Fundamental Anomalies*

Erica X.N. Li, † Guoliang Ma, ‡ Shujing Wang, § and Cindy Yu ¶

June, 2021

Abstract

This paper examines to what extent stock market anomalies are driven by firm fundamentals in an investment-based asset pricing framework. Using Bayesian Markov Chain Monte Carlo (MCMC), we estimate a two-capital q-model to match firm-level stock returns, instead of matching portfolio-level return moments. Our methodology addresses Campbell (2017)'s critique on prior studies that model parameters are chosen to fit a specific set of anomalies and different values are needed to fit each anomaly. The estimated model generates large and significant size, momentum, profitability, investment, and intangibles premiums. However, it falls short in explaining the value and accruals anomalies.

Keywords: q-theory, Investment, Profitability, Momentum, BMCMC estimation

*We have benefited from helpful comments from Li An, Hang Bai, Frederico Belo, Andrei Gonçalves, Xiaodan Gao, Bing Han, Xiaoji Lin, Clark Liu, Raymond Liaung, Xiaomeng Lu, Danqing Mei, Zhongzhi Song, Chen Xue, Lu Zhang, Jing Zhao, and seminar and conference participants at Cheung Kong Graduate School of Business, Renmin Business School, Shanghai Advanced Institute of Finance, PBCSF School of Finance, and University of Connecticut (the 6th annual finance conference). All remaining errors are our own.

†Department of Finance, Cheung Kong Graduate School of Business, Beijing, China, 100738; Tel: (86)10 85188888 ext. 3075; E-mail: xnli@ckgsb.edu.cn.

‡Department of Statistics, Iowa State University, Ames, IA 50011; E-mail: glma@iastate.edu.

§Department of Economics and Finance, School of Economics and Management, Tongji University, Shanghai, China 200092; Tel: (86) 21 65982274; E-mail: shujingwang@connect.ust.hk.

¶Department of Statistics, Iowa State University, Ames, IA 50011; Tel: 1 515 2946885; E-mail: cindyyu@iastate.edu.

Conflict-of-Interest Disclosure Statement

Erica X.N. Li

In accordance with the Journal of Finance policies and my ethical obligation as a researcher, I am reporting that I have nothing to disclose.

Guoliang Ma

In accordance with the Journal of Finance policies and my ethical obligation as a researcher, I am reporting that I have nothing to disclose.

Shujing Wang

In accordance with the Journal of Finance policies and my ethical obligation as a researcher, I am reporting that I have nothing to disclose.

Cindy Yu

In accordance with the Journal of Finance policies and my ethical obligation as a researcher, I am reporting that I have nothing to disclose.

1 Introduction

The investment-based asset pricing literature studies returns from the supply side of the economy and formulates returns based on firm fundamentals, under the assumptions that a firm operates to optimize its market equity.[1] Hayashi (1982) shows that with homogeneous degree one production technology, a firm's return on investment, which is a function of its fundamentals, equals its weighted average cost of capital (WACC). Liu, Whited and Zhang (2009) show that such a simple q-model fits value, earnings surprises, and investment anomalies well, when the parameters of the model are estimated based on this identity using the average returns of decile portfolios sorted by these anomaly variables as the target moments. Additional asset pricing anomalies can be explained in this framework as shown in subsequent studies, such as Belo, Xue and Zhang (2013) among others.

In this line of research, model parameters are estimated via the Generalized Method of Moments (GMM) with the average returns of testing portfolios as target moments. However, applying the Hayashi (1982) identity at the portfolio level is likely to miss information orthogonal to the sorting variables underlying these portfolios, and the resulting parameter estimates are generally portfolio-dependent. As Campbell (2017) (page 213) puts it: "This problem, that different parameters are needed to fit each anomaly, is a pervasive one in the q-theoretic asset pricing literature".

To address this critique, we estimate a two-capital q-model (Gonçalves, Xue and Zhang, 2020) to match the entire sample of firm-level stock returns using the Bayesian Markov Chain Monte Carlo (MCMC) method. Prior studies have used this method to match the dynamics of macroeconomic variables (Smets and Wouters, 2007; Li et al., 2019, among others) and the time series of returns on the market index (Li, Wells and Yu, 2008, among others). The resulting parameter values in our estimation are portfolio-independent and fully reflect the information in the data. As such, we can examine the capability of the estimated q-model to explain stock return anomalies in general. Using simulation studies, we show that Bayesian MCMC is able to discover the true parameter values in the context of our model framework.

Based on the estimated model parameters and observed firm fundamentals such as sales-

[1] Examples include Cochrane (1991), Berk, Green and Naik (1999), Gomes, Kogan and Zhang (2003), Carlson, Fisher and Giammarino (2004), Zhang (2005), Li, Livdan and Zhang (2009), Papanikolaou (2011), Kogan and Papanikolaou (2014), and Bazdresch et al. (2014), among others.

to-physical capital and investment-to-physical capital ratios, we compute firm-level fundamental stock returns for the sample from 1967 to 2017. The distribution of the fundamental returns closely resembles that of the realized ones in terms of the mean (15.65% versus 14.45% per annum), skewness (1.68 versus 2.15), and kurtosis (11.20 versus 11.05). The fundamental returns exhibit close to 60% of the variability of the realized ones, with a standard deviation of 34.17% compared to 60.78% in the data.[2] The time series average of the cross-sectional correlation between fundamental and realized firm-level stock returns is 0.20. After averaging out firm-level noises, the fundamental returns on the value-weighted market index closely co-move with the realized ones, with a correlation of 0.77. Overall, the model does a good job in matching the distribution and dynamics of firm-level stock returns.

To examine the capability of the q-theory in explaining stock market anomalies, we consider 12 well-documented anomalies covering six major categories:[3] size anomaly sorted on market capitalization (Size); value anomaly sorted on book-to-market equity ratio (BM); momentum anomaly sorted on the prior 11-month returns skipping the most recent month (R11); four investment anomalies sorted on asset growth (I/A), net stock issues (NSI), investment-to-assets ratio (ΔPI/A), and accruals (Accruals); three profitability anomalies sorted on return-on-equity (ROE), return-on-assets (ROA), and gross profitability (GP/A); and two intangibles anomalies sorted on R&D expense-to-market ratio (RD/M) and advertising expense-to-market ratio (Ad/M). For each of these anomalies, we compute returns on the corresponding decile portfolios using realized and fundamental stock returns separately.

The most important result of the paper is that the fundamental returns exhibit large and significant size, momentum, investment (except the accruals), profitability, and intangibles premiums. Specifically, the posterior means for each premium and its t-statistic are as follows: the fundamental size premium is -5.99% per annum (t=-5.63), the momentum premium is 11.82% (t=12.51), the I/A premium is -3.08% (t=-2.25), the NSI premium is -3.05% (t=-3.36), the ΔPI/A premium is -5.79 (t=-4.81), the ROE premium is 4.62% (t=5.72), the ROA premium is 3.80% (t=3.99), the GP/A premium is 7.26% (t=5.84), the RD/M premium is 5.24% (t=2.12), and the Ad/M premium is 7.46% (t=2.82). In addition, the difference between the realized and fundamental premiums, defined as alpha, is statistically insignificant at the 5% level for seven out of these ten anomalies, with the

[2]The statistics of fundamental stock returns reported here are the means of the posterior distributions.
[3]We follow the classification in Hou, Xue and Zhang (2020).

exception of the I/A, NSI, and GP/A premiums.

Despite the aforementioned success, the model falls short in two ways. First, the model explains the value premium only in the first half of the sample but fails in the second half: the alpha of the value premium is 3.52% (t=1.66) in the June 1967-June 1991 period but is 8.87% (t=3.06) in the July 1991-December 2017 period. Second, the model cannot generate the accruals anomaly, which is negative in the data, −5.58% (t=−3.14), but positive in the model, 4.74% (t =4.45). An explicit modeling of intangible capitals and the distinction between cash and accrual basis accountings might be needed to explain these two anomalies.

In our baseline estimation, we allow the structural parameters to vary across the Fama-French 10 industries and change over time. For robustness, we conduct three alternative estimations: parameters with industry variations only, with time variations only, and with constant values, respectively. The results show that industry and time variations significantly improve the ability of the model to match firm-level stock returns. However, these variations do not necessarily improve the overall ability of the model to explain anomalies and in some cases they even hurt the performance. The reason is that anomaly premiums are not the target of the estimations. The specification that matches the firm-level returns best does not necessarily explain anomalies better.

More importantly, the ability of the model to explain anomalies does not come from the look-ahead advantage of the in-sample estimation. We construct one-period-ahead fundamental returns using recursive estimation with expanding window. And the ability of the model to explain anomalies does not change qualitatively.

Finally, we show that the simple q-model combined with the Bayesian MCMC has reliable out-of-sample predictive power. The average realized return spread between firms with high and low predicted returns is large and significant (0.45% per month with t=2.45). Moreover, this return spread cannot be explained by the Capital Asset Pricing Model (CAPM), Fama-French factor models, nor by the Hou-Xue-Zhang q-factor model. Given that these linear risk-factor models have poor out-of-sample performance (Fama and French, 1997; Gonçalves, Xue and Zhang, 2020), our results highlight the importance of the model's simple yet powerful economic structure to its out-of-sample performance.

Our work is built directly on Liu, Whited and Zhang (2009), Liu and Zhang (2014), and Gonçalves, Xue and Zhang (2020). These papers conduct GMM estimations of various q-models using average anomaly portfolio returns as target moments. Liu, Whited and Zhang

(2009) show that a one-capital q-model can match the average returns of portfolios sorted on earnings surprises, book-to-market equity, and capital investment. Liu and Zhang (2014) use the same model and estimation procedure to explain the momentum premium. However, the parameter values vary with testing portfolios substantially. Gonçalves, Xue and Zhang (2020) estimate a two-capital q-model to match the average returns of 40 decile portfolios sorted on book-to-market equity, asset growth, return-on-equity, and momentum. They show that when fundamental returns are computed at firm level rather than at portfolio level, parameter estimates are more stable due to better aggregation. Different from these previous studies, our estimation method does not involve aggregation and portfolios.

In a similar vein, Belo, Xue and Zhang (2013) estimate a q-model by matching average q at the portfolio level, in addition to matching return moments. Belo et al. (2021) estimate a q-model with both tangible and intangible capitals by matching the time series of portfolio-level cross-sectional mean valuation ratios for a given set of testing portfolios. The estimation method in Belo et al. (2021) allows the dynamics of valuation ratios to be better captured. Our method can be easily applied to explain firm-level valuation ratios, which is a promising future direction.

Finally, our paper belongs to the growing literature in finance that combines structural models with data using various estimation methods (e.g., Eisfeldt and Muir, 2016; Bazdresch, Kahn and Whited, 2018; Taylor, 2013) to make causal and quantitative inferences about the underlying economic mechanisms. Bayesian MCMC used in this paper is a useful addition to the toolbox of structural estimation.

The rest of the paper is organized as follows. Section 2 outlines the model. Section 3 explains the data used in the estimation and the construction of anomalies. Section 4 describes the estimation procedure and verify the accuracy of Bayesian MCMC estimates under our model framework using simulation studies. Section 5 presents the estimation results and model-implied fundamental anomaly premiums. Section 6 explores the economic mechanisms behind the capability of the estimated model in explaining anomalies and the limitations of the model. Section 7 discusses the dynamic features of the fundamental factor premiums, estimation with expanding window, and out-of-sample forecasts. Section 8 concludes.

2 The Model

We adopt the two-capital model in Gonçalves, Xue and Zhang (2020), in which firms use three inputs in production: long-term physical capital (K), short-term working capital (W), and costlessly adjustable input (S) such as energy and purchased service, the prices of which are taken as given by firms. Operating profit of firm i in industry j at time t is $\Pi_{it} = \Pi(K_{it}, W_{it}, S_{it})$, which exhibits constant-return-to-scale. Under the assumption of a perfectly competitive and frictionless market for input S, S_{it} is chosen to maximize contemporaneous operating profits. With Cobb-Douglas production technology, marginal products of physical and working capital are given by $\partial\Pi_{it}/\partial K_{i,t} = \gamma_{jt}^{K} Y_{it}/K_{it}$ and $\partial\Pi_{it}/\partial W_{i,t} = \gamma_{jt}^{W} Y_{it}/W_{it}$, respectively, in which γ_{jt}^{K}, $\gamma_{jt}^{W} > 0$ are the corresponding shares of capital in sales Y_{it} with $\gamma_{jt} \equiv \gamma_{jt}^{K} + \gamma_{jt}^{W} < 1$.[4] We allow structural parameters to be industry specific and time varying.

Firms choose investments in physical and working capital to maximize the market equity. Physical capital evolves as $K_{it+1} = (1-\delta_{it})K_{it} + I_{it}$ in which I_{it} is the investment in physical capital, and δ_{it} is the depreciation rate. Investment in physical capital incurs quadratic adjustment costs:

$$\Phi_{it} \equiv \Phi(I_{it}, K_{it}) = \frac{a_{jt}}{2}\left(\frac{I_{it}}{K_{it}}\right)^{2} K_{it}\,, \tag{1}$$

where a_{jt} is the physical adjustment costs parameter. Working capital evolves as $W_{it+1} = W_{it} + \Delta W_{it}$, in which ΔW_{it} is the investment in working capital. In addition, working capital does not depreciate and is not accompanied with adjustment costs.

In addition to equity financing, firm i in industry j issues debt B_{it+1} with interest rate r_{it+1}^{B} at the beginning of time t, which is repaid at the beginning of $t+1$. At tax rate τ_t, firm i's net payout is given by $D_{it} \equiv (1-\tau_t)\left(\Pi_{it} - \Phi_{it}\right) - I_{it} - \Delta W_{it} + B_{it+1} - r_{it}^{Ba} B_{it} + \tau_t \delta_{it} K_{it}$, in which $r_{it}^{Ba} \equiv r_{it}^{B} - \tau_t(r_{it}^{B} - 1)$ is the after-tax interest rate. Taking the stochastic pricing kernel, M_{t+1}, as given, firm i chooses I_{it}, K_{it+1}, ΔW_{it}, W_{it+1}, and B_{it+1} to maximize its cum-dividend market equity, $V_{it} \equiv E_t[\sum_{s=0}^{\infty} M_{t+s} D_{t+s}]$. The first-order condition for physical investment implies that $E_t[M_{t+1} r_{it+1}^{K}] = 1$, in which r_{it+1}^{K} is the return on physical capital

[4]Section A in the Internet Appendix provides the proof.

investment:

$$r^K_{it+1} = \frac{(1-\tau_{t+1})\left[\gamma^K_{jt+1}\left(\frac{Y_{it+1}}{K_{it+1}}\right) + \frac{a_{jt+1}}{2}\left(\frac{I_{it+1}}{K_{it+1}}\right)^2\right] + \tau_{t+1}\delta_{it+1} + (1-\delta_{it+1})\left[1+(1-\tau_{t+1})a_{jt+1}\left(\frac{I_{it+1}}{K_{it+1}}\right)\right]}{1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)} . \quad (2)$$

Similarly, the first-order condition for working capital investment implies that $E_t[M_{t+1}r^W_{it+1}] = 1$, in which r^W_{it+1} is the return on working capital investment:

$$r^W_{it+1} = 1 + (1-\tau_{t+1})\gamma^W_{jt+1}\frac{Y_{it+1}}{W_{it+1}} . \quad (3)$$

Section A in the Internet Appendix shows that the weighted average of the two investment returns equals the weighted average cost of equity and the after-tax cost of debt:

$$w^K_{it}r^K_{it+1} + (1-w^K_{it})r^W_{it+1} = w^B_{it}r^{Ba}_{it+1} + (1-w^B_{it})r^S_{it+1} , \quad (4)$$

in which $w^B_{it} \equiv B_{it+1}/(V_{it}-D_{it}+B_{it+1})$ is the firm's market leverage, $r^S_{it+1} \equiv V_{it+1}/(V_{it}-D_{it})$ is the stock return, $w^K_{it} \equiv q^K_{it}K_{it+1}/(q_{it}K_{it+1}+W_{it+1})$ is the weight of firm's market value attributed to physical capital and $q^K_{it} \equiv 1 + a_{jt}(1-\tau_t)I_{it}/K_{it}$ is the marginal q of capital. The marginal q of working capital is one. The Tobin's q of firm i at time t is the weighted average of marginal q's of physical and working capitals, given by

$$q_{it} = \frac{P_{it}+B_{it+1}}{K_{t+1}+W_{it+1}} = \left[1+a_{jt}(1-\tau_t)\frac{I_{it}}{K_{it}}\right]\frac{K_{it+1}}{K_{it+1}+W_{it+1}} + \frac{W_{it+1}}{K_{it+1}+W_{it+1}} . \quad (5)$$

Solving for the stock return from equation (4) leads to the model-implied fundamental stock return of firm i from t to $t+1$:

$$\begin{aligned} r^F_{it+1} &\equiv f\left(X_{it}, X_{it+1}|\theta_t, \theta_{t+1}\right) \\ &= \left\{(1-\tau_{t+1})\left[\gamma_{jt+1}\left(\frac{Y_{it+1}}{K_{it+1}}\right) + \frac{a_{jt+1}}{2}\left(\frac{I_{it+1}}{K_{it+1}}\right)^2\right] + \tau_{t+1}\delta_{it+1}\right. \\ &\quad +(1-\delta_{it+1})\left[1+(1-\tau_{t+1})a_{jt+1}\left(\frac{I_{it+1}}{K_{it+1}}\right)\right] \\ &\quad \left.+\frac{W_{it+1}}{K_{it+1}}\right\} \Big/ \left\{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right) + \frac{W_{it+1}}{K_{it+1}}\right]\right\} - \frac{w^B_{it} r^{Ba}_{it+1}}{1-w^B_{it}}, \end{aligned} \quad (6)$$

where X_{it} is the set of accounting variables used in equation (6) that represent firm i's fundamentals, and $\theta_t \equiv \{(\gamma_{jt}, a_{jt}); j = 1, \ldots, 10\}$ is the set of model parameters at time t for Fama-French 10 industries. The equality between the realized stock return and the model-implied fundamental return, $r^S_{it+1} = r^F_{it+1}$, holds for any firm i and for any period from t to $t+1$ under this framework. Next, we estimate the two structural parameters, γ and a, based on this equality. Notice that γ^K and γ^W cannot be separately identified because r^F depends on their summation only.

3 Data

Our sample includes all common stocks traded on NYSE, Amex, and NASDAQ with available accounting and return data. We exclude firms with primary standard industrial classifications between 6000 and 6999 (financial firms), firms with negative book equity, and firms with nonpositive total assets, net property, plant, and equipment, or sales at the portfolio formation. These data items are needed to calculate firm-level fundamental returns. We obtain monthly stock return data from the Center for Research in Security Prices (CRSP). Firm-level accounting data are obtained from the annual and quarterly Standard and Poor's Compustat industrial files. Our data sample covers the period from January 1967 to June 2017.

3.1 Anomalies

We explore 12 anomalies in six categories: size anomaly sorted on market capitalization (Size); value anomaly sorted on book-to-market equity ratio (BM); momentum anomaly sorted on the prior 11-month returns skipping the most recent month (R11); four investment anomalies sorted on asset growth (I/A), net stock issues (NSI), investment-to-assets ratio (ΔPI/A), and accruals (Accruals); three profitability anomalies sorted on return-on-equity (ROE), return-on-assets (ROA), and gross profitability (GP/A); and two intangibles anomalies sorted on R&D expense-to-market ratio (RD/M) and advertising expense-to-market ratio (Ad/M). Following Hou, Xue and Zhang (2020), we select these anomalies based on the criteria that the average value-weighted returns of their high-minus-low deciles with NYSE breakpoints are significant at the 5% level, with the exception of the size anomaly. Size anomaly is included since it is one of the most studied anomalies in the literature. Section C in the Internet Appendix provides the definitions of these variables and the construction of the corresponding decile portfolios.

Table 1 presents the monthly average excess returns of the 10 decile portfolios sorted on each of the 12 anomaly variables. The t-statistics adjusted for heteroscedasticity and autocorrelations are reported in parentheses. "L" denotes the lowest decile, "H" the highest decile, and "H-L" the high-minus-low decile. As in Hou, Xue and Zhang (2020), decile portfolios are formed with NYSE breakpoints and value-weighted returns to control for microcaps. The sample period is from January 1967 to June 2017 for all anomaly variables except ROA, RD/M, and Ad/M, the samples for which start from July 1972, July 1976, and July 1973, respectively, due to data availability. All 12 anomalies except size have statistically and economically significant premiums in our sample period.

3.2 Measures and timing alignment

Model-implied fundamental returns are constructed in annual frequency because the needed fundamental variables such as investments are only available at annual frequency for the long sample starting from 1967. In the model, time-t stock variables are at the beginning of year t, and time-t flow variables are over the course of year t. Thus, time-t stock variables are obtained from the balance sheet of fiscal year $t-1$ and flow variables from the balance sheet of fiscal year t.

We adopt the same measures used by Gonçalves, Xue and Zhang (2020) for the variables needed to construct the fundamental returns. Specifically, output, Y_{it}, is measured as sales (Compustat annual item SALE). Physical capital, K_{it}, is net property, plant, and equipment (item PPENT). Short-term working capital, W_{it}, is current assets (item ACT). Total debt, B_{it+1}, is long-term debt (item DLTT, zero if missing) plus short-term debt (item DLC, zero if missing) from fiscal year t balance sheet. Tax rate τ is the statutory corporate income tax rate from the commerce clearing house's annual publications. The depreciation rate of physical capital, δ_{it}, is the amount of depreciation and amortization (item DP) minus the amortization of intangibles (item AM, zero if missing) divided by physical capital (item PPENT). Physical investment, I_{it}, is measured as $K_{it+1} - (1 - \delta_{it})K_{it}$. The market leverage, w^B_{it}, is the ratio of total debt to the sum of total debt and market equity. The pre-tax cost of debt, r^B_{it}, is the ratio of total interest and related expenses (item XINT) scaled by total debt, B_{it}. Following Gonçalves, Xue and Zhang (2020), we winsorize unbounded variables, including I_{it}/K_{it}, I_{it+1}/K_{it+1}, $\Delta W_{it}/W_{it}$, $\Delta W_{it+1}/W_{it+1}$, at the 2.5% - 97.5% level. For variables bounded below by zero, including Y_{it+1}/K_{it+1}, Y_{it+1}/W_{it+1}, $Y_{it+1}/(K_{it+1} + W_{it+1})$, W_{it+1}/K_{it+1}, δ_{it+1}, and r^B_{it+1}, we winsorize them at the 0% - 95% level. We do not winsorize variables bounded between zero and one, such as $K_{it+1}/(K_{it+1} + W_{it+1})$ or the market leverage, w^B_{it}. Summary statistics and correlation matrix of the aforementioned variables are reported in Table 2 and closely match those in Gonçalves, Xue and Zhang (2020).

In the model, the fundamental stock return of firm i from year t to $t + 1$, r^F_{it+1}, is constructed with both stock and flow variables at annual frequency. In the estimation, we match r^F_{it+1} with the observed annual return of firm i from the middle of fiscal year t to the middle of fiscal year $t + 1$, following Gonçalves, Xue and Zhang (2020). Specifically, if firm i's fiscal end of year t is month l, r^S_{it+1}, the counterpart of r^F_{it+1}, is the realized 12-month return between month $l - 5$ and $l + 6$.

To study anomalies, we construct fundamental portfolio returns based on fundamental firm-level returns. Even though firm-level fundamental returns change annually (in fiscal year), fundamental portfolio returns change monthly because fiscal year-endings vary across firms. However, the fundamental portfolio returns of a given month are based on annual accounting variables both prior to and after the month. To better align the timing and make a fair comparison, we follow Gonçalves, Xue and Zhang (2020) and compound the realized portfolio stock returns within a 12-month rolling window with the month in question in the middle of the window. Specifically, we multiply gross returns from month $t-5$ to month $t+6$ to match the fundamental returns constructed in month t. Applying this rolling procedure to the monthly portfolio returns (January 1967 to June 2017) yields the monthly observations of annualized portfolio returns from June 1967 to December 2016.

We validate our data construction and portfolio formation by successfully reproducing the realized and predicted returns (by the baseline model) of the book-to-market (BM), momentum (R11), asset growth (I/A), and return-on-equity (ROE) deciles in Gonçalves, Xue and Zhang (2020) using their estimated model parameters. The results are plotted in Figure A.1, which replicates Panel B of Figure 3 in Gonçalves, Xue and Zhang (2020).

4 Estimation methodology

Prior studies (Liu, Whited and Zhang, 2009; Gonçalves, Xue and Zhang, 2020, among others) use the General Method of Moments (GMM) to match the unconditional moments derived from equation (6): $E_T[r^S_{pt+1} - r^F_{pt+1}] = 0$ for testing portfolio p, where $E_T[\cdot]$ refers to the operation of taking time series average. We instead target the entire panel of firm-level stock returns using the full-information Bayesian Markov Chain Monte Carlo (MCMC) method. We consider four specifications in the estimation: parameters with industry and time variations, parameters with industry variations only, parameters with time variations only, and constant parameters. Next, we explain our methodology in details in terms of the specification that allows both industry and time variations in parameter values.

4.1 Bayesian MCMC

Denote the technology parameter in the production function for industry j at time t as γ_{jt} and the physical adjustment costs parameter as a_{jt}. The time series of parameter values are referred to as "latent variables" in Bayesian MCMC estimation and are assumed to evolve as random walk processes:

$$\begin{bmatrix} \gamma_{jt+1} \\ a_{jt+1} \end{bmatrix} = \begin{bmatrix} \gamma_{jt} \\ a_{jt} \end{bmatrix} + \begin{bmatrix} \sigma_\gamma & \\ & \sigma_a \end{bmatrix} \begin{bmatrix} e^\gamma_{jt+1} \\ e^a_{jt+1} \end{bmatrix}, \tag{7}$$

where e^{γ}_{jt+1} and e^{a}_{jt+1} follow standard normal distributions independently, and σ_γ and σ_a are the conditional standard deviations of latent variables γ_{jt+1} and a_{jt+1} conditioning on previous time t. Imposing a random walk process on the deep parameters not only encourages persistence, but also enables us to borrow information across time in estimation, leading to more efficient estimates.[5]

Realized stock return of firm i (in industry j) at time $t+1$ is modeled as the corresponding fundamental return plus an estimation error:

$$r^{S}_{it+1} = r^{F}_{it+1} + \varpi_{it}^{-1/2}\sigma_r e^{r}_{it+1}, \tag{8}$$

where e^{r}_{it+1} follows the standard normal distribution, σ_r is a parameter to be estimated, and the weight $\varpi_{it}^{-1/2}$ in the estimation error is specified as:

$$\varpi_{it} \equiv \frac{V_{it}}{\sum_{i=1}^{N_{jt}} V_{it}}\,, \tag{9}$$

in which N_{jt} is the number of firms at time t in industry j to which firm i belongs. By this specification, we introduce heteroskedasticity into the estimation errors of realized stock returns. The variance of a firm's estimation errors decreases with its market equity V_{it} in order to accommodate the fact that stock returns of large firms are less noisy and more reflective of their fundamentals than the returns of small firms.[6] More importantly, such specification makes the estimated model economically relevant in the sense that it captures the regularity of the majority of the economy. The same rationale motivates the use of NYSE breakpoints in constructing portfolios and regressions with weighted least squares in asset pricing studies (e.g., Hou, Xue and Zhang, 2015).[7]

For the MCMC method, prior distributions of the model parameters need to be specified. We use inverse gamma distributions for the priors of variances: $\sigma^2_\gamma \sim IG(\kappa^\gamma_1, \kappa^\gamma_2)$, $\sigma^2_a \sim$

[5] We also estimate an autoregressive process with order one. The estimated persistence parameters are very close to one for both processes of γ_{jt} and a_{jt}. Thus, we use random walk processes in our baseline model for simplicity.

[6] Large firms have more analysts following than small firms, thus their value is under much closer scrutiny (Bhushan, 1989). Moreover, stocks of large firms are generally more liquid and their market values are less likely to be manipulated or affected by a small group of investors (Amihud, 2002).

[7] Effort (not reported here) has been made to investigate other kinds of functional forms relating the variability in estimation errors of a firm's stock returns to its market equity. The relationship specified in equation (9) best fits the data in terms of mean absolute error (*m.a.e.*) of firm-level stock returns.

$IG(\kappa_1^a, \kappa_2^a)$, and $\sigma_r^2 \sim IG(\kappa_1^r, \kappa_2^r)$, where κ_1 and κ_2 are hyper-parameters of the inverse gamma distribution (shape-scale parameterizations). The values of κ_1^r, κ_1^γ, and κ_1^a are specified to be 0.01, 1, and 1, respectively; the values of κ_2^r, κ_2^γ, and κ_2^a are chosen to be 0.02, 5, and 5, respectively. The values of κ_1's are chosen relatively small so that the information from data is more likely to dominate (see Section B in the Internet Appendix). The values of κ_2's are set relatively large so that the variances of the priors are large and thus less informative. Although the choices of these values are seemingly arbitrary, as MCMC runs and information from the data gets entered into the posterior draws, these hyper-parameters weigh less and less. The information from data dominates the posterior draws when MCMC converges.

