw $\delta_y \sim \text{MVN} \left( \mu = \delta, \sigma^2 = V \right)$ for each 12-month time period, $y$. We set

$$
\delta = \begin{bmatrix}
\delta_0 = 0 \\
\delta_1 = 1 \\
\delta_2 = 1
\end{bmatrix}, \quad \text{and} \quad V = \begin{bmatrix}
1.5 & 0.5 & 0.5 \\
0.5 & 1.5 & 0.5 \\
0.5 & 0.5 & 1.5
\end{bmatrix}.
$$

2. Draw $\alpha_y \sim \text{MVN} \left( \mu = \delta_{0,y} + \delta_{1,y}x_1 + \delta_{2,y}x_2, \sigma^2 = \Sigma_\alpha \right)$, where $\alpha_y$ is a column vectors of firm-specific alphas in time period $y$. We consider two specifications for the variance-covariance matrix, $\Sigma_\alpha$, one in which the error terms are independent across firms, and one in which the error terms are correlated across firms. We examine two different levels of correlations, low to medium with correlations ranging from $-0.5$ to $+0.5$, and medium to high with correlations ranging from $-0.9$ to $+0.9$. The diagonal elements of $\Sigma_\alpha$ are set equal to $\sigma^2 = 2$.\[16\]

3. Draw $\beta_{i,y} \sim N \left( \mu = 1, \sigma^2 = 4 \right)$ for each firm $i$ in each time period $y$.

4. Generate excess monthly returns on the market: $r_{m,t,y} \sim N \left( \mu = 0.5, \sigma^2 = 25 \right)$.

5. Generate monthly excess returns for each firm in each month of each time period: $r_{t,y} = \alpha_y + \beta_y r_{m,t,y} + \epsilon_{t,y}$, where $\epsilon_{t,y} \sim \text{MVN} \left( \mu = 0, \sigma^2 = \Sigma_{ret} \right)$ and $r_{t,y}$ denotes a column vector of excess returns for all firms in month $t$ of time period $y$. $\alpha_y$ and $\beta_y$ are column vectors of firm-specific alphas and betas. The specifications for the variance-covariance matrix, $\Sigma_{ret}$, are constructed in a similar manner to those for $\Sigma_\alpha$. The only difference is that the diagonal elements of $\Sigma_{ret}$, $\sigma^2_{ret}$, are set equal to 169.

We examine seven different scenarios to investigate the sensitivity of our model to different correlation structures in the error terms of equations (5) and (6). The MCMC algorithm is run for 1,000 iterations for each scenario. The algorithm converges quickly. The posterior distributions are characterized using the final 500 iterations. We use the same seed for the random number generator for each scenario. Table A.1 reports the results from the simulation study. Regardless of the correlation structure in the error terms of equations (5) and (6), the estimation algorithm

---

\[16\] We use the following procedure to create a $1,000 \times 1,000$ variance-covariance matrix. First, create a column vector, $u$, with 1,000 draws from the Uniform(-1,1) distribution. Second, calculate $\kappa u u'$ where $\kappa = p\sigma^2_\alpha$. The parameter, $p$, is a scaling factor, between 0 and 1, for the maximum level of correlation in the error terms across firms. If $p = 0$, firm-level alphas are independent. If $p = 1$, $\kappa u u'$ correlations range from $-1$ to $+1$. For low to medium correlations we set $p = 0.5$, while for medium to high correlations we set $p = 0.9$. Finally, set $\Sigma_\alpha = \kappa u u'$ and replace the diagonal elements with $\sigma^2_\alpha = 2$.\[22\]
is able to accurately recover the aggregate-level model parameters, $\bar{\delta}$ and $V$, indicating that the approach is not sensitive to the possibility of cross-correlations across firms.
Table A.I: Model Estimation on Simulated Data.
The table presents the results from the estimation of the model described in equations (5) to (7) for simulated data. We report the posterior mean and standard deviation for the aggregate-level parameters $\delta$ and $V$. We simulate data for 1,000 firms over a 45-year period. We estimate the model using annual periods and monthly subperiods. We create seven different sets of data using the same seed for the random number generator in each scenario. Each set of data differs only with respect to the assumptions about cross-correlations in the error terms of equation (5) (monthly firm returns) and/or equation (6) (firm-year alphas). Specifically, we allow cross-correlations in each equation to take on one of three levels: zero, low, or high. The low level allows cross-correlations to range between $\pm 0.5$, while the high level allows cross-correlations to range between $\pm 0.9$. We run the Gibbs sampler for 1,000 iterations and discard the first 500 as a burn-in period. An * (**) indicates that the 95% (99%) credible interval of the posterior distribution does not include the true value.