Finally, the time series of latent variables $\boldsymbol{\theta} \equiv \{\theta_t; t = 1, \cdots, T\}$, where $\theta_t = \{(\gamma_{jt}, a_{jt}); j = 1, \ldots, 10\}$ and variance parameters $\boldsymbol{\sigma} \equiv \{\sigma_\gamma^2, \sigma_a^2, \sigma_r^2\}$ are drawn in an iterative manner from each complete conditional posterior distribution, resulting in posterior samples from the joint posterior distribution. Based on the model specifications in equation (7) and (8), the joint posterior distribution of $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$ can be written (in a proportional form) as:

$$
\begin{aligned}
\mathcal{P}(\boldsymbol{\theta}, \boldsymbol{\sigma} | \boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba}) \propto & \prod_{t=0}^{T-1} \prod_{i=1}^{N_{t+1}} \mathcal{N}\Big(r_{it+1}^S; r_{it+1}^F, \sigma_r^2\Big) \\
& \times \prod_{t=0}^{T-1} \prod_{j=1}^{N_d} \mathcal{N}\Big(\gamma_{jt+1}; \gamma_{jt}, \sigma_\gamma^2\Big) \\
& \times \prod_{t=0}^{T-1} \prod_{j=1}^{N_d} \mathcal{N}\Big(a_{jt+1}; a_{jt}, \sigma_a^2\Big) \\
& \times \mathcal{IG}\Big(\sigma_r^2; \kappa_1^r, \kappa_2^r\Big) \times \mathcal{IG}\Big(\sigma_\gamma^2; \kappa_1^\gamma, \kappa_2^\gamma\Big) \times \mathcal{IG}\Big(\sigma_a^2; \kappa_1^a, \kappa_2^a\Big),
\end{aligned}
\tag{10}
$$

where N_{t+1} is the number of firms at time $t+1$, N_d is the number of industries, and r_{it+1}^F is defined in (6). In equation (10), $\boldsymbol{X} \equiv \{X_{it}; i = 1, \cdots, N_t, t = 1, \cdots, T\}$ is the panel of fundamental observables, $\boldsymbol{r}^S$ and $\boldsymbol{r}^{Ba}$ are the panels of realized stock and bond returns, and $\mathcal{N}(\cdot\,; \mu, \sigma^2)$ and $\mathcal{IG}(\cdot\,; \kappa_1^{\cdot}, \kappa_2^{\cdot})$ refer to the probability density functions of normal distribution with mean μ and variance σ^2 and inverse gamma distribution with shape-scale parameters κ_1 and κ_2, respectively. We run 20,000 MCMC iterations and use the last 5,000 iterations to obtain posterior draws. We confirm the convergence of the posterior distributions. Section B in the Internet Appendix details the sampling algorithm and posterior derivations.

4.2 Simulation studies

In this section, we use simulation studies to examine whether Bayesian MCMC can discover the true parameter values under our model framework, which is highly nonlinear. W combine the accounting information of a subsample of firms and a pre-determined set of parameter values to generate a simulated panel of firm-level stock returns based on equations (6) and (8). For simplicity, we require the subsample used in the simulation to be a balanced panel of 1,052 firms for 15 years, which covers seven out of the ten Fama-French industries. We make sure that the simulated returns have a similar distribution as that of the realized returns. Based on the simulated data, we estimate the posterior distributions of the model parameters using Bayesian MCMC to check whether these estimates can discover the true parameter values, i.e., the pre-determined parameter values used to generate the simulated data. The details of the simulation studies are described in Section D of the Internet Appendix. It is worth noting that using firms' true accounting information to construct the simulated data increases the difficulty of the estimation due to the non-normal distributions of these accounting variables as shown in Gonçalves, Xue and Zhang (2020). Simulation studies are done for all four specifications: parameters with industry and time variations, with industry variations only, with time variations only, and with constant values.

Figure 1 plots the true values (in red solid lines), the nonlinear least square (NLS) estimates used as initial guesses (in green lines with triangle markers), and the Bayesian posterior means (in blue dashed lines) and the associated 95% credible intervals (in shaded areas) of the model parameters $\boldsymbol{\theta}$ estimated from the simulated data under the specification with industry specific and time varying parameters. Credible interval is frequently used in Bayesian framework. It refers to the interval wherein a random variable (here a parameter) falls with the specified probability. It is an interval in the domain of a posterior distribution of a parameter. Because we assume parameters to be random variables in Bayesian framework, we can calculate the probability that a parameter locates in a given interval based on its posterior distribution. Notationally, let I_p be the posterior credible interval of $\boldsymbol{\theta}$ that satisfies $P(\boldsymbol{\theta} \in I_p | \boldsymbol{X}, \boldsymbol{r^S}, \boldsymbol{r^{Ba}}) = p$, where p is the probability.[8]

Figure 1 shows that the true values of the model parameters are almost always confined in

[8]For illustration purpose, Figure 1 only includes the results of three industries: Consumer Nondurables, Manufacturing, and Business Equipment. The results of the other industries used in the simulation studies are plotted in Figure A.2 in the Internet Appendix.

the narrow credible intervals of the Bayesian MCMC posterior distributions, even when the initial guesses are far away from the true values. The posterior means imply small relative mean absolute errors (m.a.e.) of 3.59% and 3.37% on average across industries for γ and for a, respectively. Similar results are found for the other three specifications.[9]

Our MCMC estimation approach is fundamentally different from the estimation method in Liu, Whited and Zhang (2009) and Gonçalves, Xue and Zhang (2020), among others, and it offers several advantages. First, our estimates of parameter values are independent of any specific testing portfolios. We utilize the entire *distribution* of firm-level stock returns to estimate model parameters, while GMM matches the time-series averages of returns on testing portfolios. This feature is critical for addressing the critique of Campbell (2017) that the parameter values of the model are chosen to fit a specific set of anomalies and different values are required for different anomalies.

Second, our MCMC algorithm generates random draws of model parameters from their joint posterior distribution given the observations on firms' stock and bond returns and fundamentals, while GMM outputs point estimates of model parameters, which are *deterministic* given the same set of observations. One advantage of our Bayesian approach is that *probabilistic* inferences for the estimated parameters and fundamental stock returns can be easily made using posterior draws from the MCMC iterations.

Lastly, the posterior of an industry-time specific parameter, θ_{jt}, utilizes the information of the entire data sample, rather than using the information of industry j at time t only as in the frequentist method NLS. For this reason, Bayesian MCMC can accurately identify the true values of model parameters even for the specification where these parameters vary across industries and over time. In contrast, NLS often fails to discover the true parameter values that are time varying.[10] This feature can be extremely important when the modeled economy is highly heterogeneous and changing over time.

In sum, the simulation studies suggest that Bayesian MCMC performs very well for our highly nonlinear model and is able to discover the true parameter values under all four

[9]The results for the specification with time variation in parameter values only are plotted in Figure A.3. The results for the specification with industry variation in parameter values only and for the specification with constant parameter values are reported in Table A.1 in the Internet Appendix.

[10]Detailed discussion on the reasons why the posterior of an industry-time specific parameter, θ_{jt}, utilizes the information of the entire data sample, the differences between Bayesian and NLS, and the comparison of their estimation accuracy under our model framework are provided in Section E of the Internet Appendix.

specifications. This method shows great potential in estimating nonlinear models, especially when the structural parameters are time varying.

5 Estimation results

In our baseline estimation, we allow model parameters to be industry specific and time varying. For comparison, we also estimate three alternative specifications under which the time, industry, or both variations are shut down, respectively. In this section, we discuss the main findings of these estimations.

5.1 Parameter estimates

In this section, we apply our MCMC method to the real data. In the baseline estimation, we allow the technology parameter in the production function, γ_{jt}, and the physical adjustment costs parameter, a_{jt}, to be industry specific and time varying.

Within the model framework, γ_{jt} reflects industry j's profit margin as the model implies $\Pi_{it} = \gamma_{jt} Y_{it}$ for any firm i in industry j at time t, where Π_{it} and Y_{it} are the profits and sales, respectively. Therefore, variations in γ_{jt} can be driven by both technology changes and changes in market demand, the latter of which can be caused by fluctuations in consumer taste, economic conditions, market competitiveness, etc.

Equation (5) implies that Tobin's q of firm i in industry j at time t follows $q_{it} = 1 + a_{jt}(1 - \tau_t) I_{it}/K_{it} \times K_{it+1}/(K_{it+1} + W_{it+1})$. Therefore, the magnitude of a_{jt} reflects both the marginal costs and marginal benefits of investing one dollar in physical capital and has a positive relation with Tobin's q. Consequently, variations in a_{jt} can be driven by changes in production technology, price of capital goods, which is cyclical (Eisfeldt and Rampini, 2006), and opportunity costs in terms of lost output, which vary with procyclical capacity utilization. Lastly, entry and exit in an industry can also lead to changes in the estimated parameter values of this industry at a given fiscal year t. It thus makes economic sense to allow γ_{jt} and a_{jt} to be industry specific and time varying.

Our estimation generates the posterior distributions of the marginal product parameter γ and the physical investment adjustment costs parameter a for each of the Fama-French 10 industries and for each year between 1967 and 2016. For illustration, we plot the time series of the posterior means and 95% credible intervals of γ and a, averaged across Fama-French 10 industries in Figure 2 for the baseline estimation.[11] The first thing to notice is that the credible intervals around the posterior means in Figure 2 are extremely narrow. The simple model structure (only two structural parameters) and the large amount of information (136,598 firm-level observations) enable the Bayesian MCMC estimation to identify these parameter values with high precision, provided that the model is correctly specified. In addition, there are no noticeable time trends in both parameter values, although we do see fluctuations.

Table 3 presents the time-series averages of the posterior means and 95% credible intervals (CI) of γ and a for each industry. The results show that parameter estimates differ greatly across industries. The magnitude of these estimates is largely consistent with our economic intuitions. For example, γ is estimated to be 0.08 on average with a CI of $[0.08, 0.09]$ for the wholesale & retail sector, compared to 0.28 for the telecom sector, consistent with the fact that capital is less important for the wholesale & retail sector than for the telecom sector. As explained in Erickson and Whited (2000), it can be misleading to interpret the value of a in terms of adjustment costs or speeds. We thus gauge the economic magnitude of this parameter in terms of Tobin's q. For each industry j, we compute the value-weighted cross-sectional average of Tobin's q for year t and then take the time series average,

$$\overline{q}_j = \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N_{jt}} \varpi_{it-1} q_{it} \,,$$

where N_{jt} is the number of firms in industry j at time t, the weight ϖ_{it-1} is defined in equation (9), and q_{it} is the Tobin's q of firm i at time t, given by equation (5). Table 3 shows that the business equipment sector has the highest Tobin's q of 1.81 while the utilities sector has the lowest Tobin's q of 1.16. These estimates are consistent with our intuition that the business equipment sector, which includes the high-tech firms, has the highest growth potential while the regulated utilities sector has the lowest potential for growth.[12]

Lastly, the estimates in alternative specifications under which the time, industry, or both

[11] Figure A.4 in the Internet Appendix plots the means (in solid lines) and 95% credible intervals (in dotted lines) of the posterior distributions of γ and a, respectively, for each industry and for each year.

[12] Section F in the Internet Appendix provides further discussion on whether these estimated parameter values in the baseline make economic sense.

variations are shut down are close to the baseline estimates. The average (over time or across industry) production technology parameter γ is 0.18 or 0.15 when time or industry variation is shut down, while the average adjustment costs parameter a is 0.19 or 0.28.[13] When both industry and time variations are shut down, the posterior means of γ and a are 0.15 (std = 0.0003) and 0.14 (std = 0.0027), respectively. These estimates are different from the estimates in Gonçalves, Xue and Zhang (2020), where γ is 0.18 (std = 0.019) and a is 2.84 (std = 0.47). Noticeably, their estimate of a is 20 times larger than our estimate. This result implies that parameter values chosen to match a set of portfolio returns can still be very different from the values chosen to match firm-level returns even when the portfolio returns are aggregated from firm-level fundamental returns as in Gonçalves, Xue and Zhang (2020). Moreover, our estimates have much smaller standard deviations than those in Gonçalves, Xue and Zhang (2020). The reason is that Bayesian MCMC utilizes the information of the entire time series of firm-level stock returns, while the GMM estimation in Gonçalves, Xue and Zhang (2020) utilizes the average returns on the 40 testing portfolios only.

5.2 Overall fit of the estimation

Different from GMM, which gives point estimates of the parameters, Bayesian MCMC offers a probabilistic view of the parameters and thus the fundamental returns. Given firm-level accounting variables, each posterior draw of $\boldsymbol{\theta} = \{\theta_t; t = 1, \cdots, T\}$ leads to a panel of firm-level fundamental returns. Therefore, our estimation generates posterior distributions of the fundamental returns and any statistical moments of these returns. Table 4 presents the posterior means and 95% credible intervals of the mean, standard deviation, skewness, kurtosis, and mean absolute error (m.a.e.) of the fundamental firm-level stock returns and the time-series average of the cross-sectional correlations between fundamental and realized stock returns for the baseline and the three alternative estimation specifications. And the same moments of the realized stock returns are presented for comparison. The definition of m.a.e. is given by m.a.e. $\equiv \frac{1}{T}\sum_{t=0}^{T-1}\frac{1}{N_{t+1}}\sum_{i=1}^{N_{t+1}}|r_{it}^S - r_{it}^F|$, where N_{t+1} is the number of firms in period $t+1$, and r^S and r^F are the realized and fundamental stock returns, respectively.

Several observations emerge from Table 4. First, by construction, the baseline with industry-specific and time varying parameter values has the smallest m.a.e. (40.10%), fol-

[13] The estimates under the three alternative specifications are available upon request.

lowed in turn by the specifications with only time variation (40.85%), with only industry variation (41.85%), and with no variation (42.45%) in parameter values. Although the magnitude of these m.a.e.'s might seem large, it is only two-thirds of the volatility of firm-level returns. The majority of the m.a.e. comes from the difference in volatility between the realized and fundamental returns discussed below. Second, the mean, skewness and kurtosis of the fundamental returns match well with those of the realized returns across all four specifications. The mean of fundamental stock returns ranges from 14.97% to 15.65% across specifications, compared to 14.45% in the data. The skewness ranges from 1.66 to 2.12, compared to 2.15 in the data, and kurtosis ranges from 10.66 to 13.33 compared to 11.05 in the data. Third, standard deviations of the fundamental returns are much smaller than that of the realized one. The baseline generates the highest standard deviation, 34.17% compared to 60.78% in the data, while the specification with no time variation in parameter values generates the lowest standard deviation, 18.49%. Lastly, the time-series averages of the cross-sectional correlations between realized and fundamental firm-level stock returns are 0.20, 0.12, 0.12, and 0.09 for the baseline specification, the ones with industry variation only, time variation only, and without variations in parameter values, respectively. For comparison, we report the same statistics of the NLS estimation for the four specifications in Table A.2 in the Internet Appendix. Bayesian MCMC results in smaller m.a.e. in every specification, echoing the superior performance of Bayesian approach compared to NLS documented in Section 4.2.

Panel A of Figure 3 plots the histograms of realized (in blue) and fundamental (in orange) firm-level returns based on the posterior means of the parameter estimates. Consistent with what Table 4 shows, realized returns have a much wider distribution than fundamental returns at both left and right tails and thus have a larger standard deviation. Both distributions have longer right tails, resulting in positive skewness and kurtosis higher than 3.

Panel B of Figure 3 plots the time series of the value-weighted realized (in blue) and fundamental (in orange) market returns, the latter of which is constructed with the posterior means of the parameter estimates. These two time series are highly correlated with a correlation coefficient of 0.77. In particular, the fundamental market returns successfully capture the significant fluctuations of the stock markets during the sample period, such as the Internet bubble around 2000 and the financial crisis around 2008.

Overall, the baseline estimation does a good job of capturing the entire distribution of the realized firm-level stock returns in terms of the mean, skewness, and kurtosis, and in terms of the correlation between realized and fundamental returns. However, the standard deviation of the fundamental returns is only half of the magnitude in the data. The reason is that the realized firm-level returns contain large amount of noises unrelated to fundamentals and are much more volatile than the accounting variables used to construct the fundamental returns.

5.3 Posterior distributions of the fundamental factor premiums

To examine the capability of the q-theory in explaining stock market anomalies, we construct 12 anomalies, each with 10 decile portfolios and one high-minus-low decile portfolio, as explained in Section 3. Based on the posterior distributions of the fundamental firm-level returns, we construct the poster distributions of these factor premiums.

Figure 4 plots the posterior distributions of the 12 fundamental factor premiums under the baseline estimation and labels the 2.5, 50, and 97.5 percentiles of each distribution. For example, the posterior distribution in Panel "BM" indicates that, given the observed accounting variables and provided that the model is correctly specified, the fundamental value premium per annum falls in the range between 0.31% and 0.60% with 95% probability and the posterior median is 0.46% per annum. The red line in each panel presents the density function of a normal distribution with mean and standard deviation taken from the corresponding posterior distribution. Notice that the posterior distributions are very much close to normal distribution, indicating that our Bayesian MCMC algorithm converges well. In cases where the Bayesian MCMC algorithm does not converge, the posterior distributions typically would have multiple peaks or/and long and fat tails. Moreover, the credible intervals of the fundamental factor premiums are extremely narrow. As we argue before, these tight posterior distributions indicate that the simplicity of the model and the richness of firm-level information confine the parameter estimates and the fundamental factor premiums to a small set of possible values.

Table 5 presents the realized and fundamental (value-weighted) factor premiums under the baseline estimation, their t-values, and the t-values of the difference between realized and fundamental factor premiums, denoted as alpha, $\alpha \equiv r^S - r^F$.[14] For each statistic of the fundamental factor premiums, we report its posterior mean and the 95% credible interval. Note that the credible intervals and t-values of the factor premiums measure different types of variability. For example, t-value of the fundamental value premium, $t\left(r^F\right)$, measures the time-series variability of the fundamental returns on the high-minus-low book-to-market portfolio generated by a given posterior draw of parameter values. In contrast, the credible interval of the fundamental value premium gives a range into which 95% of the value premiums in the posterior draws fall.

Several observations emerge from Figure 4 and Table 5. First, the model generates positive momentum, profitability (ROE, ROA, and GP/A), and intangibles (R&D and advertising) premiums and generates negative size and investment (I/A, NSI, and ΔPI/A) premiums, all of which are statistically significant at the 5% level. The t-values of these factor premiums, including the posterior means and the 95% credible intervals, are all larger than 1.96 in absolute values. For example, in terms of the posterior means, the fundamental momentum premium is 11.82% per annum with a t-stat of 12.51, while the momentum premium is 13.75% ($t = 4.15$) in the data. In general, the t-values of the fundamental factor premiums are larger than their counterparts in the data due to the low variability of the fundamental returns. Among the aforementioned 10 factor premiums, the majority of their alphas are insignificantly different from zero except the alphas of the I/A, NSI, and GP/A premiums, which are significant at the 5% level.

Second, the model generates positive but statistically insignificant value premium. The posterior means of the fundamental value premium and its t-value are 0.46% per annum and 0.26, both of which have very narrow credible intervals (CIs being $[0.31\%, 0.60\%]$ and $[0.18, 0.35]$). It turns out that the model is able to explain the value premium for the pre-July 1991 period, but not the period after. We suspect that the failure to explain the value premium in the post-July 1991 period is due to the growing importance of intangibles in firms' market value. Detailed discussion is provided in Section 6.2.

Lastly, the model predicts significantly positive accruals premium, which is however negative in the data. The fundamental accruals premium is 4.74% ($t = 4.45$) in contrast to -5.58% ($t = -3.14$) in the data. This result is not surprising in that the concept of accruals

[14] Since r^S is deterministic, the posterior distributions of the alphas have the same shape as the distributions of the corresponding fundamental factor premiums r^F.

is absent in our model, where cash and accruals basis accountings are treated the same. High accruals mean high profitability in the model, which however is not necessarily true in the data. Detailed analysis on why the model fails to generate negative fundamental accruals premium is provided in Section 6.

Figure 5 presents a comparison between the realized and fundamental factor premiums graphically. The height of a bar in Panel A represents the magnitude of the corresponding premium. The 95% credible intervals of the fundamental factor premiums are also marked as the error bars. Panel B plots the average fundamental returns against the realized returns across the 120 decile portfolios, in which the scatter points are largely aligned with the 45-degree line. Overall, the above results show that the estimated model is able to generate sizable factor premiums that are consistent with what we observe in the data except the value and accruals premiums.

5.4 Importance of industry and time variations in parameter values

In our baseline estimation, we allow model parameters to be industry specific and time varying. In this section, we examine the importance of these variations in terms of explaining firm-level returns and anomaly premiums by comparing the performance of the baseline estimations with three alternative specifications under which the time, industry, or both variations are shut down, respectively.

To compare the performance of an alternative specification with the baseline, we construct the following statistic:

$$d^a = \frac{1}{T}\sum_{t=0}^{T-1}\frac{1}{N_{t+1}}\sum_{i=1}^{N_{t+1}}\left(\left|r^S_{it+1} - r^{F(a)}_{it+1}\right| - \left|r^S_{it+1} - r^{F(b)}_{it+1}\right|\right), \tag{11}$$

where $r^{F(a)}$ and $r^{F(b)}$ are the fundamental returns under the alternative specification, a, and the baseline, respectively. Note that although our statistic is similar in form to Diebold-Mariano (1995), they differ in nature by statistical properties. However, taking advantage of MCMC, we can still make a valid inference from this statistic. Intuitively, when the estimation errors from the alternative specification are larger in magnitude, we expect d^a to be above 0 with statistical significance. Following equation (11), we record $d^{a(m)}$ for the

m-th posterior draws of parameter values. These $5,000$ posterior draws jointly provide us with the empirical distribution of the statistic d^a. A significantly positive d^a indicates that the baseline performs significantly better than the alternative specification a in explaining firm-level stock returns.

Figure 6 plots the distribution of d^a for each of the three alternative settings all in one panel in Panel (a) and separately in Panels (b) to (d). The 2.5, 50, and 97.5 percentiles of the posterior distributions are also marked in the last three panels. We can see that the posterior distributions of these three settings all lie in the positive region. Neither of the three 95% credible intervals ([0.57, 0.67] for industry-invariant, [1.22, 1.32] for time-invariant, and [1.80, 1.90] for constant parameter settings) includes zero, indicating that the baseline performs significantly better in explaining the firm-level stock returns than these alternative settings. Moreover, the location of these distributions in Panel (a) indicates that the performance of the model with industry-invariant parameters is closest to the baseline, followed by the settings with time-invariant parameters and with constant parameters. That is, time variation is more important than industry variation in explaining firm-level stock returns.

Table 6 compares the performance of these four specifications in terms of generating factor premiums. First, for the model to match anomaly premiums, time and industry variations are not necessarily helpful. In fact, shutting down time variation decreases the magnitude of alphas for the size, GP/A, and Ad/M anomalies, while shutting down industry variation decreases the magnitude of alphas for the ROE, ROA, and NSI anomalies. Second, industry variation is critical for generating intangibles premiums while time variation is not. Without the industry variation, the RD/M and Ad/M alphas become statistically significant at the 5% level. Third, the success in generating the momentum premium crucially depends on both industry and time variations in parameter estimates. Without either of them, the momentum alpha becomes larger and statistically significant. Lastly, none of these specifications can generate negative accruals premium and significantly positive value premium. The alphas of these two anomaly premiums become even larger in absolute value once industry and/or time variations in parameter estimates are shut down.

In summary, both industry and time variations in parameter values significantly improve the ability of the model to match firm-level stock returns. However, these variations are not critical to explaining the majority of these anomaly premiums and in some cases they even hurt the performance. The specification that best matches firm-level stock returns does not necessarily explain anomalies better because these anomaly premiums are not the targets of our estimation.

6 Inspecting the economic mechanism

In this section, we conduct comparative statics to quantify the relation of next period fundamental returns r^F_{it+1} with four firm characteristics: current invest-to-physical capital ratio I_{it}/K_{it}, next period invest-to-physical capital ratio I_{it+1}/K_{it+1}, next period sales-to-physical capital ratio Y_{it+1}/K_{it+1}, and next period working-to-physical capital ratio W_{it+1}/K_{it+1}. Based on the comparative statics, we explore the economic driving forces behind these fundamental anomaly premiums and try to understand why the model fails to generate the value and accruals anomalies.

6.1 Comparative statics

The fundamental stock return in equation (6) gives important implications on the relation between firm characteristics and stock returns in equilibrium. Given equation (6), it is straightforward to derive the following partial derivatives and their signs:

$$\frac{\partial r^F_{it+1}}{\partial (I_{it}/K_{it})} = \frac{-(1-\tau_t) r^{Fwacc}_{it+1}\, a_{jt}}{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)+\frac{W_{it+1}}{K_{it+1}}\right]} < 0 \tag{12}$$

$$\frac{\partial r^F_{it+1}}{\partial (I_{it+1}/K_{it+1})} = \frac{(1-\tau_{t+1})\left(1+\frac{I_{it+1}}{K_{it+1}}-\delta_{it+1}\right) a_{jt+1}}{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)+\frac{W_{it+1}}{K_{it+1}}\right]} > 0 \tag{13}$$

$$\frac{\partial r^F_{it+1}}{\partial (Y_{it+1}/K_{it+1})} = \frac{(1-\tau_{t+1})\gamma_{jt+1}}{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)+\frac{W_{it+1}}{K_{it+1}}\right]} > 0 \tag{14}$$

$$\frac{\partial r^F_{it+1}}{\partial (W_{it+1}/K_{it+1})} = \frac{1-r^{Fwacc}_{it+1}}{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)+\frac{W_{it+1}}{K_{it+1}}\right]}, \tag{15}$$

where r^{Fwacc}_{it+1} is firm i's fundamental weighted average cost of capital, defined as $r^{Fwacc}_{it+1} \equiv (1-w^B_{it})r^F_{it+1} + w^B_{it} r^{Ba}_{it+1}$.

These partial derivatives imply that, all else equal, lower current investment-to-capital ratio, I_{it}/K_{it}, higher next period investment-to-capital ratio, I_{it+1}/K_{it+1}, and higher profitability Y_{it+1}/K_{it+1} are associated with higher next period fundamental stock return r^F_{it+1}, respectively. The relation between fundamental return and the next period working-to-physical capital ratio W_{it+1}/K_{it+1} is more complicated. On one hand, higher W_{it+1}/K_{it+1} indicates higher return on working capital investment implied by equation (3). On the other hand, higher W_{it+1}/K_{it+1} means that a larger fraction of return on assets comes from return on working capital investment, which could lead to lower stock return if return on physical capital investment is higher than the return on working capital investment. Equation (15) implies that the relation between r^F_{it+1} and W_{it+1}/K_{it+1} is mostly negative since a firm's weighted average cost of capital is in general larger than one. Moreover, r^F_{it+1} is more sensitive to variations in I_{it}/K_{it} and I_{it+1}/K_{it+1} when the adjustment costs parameter a is larger and r^F_{it+1} is more sensitive to variations in Y_{it+1}/K_{it+1} when the marginal productivity parameter γ is larger on average.[15]

To conduct comparative statics on I_{it}/K_{it}, we set I_{it}/K_{it} to be its cross-sectional median at period t across all firms and use the parameter estimates from the baseline estimation to reconstruct the fundamental returns. We then recalculate the fundamental factor premiums for the 12 anomalies and the corresponding alphas. If the resulting alphas are large relative to those from the baseline estimation, we can infer that the I_{it}/K_{it} spread is quantitatively important to explain the average return spreads. The comparative statics with respective to I_{it+1}/K_{it+1}, Y_{it+1}/K_{it+1}, and W_{it+1}/K_{it+1} are designed analogously and the results are reported in Table 7. We summarize the findings below.

First, for all the anomalies except the $\Delta PI/A$, Y_{it+1}/K_{it+1} is the most important driver of these anomaly spreads. For example, our model with the baseline estimation explains the momentum premium very well, with an alpha of 1.87% ($t = 0.57$). The alpha increases to 14.88% ($t = 3.69$) when variations in Y_{it+1}/K_{it+1} across firms are shut down, compared to alphas of -0.63% ($t = -0.19$), 8.29% ($t = 2.24$), and 1.19% ($t = 0.32$) when variations in I_{it}/K_{it}, I_{it+1}/K_{it+1}, and W_{it+1}/K_{it+1} are shut down, respectively. Our finding is in sharp contrast with those of Gonçalves, Xue and Zhang (2020), who find that I_{it}/K_{it} and I_{it+1}/K_{it+1} are the most important drivers. As discussed in Section 5.1, our estimates of the adjustment costs parameter a are much smaller than the value estimated by Gonçalves,

[15] Section 6 in the Internet Appendix provides the proof.

Xue and Zhang (2020), which weakens the impacts of investment-to-physical capital ratio on fundamental returns.

Second, the comparative statics are largely consistent with the aforementioned partial derivative analysis. For example, Table 7 shows that fixing I_{it}/K_{it} across firms leads to higher momentum premium, which is consistent with the fact that winners have higher I_{it}/K_{it} compared to losers and, as the partial derivative analysis shows, higher I_{it}/K_{it} leads to lower r^F_{it+1}.[16] Similarly, fixing I_{it+1}/K_{it+1}, Y_{it+1}/K_{it+1}, and W_{it+1}/K_{it+1} across firms leads to lower, lower, and higher momentum premiums, respectively, given that winners have higher values in these three characteristics compared to losers. Next, we try to understand why the model fails to generate the value and accruals anomalies based on the above comparative statics analysis.

6.2 The value premium

Prior studies with similar models (Liu, Whited and Zhang, 2009; Gonçalves, Xue and Zhang, 2020) find that the differences in physical investment-to-capital ratio between value and growth firms contribute the most to the value premium. In our case, the sales-to-physical capital ratio is the most important driver of the return spread between value and growth firms, as shown in Table 7. Table A.4 in the Internet Appendix shows that value firms on average have lower I_{it}/K_{it}, I_{it+1}/K_{it+1}, Y_{it+1}/K_{it+1}, and W_{it+1}/K_{it+1}, which in turn lead to higher, lower, lower, and higher fundamental returns according to equations (12) to (15), respectively. Moreover, the difference in I_{it}/K_{it} between value and growth firms is the largest among the differences in these four characteristics. Because our estimated adjustment costs parameter, a, is much smaller than those in prior studies[17], the effect of I_{it}/K_{it} is not large enough to dominate the countervailing effect of Y_{it+1}/K_{it+1}, resulting in an insignificant fundamental value premium.

Data shows that the poor performance of the model mainly comes from the second half of the sample. The model generates sizable value premium 2.94% (t=1.09), and more importantly, insignificant alpha, $\alpha = 3.52\%$ (t=1.66), for the June 1967 to June 1991 period.

[16]Table A.4 in the Internet Appendix reports the average I_{it}/K_{it}, I_{it+1}/K_{it+1}, Y_{it+1}/K_{it+1}, and W_{it+1}/K_{it+1} for firms in each decile portfolio.

[17]The values of a range from 0.25 to 1.78 on average across the Fama-French 10 industries in our baseline estimation, compared to 2.84 in Gonçalves, Xue and Zhang (2020), 22.3 in Liu, Whited and Zhang (2009) when the GMM moment is the average value premium.

However, for the July 1991 to December 2016 period, the model generates negative value premium, −1.87% (t=−0.93), and large and significant alpha, 8.87% (t=3.06). Several recent papers (for example, Eisfeldt, Kim and Papanikolaou, 2020; Belo et al., 2021) show that intangible capitals become increasingly important for cross-sectional return and valuation differences. However, intangible capitals are not modeled in our framework. It is reasonable that our model fails to generate the value premium if intangible capitals become more and more critical for the return spread between value and growth firms. A framework with explicit modeling of intangibles might be the key to explain the value premium, especially for the sample period after June 1991.

6.3 The accruals anomaly

Table 7 shows that W_{it+1}/K_{it+1} and Y_{it+1}/K_{it+1} are the most important drivers of the return spreads across firms sorted on accruals. According to Table A.4 in the Internet Appendix, high-accruals firms have higher W_{it+1}/K_{it+1} and Y_{it+1}/K_{it+1} than low-accruals firms do. Our partial derivative analysis shows that higher W_{it+1}/K_{it+1} leads to lower, while higher Y_{it+1}/K_{it+1} leads to higher, fundamental return. Comparative statics in Table 7 show that the impact of Y_{it+1}/K_{it+1} dominates that of W_{it+1}/K_{it+1}, resulting in a counterfactual positive fundamental accruals premium.

The negative effect of W_{it+1}/K_{it+1} on the accruals premium echoes the argument in Wu, Zhang and Zhang (2010) and Zhang (2013) that the accruals anomaly is a manifestation of the investment anomaly in the sense that accruals capture fundamental investment in working capital and high investment is associated with low return. In the absence of the spread in Y_{it+1}/K_{it+1} between high- and low-accruals firms, our model would have generated a negative accruals anomaly, consistent with the data.

However, differences in Y_{it+1}/K_{it+1} between high- and low-accruals firms are likely overestimated. Because our model does not distinguish cash basis from accruals basis accountings, earnings of these firms are regarded as having the same quality. In contrast, data shows that subsequent write-offs of account receivables happen more often to high-accruals firms (Dechow and Dichev, 2002, among others). Therefore, the quality-adjusted difference in Y_{it+1}/K_{it+1} between high- and low-accruals firms will be smaller and imply a smaller or even negative fundamental accruals premium. A model that incorporates earnings quality is an interesting topic for future research.

7 Discussions

In this section, we explore whether the model can explain the dynamics of these 12 anomalies, such as the correlation between the realized and fundamental factor portfolios, the persistence of their returns, and the dependences of these factor premiums on market states. In addition, we discuss whether the in-sample feature of the baseline estimation is critical for the model's ability to explain anomalies. Finally, we examine the model's out-of-sample predictive power on stock returns at the cross section. Fundamental stock returns in this section are computed based on the posterior means of parameter values under the baseline specification. Given that the posterior distribution is extremely narrow, the results in this section can be largely carried over to any set of parameter values within the 95% credible interval of the posterior distribution.

7.1 Correlation between realized and fundamental portfolio returns

Table 8 reports the contemporaneous correlations between the realized and fundamental returns on the 120 decile portfolios and the 12 high-minus-low decile portfolios for the 12 anomalies. The fundamental and realized portfolio returns are all highly correlated and the correlation coefficients are all significant at the 1% level. The average correlation is 0.69 for decile portfolios and 0.43 for the high-minus-low deciles.

Notice that although the model cannot generate significant value premium and accruals premium with the correct sign, the fundamental returns on the BM and accruals deciles are still highly correlated with the realized ones. The correlations of these 20 deciles range between 0.63 and 0.78, while the correlations of the high-minus-low deciles are 0.53 for BM and 0.41 for accruals.

Figure 7 plots the time-series returns of the 12 high-minus-low deciles. Although the fundamental and realized return series do show strong comovements, the realized portfolio returns are much more volatile than the fundamental ones. In particular, the model fails to generate the extreme movements observed in the data, such as the momentum crash of

2009. The lack of large variations in fundamental returns is driven by the fact that firm fundamentals, measured by accounting information, do not exhibit fluctuations as large as those in realized stock returns.

7.2 Factor premiums and market states

The performance of long-short anomaly strategies often varies with the market conditions due to cyclical changes in firm fundamentals and market risk premiums. For example, Gonçalves, Xue and Zhang (2020) show that value and investment premiums are counter-cyclical while momentum and profitability premiums are procyclical. Following Gonçalves, Xue and Zhang (2020), we define up market as periods following nonnegative prior 36-month market returns and down market as periods following negative prior 36-month market returns, and examine the cyclicality of the 12 factor premiums.[18]

Table 9 shows that the momentum, ROE, NSI, GP/A, and ROA premiums exhibit strong pro-cyclicality, while the BM, I/A, size, ΔPI/A, and Ad/M premiums exhibit counter-cyclicality. In contrast, the fundamental premiums show less variations between up and down states, but they do exhibit the same cyclicality as those of the realized premiums. For example, the momentum premium is 18.51% following up markets but −12.99% following down markets. The contrast is 12.43% versus 8.77% for the fundamental momentum premium.

The realized RD/M premium does not show significant dependence on market states, being 8.61% following up markets and 9.53% following down markets. However, the predicted RD/M premium exhibits strong pro-cyclicality, being 6.12% versus −1.54%. This discrepancy highlights the importance of modeling R&D explicitly in order to capture the time-series dynamics of the RD/M premium. In the current model, R&D investment is not directly modeled and its influence on stock returns is bridged by its correlations with profitability and investments in physical and working capitals. Finally, the predicted accruals premium continues to show opposite signs as those of the realized one, following both up and down markets.

[18]Results do not change qualitatively when up and down markets are defined based on prior 12- or 24-month market returns.

7.3 Persistence of factor premiums

One important aspect of a factor premium is its persistence, which varies greatly across anomalies. Figure 8 presents the event-time dynamics of the realized (top of each panel) and fundamental returns (bottom of each panel) for the high and low deciles during the 36-month period after the portfolio formation for each anomaly. The momentum, ROE, and ROA premiums diminish within 12 months after the portfolio formation, while the other premiums subsist much longer. The model succeeds in reproducing the short-lived nature of the momentum, ROE, and ROA premiums, as well as the long-lived nature of the rest with the exception of the accruals premium. The accruals premium lasts for 18 months in the data while there is no noticeable decrease in the fundamental premium after 36 months. Given that the model cannot get the sign of the accruals anomaly right, it is not surprising that the model cannot explain the persistence either.

7.4 Recursive estimation of parameters with expanding window

Our baseline estimation utilizes the information of the entire sample and in principle should generate parameter estimates closest to their true values if the model is correct. However, one may be concerned that the performance of the model comes from the look-ahead advantage of the in-sample estimation. In this section, we recursively estimate the model's parameters with expanding windows and compute the one-year-ahead fundamental returns. This procedure, which combines the recursive parameters with realized accounting variables (instead of their forecasts), is in the same spirit as in Fama and French (1997).

Starting from October 1980, we recursively estimate the model parameters from an expanding window that starts in June 1967 and ends in May of each year from 1980 to 2018. The latest accounting variables in the first recursive estimation must come no later than May 1980, the latest month for fiscal year 1979. Following Gonçalves, Xue and Zhang (2020), we impose a 4-month lag to ensure no look-ahead bias. For example, with the parameters estimated at the end of May 1980, we compute the one-year-ahead fundamental returns from October 1980 to September 1981. We expand the recursive windows one year at a time until May 2017. To compare with realized returns that need to be smoothed within a 12-month window, we evaluate the fit of the recursive estimation for the sample between March 1981 and December 2016. We allow parameters to vary across industry within each estimation window.

Table 10 compares the high-minus-low alpha ($\alpha_{\text{H-L}}$), i.e., the alpha of the factor premium, and the average absolute decile alpha ($\overline{|\alpha_{\text{D}}|}$) constructed from recursive estimations with those from the baseline estimation for each anomaly between March 1981 to December 2016. As we expect, the average absolute decile alpha becomes larger in the out-of-sample (OOS) estimation than that of the baseline for most anomalies except for the momentum and size because the in-sample estimate matches stock returns better on average. However, in terms of matching the anomaly premiums, the OOS estimation does not perform worse than the baseline estimation. Each specification fails to explain two anomalies, in addition to the value and accruals anomalies. Neither of them can explain the net stock issues (NSI) premium, the high-minus-low alpha of which is significant at the 5% level under both specifications. In addition, the baseline estimation cannot explain the gross profitability (GP/A) premium while the OOS estimation cannot explain the asset growth (I/A) premium.

In sum, whether parameters are estimated in sample or out of sample is not critical for the model's ability to explain factor premiums in general. The fact that our estimation targets firm-level returns, not the average anomaly premiums, might be a key reason behind this result.

7.5 Out-of-sample (OOS) return forecast

Traditional forecasts on cross-sectional stock returns rely on linear models to organize information. Gu, Kelly and Xiu (2020) show that machine learning methods can significantly improve the OOS forecasting performance of traditional linear models.[19] However, machine learning methods lack economic structures, similar to linear risk-factor models. One advantage of our estimation is that it combines the Bayesian MCMC method with a simple yet powerful economic structure. We examine its OOS forecasting performance in this section.

To forecast stock returns, we need to forecast the firm fundamentals used in Equation (6) in addition to recursively estimating parameter values as in Section 7.4.[20] To reduce

[19]The set of machine learning methods studied in Gu, Kelly and Xiu (2020) includes generalized linear models with penalization, dimension reduction via principal components regression (PCR) and partial least squares (PLS), regression trees (including boosted trees and random forests), and neural networks.

[20]We have tried three specifications of recursive estimation: (1) allow both industry and time variations within each estimation window, and use the parameter estimates at the end of the expanding window to construct the one-year-ahead fundamental returns; (2) allow both industry and time variations within each

measurement errors, we set the expected r^{Ba}_{it+1}, τ_{t+1}, and δ_{it+1} values to their current values from the most recent fiscal year-end at least four months ago. In addition, values of physical and working capital stocks, K_{it+1} and W_{it+1}, are known at the beginning of time $t+1$. The key is to forecast Y_{it+1} and I_{it+1}. Following Gonçalves, Xue and Zhang (2020), we forecast I_{it+1}/K_{it+1} on lagged Tobin's Q_{it}, sales-to-total capital, $Y_{it}/(K_{it}+W_{it})$, and investment-to-physical capital, I_{it}/K_{it}, and forecast annual sales growth, Y_{it+1}/Y_{it}, on the year-over-year quarterly sales growth rates of the prior four quarters. We winsorize the sales growth rates at the 2.5%-97.5% level.

At the beginning of each month t from October 1980 to September 2018, we use the prior 120-month rolling window to estimate the cross-sectional forecasting regressions of I_{it+1}/K_{it+1} and Y_{it+1}/Y_{it}. Monthly Fama and MacBeth (1973) cross-sectional weighted least squares regressions are used for the forecast. The I_{it+1} and Y_{it+1} data are obtained from the most recent fiscal year ending at least four months prior to month t, and the predictors in the forecasting regressions are further lagged accordingly. We then construct predicted returns using forecasted fundamentals and recursively estimated parameters based on Equation (6) in Section 7.4.

At the beginning of each month t from October 1980 to September 2018, we form deciles based on the predicted stock returns and NYSE breakpoints and hold them for one month. Table 11 presents the realized average monthly excess returns, the CAPM alpha, the Fama-French three-factor, Carhart four-factor, and Fama-French five-factor alphas, and the Hou, Xue and Zhang (2020) q-factor alpha of the 10 deciles and the high-minus-low decile. First of all, our model shows strong and reliable forecast capability, with the realized average monthly excess return of the high-minus-low decile being 0.45% (t=2.45). Second, and more importantly, this realized return spread between firms with the highest and lowest predicted returns cannot be explained by the commonly used risk factors. In fact, the risk-adjusted alphas are even larger and more significant than the average excess return in some cases. The CAPM alpha, the Fama-French three-factor, Carhart four-factor, and Fama-French five-

estimation window, and use the average of the time-series parameter estimates to construct the one-year-ahead fundamental returns; (3) allow industry but not time variations, and use the estimates to construct the one-year-ahead fundamental returns. The last specification, which is the one used in Section 7.4, gives us the highest prediction power. Estimates of the third specification better utilize the information of the entire prior sample in a structural way. The only scenario where the first specification would perform better is when there is a trend in the time series of parameter estimates, which is not the case here as shown in Figure 2.

factor alphas, and the Hou, Xue and Zhang (2020) q-factor alpha are 0.43% (t=2.38), 0.58% (t=3.25), 0.52% (t=2.87), 0.61% (t=3.08), and 0.47% (t=2.22), respectively. The fact that these linear factor models cannot explain the return spread between firms with the highest and lowest predicted returns suggests that the nonlinear structure imposed by the simple q-model plays a critical role in explaining the cross-sectional return differences.

8 Conclusion

Can stock market anomalies be explained within an investment-based asset pricing framework? To answer this question, prior studies often choose model parameters to fit a specific set of anomalies and different values are needed to fit each anomaly. Using Bayesian MCMC, this paper estimates a simple two-capital q-model to match firm-level stock returns. The estimated model generates large and significant size, momentum, investment, profitability, and intangibles premiums. The results hold for a variety of alternative specifications, including in-sample and out-of-sample estimations, and estimations with industry and time variations in parameter values, with time variations only, with industry variations only, and with constant values. Moreover, the model exhibits reliable out-of-sample forecasts on stock returns in the cross section, which can not be explained by the commonly used linear factor models. Our results show that the simple q-model goes a long way in explaining stock market anomalies.

References

Amihud, Yakov. 2002. "Illiquidity and stock returns: cross-section and time-series effects." Journal of Financial Markets 5(1):31–56.

Bazdresch, Santiago, Frederico Belo, and Xiaoji Lin. 2014. "Labor Hiring, investment and stock return predictability in the cross section." Journal of Political Economy 122(1):129–177.

Bazdresch, Santiago, R. Jay Kahn and Toni M. Whited. 2018. "Estimating and testing dynamic corporate finance models." Review of Financial Studies 31:322–361.

Belo, Frederico, Chen Xue and Lu Zhang. 2013. "A supply approach to valuation." Review of Financial Studies 26:3029–3067.

Belo, Frederico, Vito Gala, Juliana Salomao and Maria Ana Vitorino. 2021. "Decomposing firm value." Journal of Financial Economics, forthcoming.

Berk, Jonathan, Richard Green and Vasant Naik. 1999. "Optimal investment, growth options, and security returns." Journal of Finance 54:1553–1607.

Bhushan, Ravi. 1989. "Firm characteristics and analyst following." Journal of Accounting and Economics 11(2-3):255–274.

Campbell, John Y. 2017. Financial Decisions and Markets: A Course in Asset Pricing. Princeton University Press.

Carlson, Murray, Adlai Fisher and Ron Giammarino. 2004. "Corporate investment and asset price dynamics: implications for the cross-section of returns." Journal of Finance 59:2577–2603.

Cochrane, John H. 1991. "Production-based asset pricing and the link between stock returns and economic fluctuations." Journal of Finance 46:209–237.

Dechow, Patricia M and Ilia D. Dichev. 2002. "The quality of accruals and earnings: The role of accrual errors." The Accounting Review 77:35–59.

Eisfeldt, Andrea L. and Adriano A. Rampini. 2006. "Capital reallocation and liquidity." Journal of Monetary Economics 53(3):369–399.

Eisfeldt, Andrea L., Edward Kim and Dimitris Papanikolaou. 2020. Intangible value. Working paper, UCLA Anderson School of Management and NBER.

Eisfeldt, Andrea L. and Tyler Muir. 2016. "Aggregate external financing and savings waves." Journal of Monetary Economics 84:116–133.

Erickson, Timothy and Toni Whited. 2000. "Measurement error and the relationship between investment and q." Journal of Political Economy 108(5):1027–1057.

Fama, Eugene F. and James D. MacBeth. 1973. "Risk, return, and equilibrium: Empirical tests." Journal of Political Economy 81(3):607–636.

Fama, Eugene F. and Kenneth R. French. 1997. "Industry cost of equity." Journal of Finance 43:153–193.

Gomes, Joao F, Leonid Kogan and Lu Zhang. 2003. "Equilibrium cross section of returns." Journal of Political Economy 111:693–732.

Gonçalves, Andrei, Chen Xue and Lu Zhang. 2020. "Aggregation, capital heterogeneity, and the investment CAPM." Review of Financial Studies 33:2728–2771.

Gu, Shihao, Bryan Kelly and Dacheng Xiu. 2020. "Empirical asset pricing via machine learning." Review of Financial Studies 33(5):2223–2273.

Hayashi, Fumio. 1982. "Tobin's marginal q and average q: A neoclassical interpretation." Econometrica 8:213–224.

Hou, Kewei, Chen Xue and Lu Zhang. 2015. "Digesting anomalies: An investment approach." Review of Financial Studies 28(3):650–705.

Hou, Kewei, Chen Xue and Lu Zhang. 2020. "Replicating anomalies." Review of Financial Studies 33(5):2019–2133.

Kogan, Leonid and Dimitris Papanikolaou. 2014. "Growth Opportunities, Technology Shocks, and Asset Prices." Journal of Finance 69:675–718.

Li, Erica X.N., Dmitry Livdan and Lu Zhang. 2009. "Anomalies." Review of Financial Studies 22:4301–4334.

Li, Erica X.N., Haitao Li, Shujing Wang and Cindy Yu. 2019. "Macroeconomic risks and asset pricing: Evidence from a dynamic stochastic general equilibrium model." Management Science 65(8):3449–3947.

Li, Haitao, Martin T. Wells and Cindy L. Yu. 2008. "A Bayesian analysis of return dynamics with Lévy jumps." Review of Financial Studies 21(2):2345–2378.

Liu, Laura X. L. and Lu Zhang. 2014. "A neoclassical interpretation of momentum." Journal of Monetary Economics 67:109–128.

Liu, Laura X.L., Toni W. Whited and Lu Zhang. 2009. "Investment-based expected stock returns." Journal of Political Economy 117(6):1105–1139.

Papanikolaou, Dimitris. 2011. "Investment Shocks and Asset Prices." Journal of Political Economy 119(4):639–685.

Smets, Frank and Rafael Wouters. 2007. "Shocks and frictions in US business cycles: A Bayesian DSGE approach." American Economic Review 97(3):586–606.

Taylor, Lucian A. 2013. "CEO wage dynamics: Estimates from a learning model." Journal of Financial Economics 108(1):79–98.

Wu, Jin (Ginger), Lu Zhang and X. Frank Zhang. 2010. "The q-theory approach to understanding the accrual anomaly." Journal of Accounting Research 48(1):177–223.

Zhang, Lu. 2005. "The value premium." Journal of Finance 60(1):67–103.

Zhang, X. Frank. 2013. "Accruals, investment, and the accrual anomaly." The Accounting Review 82(5):1333–1363.

Figure 1: **Simulation study**

This figure plots the true values (in red solid lines), the NLS estimates (in green lines with triangle markers), the Bayesian MCMC posterior means (in blue dashed lines) and the 95% prediction intervals (in shaded areas) of the model parameters estimated from the simulated data. The marginal product and adjustment costs parameters of Consumer Nondurables, Manufacturing, and Business Equipment industries are denoted as γ_i and a_i for $i = 1, 2, 3$, respectively..

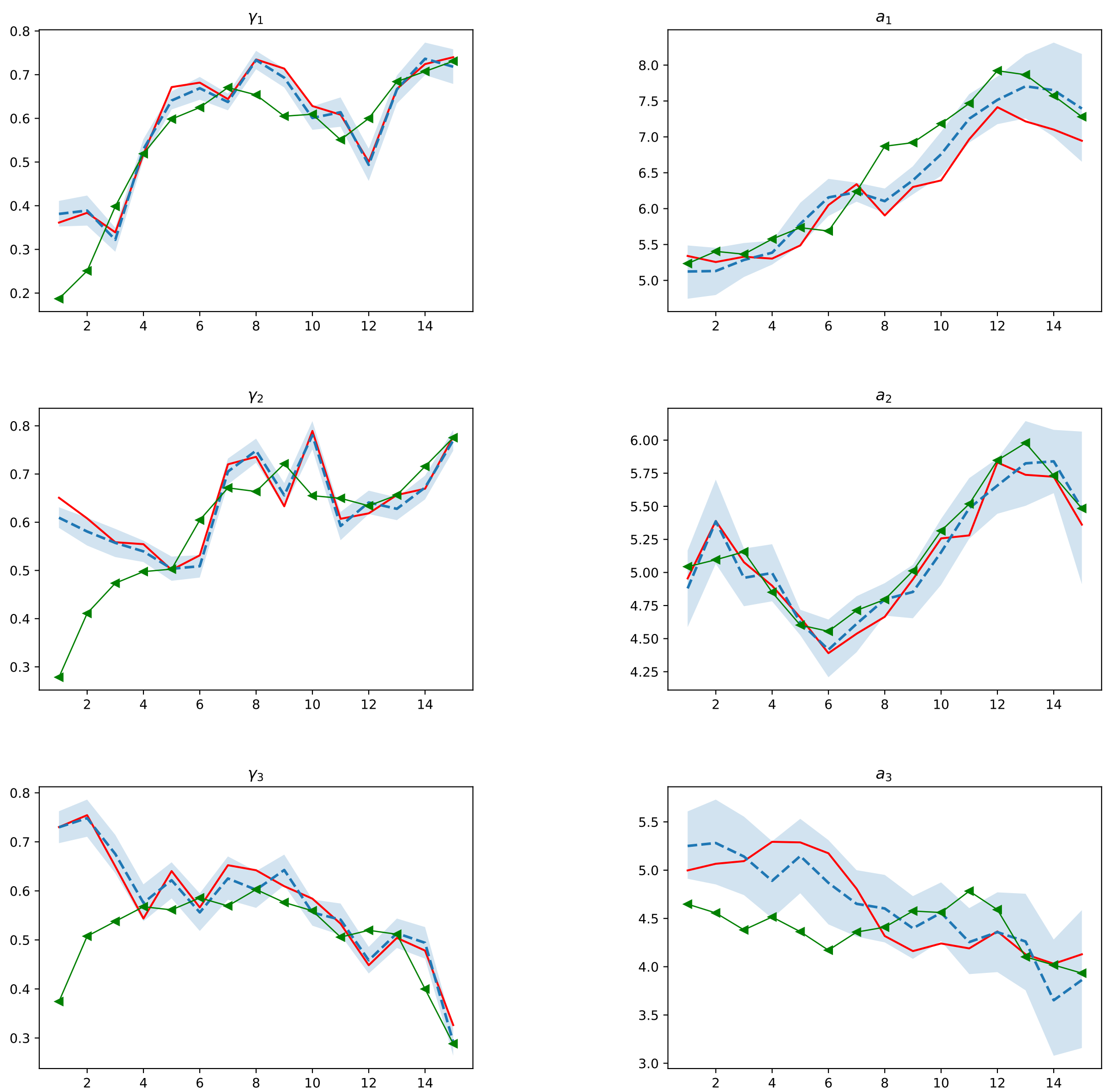
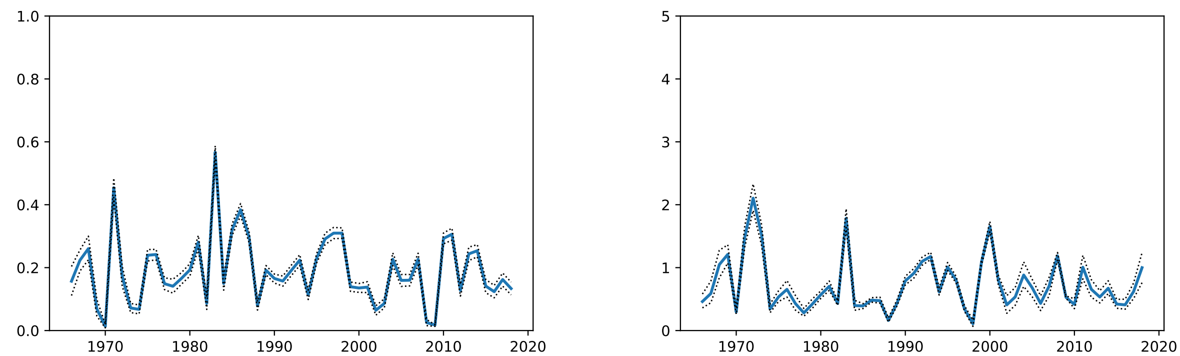


Figure 2: **Time series of estimated parameters**

This figure shows the time series (1967-2016) of the posterior means of the marginal product parameter, γ, and physical investment adjustment costs parameter a, averaged across the Fama-French 10 industries. Parameter values are plotted in solid lines and the 95% credible intervals are in dotted lines.

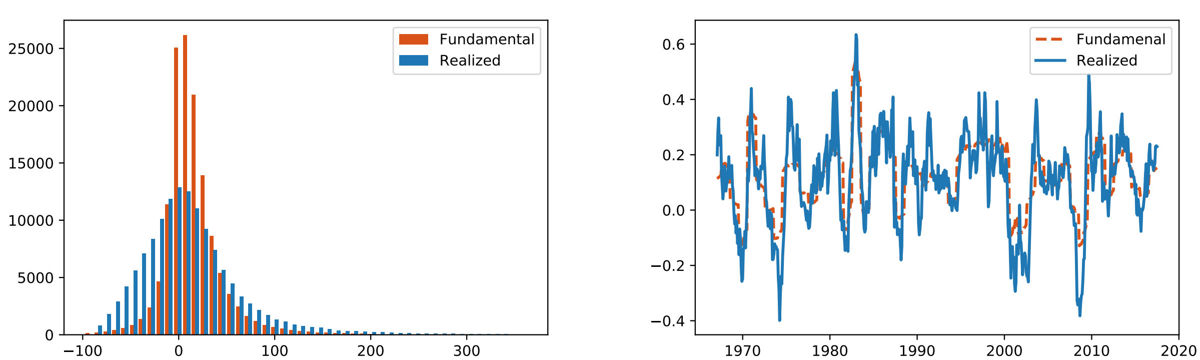


(a) Marginal productivity parameter: γ

(b) Investment adjustment costs parameter: a

Figure 3: **Distribution of firm-level returns: realized vs. fundamental**

Panel (a) presents the histograms of the realized (in blue) and fundamental (in orange) firm-level stock returns from June 1967 to December 2016. Panel (b) presents the time series of the value-weighted returns on the market. Returns are in percentage per annum. The number of observations is 136,598. Observations in Panel (a) are trimmed at 0.5 and 99.5 percentiles for illustration purposes.



(a) Histogram of firm-level stock returns

(b) Return on the market index

Figure 4: **Posterior distributions of the fundamental factor premiums**

This figure plots the posterior probability density functions of the fundamental factor premiums formed on book-to-market (BM), momentum (R11), asset growth (I/A), return-on-equity (ROE), size (Size), accruals (Accruals), net share issues (NSI), investment-to-assets ratio (ΔPI/A), gross profitability (GP/A), return-on-assets (ROA), R&D-to-market ratio (RD/M), and advertising-to-market ratio (Ad/M). The red lines represent normal distributions with means and standard deviations being the posterior means and standard deviations of the corresponding factor premiums. The 2.5%, 50%, and 97.5% percentiles of each posterior distribution are labeled below the horizontal axes.