<table>
<thead>
<tr>
<th>Case</th>
<th>Cross-Correlation in Return Error (Equation (5))</th>
<th>Cross-Correlation in Alpha Error (Equation (6))</th>
<th>$\bar{\delta}_0$</th>
<th>$\bar{\delta}_1$</th>
<th>$\bar{\delta}_2$</th>
<th>$V_{11}$</th>
<th>$V_{22}$</th>
<th>$V_{33}$</th>
<th>$V_{12}$</th>
<th>$V_{13}$</th>
<th>$V_{23}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Values</td>
<td></td>
<td></td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.50</td>
<td>1.50</td>
<td>1.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Case 1</td>
<td>None</td>
<td>None</td>
<td>-0.22</td>
<td>0.94</td>
<td>1.07</td>
<td>1.86</td>
<td>1.74</td>
<td>1.67</td>
<td>0.57</td>
<td>0.68</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>None</td>
<td>-0.21</td>
<td>0.87</td>
<td>1.06</td>
<td>1.86</td>
<td>1.79</td>
<td>1.65</td>
<td>0.51</td>
<td>0.67</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>None</td>
<td>-0.20</td>
<td>0.84</td>
<td>1.08</td>
<td>1.87</td>
<td>1.79</td>
<td>1.65</td>
<td>0.49</td>
<td>0.66</td>
<td>0.78</td>
</tr>
<tr>
<td>Case 4</td>
<td>None</td>
<td>Low</td>
<td>-0.22</td>
<td>0.93</td>
<td>1.07</td>
<td>1.86</td>
<td>1.74</td>
<td>1.72</td>
<td>0.56</td>
<td>0.68</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>High</td>
<td>-0.22</td>
<td>0.93</td>
<td>1.07</td>
<td>1.86</td>
<td>1.74</td>
<td>1.72</td>
<td>0.56</td>
<td>0.67</td>
<td>0.80</td>
</tr>
<tr>
<td>Case 5</td>
<td>None</td>
<td>Low</td>
<td>-0.21</td>
<td>0.86</td>
<td>1.06</td>
<td>1.86</td>
<td>1.82</td>
<td>1.70</td>
<td>0.51</td>
<td>0.66</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>High</td>
<td>-0.20</td>
<td>0.83</td>
<td>1.08</td>
<td>1.88</td>
<td>1.82</td>
<td>1.69</td>
<td>0.49</td>
<td>0.64</td>
<td>0.84</td>
</tr>
</tbody>
</table>
B Data Appendix

We obtain accounting data from the Compustat Fundamentals Annual files and stock return data from the CRSP monthly return files. Each of the anomaly variables is measured once a year at the end of June in calendar year $y$. The variables are matched to returns from July of calendar year $y$ to June of calendar year $y+1$. To ensure that the accounting data are known prior to the returns they are used to forecast, we lag all accounting variables by six months. The sample includes all NYSE, Amex, and NASDAQ ordinary common stocks with the data required to compute at least one of the following anomaly variables:

1. **$M$ (Size)**: The natural log of price per share times the number of shares outstanding at the end of June of year $y$.