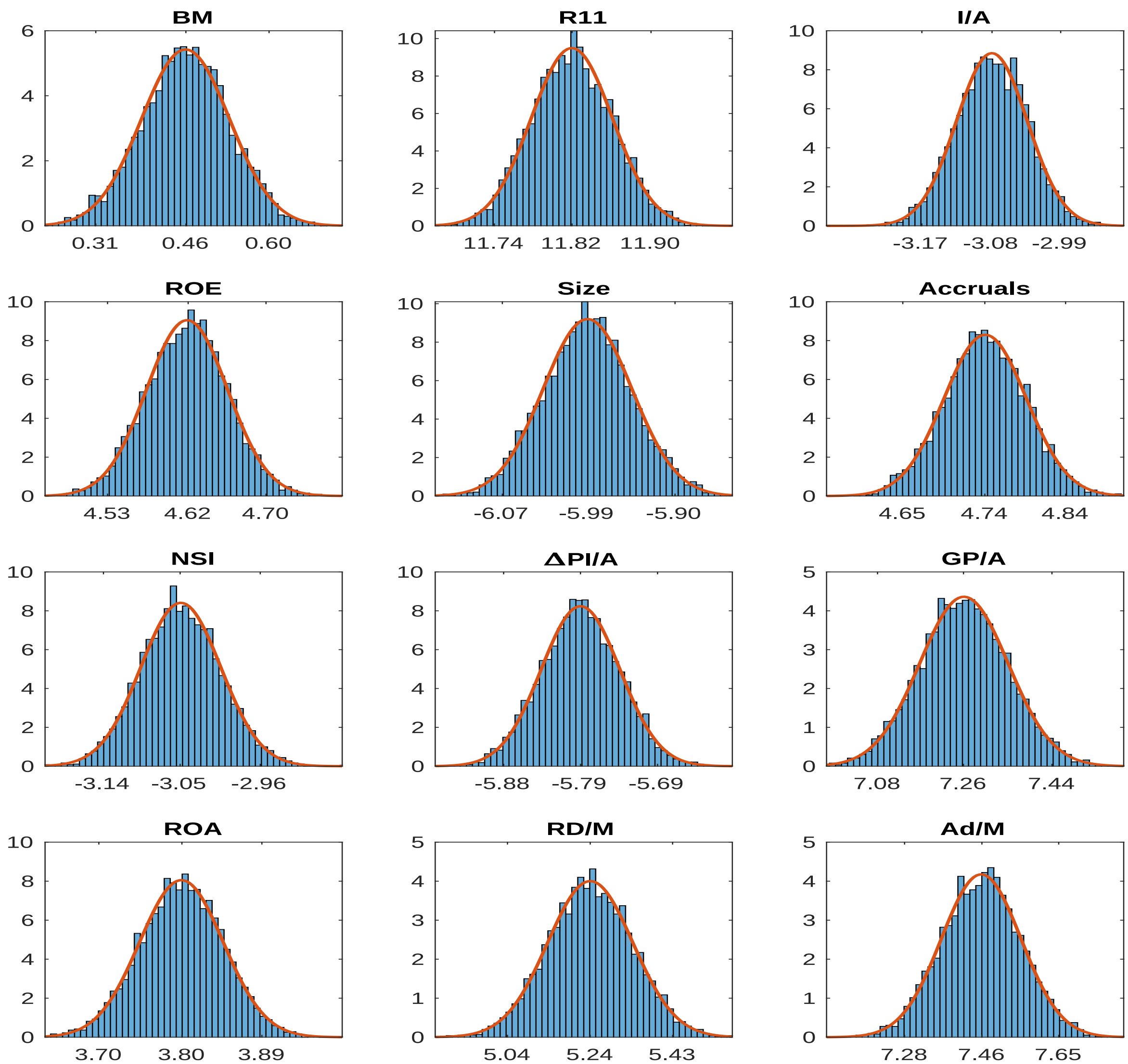
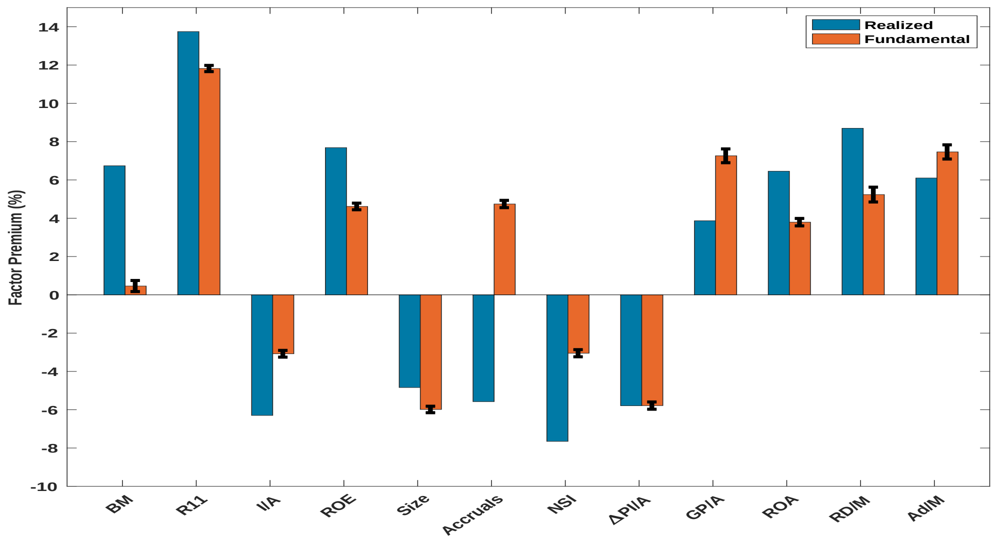
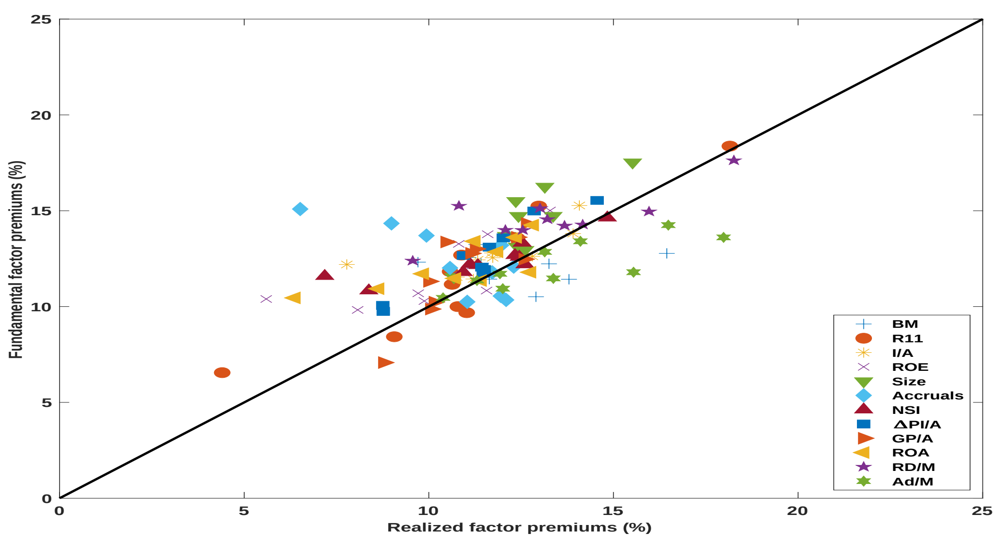


Figure 5: **Factor premiums: realized vs. fundamental**

The realized factor premiums and the posterior means of the fundamental factor premiums are plotted in Panel (a). The error bars represent the 95% credible intervals of the posterior distributions. The average fundamental returns of the decile portfolios sorted on the 12 anomaly variables are plotted against the corresponding realized ones in Panel (b). Returns are in percentage per annum.



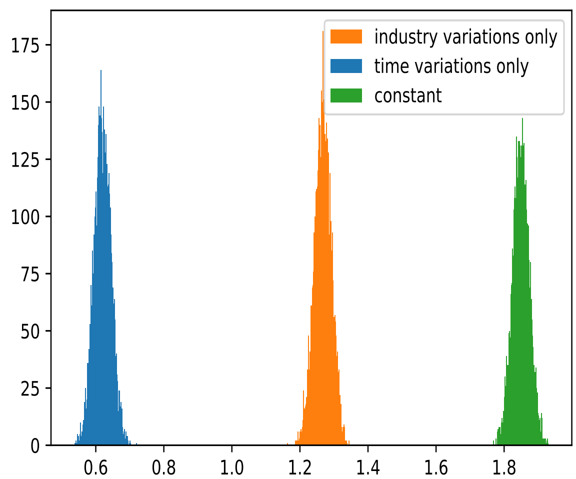
(a) Factor premiums



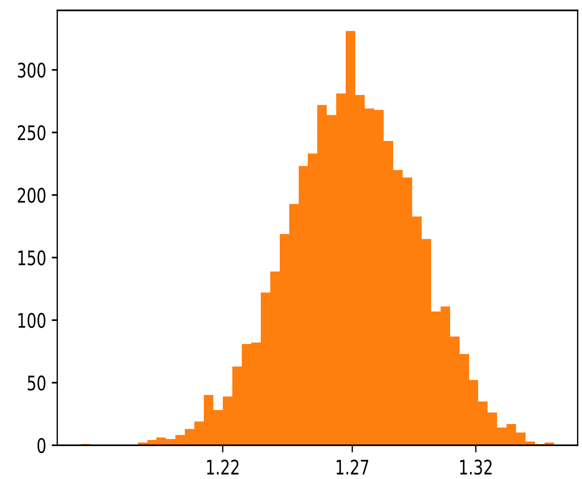
(b) Returns of decile portfolios

Figure 6: **Alternative estimation specifications**

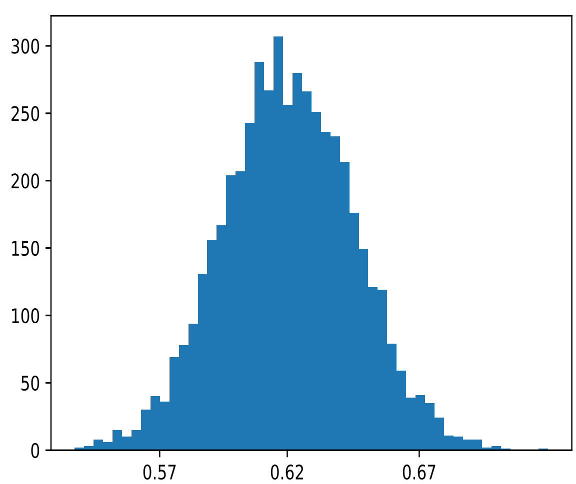
This figure plots the posterior distributions of $d^a = \dfrac{1}{T}\sum_{t=1}^{T}\dfrac{1}{N_t}\sum_{i=1}^{N_t} d_{it}^a$ for three alternative settings: parameters with industry variations only, time variations only, and constant values, respectively, where $d_{it}^a = \left|r_{it+1}^S - r_{it+1}^{F(a)}\right| - \left|r_{it+1}^S - r_{it+1}^{F(b)}\right|$, $r^{F(a)}$ and $r^{F(b)}$ are the fundamental returns under the alternative specification, a, and the baseline, b, and d^a is in percentage per annum.



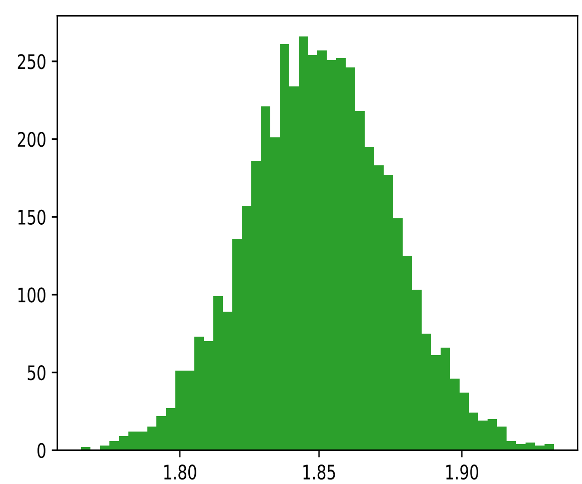
(a) Three alternative estimations



(b) Industry variations only



(c) Time variations only



(d) Constant values

Figure 7: **Time series of factor premiums: realized vs. fundamental**

This figure plots the time series of the realized (in blue solid lines) and fundamental (in red dotted lines) factor premiums. Returns are in percentage per annum and in monthly frequency.

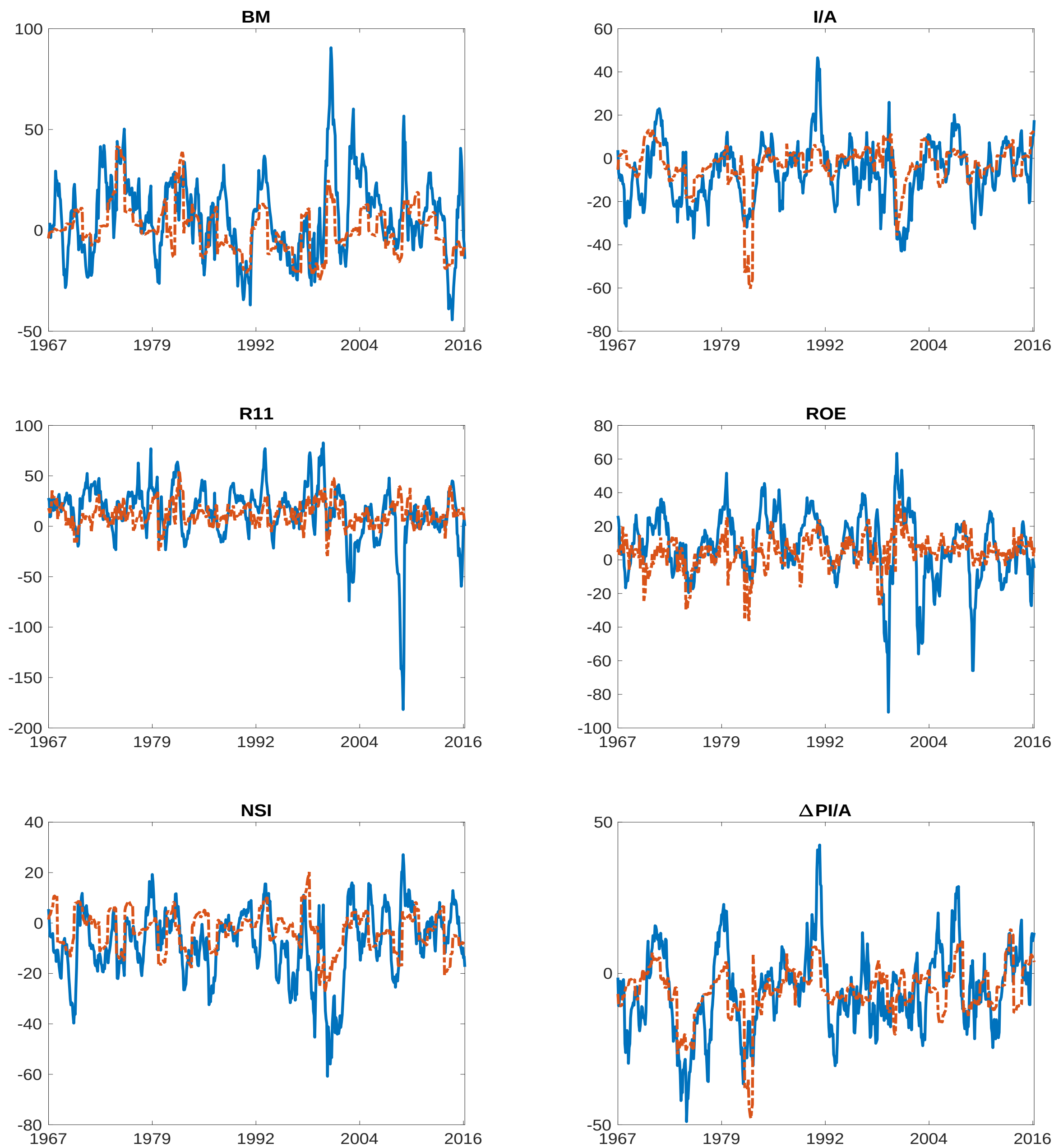


Figure 7: **Time series of factor premiums: realized vs. fundamental (continued)**

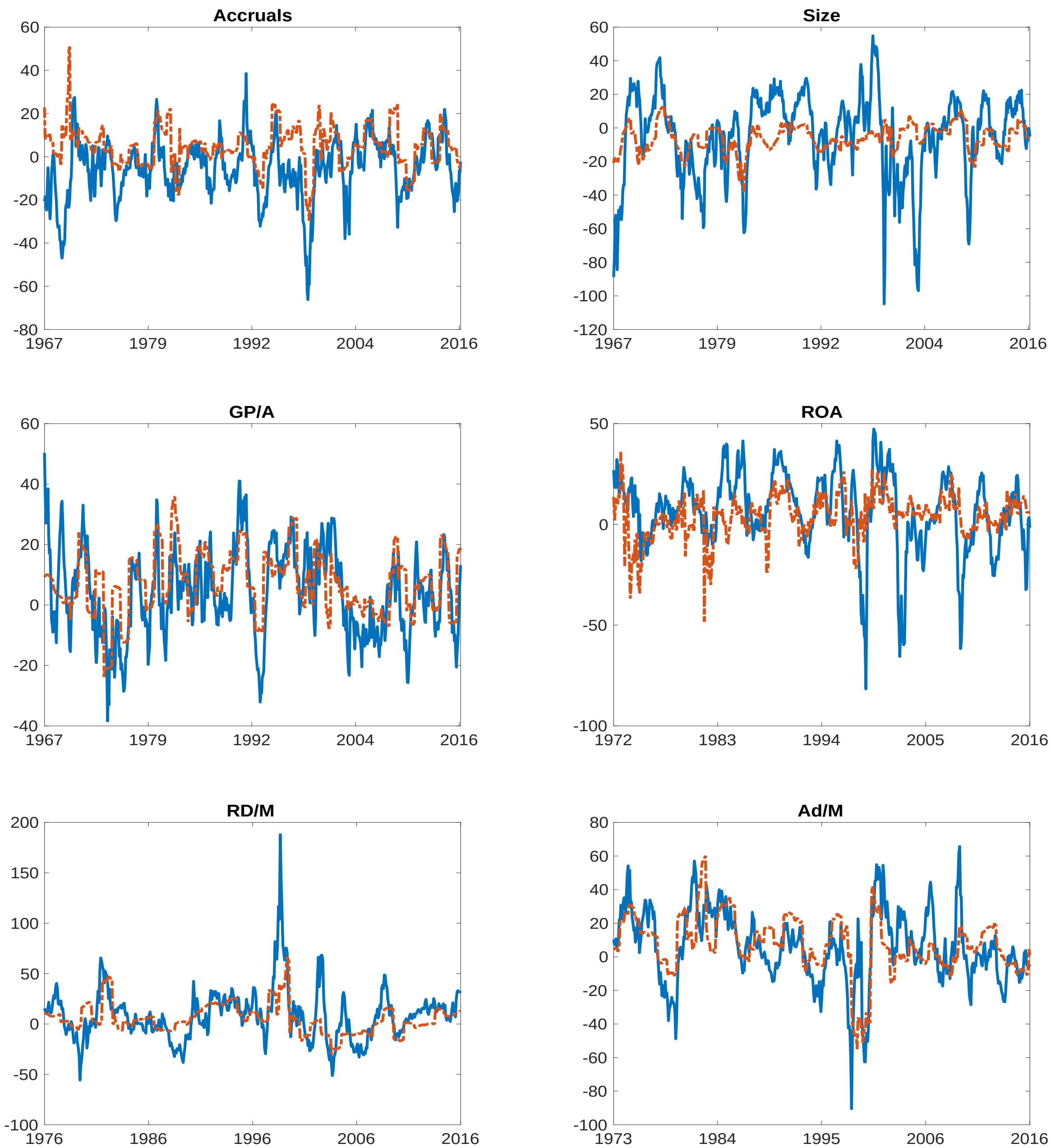
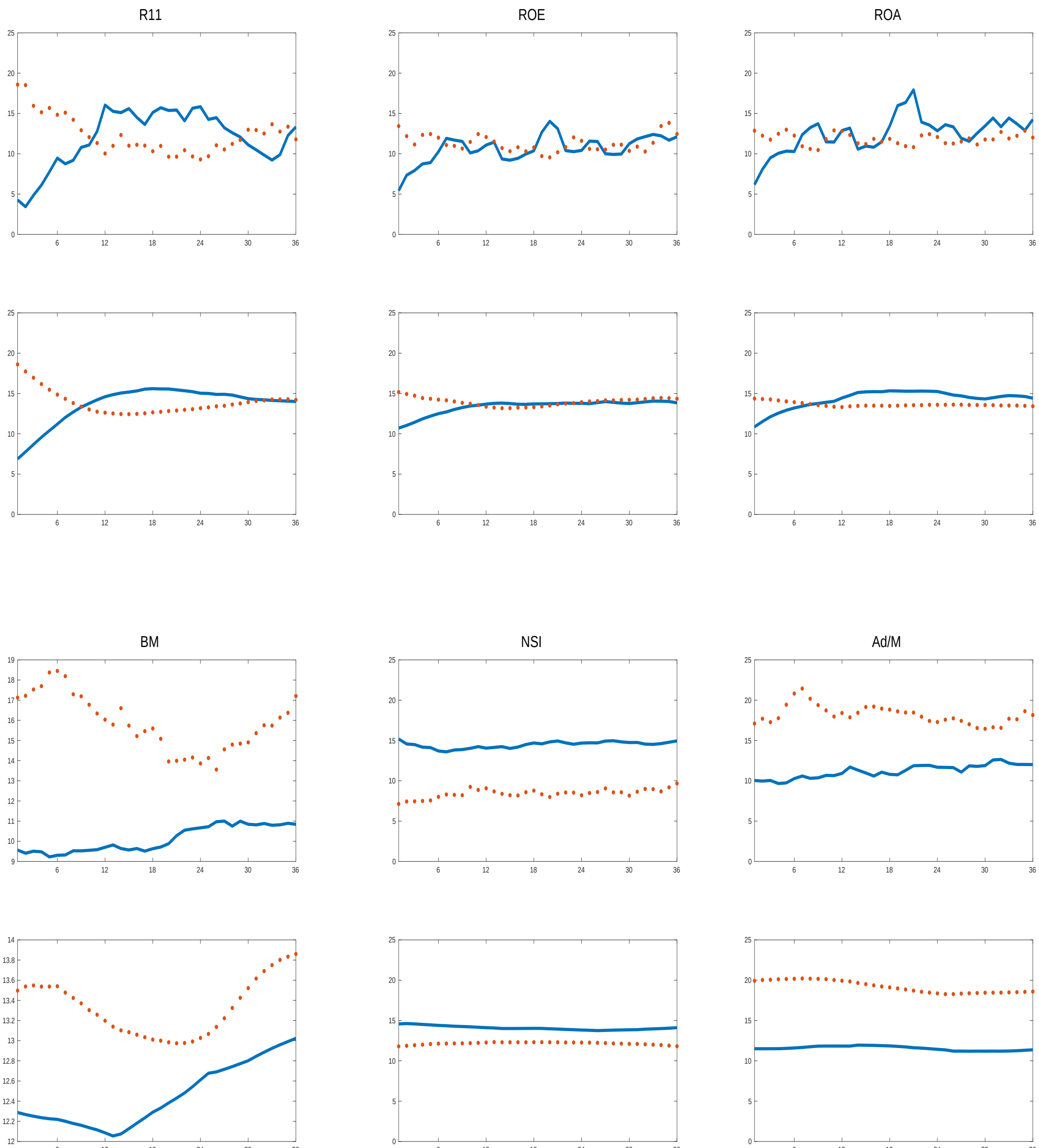


Figure 8: **Persistence of factor premiums: realized vs. fundamental**

This figure plots the realized (top of each panel) and fundamental returns (top of each panel) on the low (blue solid lines) and high (red dotted lines) deciles for 36 months after the portfolio formation for each anomaly. Returns are in percentage per annum and in monthly frequency.



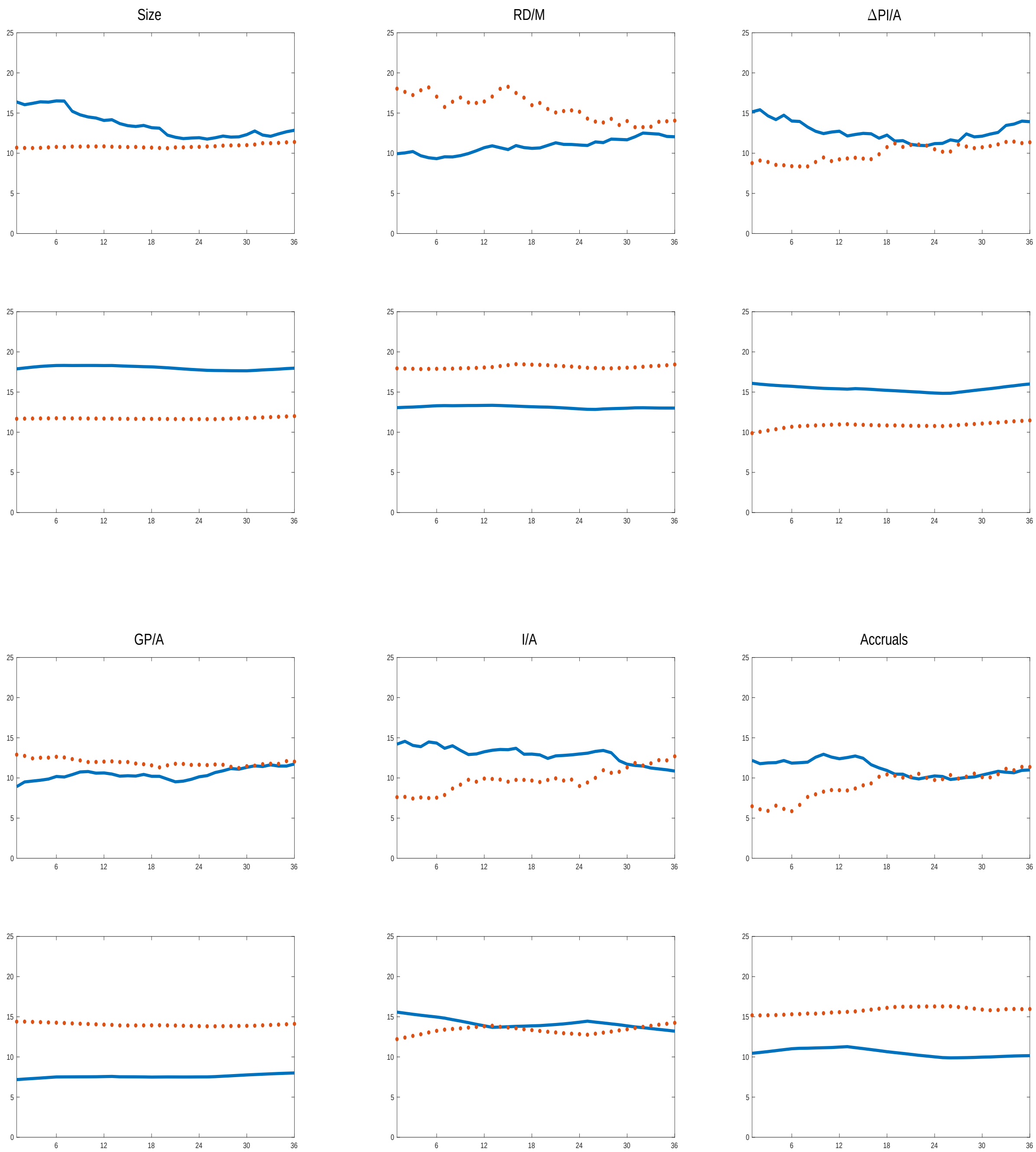

Size
RD/M
ΔPI/A
GP/A
I/A
Accruals

Table 1: **Descriptive properties of decile portfolios**

This table reports the monthly average excess returns of decile portfolios for 12 anomaly variables, including book-to-market equity ratio (BM), momentum (R11), asset growth (I/A), return-on-equity (ROE), size (Size), accruals (Accruals), net share issues (NSI), investment-to-assets ratio (ΔPI/A), gross profitability (GP/A), return-on-assets (ROA), R&D-to-market ratio (RD/M), and advertising-to-market ratio (Ad/M). The *t*-statistics adjusted for heteroscedasticity and autocorrelations are reported in parentheses. Decile portfolios are formed with NYSE breakpoints and value-weighted returns. L denotes the low decile, H the high decile, and H-L the high-minus-low decile. The sample period is from January 1967 to June 2017 for all anomaly variables except for ROA, RD/M, and Ad/M. The sample starts at July 1972, July 1976, and July 1973 for ROA, RD/M, and Ad/M, respectively, due to data availability.

	L	2	3	4	5	6	7	8	9	H	H-L
BM	0.40	0.53	0.60	0.51	0.49	0.54	0.65	0.63	0.72	0.90	0.50
	(1.83)	(2.67)	(3.14)	(2.48)	(2.64)	(2.92)	(3.41)	(3.33)	(3.58)	(3.77)	(2.39)
R11	-0.01	0.34	0.52	0.48	0.48	0.46	0.49	0.60	0.64	1.04	1.05
	(-0.04)	(1.25)	(2.40)	(2.39)	(2.54)	(2.38)	(2.84)	(3.30)	(3.11)	(3.91)	(3.51)
I/A	0.71	0.70	0.63	0.50	0.55	0.54	0.56	0.50	0.54	0.26	-0.45
	(2.97)	(3.64)	(3.74)	(2.92)	(3.19)	(2.96)	(3.06)	(2.52)	(2.27)	(1.02)	(-3.06)
ROE	0.04	0.27	0.42	0.42	0.55	0.48	0.53	0.49	0.55	0.68	0.64
	(0.12)	(1.08)	(1.95)	(2.30)	(3.04)	(2.47)	(2.76)	(2.55)	(2.88)	(3.28)	(3.10)
Size	0.84	0.70	0.65	0.64	0.69	0.61	0.61	0.62	0.53	0.46	-0.37
	(2.63)	(2.49)	(2.50)	(2.61)	(2.83)	(2.72)	(2.69)	(2.96)	(2.75)	(2.66)	(-1.55)
Accruals	0.58	0.51	0.57	0.47	0.58	0.58	0.59	0.44	0.33	0.19	-0.40
	(2.31)	(2.48)	(3.28)	(2.62)	(2.96)	(3.06)	(2.96)	(2.09)	(1.39)	(0.69)	(-2.61)
NSI	0.80	0.61	0.53	0.56	0.50	0.51	0.62	0.58	0.30	0.19	-0.62
	(4.35)	(3.40)	(2.77)	(3.00)	(2.61)	(2.63)	(3.00)	(2.63)	(1.33)	(0.78)	(-4.07)
ΔPI/A	0.75	0.65	0.59	0.54	0.48	0.53	0.53	0.54	0.33	0.34	-0.41
	(3.25)	(3.37)	(3.40)	(2.98)	(2.53)	(2.95)	(2.70)	(2.61)	(1.47)	(1.31)	(-2.96)
GP/A	0.28	0.39	0.41	0.42	0.63	0.55	0.47	0.52	0.61	0.63	0.35
	(1.41)	(2.05)	(2.02)	(2.04)	(3.18)	(2.78)	(2.22)	(2.46)	(3.02)	(3.15)	(2.36)
ROA	0.01	0.23	0.49	0.45	0.38	0.59	0.52	0.51	0.57	0.59	0.58
	(0.03)	(0.83)	(2.07)	(2.22)	(1.85)	(2.78)	(2.57)	(2.47)	(2.71)	(2.67)	(2.73)
RD/M	0.37	0.56	0.45	0.62	0.61	0.66	0.71	0.88	0.74	1.02	0.65
	(1.85)	(2.53)	(1.98)	(2.72)	(2.64)	(3.09)	(3.16)	(3.62)	(2.95)	(3.11)	(2.41)
Ad/M	0.42	0.56	0.51	0.64	0.54	0.80	0.64	0.71	0.94	0.91	0.48
	(1.62)	(2.32)	(2.18)	(2.84)	(2.55)	(3.86)	(3.02)	(3.14)	(3.58)	(3.23)	(2.08)

Table 2: **Descriptive statistics of firm-level accounting variables in the fundamental returns**

This table reports the time series averages of the cross-sectional summary statistics (Panel A), including mean, standard deviation, percentiles (5th, 25th, 50th, 75th, and 95th), and pairwise correlations (Panel B) for firm-level annual accounting variables. The sample period of fundamental returns is from June 1967 to December 2016.

Panel A. Summary statistics

	Mean	StdDev	p5	p25	p50	p75	p95
I_{it}/K_{it}	0.37	0.48	-0.04	0.11	0.22	0.43	1.43
$\Delta W_{it}/W_{it}$	0.13	0.36	-0.34	-0.06	0.07	0.23	0.88
Y_{it}/K_{it}	7.82	8.72	0.44	2.23	4.91	9.42	31.94
Y_{it}/W_{it}	3.09	1.79	0.70	1.86	2.72	3.95	7.47
$Y_{it}/(K_{it}+W_{it})$	1.64	0.94	0.30	0.97	1.53	2.15	3.78
$K_{it}/(K_{it}+W_{it})$	0.39	0.25	0.06	0.19	0.34	0.57	0.88
w_{it}^B	0.29	0.23	0.01	0.09	0.24	0.45	0.74
δ_{it+1}	0.19	0.13	0.05	0.10	0.15	0.24	0.53
r_{it+1}^B	0.10	0.06	0.02	0.06	0.09	0.12	0.29

Panel B. Pairwise correlations

	$\frac{I_{it+1}}{K_{it+1}}$	$\frac{\Delta W_{it}}{W_{it}}$	$\frac{\Delta W_{it+1}}{W_{it+1}}$	$\frac{Y_{it+1}}{K_{it+1}}$	$\frac{Y_{it+1}}{W_{it+1}}$	$\frac{Y_{it+1}}{(K_{it+1}+W_{it+1})}$	$\frac{K_{it+1}}{(K_{it+1}+W_{it+1})}$	w_{it}^B	δ_{it+1}	r_{it+1}^B
I_{it}/K_{it}	0.32	0.32	0.09	0.16	-0.07	0.05	-0.17	-0.18	0.30	0.10
I_{it+1}/K_{it+1}		0.24	0.32	0.37	0.02	0.21	-0.27	-0.27	0.54	0.26
$\Delta W_{it}/W_{it}$			0.03	0.07	-0.05	0.00	-0.07	-0.09	0.06	0.04
$\Delta W_{it+1}/W_{it+1}$				0.09	0.27	0.20	0.08	-0.13	0.08	0.17
Y_{it+1}/K_{it+1}					0.07	0.60	-0.65	-0.16	0.54	0.10
Y_{it+1}/W_{it+1}						0.57	0.47	0.18	-0.19	0.05
$Y_{it+1}/(K_{it+1}+W_{it+1})$							-0.35	-0.07	0.21	0.13
$K_{it+1}/(K_{it+1}+W_{it+1})$								0.32	-0.56	-0.08
w_{it}^B									-0.32	-0.11
δ_{it+1}										0.13

Table 3: **Parameter estimates**

Column $\overline{\gamma}$ reports the times-series average of the posterior means of the marginal product parameter γ for each of the Fama-French 10 industries; column $\overline{\text{CI}}_{\gamma}$ reports the time-series average of the 95% credible intervals of γ; column $\sigma(\gamma)$ reports the time-series standard deviation of the posterior means of γ; and similar definitions apply to columns $\overline{a}$, $\overline{\text{CI}}_{a}$, and $\sigma(a)$ for the adjustment costs parameter a. Column $\overline{q}$ reports the model implied average Tobin's q for each industry, defined as $\overline{q}_j = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_{jt}}\varpi_{it-1}\left\{\left[1+a_{jt}(1-\tau_t)\frac{I_{it}}{K_{it}}\right]\frac{K_{it+1}}{K_{it+1}+W_{it+1}}+\frac{W_{it+1}}{K_{it+1}+W_{it+1}}\right\}$, where $\varpi_{it-1} = \frac{V_{it-1}}{\sum_{i=1}^{N_{jt-1}}V_{it-1}}$, N_{jt} is the number of firms in industry j in time t, V_{it}, I_{it}, K_{it}, and W_{it} are the market equity, investment in physical capital, physical capital, and working capital of firm i at time t, respectively.

Industry	$\overline{\gamma}$	$\overline{\text{CI}}_{\gamma}$	$\sigma(\gamma)$	$\overline{a}$	$\overline{\text{CI}}_{a}$	$\sigma(a)$	$\overline{q}$
Consumer Nondurables	0.13	[0.11, 0.14]	0.09	0.41	[0.29, 0.55]	0.43	1.60
Consumer Durables	0.16	[0.14, 0.19]	0.18	1.15	[0.83, 1.47]	1.13	1.71
Manufacturing	0.16	[0.15, 0.17]	0.11	0.58	[0.51, 0.65]	0.98	1.60
Energy	0.20	[0.18, 0.22]	0.13	0.45	[0.40, 0.48]	0.54	1.33
Business Equipment	0.23	[0.21, 0.24]	0.19	1.78	[1.67, 1.83]	2.00	1.81
Telecom	0.28	[0.25, 0.30]	0.21	0.71	[0.65, 0.76]	0.66	1.33
Wholesale & Retail	0.08	[0.07, 0.09]	0.06	0.87	[0.77, 0.96]	0.98	1.62
Healthcare	0.19	[0.17, 0.21]	0.16	0.60	[0.44, 0.73]	0.68	1.68
Utilities	0.29	[0.25, 0.32]	0.17	0.25	[0.20, 0.32]	0.32	1.16
Others	0.17	[0.15, 0.18]	0.13	0.48	[0.47, 0.54]	0.52	1.47

Table 4: **Summary statistics of the realized and fundamental firm-level stock returns**

This table reports the following key statistics for the realized (r^S) and fundamental (r^F) firm-level stock returns: mean, standard deviation, skewness, kurtosis, mean absolute error (m.a.e.) of the fundamental returns, and the time series average of cross-sectional correlations between the realized and fundamental returns. The m.a.e. is defined as m.a.e. $\equiv \frac{1}{T}\sum_{t=0}^{T-1}\frac{1}{N_{t+1}}\sum_{i=1}^{N_{t+1}}|r_{it}^S - r_{it}^F|$, where N_{t+1} is the number of firms in period $t+1$. For fundamental returns, both the posterior means and the 95% credible intervals (in square brackets) of these statistics are reported. Both realized and fundamental returns are winsorized at 0.5 and 99.5 percentiles. The fundamental stock returns are computed based on four model setups: the baseline setup (under column $\boldsymbol{\theta}_{jt}$) in which the estimated parameters are industry specific and time varying; the setup ($\boldsymbol{\theta}_j$) in which the estimated parameters are industry specific but constant over time; the setup ($\boldsymbol{\theta}_t$) in which the estimated parameters are time varying but constant across industries; and the setup ($\boldsymbol{\theta}$) in which the estimated parameters are constant over time and across industries. The sample period is from June 1967 to December 2016.

	Data	$\boldsymbol{\theta_{jt}}$	$\boldsymbol{\theta_j}$	$\boldsymbol{\theta_t}$	$\boldsymbol{\theta}$
Mean	14.45	15.65	15.47	14.97	15.47
		[15.55, 15.75]	[15.36, 15.57]	[14.87, 15.06]	[15.60, 15.83]
StdDev	60.78	34.17	18.49	27.36	19.76
		[34.06, 34.27]	[18.39, 18.59]	[27.26, 27.46]	[19.67, 19.85]
Skewness	2.15	1.68	1.68	1.66	2.12
		[1.67, 1.70]	[1.66, 1.70]	[1.64, 1.67]	[2.11, 2.14]
Kurtosis	11.05	11.20	10.66	10.74	13.33
		[11.11, 11.29]	[10.59, 10.73]	[10.66, 10.82]	[13.26, 13.41]
Correlation	na	0.20	0.12	0.12	0.09
		[0.20, 0.20]	[0.12, 0.12]	[0.12, 0.12]	[0.09, 0.10]
m.a.e	na	40.10	41.85	40.85	42.45
		[40.06, 40.13]	[41.82, 41.89]	[40.82, 40.88]	[42.42, 42.48]

Table 5: **Posterior summary of anomaly premiums under the baseline estimation**

For each anomaly premium, this table reports the average annualized returns of the 12 anomaly premiums, r^S, and their corresponding t-values in the data, the posterior means of the fundamental premiums, r^F, the t-values of r^F, and the t-values of alphas, defined as $\alpha \equiv r^S - r^F$, in the baseline estimation. The t-values are adjusted for heteroscedasticity and autocorrelations with lags up to 24 months. We report the 95% credible intervals for r^F, $t(r^F)$, and $t(\alpha)$ in square brackets. The posterior distributions are based on 5,000 Bayesian MCMC draws. Returns are in percentage per annum. The sample period is from June 1967 to December 2016 for all anomaly variables except for ROA, RD/M, and Ad/M, for which The sample starts at December 1972, December 1976, and December 1973, respectively, due to data availability.

Anomaly	r^S	$t(r^S)$	r^F	$t(r^F)$	$t(\alpha)$
BM	6.74	2.57	0.46	0.26	3.33
			[0.31, 0.60]	[0.18, 0.35]	[3.24, 3.42]
R11	13.75	4.15	11.82	12.51	0.78
			[11.74, 11.90]	[12.38, 12.65]	[0.75, 0.81]
I/A	-6.30	-3.23	-3.08	-2.25	-2.10
			[-3.17, -2.99]	[-2.32, -2.18]	[-2.16, -2.04]
ROE	7.69	3.06	4.62	5.72	1.81
			[4.53, 4.70]	[5.58, 5.85]	[1.76, 1.86]
Size	-4.84	-1.37	-5.99	-5.63	0.34
			[-6.07, -5.90]	[-5.73, -5.54]	[0.31, 0.37]
Accruals	-5.58	-3.14	4.74	4.45	-6.28
			[4.65, 4.84]	[4.34, 4.56]	[-6.37, -6.19]
NSI	-7.65	-4.26	-3.05	-3.36	-2.93
			[-3.14, -2.96]	[-3.48, -3.25]	[-2.99, -2.86]
ΔPI/A	-5.79	-2.85	-5.79	-4.81	-0.00
			[-5.88, -5.69]	[-4.93, -4.71]	[-0.07, 0.07]
GP/A	3.87	2.00	7.26	5.84	-2.63
			[7.08, 7.44]	[5.63, 6.07]	[-2.78, -2.48]
ROA	6.46	2.52	3.80	3.99	1.48
			[3.70, 3.89]	[3.86, 4.11]	[1.43, 1.53]
RD/M	8.70	2.26	5.24	2.12	1.42
			[5.04, 5.43]	[2.04, 2.21]	[1.33, 1.50]
Ad/M	6.10	1.87	7.46	2.82	-0.58
			[7.28, 7.65]	[2.74, 2.90]	[-0.66, -0.50]

Table 6: **Anomaly premiums under alternative estimation specifications**

This table reports the posterior means of the fundamental factor premiums (r^F) and the alphas ($\alpha = r^S - r^F$) of the 12 anomalies, with the posterior means of the corresponding t-statistics in parentheses. Fundamental stock returns are computed based on four estimation specifications: the baseline specification (under column $\boldsymbol{\theta_{jt}}$) with industry specific and time varying parameter values; the specification ($\boldsymbol{\theta_j}$) with industry variations in parameter values only; the specification ($\boldsymbol{\theta_t}$) with time variations only; and the setup ($\boldsymbol{\theta}$) with constant parameter values. The fundamental premiums and alphas that are significant at the 1%, 5%, and 10% levels are denoted with three stars, two stars, and one star, respectively. The sample period is from June 1967 to December 2016 for all anomaly variables except for ROA, RD/M, and Ad/M, for which the sample starts at December 1972, December 1976, and December 1973, respectively, due to data availability.

	r^F				α			
	$\boldsymbol{\theta_{jt}}$	$\boldsymbol{\theta_j}$	$\boldsymbol{\theta_t}$	$\boldsymbol{\theta}$	$\boldsymbol{\theta_{jt}}$	$\boldsymbol{\theta_j}$	$\boldsymbol{\theta_t}$	$\boldsymbol{\theta}$
BM	0.46	-0.64	-3.78**	-3.25*	6.28***	7.38***	10.52***	9.99***
	(0.26)	(-0.52)	(-2.17)	(-1.95)	(3.33)	(3.25)	(4.20)	(4.23)
R11	11.82***	3.59***	5.75***	3.82***	1.93	10.16***	8.00***	9.92***
	(12.51)	(7.75)	(6.88)	(6.38)	(0.78)	(4.06)	(2.93)	(3.82)
I/A	-3.08**	-0.06	-0.68	0.48	-3.22**	-6.24***	-5.62***	-6.78***
	(-2.25)	(-0.09)	(-0.43)	(0.48)	(-2.10)	(-3.42)	(-3.18)	(-4.03)
ROE	4.62***	3.89***	5.32***	4.72***	3.07*	3.80**	2.36	2.97
	(5.72)	(9.65)	(9.26)	(9.30)	(1.81)	(2.10)	(1.29)	(1.60)
Size	-5.99***	-5.82***	-7.57***	-8.10***	1.15	0.98	2.73	3.26
	(-5.63)	(-8.62)	(-6.05)	(-8.84)	(0.34)	(0.28)	(0.73)	(0.93)
Accruals	4.74***	4.90***	8.80***	9.01***	-10.32***	-10.48***	-14.38***	-14.59***
	(4.45)	(7.96)	(14.55)	(16.76)	(-6.28)	(-5.96)	(-7.68)	(-8.08)
NSI	-3.05***	-2.80***	-4.59***	-4.70***	-4.60***	-4.85**	-3.07	-2.96
	(-3.36)	(-4.28)	(-6.70)	(-7.18)	(-2.93)	(-2.44)	(-1.49)	(-1.51)
ΔPI/A	-5.79***	-4.21***	-3.12**	-2.29**	-0.01	-1.58	-2.67	-3.50*
	(-4.81)	(-6.99)	(-2.07)	(-2.26)	(-0.00)	(-0.87)	(-1.48)	(-1.90)
GP/A	7.26***	6.77***	15.48***	15.28***	-3.39***	-2.90	-11.61***	-11.42***
	(5.84)	(13.05)	(14.63)	(19.27)	(-2.63)	(-1.47)	(-5.41)	(-5.20)
ROA	3.80***	3.06***	4.10***	3.84***	2.66	3.40*	2.35	2.62
	(3.99)	(6.25)	(5.80)	(6.85)	(1.48)	(1.77)	(1.24)	(1.34)
RD/M	5.24**	2.52*	0.47	0.52	3.46	6.18	8.22**	8.18**
	(2.12)	(1.92)	(0.50)	(0.49)	(1.42)	(1.49)	(2.07)	(1.97)
Ad/M	7.46***	6.39***	13.89***	14.14***	-1.36	-0.29	-7.79**	-8.04***
	(2.82)	(5.75)	(7.32)	(8.96)	(-0.58)	(-0.10)	(-2.42)	(-2.71)

Table 7: **Comparative statics of anomaly premiums under the baseline estimation**

This table reports the fundamental factor premium, denoted r^F, and alpha, defined as $\alpha \equiv r^S - r^F$, from the baseline estimation and four comparative statics for the 12 anomalies. In the comparative static analysis denoted $\overline{I_{it}/K_{it}}$, I_{it}/K_{it} is set to be its cross-sectional median at period t across all the firms. The parameters from the baseline estimation are used to construct the fundamental returns, with all the other firm characteristics remain unchanged. The other three comparative static analyses, $\overline{I_{it+1}/K_{it+1}}$, $\overline{Y_{it+1}/K_{it+1}}$, and $\overline{W_{it+1}/K_{it+1}}$, are designed similarly. The t-values reported in parentheses are adjusted for heteroscedasticity and autocorrelations of up to 24 lags.

	BM		R11		I/A		ROE	
	r^F	α	r^F	α	r^F	α	r^F	α
Baseline	0.47	6.27	11.88	1.87	-3.14	-3.16	4.64	3.05
	(0.27)	(3.33)	(11.03)	(0.57)	(-2.29)	(-2.06)	(4.26)	(1.33)
$\overline{I_{it}/K_{it}}$	-6.62	13.36	14.38	-0.63	3.58	-9.88	5.34	2.35
	(-3.43)	(5.90)	(10.49)	(-0.19)	(2.26)	(-5.47)	(4.69)	(0.96)
$\overline{I_{it+1}/K_{it+1}}$	6.86	-0.12	5.45	8.29	-7.22	0.92	2.33	5.36
	(3.58)	(-0.06)	(4.74)	(2.24)	(-5.22)	(0.55)	(1.41)	(1.93)
$\overline{Y_{it+1}/K_{it+1}}$	69.00	-62.26	-1.13	14.88	-11.83	5.53	-17.83	25.52
	(8.97)	(-7.47)	(-0.41)	(3.69)	(-4.83)	(1.72)	(-5.87)	(7.62)
$\overline{W_{it+1}/K_{it+1}}$	-5.51	12.25	12.56	1.19	-4.50	-1.79	5.17	2.52
	(-2.74)	(6.03)	(6.54)	(0.32)	(-1.09)	(-0.48)	(4.08)	(1.06)
	Size		Accruals		NSI		$\Delta PI/A$	
	r^F	α	r^F	α	r^F	α	r^F	α
Baseline	-5.98	1.14	4.75	-10.33	-3.09	-4.56	-5.83	0.04
	(-5.62)	(0.34)	(4.46)	(-6.29)	(-3.41)	(-2.90)	(-4.82)	(0.03)
$\overline{I_{it}/K_{it}}$	-6.26	1.42	3.97	-9.54	-0.74	-6.91	-0.06	-5.73
	(-5.36)	(0.40)	(3.06)	(-5.29)	(-0.69)	(-3.64)	(-0.04)	(-3.92)
$\overline{I_{it+1}/K_{it+1}}$	-5.35	0.51	6.09	-11.67	-4.40	-3.26	-9.54	3.75
	(-5.23)	(0.15)	(5.53)	(-7.01)	(-4.16)	(-2.15)	(-8.12)	(2.31)
$\overline{Y_{it+1}/K_{it+1}}$	9.87	-14.71	-16.74	11.17	28.67	-36.32	2.54	-8.33
	(4.87)	(-4.23)	(-6.83)	(4.08)	(3.96)	(-5.14)	(0.84)	(-2.36)
$\overline{W_{it+1}/K_{it+1}}$	-20.67	15.83	14.02	-19.60	-3.61	-4.04	-15.20	9.41
	(-9.73)	(3.95)	(7.33)	(-7.90)	(-2.02)	(-2.00)	(-3.56)	(2.48)
	GP/A		ROA		RD/M		Ad/M	
	r^F	α	r^F	α	r^F	α	r^F	α
Baseline	7.26	-3.39	3.83	2.63	5.23	3.47	7.49	-1.39
	(5.84)	(-2.64)	(2.84)	(1.11)	(2.12)	(1.42)	(2.83)	(-0.59)
$\overline{I_{it}/K_{it}}$	10.94	-7.07	4.67	1.79	4.80	3.90	2.64	3.46
	(7.95)	(-4.69)	(3.70)	(0.73)	(2.14)	(1.17)	(0.99)	(1.38)
$\overline{I_{it+1}/K_{it+1}}$	3.47	0.40	2.10	4.36	3.56	5.13	11.52	-5.42
	(2.46)	(0.29)	(1.24)	(1.59)	(1.48)	(2.07)	(4.82)	(-2.16)
$\overline{Y_{it+1}/K_{it+1}}$	-105.18	109.05	-24.74	31.19	-19.10	27.80	-0.81	6.91
	(-10.05)	(11.25)	(-5.75)	(6.67)	(-4.27)	(8.17)	(-0.29)	(2.54)
$\overline{W_{it+1}/K_{it+1}}$	11.45	-7.58	3.43	3.03	11.51	-2.81	6.54	-0.44
	(6.70)	(-3.68)	(2.58)	(1.35)	(5.32)	(-0.83)	(1.48)	(-0.13)

Table 8: **Correlation between realized stock returns and fundamental returns**

This table reports the contemporaneous correlations between the realized and fundamental returns on the decile portfolios and high-minus-low decile portfolios constructed from the 12 anomaly variables. All correlation coefficients are significant at the 1% level. The sample period is from June 1967 to December 2016 for all anomaly variables except for ROA, RD/M, and Ad/M. The sample starts at December 1972, December 1976, and December 1973 for ROA, RD/M, and Ad/M, respectively, due to data availability.

	L	2	3	4	5	6	7	8	9	H	H-L
BM	0.78	0.77	0.74	0.73	0.72	0.69	0.64	0.68	0.63	0.63	0.53
R11	0.45	0.48	0.63	0.67	0.63	0.67	0.68	0.72	0.67	0.70	0.23
I/A	0.67	0.73	0.72	0.62	0.73	0.73	0.76	0.71	0.73	0.67	0.49
ROE	0.59	0.59	0.63	0.65	0.63	0.60	0.67	0.73	0.74	0.71	0.28
Size	0.57	0.64	0.65	0.68	0.69	0.71	0.71	0.71	0.71	0.77	0.36
Accruals	0.71	0.73	0.70	0.72	0.71	0.68	0.78	0.77	0.69	0.68	0.41
NSI	0.69	0.71	0.73	0.72	0.69	0.67	0.74	0.71	0.76	0.67	0.39
$\Delta PI/A$	0.68	0.72	0.73	0.71	0.73	0.74	0.74	0.77	0.72	0.64	0.52
GP/A	0.64	0.71	0.72	0.67	0.70	0.72	0.73	0.75	0.80	0.76	0.56
ROA	0.62	0.65	0.62	0.62	0.64	0.72	0.70	0.70	0.75	0.79	0.27
RD/M	0.72	0.80	0.71	0.68	0.75	0.67	0.72	0.68	0.68	0.62	0.57
Ad/M	0.72	0.69	0.65	0.67	0.66	0.72	0.66	0.71	0.68	0.58	0.56

Table 9: **Market states and factor premium**

For each month, we categorize the market state as Up if the value-weighted market returns from month t-36 to t-1 are nonnegative and as Down if negative. We report the high-minus-low decile returns averaged across Up and Down states, respectively. r^S denotes the stock returns, and r^F the fundamental returns. The t-values are in parentheses and adjusted for heteroscedasticity and autocorrelations with lags up to 24 months. The sample period is from June 1967 to December 2016 for all anomaly variables except for ROA, RD/M, and Ad/M. The sample starts at December 1972, December 1976, and December 1973 for ROA, RD/M, and Ad/M, respectively, due to data availability.

	BM		R11		I/A		ROE	
Market State	r^S	r^F	r^S	r^F	r^S	r^F	r^S	r^F
Down	14.24	7.68	-12.99	8.77	-12.97	-3.50	-6.67	1.01
	(5.19)	(1.83)	(-1.03)	(3.05)	(-6.07)	(-1.66)	(-1.31)	(0.39)
Up	5.40	-0.82	18.51	12.43	-5.11	-3.08	10.25	5.29
	(1.81)	(-0.52)	(8.30)	(11.74)	(-2.58)	(-2.08)	(4.35)	(5.09)
	Size		Accruals		NSI		ΔPI/A	
Market State	r^S	r^F	r^S	r^F	r^S	r^F	r^S	r^F
Down	-22.71	-8.31	-7.55	3.15	-4.99	-1.70	-14.60	-9.13
	(-3.40)	(-3.56)	(-2.71)	(1.18)	(-1.09)	(-1.18)	(-3.76)	(-3.27)
Up	-1.66	-5.57	-5.23	5.04	-8.13	-3.34	-4.22	-5.24
	(-0.47)	(-5.09)	(-2.65)	(4.87)	(-4.57)	(-3.49)	(-2.11)	(-4.38)
	GP/A		ROA		RD/M		Ad/M	
Market State	r^S	r^F	r^S	r^F	r^S	r^F	r^S	r^F
Down	-4.66	3.02	-7.19	-2.21	9.35	-1.54	17.17	8.28
	(-2.06)	(0.96)	(-1.04)	(-0.55)	(1.87)	(-0.63)	(4.05)	(1.88)
Up	5.39	8.02	8.89	4.90	8.61	6.12	4.07	7.34
	(2.85)	(6.34)	(3.85)	(4.12)	(1.99)	(2.28)	(1.15)	(2.48)

Table 10: **Out-of-sample (OOS) prediction with expanding-window estimates**

Out-of-sample predicted returns are constructed using parameter estimates from the expanding window starting in June 1967. Within the expanding window, parameter estimates vary across industries but stay constant over time. The prediction period is from March 1981 to December 2016. $\alpha_{\text{H-L}}$ is the high-minus-low alpha and $\overline{|\alpha_{\text{D}}|}$ is the average absolute alpha across the 10 decile portfolios of each anomaly. The t-values are in parentheses and adjusted for heteroscedasticity and autocorrelations with lags up to 24 months. Significance at the 1%, 5%, and 10% levels are denoted with three stars, two stars, and one star, respectively. Returns are in percentage per annum.