2. **$BM$ (Book-to-market)**: The natural log of the ratio of book value of equity to market value of equity. Following Fama and French (2008), we define the book value of equity as total assets ($at$), minus total liabilities ($lt$), plus balance sheet deferred taxes and investment tax credits ($txditc$) if available, minus the book value of preferred stock if available. Depending on availability, we use liquidating value ($pstkl$), redemption value ($pstkrv$), or carrying value ($upstk$) for the book value of preferred stock. The market value of equity is price per share times the number of shares outstanding at the end of December of year $y-1$.

3. **$MOM$ (Momentum)**: The continuously compounded stock return from January to June of year $y$. We require a firm to have a price for the end of December of year $y-1$ and a good return for June of year $y$.

4. **$REV$ (Reversal)**: The continuously compounded stock return from July of year $y-5$ to June of year $y-1$. We require a firm to have a price for the end of June of year $y-5$ and a good return for June of year $y-1$.

5. **$ROA$ (Profitability)**: Income before extraordinary items ($ib$), minus dividends on preferred ($dvp$) if available, plus income statement deferred taxes ($txdi$) if available divided by total assets ($at$).

6. **$AG$ (Asset growth)**: Total assets ($at$) at the fiscal year end in year $y-1$, minus total assets at the fiscal year end in year $y-2$ divided by total assets at the fiscal year end in year $y-2$. We also require a firm to have non-zero total assets in both year $y-1$ and $y-2$. 

25
7. **NS** (Net stock issues): The natural log of the ratio of split-adjusted shares at the fiscal year end in year \(y - 1\) divided by split-adjusted shares at the fiscal year end in year \(y - 2\). The number of split-adjusted shares outstanding is common shares outstanding from Compustat (csho) times the cumulative adjustment factor by ex-date (adjex.f).

8. **ACC** (Accruals): The change in current assets (act) from the fiscal year end in year \(y - 2\) to \(y - 1\), minus the change in current liabilities (lct), minus the change in cash and short-term investments (che), plus the change in debt in current liabilities (dlc), minus depreciation (dp) in fiscal year \(y - 1\) divided by total assets (at) from the fiscal year end in year \(y - 2\).

9. **OS** (Financial distress): Ohlson’s (1980) \(O\)-score:

\[
O\text{-score} = \frac{1}{1 + \exp(-x)},
\]

where

\[
x = -1.32 - 0.407 \times (SIZE) + 6.03 \times (TLTA) - 1.43 \times (WCTA) \\
+ 0.076 \times (CLCA) - 1.72 \times (OENEG) - 2.37 \times (NITA) - 1.83 \times (FUTL) \\
+ 0.285 \times (INTWO) - 0.521 \times (CHIN),
\]

where \(SIZE\) is the log of the ratio of total assets (at) to the GNP price-level index, \(TLTA\) is the ratio of total liabilities (lt) to total assets, \(WCTA\) is the ratio of working capital (act – lct) to total assets, \(CLCA\) is the ratio of current liabilities (lct) to current assets (act), \(OENEG\) is a dummy variable equal to one if total liabilities exceeds total assets and zero otherwise, \(NITA\) is the ratio of net income (ni) to total assets, \(FUTL\) is the ratio of funds from operations (pi) to total liabilities, \(INTWO\) is a dummy variable equal to one if total net income was negative for the past two years and zero otherwise, and \(CHIN\) is the change in net income from fiscal year \(y - 2\) to \(y - 1\) divided by the sum of the absolute values of net income in fiscal years \(y - 2\) and \(y - 1\). Data on the GNP price-level index are from the Federal Reserve Bank of St. Louis website. Following Ohlson (1980), we assign the index a value of 100 in 1968, and the index year is as of the year prior to the year of the balance sheet date.

\[\text{http://research.stlouisfed.org/fred2/}.\]
We exclude financial firms (SIC codes between 6000 and 6999) and firms with negative book equity. The sample period is July 1963 to June 2008. To alleviate the influence of outliers, we winsorize ROA, AG, NS, and ACC at the 1st and 99th percentiles. For cases in which a firm is delisted from an exchange during a given month, we replace any missing returns with the delisting returns provided by CRSP.
References