	$\alpha_{\text{H-L}}$		$\overline{\lvert\alpha_{\text{D}}\rvert}$	
	Baseline	OOS	Baseline	OOS
BM	7.27***	9.63***	1.70	1.90
	(2.86)	(3.25)		
R11	-2.81	6.63	1.84	1.67
	(-0.71)	(1.69)		
I/A	-1.11	-5.27**	1.14	1.16
	(-0.57)	(-2.21)		
ROE	-0.14	0.21	1.28	1.54
	(-0.05)	(0.06)		
Size	2.16	0.88	1.90	1.72
	(0.56)	(0.23)		
Accruals	-8.14***	-7.72***	1.88	2.00
	(-5.11)	(-4.08)		
NSI	-3.87**	-5.31**	0.89	1.13
	(-2.00)	(-2.00)		
ΔPI/A	1.28	-0.73	1.25	1.62
	(0.85)	(-0.38)		
GP/A	-3.78***	-2.49	1.20	1.26
	(-2.91)	(-1.12)		
ROA	-0.74	0.38	1.33	1.70
	(-0.25)	(0.12)		
RD/M	3.38	8.06	1.72	2.42
	(1.45)	(1.76)		
Ad/M	-1.09	1.33	1.01	1.56
	(-0.33)	(0.37)		

Table 11: **Realized returns on decile portfolios formed on the predicted returns, October 1980 - September 2018**

This table reports the realized average excess returns, the CAPM alphas (α_M), the Fama-French three-factor alphas (α_{FF3}), the Carhart four-factor alphas ($\alpha_{Carhart}$), the Fama-French five-factor alphas (α_{FF5}) and the Hou-Xue-Zhang q-factor alphas (α_{q4}) of decile portfolios sorted by predicted returns. L denotes the decile with lowest predicted returns and H denotes the decile with highest predicted returns. The deciles are rebalanced monthly. The predicted returns are calculated from the parameter estimates from expanding windows and forecasts on sales growth and investment-to-capital ratios constructed following Gonçalves, Xue and Zhang (2020). The t-values are in parentheses and adjusted for heteroscedasticity and autocorrelations with lags up to 24 months. Returns are in percentage per month.

	L	2	3	4	5	6	7	8	9	H	H-L
R-rf	0.36	0.69	0.69	0.76	0.62	0.76	0.72	0.77	0.54	0.80	0.45
	(1.26)	(3.16)	(3.54)	(3.70)	(3.14)	(3.81)	(3.39)	(3.63)	(2.27)	(3.20)	(2.45)
α_{CAPM}	-0.36	0.09	0.10	0.15	0.05	0.17	0.09	0.11	-0.17	0.07	0.43
	(-2.40)	(1.04)	(1.47)	(1.79)	(0.72)	(1.94)	(1.12)	(1.31)	(-1.72)	(0.61)	(2.38)
α_{FF3}	-0.46	0.05	0.11	0.17	0.05	0.19	0.11	0.21	-0.14	0.12	0.58
	(-3.02)	(0.54)	(1.56)	(2.21)	(0.77)	(2.22)	(1.34)	(2.54)	(-1.53)	(1.00)	(3.25)
$\alpha_{Carhart}$	-0.39	0.04	0.11	0.16	0.08	0.17	0.09	0.19	-0.09	0.13	0.52
	(-2.59)	(0.46)	(1.64)	(2.06)	(1.12)	(1.93)	(1.11)	(1.93)	(-0.92)	(1.06)	(2.87)
α_{FF5}	-0.41	0.05	0.04	0.07	-0.07	0.02	0.00	0.11	-0.14	0.20	0.61
	(-2.63)	(0.53)	(0.50)	(0.85)	(-1.09)	(0.28)	(-0.05)	(1.30)	(-1.26)	(1.45)	(3.08)
α_{q4}	0.13	0.40	0.38	0.44	0.25	0.35	0.30	0.50	0.28	0.60	0.47
	(0.87)	(4.02)	(4.98)	(4.54)	(3.27)	(4.01)	(3.18)	(4.76)	(2.46)	(3.99)	(2.22)

Internet Appendix: “Fundamental Anomalies” (for Online Publication Only)

Erica X.N. Li, [*] Guoliang Ma, [†] Shujing Wang, [‡] and Cindy Yu [§]

June, 2021

Abstract

This Internet Appendix furnishes supplementary materials for our manuscript “Fundamental Anomalies”.

[*]Department of Finance, Cheung Kong Graduate School of Business, Beijing, China, 100738; Tel: (86)10 85188888 ext. 3075; E-mail: xnli@ckgsb.edu.cn.

[†]Department of Statistics, Iowa State University, Ames, IA 50011; E-mail: glma@iastate.edu.

[‡]Department of Economics and Finance, School of Economics and Management, Tongji University, Shanghai, China 200092; Tel: (86) 21 65982274; E-mail: shujingwang@connect.ust.hk.

[§]Department of Statistics, Iowa State University, Ames, IA 50011; Tel: 1 515 2946885; E-mail: cindyyu@iastate.edu.

A Derivations in the two-capital Model

A.1

Let the production function be $Y_{it} \equiv Y(K_{it}, W_{it}, S_{it}, X_{it}) = X_{it} K_{it}^{\gamma_K} W_{it}^{\gamma_W} S_{it}^{1-\gamma_K-\gamma_W}$, in which S_{it} represents any costly adjustable intermediate inputs and its price p_t^S is taken as given. Y_{it} is of constant returns to scale in physical capital, working capital, and intermediate inputs with their shares given by γ_K, γ_W, and $1 - \gamma_K - \gamma_W$, respectively. The operating profits function solves the static optimization problem:

$$\Pi(K_{it}, W_{it}, X_{it}) = \max_{\{S_{it}\}} X_{it} K_{it}^{\gamma_K} W_{it}^{\gamma_W} S_{it}^{1-\gamma_K-\gamma_W} - p_t^S S_{it}.$$

The first-order condition with respect to S_{it} is $(1 - \gamma_K - \gamma_W) Y_{it}/S_{it} = p_t^S$. Solving for S_{it} yields

$$S_{it} = \left[\frac{(1 - \gamma_K - \gamma_W) X_{it} K_{it}^{\gamma_K} W_{it}^{\gamma_W}}{p_t^S} \right]^{\frac{1}{\gamma_K + \gamma_W}}.$$

Plugging the first-order condition back to $\Pi(K_{it}, W_{it}, X_{it})$ yields $\Pi_{it} = (\gamma_K + \gamma_W) Y_{it}$. Plugging the optimal S_{it} into Y_{it} to rewrite Π_{it} only in terms of K_{it} and W_{it} yields

$$\Pi(K_{it}, W_{it}, X_{it}) = (\gamma_K + \gamma_W) X_{it}^{\frac{1}{\gamma_K + \gamma_W}} \left(\frac{1 - \gamma_K - \gamma_W}{p_t^S} \right)^{\frac{1-\gamma_K-\gamma_W}{\gamma_K+\gamma_W}} K_{it}^{\frac{\gamma_K}{\gamma_K+\gamma_W}} W_{it}^{\frac{\gamma_K}{\gamma_K+\gamma_W}}.$$

As such, $\Pi(K_{it}, W_{it}, X_{it})$ is of constant returns to scale in K_{it} and W_{it}, and their shares, given by $\gamma_K/(\gamma_K + \gamma_W)$ and $\gamma_W/(\gamma_K + \gamma_W)$, respectively, sum to one. In particular, $\partial \Pi_{it}/\partial K_{it} = [\gamma_K/(\gamma_K + \gamma_W)](\Pi_{it}/K_{it}) = \gamma_K Y_{it}/K_{it}$. Similarly, $\partial \Pi_{it}/\partial W_{it} = \gamma_W Y_{it}/W_{it}$.

A.2

The optimization problem can be written as:

$$\begin{aligned} V_{it}(X_{it}, K_{it}, W_{it}, B_{it}) = & \max_{I_{it}, \Delta W_{it}, K_{it+1}, W_{it+1}, B_{it+1}} \{D_{it} + E_t[M_{t+1} V_{it}(X_{it+1}, K_{it+1}, W_{it+1}, B_{it+1})]\} \\ & s.t.\ K_{it+1} = I_{it} + (1 - \delta_{it}) K_{it} \\ & \quad W_{it+1} = \Delta W_{it} + W_{it} \end{aligned} \tag{1}$$

where $D_{it} = (1 - \tau_t)(\Pi_{it} - \Phi_{it}) - I_{it} - \Delta W_{it} + B_{it+1} - r_{it}^B B_{it} + \tau_t \delta_{it} K_{it} + \tau_t (r_{it}^B - 1) B_{it}$. Let q_{it}^K and q_{it}^W

be the Lagrangian multipliers associated with $K_{it+1} = I_{it} + (1 - \delta_{it})K_{it}$ and $W_{it+1} = W_{it} + \Delta W_{it}$, respectively. The Lagrangian function can be written as:

$$L_{it} = D_{it} + E_t[M_{t+1}V_{it}] - q_{it}^K(K_{it+1} - (1 - \delta_{it})K_{it} - I_{it}) - q_{it}^W(W_{it+1} - W_{it} - \Delta W_{it}) \,. \tag{2}$$

Taking the first-order derivatives of L_{it} with respect I_{it}, ΔW_{it}, K_{it}, W_{it+1}, and B_{it+1} to zero and apply the envelop theorem gives the following:

$$q_{it}^K = 1 + (1 - \tau_t)\frac{\partial \Phi_{it}}{\partial I_{it}} \tag{3}$$

$$q_{it}^W = 1 \tag{4}$$

$$q_{it}^K = E_t\left[M_{t+1}\left[(1 - \tau_{t+1})\left(\frac{\partial \Pi_{it+1}}{\partial K_{it+1}} - \frac{\partial \Phi_{it+1}}{\partial K_{it+1}}\right) + \tau_{t+1}\delta_{it+1} + (1 - \delta_{it+1})q_{it+1}^K\right]\right] \tag{5}$$

$$q_{it}^W = E_t\left[M_{t+1}\left[(1 - \tau_{t+1})\frac{\partial \Pi_{it+1}}{\partial W_{it+1}} + q_{it+1}^W\right]\right] \tag{6}$$

$$1 = E_t[M_{t+1}(r_{it+1}^B - (r_{it+1}^B - 1)\tau_{t+1})] = E_t[M_{t+1}r_{it+1}^{Ba}] \tag{7}$$

Combining equations (3) and (5) leads to

$$E_t[M_{t+1}r_{it+1}^K] = 1$$

where

$$r_{it+1}^K = \frac{(1 - \tau_{t+1})\left(\frac{\partial \Pi_{it+1}}{\partial K_{it+1}} - \frac{\partial \Phi_{it+1}}{\partial K_{it+1}}\right) + \tau_{t+1}\delta_{it+1} + (1 - \delta_{it+1})q_{it+1}^K}{q_{it}^K} \,.$$

Similarly, combining equations (4) and (6) leads to

$$E_t[M_{t+1}r_{it+1}^W] = 1$$

where

$$r_{it+1}^W = 1 + (1 - \tau_{t+1})\frac{\partial \Pi_{it+1}}{\partial W_{it+1}} \,.$$

A.3

To prove equation (4), i.e.,

$$w_{it}^K r_{it+1}^K + (1 - w_{it}^K)r_{it+1}^W = w_{it}^B r_{it+1}^{Ba} + (1 - w_{it}^B)r_{it+1}^S \,,$$

we proceed in three steps:

1. Show that firm asset value V_{it}^a can be written as

$$V_{it}^a = P_{it} + B_{it+1} = \mathbb{E}_t \left[\sum_{s=1}^{\infty} M_{t+s} D_{it+s}^a \right]$$

where

$$D_{it+1}^a \equiv (1-\tau_{t+1})(\Pi_{it+1} - \Phi_{it+1}) + \tau_{t+1}\delta_{t+1}K_{it+1} - I_{it+1} - \Delta W_{it+1} \, .$$

Proof:

$$\begin{aligned} V_t^a &= P_t + B_{t+1} = \mathbb{E}_t \left[M_{t+1} \left(D_{t+1} + P_{t+1} \right) \right] + B_{t+1} \\ &= \mathbb{E}_t \Big[M_{t+1} \Big[(1-\tau_{t+1})(\Pi_{t+1} - \Phi_{t+1}) + \tau_{t+1}\delta_{t+1}K_{t+1} - I_{t+1} - \Delta W_{it+1} + B_{t+2} - r_{t+1}^B B_{t+1} \\ &\quad + \tau_{t+1}(r_{t+1}^B - 1)B_{t+1} + P_{t+1} \Big] \Big] + B_{t+1} \end{aligned} \tag{8}$$

The optimality w.r.t. B_{t+1}, equation (7), implies

$$B_{t+1} = \mathbb{E}_t \left[M_{t+1} \left[r_{t+1}^B B_{t+1} - \tau_{t+1}(r_{t+1}^B - 1) B_{t+1} \right] \right] \, .$$

Substitute the above equation into equation (8) and get

$$\begin{aligned} V_t^a &= \mathbb{E}_t \left[M_{t+1} \left[(1-\tau_{t+1})(\Pi_{t+1} - \Phi_{t+1}) + \tau_{t+1}\delta_{t+1}K_{t+1} - I_{t+1} - \Delta W_{it} + P_{t+1} + B_{t+2} \right\} \right] \\ &= \mathbb{E}_t \left[M_{t+1} \left[D_{t+1}^a + P_{t+1} + B_{t+2} \right] \right] \\ &= \mathbb{E}_t \left[\sum_{s=1}^{\infty} M_{t+s} D_{t+s}^a \right] \, . \end{aligned}$$

Q.E.D.

2. Show that $q_{it}^K K_{it+1} + W_{it+1} = P_t + B_{t+1}$.

Proof: Using equations (5) and (6), we have

$$
\begin{aligned}
q_{it}^K K_{it+1} + W_{it+1} &= q_{it}^K K_{it+1} + q_{it}^W W_{it+1} \\
&= E_t \left[M_{t+1} \left[(1-\tau_{t+1}) \left(K_{it+1} \frac{\partial \Pi_{it+1}}{\partial K_{it+1}} - K_{it+1} \frac{\partial \Phi_{it+1}}{\partial K_{it+1}} \right) + \tau_{t+1} \delta_{it+1} K_{it+1} \right.\right. \\
&\quad \left.\left. + (1-\delta_{it+1}) q_{it+1}^K K_{it+1} + (1-\tau_{t+1}) W_{it+1} \frac{\partial \Pi_{it+1}}{\partial W_{it+1}} + q_{it+1}^W W_{it+1} \right]\right] \\
&= E_t \left[M_{t+1} \left[(1-\tau_{t+1}) (\Pi_{it+1} - \Phi_{it+1}) + \tau_{t+1} \delta_{it+1} K_{it+1} + q_{it+1}^K K_{it+2} \right.\right. \\
&\quad \left.\left. - I_{t+1} + W_{it+1} \right]\right] \\
&= E_t \left[M_{t+1} \left[D_{it+1}^a + q_{it+1}^K K_{it+2} + W_{t+2} \right]\right] \\
&= \mathbb{E}_t \left[\sum_{s=1}^{\infty} M_{t+s} D_{t+s}^a \right] \\
&= P_t + B_{t+1} \, ,
\end{aligned}
$$

where the second equality is derived using equations (3), (4), and the following two identities

$$
\begin{aligned}
\Phi_{it} &= I_{it} \partial \Phi_{it} / \partial I_{it} + K_{it} \partial \Phi_{it} / \partial K_{it} \, , \\
\Pi_{it} &= K_{it} \partial \Pi_{it} / \partial K_{it} + W_{it} \partial \Pi_{it} / \partial W_{it} \, .
\end{aligned}
$$

Q.E.D.

3. From the definitions of r_{it+1}^K and r_{it+1}^W and the proof in Step 2, it is straightforward to show that

$$
\begin{aligned}
w_{it}^K r_{it+1}^K + (1-w_{it}^K) r_{it+1}^W &= \frac{q_{it}^K K_{it+1} r_{it+1}^K + W_{it+1} r_{it+1}^W}{q_{it}^K K_{it+1} + W_{it+1}} \\
&= \frac{D_{it+1}^a + q_{it+1}^K K_{it+2} + W_{t+2}}{q_{it}^K K_{it+1} + W_{it+1}} \\
&= \frac{D_{it+1}^a + P_{t+1} + B_{t+2}}{P_t + B_{t+1}} \\
&= w_{it}^B r_{it+1}^{Ba} + (1-w_{it}^B) r_{it+1}^S \, .
\end{aligned}
$$

Q.E.D.

B Bayesian MCMC

Estimation of the model parameters $\boldsymbol{\sigma}$ and latent variables $\boldsymbol{\theta}$ in the baseline model is very difficult due to the high dimensionality. The total dimension of $\boldsymbol{\sigma}$ and $\boldsymbol{\theta}$ that needs to be estimated is 1063 (that is, $\dim(\boldsymbol{\theta}) + \dim(\boldsymbol{\sigma}) = 2 \times 10 \times 53 + 3$), which makes it impractical to use moment based or maximum likelihood methods. We use the Bayesian MCMC method to overcome this estimation difficulty. The main objective of Bayesian analysis is to make inferences about model parameters $\boldsymbol{\sigma}$ and latent variables $\boldsymbol{\theta}$ based on observations: $\boldsymbol{X}, \boldsymbol{r}^S$, and $\boldsymbol{r}^{Ba}$. That is, we need to estimate $\mathcal{P}(\boldsymbol{\sigma}, \boldsymbol{\theta} | \boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba})$, the so called joint posterior distribution of $(\boldsymbol{\sigma}, \boldsymbol{\theta})$ given $(\boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba})$.

According to Bayes' rule, the joint posterior distribution is

$$\begin{aligned} &\mathcal{P}(\boldsymbol{\theta}, \boldsymbol{\sigma} | \boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba}) \\ =&\frac{\mathcal{P}(\boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba})}{\mathcal{P}(\boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba})} \\ \propto&\mathcal{P}(\boldsymbol{r}^S | \boldsymbol{X}, \boldsymbol{r}^{Ba}; \boldsymbol{\theta}, \boldsymbol{\sigma}) \mathcal{P}(\boldsymbol{\theta}, \boldsymbol{\sigma}) \\ =&\mathcal{P}(\boldsymbol{r}^S | \boldsymbol{X}, \boldsymbol{r}^{Ba}; \boldsymbol{\theta}, \boldsymbol{\sigma}) \mathcal{P}(\boldsymbol{\theta} | \boldsymbol{\sigma}) \pi(\boldsymbol{\sigma}), \end{aligned} \tag{9}$$

where $\mathcal{P}(\boldsymbol{r}^S | \boldsymbol{X}, \boldsymbol{r}^{Ba}; \boldsymbol{\theta}, \boldsymbol{\sigma})$ is the conditional distribution of returns given fundamental variables, latent variables and parameters, $\mathcal{P}(\boldsymbol{\theta} | \boldsymbol{\sigma})$ is the conditional distribution of latent variables given parameters $\boldsymbol{\sigma}$, and $\pi(\boldsymbol{\sigma})$ is the joint prior distribution of $\boldsymbol{\sigma}$.

More specifically, we define *weighted scaled asset return* [1]

$$ret_{it+1} = \varpi_{it}^{1/2} \times \left(r_{it+1}^S + \frac{w_{it}^B r_{it+1}^{Ba}}{1 - w_{it}^B} \right) = \frac{\varpi_{it}^{1/2}}{1 - w_{it}^B} r_{it+1}^K + \sigma_r e_{it+1}^r. \tag{10}$$

The newly defined ret can be seen as a function of latent variables, which we denote functionally, for firm i that belongs to industry j at time $t+1$ but to industry j' at time t, as

$$ret_{it+1} \equiv \Lambda_{it+1}\Big(\gamma_{jt+1}, a_{jt+1}, a_{j't}\Big) + \sigma_r e_{it+1}^r, \tag{11}$$

where $\Lambda_{it+1}(\gamma_{jt+1}, a_{jt+1}, a_{j't}) = \frac{\varpi_{it}^{1/2}}{1-w_{it}^B} r_{it+1}^K$ and r_{it+1}^K is defined in (2).

We further assign conjugate inverse gamma distributions as priors for the parameters: $\sigma_r^2 \sim \mathcal{IG}(\kappa_1^r, \kappa_2^r)$, $\sigma_\gamma^2 \sim \mathcal{IG}(\kappa_1^\gamma, \kappa_2^\gamma)$ and $\sigma_a^2 \sim \mathcal{IG}(\kappa_1^a, \kappa_2^a)$. With this variable transformation (ret_{it+1}) and

[1] We use the weighted scaled asset returns instead of stock returns to facilitate discussion of posterior distributions. Another benefit of using the weighted scaled asset returns is that they are homogeneous (of equal variance). When the estimation is finished, we convert ret back to stock returns.

prior specifications, equation (9) can be written in proportion as:

$$\begin{aligned}\mathcal{P}(\boldsymbol{\theta},\boldsymbol{\sigma}|\boldsymbol{X},\boldsymbol{r}^S,\boldsymbol{r}^{Ba}) \propto & \prod_{t=0}^{T-1}\prod_{i=1}^{N_{t+1}} \mathcal{N}\Big(ret_{it+1};\Lambda_{it+1},\sigma_r^2\Big) \\ & \cdot\prod_{t=0}^{T-1}\prod_{j=1}^{N_d}\mathcal{N}\Big(\gamma_{jt+1};\gamma_{jt},\sigma_\gamma^2\Big)\cdot\prod_{t=0}^{T-1}\prod_{j=1}^{N_d}\mathcal{N}\Big(a_{jt+1};a_{jt},\sigma_a^2\Big) \\ & \cdot\mathcal{IG}\Big(\sigma_r^2;\kappa_1^r,\kappa_2^r\Big)\cdot\mathcal{IG}\Big(\sigma_\gamma^2;\kappa_1^\gamma,\kappa_2^\gamma\Big)\cdot\mathcal{IG}\Big(\sigma_a^2;\kappa_1^a,\kappa_2^a\Big),\end{aligned} \tag{12}$$

where N_t is the number of firms at time t, N_d is the number of industries, and T is the length of the observation period.[2]

Given the high dimensionality of parameters and latent variables, it's impossible to draw directly from this joint posterior distribution. However, the Clifford-Hammersley theorem indicates that the joint posterior is equivalent to its complete conditionals. In other words, instead of drawing directly from the 1063-dimensional joint posterior distribution, MCMC draws iteratively from 1063 one-dimensional complete conditionals individually, resulting in legitimate draws from the target joint posterior distribution.

Specifically in our model, the joint posterior distribution of parameters $\boldsymbol{\sigma}$ and latent variables $\boldsymbol{\theta}$ given returns and fundamental variables, the target, is equivalently characterized by its complete conditional posteriors:

$$\mathcal{P}(\boldsymbol{\theta},\boldsymbol{\sigma}|\boldsymbol{X},\boldsymbol{r}^S,\boldsymbol{r}^{Ba}) \Longleftrightarrow \mathcal{P}(\boldsymbol{\theta}|\boldsymbol{\sigma},\boldsymbol{X},\boldsymbol{r}^S,\boldsymbol{r}^{Ba}) \text{ and } \mathcal{P}(\boldsymbol{\sigma}|\boldsymbol{\theta},\boldsymbol{X},\boldsymbol{r}^S,\boldsymbol{r}^{Ba}). \tag{13}$$

Therefore, we simulate the posterior samples of each parameter and latent variable (of $\boldsymbol{\sigma}$ and $\boldsymbol{\theta}$) from the complete conditionals as follows iteratively. Given initial values $(\boldsymbol{\sigma}^{(0)},\boldsymbol{\theta}^{(0)})$, for the current $(g+1)^{th}$ iteration:

- draw $\boldsymbol{\theta}^{(g+1)} \sim \mathcal{P}(\boldsymbol{\theta}|\boldsymbol{\sigma}^{(g)},\boldsymbol{X},\boldsymbol{r}^S,\boldsymbol{r}^{Ba})$;
- draw $\boldsymbol{\sigma}^{(g+1)} \sim \mathcal{P}(\boldsymbol{\sigma}|\boldsymbol{\theta}^{(g+1)},\boldsymbol{X},\boldsymbol{r}^S,\boldsymbol{r}^{Ba})$,

where $\boldsymbol{\sigma}^{(g)}$ is the MCMC draw from the previous iteration.

It is worth noting that there are two advantages of using MCMC algorithms to implement the above iterative procedure: (1) MCMC samplers do not require a closed form of the poste-

[2] ϖ is not present in the formula to be consistent with equation (9). Besides, it has no influence in the following derivation. Also note that in equation (12), no prior distributions for latent variables are assigned because we treat the initial latent variables γ_{j0} and a_{j0} as unknown constants. The driver for the evolvement of latent variables is fully explained by the variances of e^{γ}_{jt+1} and e^{a}_{jt+1} so we do not assign priors to the other latent variables, either.

rior distribution and (2) MCMC samplers need only the conditional posterior up to a constant proportion. In implementing MCMC, Metropolis-Hastings embedded Gibbs sampler is used for estimation in our paper. Whenever the closed form for complete conditional posterior distribution is not directly attainable, we use Metropolis-Hastings algorithm. For a thorough discussion of Gibbs sampling and Metropolis-Hastings, see Robert and Casella (2013).

For time $t+1$, $t \in [0, T-1]$ and industry $j \in [1, N_d]$, let D_{jt+1} be the set of firms that belong to industry j at time $t+1$ and let E_{jt+1} be the set of firms that belong to industry j at time t and exist at time $t+1$. We derive the complete conditional posterior distributions of latent variables γ_{jt+1} and a_{jt+1} and parameters σ_r^2, σ_γ^2 and σ_a^2 (in a proportional form) as follows:

Posterior for γ_{jt+1}:

For the latent variables γ_{jt+1}, the posterior is normal. Let $\mathbb{1}_{condition}$ be the indicator function, i.e., $\mathbb{1} = 1$ when *condition* holds, and otherwise, $\mathbb{1} = 0$:

$$p\left(\gamma_{jt+1} \middle| \left\{\gamma_{jt}\right\}, \left\{a_{jt}\right\}, \sigma_r^2, \sigma_\gamma^2, \sigma_a^2\right) \propto \mathcal{N}\left(\gamma_{jt+1}; \frac{v_1}{u_1}, \frac{1}{u_1}\right), \tag{14}$$

where

$$\begin{aligned}
u_1 :=& \frac{1}{\sigma_r^2} \sum_{i \in D_{jt+1}} A_{it+1}^2 + \frac{1 + \mathbb{1}_{t+1 \notin \{1,T\}}}{\sigma_\gamma^2}, \\
v_1 :=& \frac{1}{\sigma_r^2} \sum_{i \in D_{jt+1}} \varphi_{it+1} A_{it+1} + \frac{1}{\sigma_\gamma^2}(\gamma_{jt}\mathbb{1}_{t \geq 0} + \gamma_{jt+2}\mathbb{1}_{t+2 \leq T}), \\
\varphi_{it+1} :=& ret_{it+1} - \varpi_{it}^{1/2} \times \frac{\tau_{it+1}\delta_{it+1} + \frac{W_{it+1}}{K_{it+1}} + (1 - \delta_{it+1})}{(1 - w_{it}^B)\left[1 + (1 - \tau_{it}) a_{it} \frac{I_{it}}{K_{it}} + \frac{W_{it+1}}{K_{it+1}}\right]} \\
& - \varpi_{it}^{1/2} \times \frac{\frac{1}{2}(1 - \tau_{it+1})\left(\frac{I_{it+1}}{K_{it+1}}\right)^2 + (1 - \delta_{it+1})(1 - \tau_{it+1})\frac{I_{it+1}}{K_{it+1}}}{(1 - w_{it}^B)\left[1 + (1 - \tau_{it}) a_{it} \frac{I_{it}}{K_{it}} + \frac{W_{it+1}}{K_{it+1}}\right]} a_{jt+1}, \\
\text{and } A_{it+1} :=& \varpi_{it}^{1/2} \times \frac{(1 - \tau_{it+1})\frac{Y_{it+1}}{K_{it+1}}}{(1 - w_{it}^B)\left[1 + (1 - \tau_{it}) a_{it} \frac{I_{it}}{K_{it}} + \frac{W_{it+1}}{K_{it+1}}\right]}.
\end{aligned}$$

Posterior for a_{jt+1}:

For adjustment costs parameters a_{jt+1}, there are no clear closed form posterior distributions. We implement Metropolis-Hastings. It is a propose-reject method which first proposes a *candidate* draw and then decide whether a *jump* is made from the current state to the proposed. Depending on the difference of proposal distributions, there are many variations under this generic heading. In our paper, the candidate is chosen in a manner that exploits as much information from the posterior distributions as possible.

We first consider the posterior of a_{jt+1} although it is not clear what distribution it follows:

$$p\left(a_{jt+1} \middle| \{\gamma_{jt+1}\}, \{a_{jt+1}\}, \sigma_r^2, \sigma_\gamma^2, \sigma_a^2\right) \propto \prod_{t=0}^{T-1} \prod_{i=1}^{N} \mathcal{N}\left(ret_{it+1}; \Lambda_{it+1}, \sigma_r^2\right) \prod_{t=0}^{T-1} \mathcal{N}\left(a_{jt+1}; a_{jt}, \sigma_a^2\right). \quad (15)$$

We propose from

$$\mathcal{N}\left(a_{jt+1}; \frac{v_2}{u_2}, \frac{1}{u_2}\right) \quad (16)$$

where

$$\begin{aligned}
u_2 :=& \frac{1}{\sigma_r^2} \sum_{i \in D_{jt+1}} B_{it+1}^2 + \frac{1 + \mathbb{1}_{t+1 \notin \{1,T\}}}{\sigma_a^2}, \\
v_2 :=& \frac{1}{\sigma_r^2} \sum_{i \in D_{jt+1}} \psi_{it+1} B_{it+1} + \frac{1}{\sigma_a^2} \left(a_{jt} \mathbb{1}_{t \geq 0} + a_{jt+2} \mathbb{1}_{t+2 \leq T}\right), \\
\psi_{it+1} :=& ret_{it+1} - \varpi_{it}^{1/2} \times \frac{\tau_{it+1}\delta_{it+1} + \frac{W_{it+1}}{K_{it+1}} + (1 - \delta_{it+1})}{(1 - w_{it}^B)\left[1 + (1 - \tau_{it}) a_{it} \frac{I_{it}}{K_{it}} + \frac{W_{it+1}}{K_{it+1}}\right]}, \\
& - \varpi_{it}^{1/2} \times \frac{(1 - \tau_{it+1}) \frac{Y_{it+1}}{K_{it+1}}}{(1 - w_{it}^B)\left[1 + (1 - \tau_{it}) a_{it} \frac{I_{it}}{K_{it}} + \frac{W_{it+1}}{K_{it+1}}\right]} \gamma_{jt+1}, \\
\text{and } B_{it+1} :=& \varpi_{it}^{1/2} \times \frac{\frac{1}{2}(1 - \tau_{it+1}) \left(\frac{I_{it+1}}{K_{it+1}}\right)^2 + (1 - \delta_{it+1})(1 - \tau_{it+1}) \frac{I_{it+1}}{K_{it+1}}}{(1 - w_{it}^B)\left[1 + (1 - \tau_{it}) a_{it} \frac{I_{it}}{K_{it}} + \frac{W_{it+1}}{K_{it+1}}\right]}.
\end{aligned}$$

To decide whether to accept the candidate, let [3]

$$\begin{aligned}
\pi(x) = & \prod_{i \in D_{jt+1}} \mathcal{N}\left(ret_{it+1}; \Lambda_{it+1}\left(\gamma_{jt+1}, x, a_{it}\right), \sigma_r^2\right) \cdot \prod_{i \in E_{jt+2}} \mathcal{N}\left(r_{it+2}; \Lambda_{it+2}\left(\gamma_{it+2}, a_{it+2}, x\right), \sigma_r^2\right) \\
& \cdot \mathcal{N}\left(x; a_{jt+2}, \sigma_a^2\right) \cdot \mathcal{N}\left(x; a_{jt}, \sigma_a^2\right).
\end{aligned}$$

[3] Here a slight abuse of notation for generality is that we change $a_{j't}$ to a_{it} to indicate that firms belong to different industries at time t. Similarly, we use γ_{it} to indicate firms belong to different industries at time t.

The acceptance rate α is then

$$\alpha = \frac{\pi(a_{jt+1}^{prop})}{\pi(a_{jt+1})} \cdot \frac{\mathcal{N}\left(a_{jt+1}; \frac{v_2}{u_2}, \frac{1}{u_2}\right)}{\mathcal{N}\left(a_{jt+1}^{prop}; \frac{v_2}{u_2}, \frac{1}{u_2}\right)} = \frac{\prod_{i \in E_{jt+2}} \mathcal{N}\left(r_{it+2}; \Lambda_{it+2}\left(\gamma_{it+2}, a_{it+2}, a_{jt+1}^{prop}\right), \sigma_r^2\right)}{\prod_{i \in E_{jt+2}} \mathcal{N}\left(r_{it+2}; \Lambda_{it+2}\left(\gamma_{it+2}, a_{it+2}, a_{jt+1}^{(g-1)}\right), \sigma_r^2\right)}.$$

Posteriors for σ_r^2, σ_γ^2 and σ_a^2:

The posterior distributions for parameters σ_r^2, σ_γ^2 and σ_a^2 are:

$$p\left(\sigma_r^2 \middle| \left\{\gamma_{jt}\right\}, \left\{a_{jt}\right\}, \sigma_\gamma^2, \sigma_a^2\right) \sim \mathcal{IG}\left(\kappa_1^r + \frac{\sum_{t=0}^{T-1} N_{t+1}}{2}, \kappa_2^r + \frac{1}{2}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\left(ret_{it+1} - \Lambda_{it+1}\right)^2\right). \tag{17}$$

$$p\left(\sigma_\gamma^2 \middle| \left\{\gamma_{jt}\right\}, \left\{a_{jt}\right\}, \sigma_r^2, \sigma_a^2\right) \sim \mathcal{IG}\left(\kappa_1^\gamma + \frac{N_d T}{2}, \kappa_2^\gamma + \frac{1}{2}\sum_{t=0}^{T-1}\sum_{k=1}^{N_d}\left(\gamma_{jt+1} - \gamma_{jt}\right)^2\right). \tag{18}$$

$$p\left(\sigma_a^2 \middle| \left\{\gamma_{jt}\right\}, \left\{a_{jt}\right\}, \sigma_r^2, \sigma_\gamma^2\right) \sim \mathcal{IG}\left(\kappa_1^a + \frac{N_d T}{2}, \kappa_2^a + \frac{1}{2}\sum_{t=0}^{T-1}\sum_{k=1}^{N_d}\left(a_{jt+1} - a_{jt}\right)^2\right). \tag{19}$$

where κ_1^r, κ_2^r, κ_1^γ, κ_2^γ, κ_1^a and κ_2^a are prior parameters for prior inverse gamma distributions.

In each MCMC iteration, a systematic scan is used, i.e., we sample by a pre-specified order the parameters/latent variables from the above posterior distribution conditional on the most updated information. After all the parameters and latent variables are updated, a new iteration is started. We run 20,000 iterations in total and use the last 5,000 iterations to obtain posterior means and 95% credible intervals.

C Definition of Sorting Variables

BM (Davis, Fama and French, 2000) Book-to-market equity ratio, defined as the book value of equity for fiscal year end in the previous calendar year $t-1$ divided by the market value of equity at the end of December of the previous calendar year $t-1$. We measure book equity as stockholders' book equity, plus balance sheet deferred taxes and investment tax credit (item TXDITC or the sum of item TXDB and item ITCB) if available, minus the book value of preferred stock. Stockholders' equity is the value reported by Compustat (item SEQ) if it is available. If not, we measure stockholders' equity as the book value of common equity (item CEQ) plus the par value of preferred stock (item PSTK), or the book value of assets (item AT) minus total liabilities (item LT). Depending on availability, we use redemption (item PSTKRV), liquidating (item PSTKL), or par value (item PSTK) for the book value of preferred stock.

R11 (Fama and French, 1996; Carhart, 1997) Prior 11-month returns from month t-12 to t-2.

I/A (Cooper, Gulen and Schill (2008)) We measure I/A as change in total assets (Compustat annual item AT) scaled by lagged total assets. At the end of June of each year t, we use NYSE breakpoints to split stocks into deciles based on I/A for the fiscal year ending in calendar year t-1 and calculate monthly value-weighted decile returns from July of year t to June of t+1.

ROE (Hou, Xue and Zhang, 2020) ROE is income before extraordinary items (Compustat quarterly item IBQ) divided by 1-quarter-lagged book equity. From 1972 onward, quarterly book equity is shareholders' equity, plus balance sheet deferred taxes and investment tax credit (item TXDITCQ) if available, minus the book value of preferred stock (item PSTKQ). Depending on availability, we use stockholders' equity (item SEQQ), or common equity (item CEQQ) plus the book value of preferred stock (item PSTKQ), or total assets (item ATQ) minus total liabilities (item LTQ) in that order as shareholders' equity. Prior to 1972, we expand the sample coverage by using book equity from Compustat annual files and imputing quarterly book equity with clean surplus accounting.

At the beginning of each month t, we sort stocks into deciles on their most recent ROE. Before 1972, we use the most recent ROE computed with quarterly earnings from the fiscal quarter ending at least four months ago. From 1972 onward, we use ROE computed with quarterly earnings from the most recent quarterly earnings announcement date (item RDQ). For a firm to enter the portfolio formation, we require the end of the fiscal quarter corresponding to its most recent ROE to be within six months prior to the portfolio formation and its earnings announcement date to be after the corresponding fiscal quarter end. Monthly decile returns are calculated for the current month t, and the deciles are rebalanced at the beginning of month t + 1.

Size (Fama and French, 1992) Size is price times shares outstanding from CRSP. At the end of June of each year t, we use NYSE breakpoints to sort stocks into deciles based on the June-end $Size$, and calculate monthly value-weighted decile returns from July of year t to June of t+1.

Accruals (Sloan, 1996) We measure Accruals as $\frac{\Delta ACT - \Delta CHE - \Delta LCT + \Delta DLC + \Delta TXP - \Delta DP}{(AT + AT_{-1})/2}$, where ΔACT is the annual change in total current assets, ΔCHE is the annual change in total cash and short-term investments, ΔLCT is the annual change in current liabilities, ΔDLC is the annual change in debt in current liabilities, ΔTXP is the annual change in income taxes payable, ΔDP is the annual change in depreciation and amortization, and $(AT + AT_{-1})/2$ is average total assets over the last two years. At the end of June of each year t, we use NYSE breakpoints to sort all stocks into deciles based on Accruals for the fiscal year ending in calendar year t-1, and calculate monthly value-weighted decile returns from July of year t to June of t+1.

NSI (Fama and French, 2008) We measure net stock issues (NSI) as the natural log of the ratio of the split-adjusted shares outstanding scaled by lagged split-adjusted shares outstanding. The split-adjusted shares outstanding is shares outstanding (Compustat annual item CSHO) times the adjustment factor (item AJEX). At the end of June of each year t, we use NYSE breakpoints to sort all stocks into deciles based on NSI for the fiscal year ending in calendar year t-1, and calculate monthly value-weighted decile returns from July of year t to June of t+1.

ΔPI/A (Lyandres, Sun and Zhang, 2008) We measure ΔPI/A as changes in gross property, plant, and equipment (Compustat annual item PPEGT) plus changes in inventory (item INVT) scaled by lagged total assets (item AT). At the end of June of each year t, we use NYSE breakpoints to sort stocks into deciles based on ΔPI/A for the fiscal year ending in calendar year t-1, and calculate monthly value-weighted decile returns from July of year t to June of t+1.

GP/A (Novy-Marx, 2013) We measure GP/A as total revenue (Compustat annual item REVT) minus cost of goods sold (item COGS) divided by current total assets (item AT). At the end of June of each year t, we use NYSE breakpoints to sort stocks into deciles based on GP/A for the fiscal year ending in calendar year t-1, and calculate monthly value-weighted decile returns from

July of year t to June of t+1.

ROA (Balakrishnan, Bartov and Faurel, 2010; Hou, Xue and Zhang, 2020) We measure ROA as income before extraordinary items (Compustat quarterly item IBQ) divided by 1-quarter-lagged total assets (item ATQ). At the beginning of each month t, we use NYSE breakpoints to sort all stocks into deciles based on ROA computed with the most recently announced quarterly earnings. Monthly value-weighted decile returns are calculated for month t, and the deciles are rebalanced at the beginning of t+1. For a firm to enter the portfolio formation, we require the end of the fiscal quarter that corresponds to the most recently announced earnings to be within six months prior to the portfolio formation to exclude stale earnings information.

RD/M (Chan, Lakonishok and Sougiannis, 2001; Hou, Xue and Zhang, 2020) We measure RD/M as R&D expenses (Compustat annual item XRD) divided by market equity. At the end of June of each year t, we use NYSE breakpoints to split stocks into deciles based on RD/M, which is R&D expenses for the fiscal year ending in calendar year t-1 divided by the market equity at the end of December of t-1, and calculate monthly value-weighted decile returns from July of year t to June of t+1. We keep only firms with positive R&D expenses. Because the accounting treatment of R&D expenses was standardized in 1975, the RD/M decile returns start in July 1976.

Ad/M (Chan, Lakonishok and Sougiannis, 2001; Hou, Xue and Zhang, 2020) We measure Ad/M as advertising expenses (Compustat annual item XAD) divided by market equity. At the end of June of each year t, we use NYSE breakpoints to split stocks into deciles based on Ad/M, which is advertising expenses for the fiscal year ending in calendar year t-1 divided by the market equity at the end of December of t-1, and calculate monthly value-weighted decile returns from July of year t to June of t+1. We keep only firms with positive advertising expenses. Because sufficient XAD data start in 1972, the Ad/M decile returns start in July 1973.

D Simulation study

The simulation study is implemented in four steps.

1. We select a balanced panel of $1,052$ firms from seven industries between 1991 to 2005, all of which have no missing variables needed to construct fundamental returns during the 15-year period. These seven industries are Consumer nondurables, Manufacturing, Business equipment, Wholesale, Healthcare, Utilities, and Others. Note that our methodology does not require a balanced panel. The only requirement is to have financial and accounting information of a firm for at least three consecutive years from year $t-1$ to $t+1$, which is required to compute the fundamental return at year t. The choice of a balanced panel is for simplicity.

2. We generate the time series of the latent variables by simulating random walk processes according to equation (7) for each of the seven industries, denoted as $\boldsymbol{\theta}$. The standard deviations of these random walk processes, σ_γ and σ_a, are chosen to be 0.1 and 0.3, which are close to the estimated magnitudes. The time 0 value of the technology parameter γ_{j0} of industry j is randomly chosen from a logic-transformed-normal distribution to ensure that the technology parameter falls into the range between 0 and 1. The time 0 value of the adjustment cost parameter a_{j0} is drawn from a normal distribution with mean and standard deviation being 5 and 0.3. The mean of the distribution is close to the estimates in Gonçalves, Xue and Zhang (2020).

3. We generate stock returns for firm i in the selected subgroup based on equation (8) added with white noises, i.e.,

$$r^S_{it+1} = f\left(\boldsymbol{X}_{it}, \boldsymbol{X}_{it+1} | \boldsymbol{\theta}_t, \boldsymbol{\theta}_{t+1}\right) + \varpi^{-1/2}_{it+1} \sigma_r e^r_{it+1}\,,$$

   where $\boldsymbol{X}_{it}$ is the accounting information of firm i at time t, $\varpi^{-1/2}_{it}$ is computed based on equation (9) using firm i's financial information, σ_r is set to be 5%, and e^r_{it+1} follows the standard normal distribution so that the volatility of the simulated returns r^S_{it+1} is comparable to the volatility of the corresponding observed stock returns. The simulated sample of firm-level stock returns has mean, standard deviation, and skewness of 0.66, 1.09, and 7.63, compared with 0.21, 0.71, and 9.78 in the data for the sample.

4. Using the Bayesian MCMC method in Section 4, we draw from the posterior distributions of the latent variables given the financial and accounting information X and the simulated

stock return r^S of this subgroup of firms. The initial guesses, $\boldsymbol{\theta}_{jt+1}^{(0)}$, for industry j at time $t+1$ are the minimizers of the residual sum of squares (RSS) of firm-level stock returns of firms in industry j at time $t+1$ given $\boldsymbol{\theta}_t^{(0)}$, defined as

$$\boldsymbol{\theta}_{jt+1}^{(0)} = \operatorname{argmin} \sum_{i=1}^{N_j} \left[f\left(\boldsymbol{X}_{it}, \boldsymbol{X}_{it+1} | \boldsymbol{\theta}_t^{(0)}, \boldsymbol{\theta}_{t+1} \right) - r_{it+1}^S \right]^2 .$$

Assuming that $\boldsymbol{\theta}_{j0}^{(0)} = \boldsymbol{\theta}_{j1}^{(0)}$, the initial guesses for $t = 1, \cdots, T$ can be estimated sequentially. We have tried constant initial guesses and the estimation converges to the same posterior distributions. It shows that our method is robust to the choice of initials.

E Comparison of Bayesian and NLS via simulation studies

Frequentist methods, such as Nonlinear Least Squares (NLS), can also be used to match firm-level stock returns. Under NLS, parameter values are chosen to minimize the sum of squared estimates of errors sequentially as follows. For parameter θ_{jt+1}, for $j = 1, \ldots, N_d$ and $t = 1, \ldots, T-1$:[4]

$$\hat{\theta}_{jt+1}^{NLS} = \arg\min_{\theta_{jt+1}} \sum_{i=1}^{N_{jt+1}} \varpi_{it} \left[f\left(X_{it}, X_{it+1} | \hat{\theta}_{jt}^{NLS}, \theta_{jt+1} \right) - r_{it+1}^{S} \right]^2 , \tag{20}$$

where N_d is the number of industries, N_{jt+1} is the number of firms in industry j at time $t+1$, $\hat{\theta}_{jt}^{NLS}$ is the estimated parameters for industry j at t, and ϖ_{it-1}, which is proportional to the market equity V_{it-1} as defined in equation (9), is used to be consistent with our Bayesian MCMC estimates. For $t = 0$, we assume that $\theta_{j0} = \theta_{j1}$ so the NLS estimate is

$$\hat{\theta}_{j1}^{NLS} = \arg\min_{\theta_{j1}} \sum_{i=1}^{N_{j1}} \varpi_{i0} \left[f\left(X_{i0}, X_{i1} | \theta_{j1}, \theta_{j1} \right) - r_{i1}^{S} \right]^2 , \text{ for } j = 1, \ldots, N_d.$$

In this section, we use simulation studies to examine the advantages of Bayesian MCMC over NLS.

Figure 1 plots the true values (in red solid lines), the NLS estimates (in green lines with triangle markers), and the Bayesian posterior means (in blue dashed lines) and the associated 95% credible intervals (in shaded areas) of the model parameters $\boldsymbol{\theta}$ estimated from the simulated data under the specification with industry specific and time varying parameters. Credible interval is frequently used in Bayesian framework. It refers to the interval wherein a random variable (here a parameter) falls with the specified probability. It is an interval in the domain of a posterior

[4]The NLS estimates with industry variations only, time variations only, and the estimates with constant values are obtained, respectively, as follows:

$$\hat{\theta}_{j}^{NLS} = \arg\min_{\theta_j} \sum_{t=0}^{T-1} \sum_{i=1}^{N_{jt}} \varpi_{it} \left[f\left(X_{it}, X_{it+1} | \theta_j, \theta_j \right) - r_{it+1}^{S} \right]^2 , \text{ for } j = 1, \ldots, N_d ,$$

$$\hat{\theta}_{t+1}^{NLS} = \arg\min_{\theta_{t+1}} \sum_{j=1}^{N_d} \sum_{i=1}^{N_{jt}} \varpi_{it} \left[f\left(X_{it}, X_{it+1} | \hat{\theta}_{t}^{NLS}, \theta_{t+1} \right) - r_{it+1}^{S} \right]^2 , \text{ for } t = 0, \ldots, T-1 ,$$

and

$$\hat{\theta}^{NLS} = \arg\min_{\theta} \sum_{t=0}^{T-1} \sum_{j=1}^{N_d} \sum_{i=1}^{N_{jt}} \varpi_{it} \left[f\left(X_{it}, X_{it+1} | \theta, \theta \right) - r_{it+1}^{S} \right]^2 .$$

distribution of a parameter. Because we assume parameters to be random variables in Bayesian framework, we can calculate the probability that a parameter locates in a given interval based on its posterior distribution. Notationally, let I_p be the posterior credible interval of $\boldsymbol{\theta}$ that satisfies $P(\boldsymbol{\theta} \in I_p | \boldsymbol{X}, \boldsymbol{r^S}, \boldsymbol{r^{Ba}}) = p$, where p is the probability.[5]

Figure 1 shows that the NLS estimates are often very far from the true. On the contrary, the true values of the model parameters are almost always confined in the narrow credible intervals of the Bayesian MCMC posterior distributions. The posterior means imply small relative mean absolute errors (m.a.e.) of 3.59% and 3.37% on average across industries for γ and for a, respectively. Similar results are found in the specification with time variation in parameter values only and the results are plotted in Figure A.3 in the Internet Appendix.

Table A.1 reports the true values, the NLS estimates, and the Bayesian posterior means and associated credible intervals of the model parameters under the specification with only industry variation in Panel A and under the specification with constant parameter values in Panel B. As under the specifications with time-varying parameter values, the 95% credible intervals from the Bayesian estimation always cover the corresponding true values. Bayesian estimates again have smaller estimation errors in general, although the differences between the NLS and Bayesian estimates are smaller when parameters are not time varying. For example, with constant parameter values, the Bayesian posterior means of γ and a are 0.1500 and 0.1300, which are identical to the true values (up to the fourth digit), while the the corresponding NLS estimates are 0.1501 and 0.1280.

Bayesian MCMC estimation approach is fundamentally different from NLS and GMM. Bayesian MCMC is able to extract more information from the data than these two frequentist methods. In essence, these frequentist methods choose model parameters to match a given set of moments. In the case of NLS, the matching moments are

$$\sum_{i=1}^{N_{jt+1}} \varpi_{it} \frac{\partial f\left(X_{it}, X_{it+1} | \hat{\theta}_{jt}^{NLS}, \theta_{jt+1}\right)}{\partial \theta_{jt+1}} \left[f\left(X_{it}, X_{it+1} | \hat{\theta}_{jt}^{NLS}, \theta_{jt+1}\right) - r_{it+1}^{S} \right] = \mathbf{0}\,,$$

for $j = 1, \ldots, N_d$ and $t = 1, \ldots, T-1$ and

$$\sum_{i=1}^{N_{j1}} \varpi_{i0} \frac{\partial f(X_{i0}, X_{i1} | \theta_{j1}, \theta_{j1})}{\partial \theta_{j1}} \left[f(X_{i0}, X_{i1} | \theta_{1j}, \theta_{1j}) - r_{i1}^{S} \right] = \mathbf{0}, \text{ for } j = 1, \ldots, N_d,$$

[5]For illustration purpose, Figure 1 only includes the results of three industries: Consumer Nondurables, Manufacturing, and Business Equipment. The results of the other industries used in the simulation study are plotted in Figure A.2 in the Internet Appendix.

assuming $\theta_{j0} = \theta_{j1}$, where $\mathbf{0}$ are vectors of zeros of corresponding dimensions.[6] In the case of GMM used in Liu, Whited and Zhang (2009) and Gonçalves, Xue and Zhang (2020) among others, the matching moments are the average returns of the testing portfolios. By matching moments only, these frequentist methods fail to capture the detailed information in each firm-year observation, which, on the contrary, is utilized in Bayesian MCMC. The posterior likelihood in equation (10) captures the entire posterior distributions of the firm-level stock returns.

Another critical difference is that the posterior of any specific industry-time parameter, θ_{jt}, utilizes the information of not only industry j at time t, but the entire data sample. This feature is critical for the superior performance of Bayesian MCMC relative to NLS in our simulation studies when true parameter values are time varying. The reasons are as follows. First, the random walk process imposed on parameters in equation (7) connects information across different points in time. Second, physical adjustment costs parameter a_{jt} enters into the probability distributions of both r_{it} and r_{it+1} of firm i in industry j as shown in equations (2) and (10), which also connects information in returns across time. Third, due to entry and exit, the probability distribution of stock return r_{it+1} can also connect information in returns across industries if firm i switches from industry j to k at time $t+1$. Consequently, the identification of any specific latent variables a_{jt} and γ_{jt} utilizes the information of the entire data sample.[7]

[6]The matching moments under the specification with industry variations only, time variations only, and with constant parameter values are given, respectively, by:

$$\sum_{t=0}^{T-1} \sum_{i=1}^{N_{jt}} \varpi_{it} \frac{\partial f\left(X_{it}, X_{it+1} | \theta_j, \theta_j\right)}{\partial \theta_j} \left[f\left(X_{it}, X_{it+1} | \theta_j, \theta_j\right) - r_{it+1}^S \right] = \mathbf{0}\,, \text{ for } j = 1, \ldots, N_d\,,$$

$$\sum_{j=1}^{N_d} \sum_{i=1}^{N_{jt}} \varpi_{it} \frac{\partial f\left(X_{it}, X_{it+1} | \hat{\theta}_t^{NLS}, \theta_{t+1}\right)}{\partial \theta_{t+1}} \left[f\left(X_{it}, X_{it+1} | \hat{\theta}_t^{NLS}, \theta_{t+1}\right) - r_{it+1}^S \right] = \mathbf{0}\,, \text{ for } t = 0, \ldots, T-1\,,$$

and

$$\sum_{t=0}^{T-1} \sum_{j=1}^{N_d} \sum_{i=1}^{N_{jt}} \varpi_{it} \frac{\partial f\left(X_{it}, X_{it+1} | \theta, \theta\right)}{\partial \theta} \left[f\left(X_{it}, X_{it+1} | \theta, \theta\right) - r_{it+1}^S \right] = \mathbf{0}\,.$$

[7]Even though γ_{jt} enters into the probability distribution of stock return r_{it} only for firm i in industry j at time t, γ_{jt} is identified together with a_{jt}. Therefore, the value of γ_{jt} also reflects the information of the entire data sample.

F Do industry and time variations in parameter estimates make economic sense?

In the baseline estimation, we allow the technology parameter in the production function, γ_{jt}, and the physical adjustment costs parameter, a_{jt}, to be industry specific and time varying. In this section, we explain the economic forces that drive the variations in parameter values across industries and times, and then test whether the estimated values are consistent with these underlying economics.

Within the model framework, γ_{jt} reflects industry j's profit margin as the model implies $\Pi_{it} = \gamma_{jt} Y_{it}$ for any firm i in industry j at time t, where Π_{it} and Y_{it} are the profits and sales, respectively. Therefore, variations in γ_{jt} can be driven by both technology changes and changes in market demand, the latter of which can be caused by fluctuations in consumer taste, economic conditions, market competitiveness, etc.

Equation (5) implies that Tobin's q of firm i in industry j at time t follows $q_{it} = 1 + a_{jt}(1 - \tau_t) I_{it}/K_{it} \times K_{it+1}/(K_{it+1} + W_{it+1})$. Therefore, the magnitude of a_{jt} reflects both the marginal costs and marginal benefits of investing one dollar in physical capital and has a positive relation with Tobin's q. Consequently, variations in a_{jt} can be driven by changes in production technology, price of capital goods, which is cyclical (Eisfeldt and Rampini, 2006), and opportunity costs in terms of lost output, which vary with procyclical capacity utilization. Lastly, entry and exit in an industry can also lead to changes in the estimated parameter values of this industry at a given fiscal year t. It thus makes economic sense to allow γ_{jt} and a_{jt} to be industry specific and time varying.

Even though γ and a are estimated to match firm-level stock returns, their values should be consistent with the underlying economics that they reflect. The above discussion implies that variations in the estimated values of γ should positively correlated with the variations in operating profits-to-sales ratio across industries and times. Specifically, the following regression should yield a positive and significant coefficient on γ_{jt}:

$$\overline{\Pi/Y}_{jt} = c_\gamma + b_\gamma\, \gamma_{jt} + \epsilon^\gamma_{jt}\,,$$

where the dependent variable is the value-weighted operating profits-to-sales ratio for industry j at time t, defined as $\overline{\Pi/Y}_{jt} \equiv \sum_{i=1}^{N_{jt}} \varpi_{it-1}(\Pi_{it}/Y_{it})$, the independent variable is the estimated value of γ for the same industry and time, and c_γ and b_γ are regression coefficients. Operating profits is measured by operating income before depreciation (item OIBDP). The weight ϖ_{it-1} is proportional to the market equity V_{it-1} as defined in equation (9), and is used to be consistent with the fact that the variance of estimation error is assumed to be proportional to the inverse of ϖ_{it-1} in equation (8). If our model is the true model, we would expect c_γ and b_γ to be zero and one, respectively.

Similarly, variations in the estimated values of a should positively correlates with the variations in average Tobin's q across industries and times. Specifically, the following regression should yield a positive and significant coefficient on a_{jt}:

$$\bar{q}_{jt} = c_a + b_a\, a_{jt} \times \overline{I/K}_{jt} + \epsilon^a_{jt}\,.$$

where the dependent variable is the value-weighted Tobin's q for industry j at time t, defined as $\sum_{i=1}^{N_{jt}} \varpi_{it-1} q_{it}$, the independent variable is an interaction term between the estimated value of a for the same industry and time and the weighted industry average investment rate, defined as $\overline{I/K}_{jt} \equiv \sum_{i=1}^{N_{jt}} \varpi_{it-1}(1-\tau_t) \times K_{it+1}/(K_{it+1}+W_{it+1}) \times (I_{it}/K_{it})$, and c_a and b_a are regression coefficients. If our model is the true model, we would expect c_a and b_γ to be one.

Table A.3 shows that b_γ is 0.19 with t-stat being 6.10 and b_a is 1.97 with t-stat being 3.33, both of which are positive and highly significant. The constant term c_γ is 0.18 (t=34.83) and c_a is 2.16 (t = 34.49). On one hand, these results confirm that our estimation generates economically sensible values for the structure parameters γ and a. On the other hand, values of the regression coefficients are different from the values if our model is the true model. Possible reasons behind this mismatch include measure errors in operating profits and Tobin's q, and model misspecification error. The existence of model misspecification error is not surprising, as we intentionally choose such a simple model. We show in Section 6.2 that adding intangible assets to the model is a promising direction for future research.

G Comparative statistics

Given equation (6), it is straightforward to show that:

$$\frac{\partial r^F_{it+1}}{\partial(I_{it}/K_{it})} = \frac{-(1-\tau_t)r^{Fwacc}_{it+1}\,a_{jt}}{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)+\frac{W_{it+1}}{K_{it+1}}\right]} < 0$$

$$\frac{\partial r^F_{it+1}}{\partial(I_{it+1}/K_{it+1})} = \frac{(1-\tau_{t+1})\left(1+\frac{I_{it+1}}{K_{it+1}}-\delta_{it+1}\right)a_{jt+1}}{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)+\frac{W_{it+1}}{K_{it+1}}\right]} > 0$$

$$\frac{\partial r^F_{it+1}}{\partial(Y_{it+1}/K_{it+1})} = \frac{(1-\tau_{t+1})\gamma_{jt+1}}{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)+\frac{W_{it+1}}{K_{it+1}}\right]} > 0$$

$$\frac{\partial r^F_{it+1}}{\partial(W_{it+1}/K_{it+1})} = \frac{1-r^{Fwacc}_{it+1}}{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)+\frac{W_{it+1}}{K_{it+1}}\right]},$$

where r^{Fwacc}_{it+1} is firm i's fundamental weighted average cost of capital, defined as $r^{Fwacc}_{it+1} \equiv (1-w^B_{it})r^F_{it+1} + w^B_{it} r^{Ba}_{it+1}$.

Since the denominator of all the above derivatives is positive, the signs of these partial derivatives are determined by the numerator. The signs of the derivative of r^F_{it+1} w.r.t. Y_{it+1}/K_{it+1}, I_{it}/K_{it}, and I_{it+1}/K_{it+1} are clearly negative, positive, and positive, respectively. Fundamental return decreases with W_{it+1}/K_{it+1} if $r^{Fwacc}_{it+1} > 1$, and vice versa. Since cost of capital is in general positive, i.e., $r^{Fwacc}_{it+1} > 1$, we expect the relation between r^F_{it+1} and W_{it+1}/K_{it+1} to be mostly negative.

It is straightforward to see that the magnitude of $\frac{\partial r^F_{it+1}}{\partial(Y_{it+1}/K_{it+1})}$ increases with the value of γ, that is, a unit differences in Y_{it+1}/K_{it+1} leads to larger fundamental return spread when the magnitude of γ is larger. The relation of the other three derivatives with model parameters a and γ depends on the values of firm characteristics such as investment rate and sales-to-capital ratio, and thus varies across firms in general.

For illustration purpose, we derive the relation of the other three derivatives with constant model parameters a and γ at the steady state where firm characteristics equal the sample averages, i.e., $I_{it}/K_{it} = i_k$, $Y_{it+1}/K_{it+1} = y_k$, $W_{it+1}/K_{it+1} = w_k$, $w^B_{it} = w^B$, $\tau_t = \tau$, and $\delta_{it} = \delta$. We can show that

$$\frac{\partial r^{Fwacc}_{t+1}}{\partial\, a} = -\frac{(1-\tau)i_k[(w_k+\tau)\delta + (1-\tau)\gamma y_k - (1+w_k)i_k/2]}{[1+w_k+(1-\tau)i_k a]^2}$$

and

$$\frac{\partial}{\partial\, a}\left|\frac{\partial r^F_{t+1}}{\partial (I_t/K_t)}\right| = \left(\frac{1-\tau}{1-w^B}\right)\Big\{[(1+w_k)-ai_k(1-\tau)]\gamma y_k + (1-\tau)i_k a[(1-\tau)(1-\delta)+(w_k+\tau)(1-2\delta)]$$
$$+ (1+w_k)[\tau\delta+(1-\delta)+w_k]\Big\}\Big/\,[1+w_k+(1-\tau)i_k a]^3 > 0$$
$$\frac{\partial}{\partial\, a}\frac{\partial r^F_{t+1}}{\partial (I_{t+1}/K_{t+1})} = \left(\frac{1-\tau}{1-w^B}\right)\frac{(1+w_k)(1+i_k-\delta)}{[1+w_k+(1-\tau)i_k a]^2} > 0\,.$$

The signs of the above derivatives hold when the values of firm characteristics are at the sample averages, that is, $i_k = 0.37$, $w_k = 3.60$, $y_k = 3.09$, $\gamma = 0.15$, $\delta = 0.19$, and $\tau = 0.39$.

Figure A.1: **Replication of Panel B Figure 3 in Gonçalves, Xue and Zhang (2020)**

Both the fundamental and realized decile returns are in percentage per annum. The book-to-market (BM) deciles (except for the two extreme deciles) are in blue circles, the momentum (R11) deciles in red squares, the asset growth (I/A) deciles in green diamonds, and the return-on-equity (ROE) deciles in black triangles. The low BM decile is denoted by "L" and the high BM decile by "H".

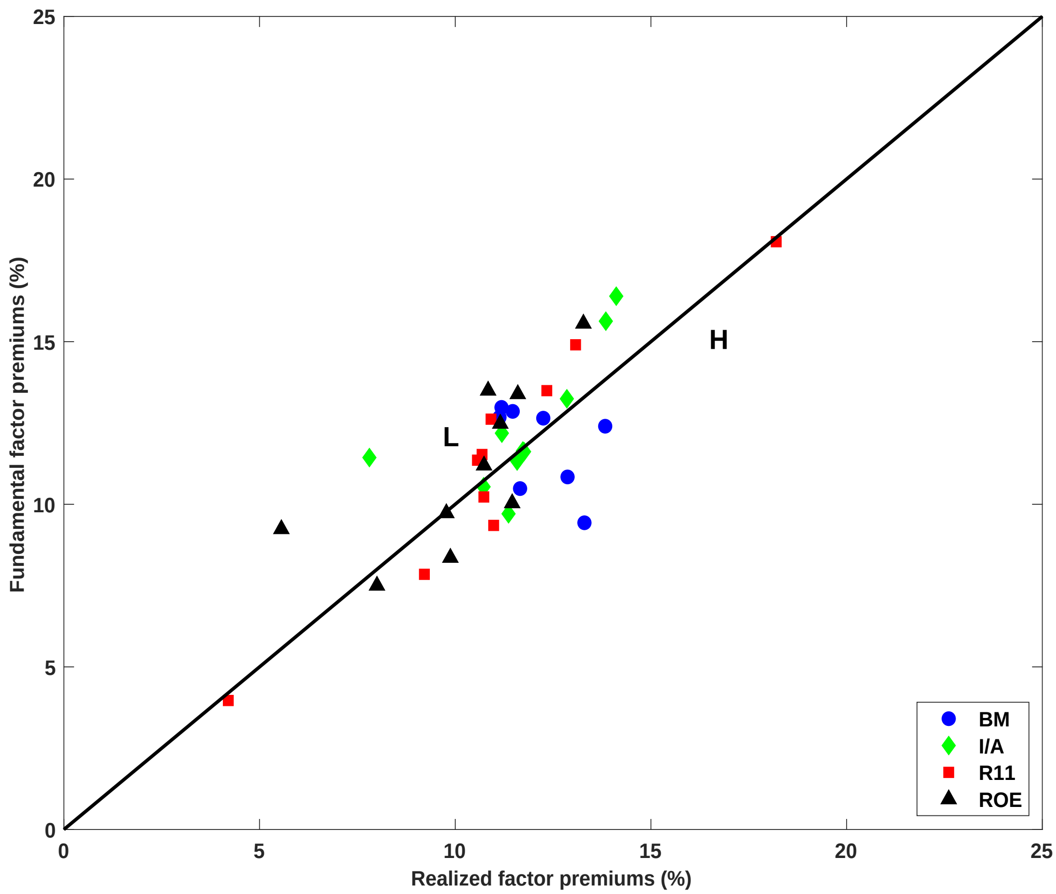


Figure A.2: **Simulation study: parameters with industry and time variations**

This figure plots the time series of the true values (in red solid lines), the NLS estimates (in green lines with triangle markers), the means (in blue dashed lines) and the 95% credible intervals (in shaded areas) of the model parameters estimated from the simulated data. The marginal product and adjustment costs parameters of Wholesale & Retail, Healthcare, Utilities, and Others industries are denoted as γ_i and a_i for $i = 4, 5, 6, 7$, respectively.

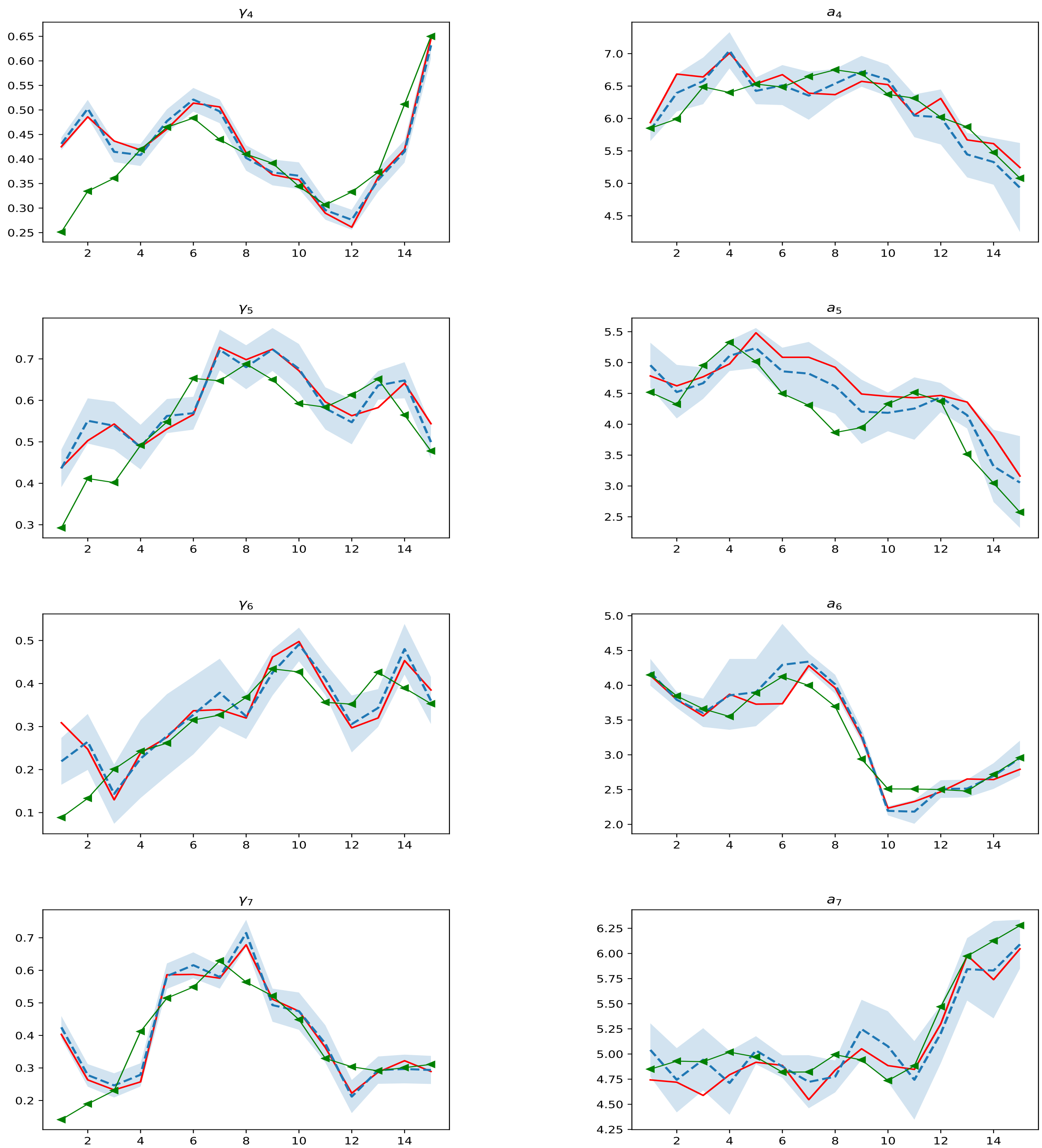
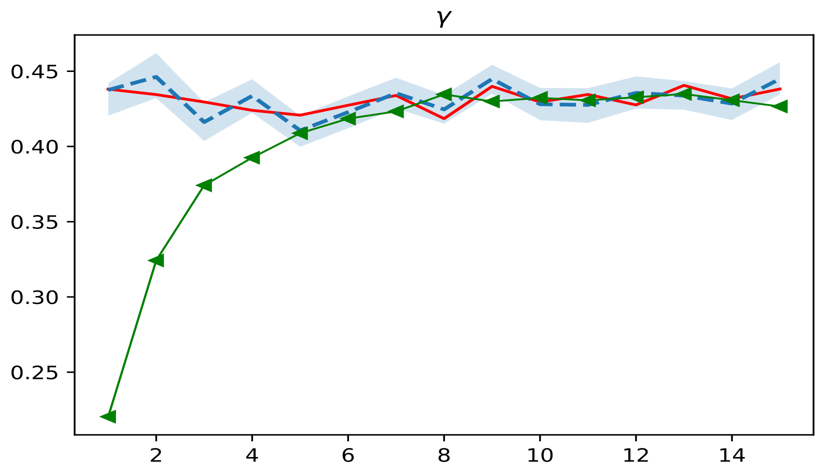


Figure A.3: **Simulation study: parameters with time variations only**

This figure plots the time series of the true values (in red solid lines), the NLS estimates (in green lines with triangle markers), the Bayesian MCMC means (in blue dashed lines) and the 95% credible intervals (in shaded areas) of the model parameters estimated from the simulated data. The marginal product parameter is denoted as γ and the adjustment costs parameter is denoted as a.



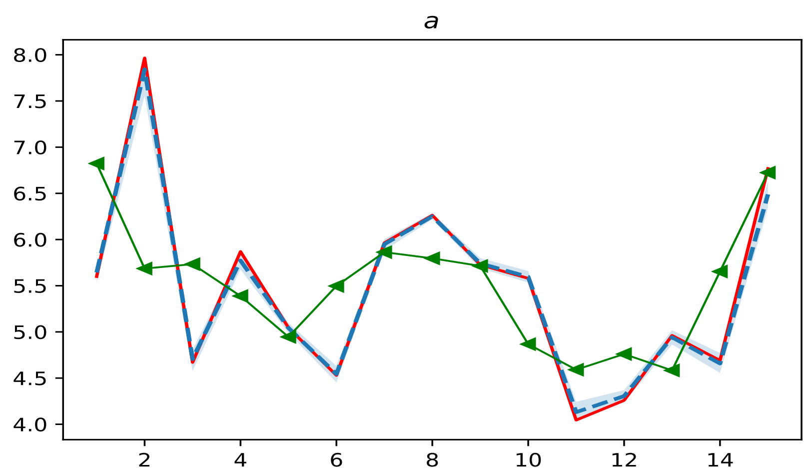
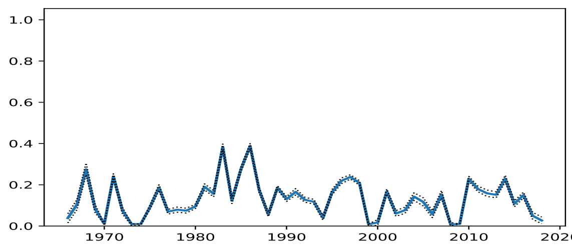
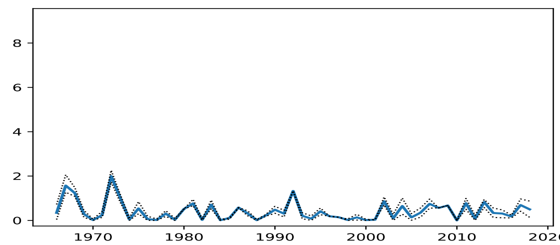


Figure A.4: **Time series of parameter estimates**

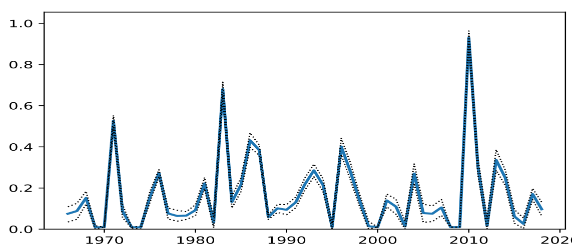
This figure presents the time series of the posterior means (in solid line) and 95% credible intervals (in dotted line) of the marginal product parameter, γ, and physical adjustment costs parameter, a, for Fama-French 10 industries.



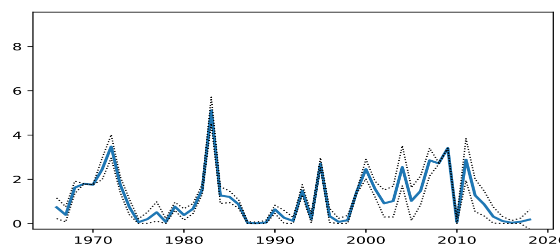
(a) Consumer Nondurables: γ



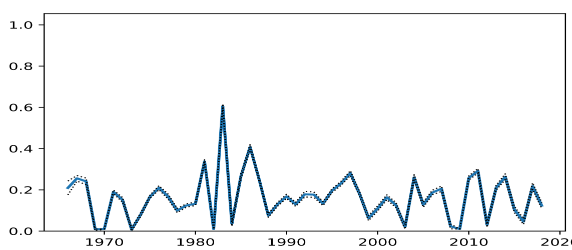
(b) Consumer Nondurables: a



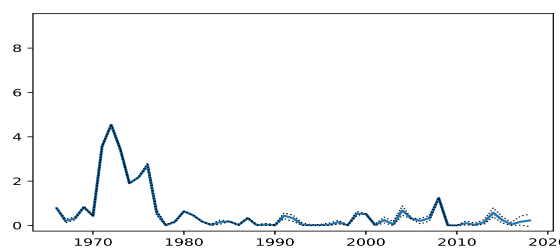
(c) Consumer Durables: γ



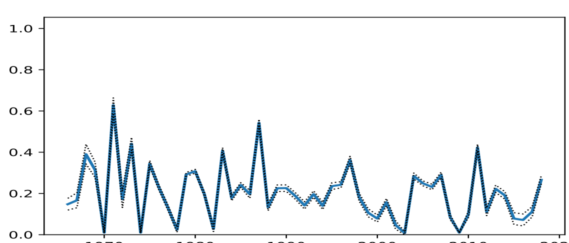
(d) Consumer Durables: a



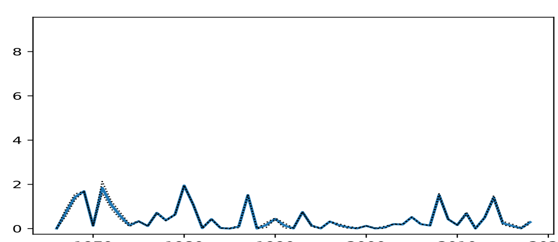
(e) Manufacturing: γ



(f) Manufacturing: a

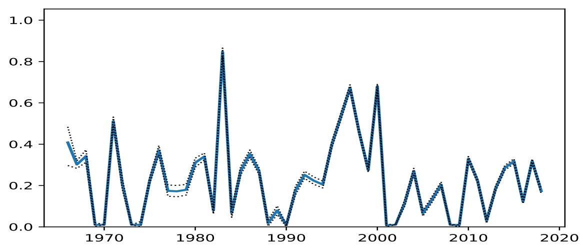


(g) Energy: γ

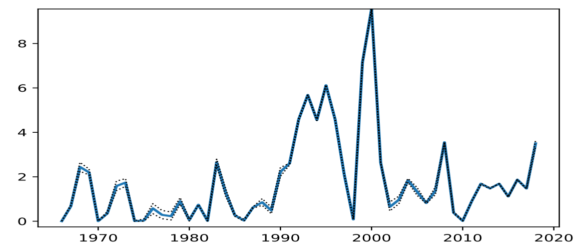


(h) Energy: a

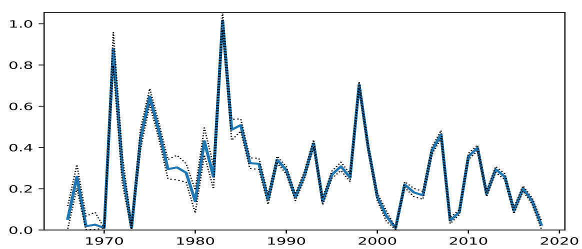
Figure A.4: **Time series of parameter estimates (continued)**



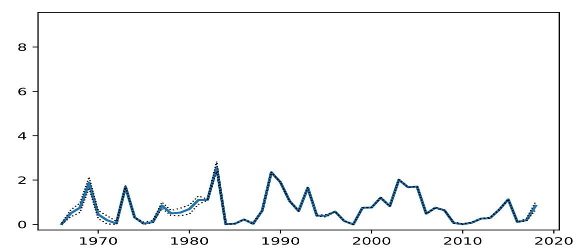
(i) Business Equipment: γ



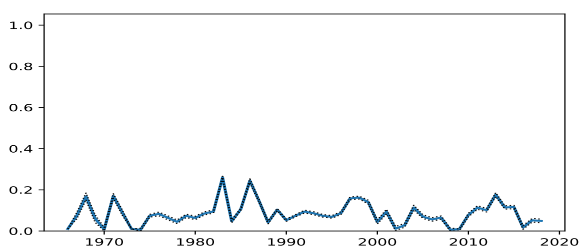
(j) Business Equipment: a



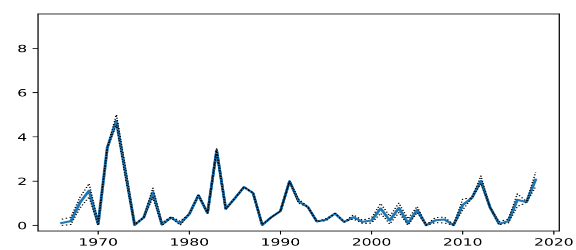
(k) Telecom: γ



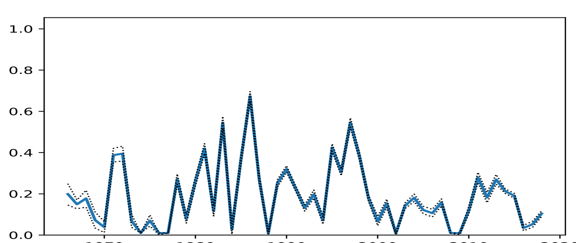
(l) Telecom: a



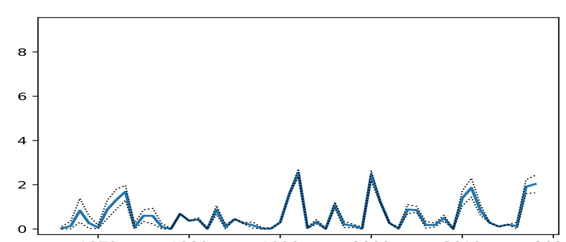
(m) Wholesale & Retail: γ



(n) Wholesale & Retail: a

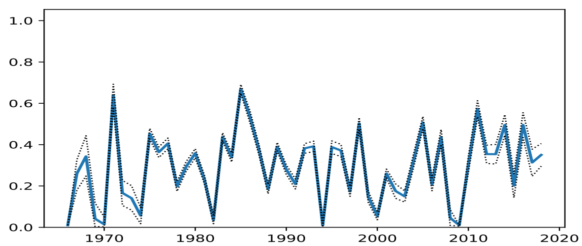


(o) Healthcare: γ

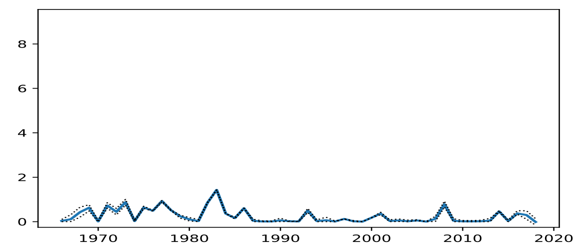


(p) Healthcare: a

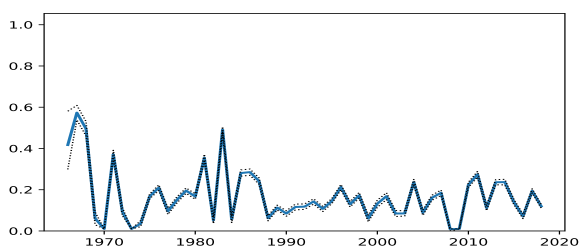
Figure A.4: **Time series of parameter estimates (continued)**



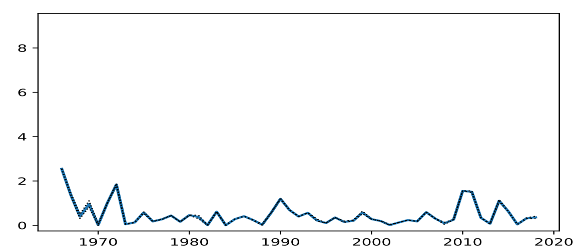
(q) Utilities: γ



(r) Utilities: a



(s) Others: γ



(t) Others: a

Table A.1: **Simulation study: Bayesian MCMC vs. NLS**

This table reports the true values, the NLS estimates, and Bayesian posterior means and credible intervals (in square brackets) of the model parameters, γ and a, under the specification with only industry variation in Panel A and under the specification with constant parameter values in Panel B.

	γ			a		
	True	NLS	Bayesian	True	NLS	Bayesian
Panel A: Parameters with industry variations only						
Consumer Nondurables	0.3295	0.3293	0.3293	5.4954	5.5691	5.5625
			[0.3241,0.3347]			[5.4930,5.6502]
Manufacturing	0.7138	0.7107	0.7112	3.9548	3.9581	3.9596
			[0.7065,0.7158]			[3.9129,4.0013]
Business Equipment	0.6348	0.6378	0.6356	6.0973	6.0769	6.0235
			[0.6281,0.6421]			[5.8258,6.1722]
Wholesale & Retail	0.5154	0.5178	0.5180	5.0573	5.0686	5.0888
			[0.5138,0.5219]			[4.9892,5.1812]
Healthcare	0.3791	0.3813	0.3800	5.5589	5.5051	5.5078
			[0.3723,0.3875]			[5.4267,5.5977]
Utilities	0.3068	0.2902	0.2901	4.8656	4.8718	4.8729
			[0.2769,0.3030]			[4.8470,4.8986]
Other	0.4607	0.4666	0.4661	6.1648	6.1627	6.1626
			[0.4559,0.4763]			[6.1346,6.1926]
Panel B: Parameters with constant values						
	0.1500	0.1501	0.1500	0.1300	0.1280	0.1300
			[0.1487,0.1520]			[0.1297,0.1407]

Table A.2: **Summary statistics of the realized and fundamental firm-level stock returns under NLS estimation**

This table reports the following key statistics for the realized (r^S) and fundamental (r^F) firm-level stock returns: mean, standard deviation, skewness, kurtosis, mean absolute error (m.a.e.) of the fundamental returns, and the time series average of cross-sectional correlations between the realized and fundamental returns. The m.a.e. is defined as m.a.e. $\equiv \frac{1}{T}\sum_{t=0}^{T-1}\frac{1}{N_{t+1}}\sum_{i=1}^{N_{t+1}}|r_{it}^S - r_{it}^F|$, where N_{t+1} is the number of firms in period $t+1$. For fundamental returns, both the posterior means and the 95% credible intervals (in square brackets) of these statistics are reported. Both realized and fundamental returns are winsorized at 0.5 and 99.5 percentiles. The fundamental stock returns are computed based on four model setups: the baseline setup (under column $\boldsymbol{\theta}_{jt}$) in which the estimated parameters are industry specific and time varying; the setup ($\boldsymbol{\theta}_j$) in which the estimated parameters are industry specific but constant over time; the setup ($\boldsymbol{\theta}_t$) in which the estimated parameters are time varying but constant across industries; and the setup ($\boldsymbol{\theta}$) in which the estimated parameters are constant over time and across industries. The sample period is from June 1967 to December 2016.

	Data	$\boldsymbol{\theta}_{jt}$	$\boldsymbol{\theta}_j$	$\boldsymbol{\theta}_t$	$\boldsymbol{\theta}$
mean	14.45	14.71	20.77	14.09	14.89
StdDev	60.78	32.82	23.69	30.12	19.75
Skewness	2.15	1.3384	2.07	1.05	2.13
Kurtosis	11.05	11.84	10.98	11.40	13.36
Correlation	na	0.17	0.09	0.15	0.09
m.a.e.	na	43.41	44.25	43.66	42.45

Table A.3: **Economic meanings of industry and time variations in parameter estimates**

This table investigates the link between operating profits-to-sales ratio (Tobin's q) and γ_{jt} (a_{jt}). Columns (1) and (2) report the the following two industry-level ordinary least squares (OLS) regressions, respectively:

$$\overline{\Pi/Y}_{jt} = c_\gamma + b_\gamma \, \gamma_{jt} + \epsilon^{\gamma}_{jt} \, ,$$

$$\overline{q}_{jt} = c_a + b_a \, a_{jt} \times \overline{I/K}_{jt} + \epsilon^{a}_{jt} \, ,$$

where $\overline{\Pi/Y}_{jt} \equiv \sum_{i=1}^{N_{jt}} \varpi_{it-1}\Pi_{it}/Y_{it}$, $\overline{q}_{jt} \equiv \sum_{i=1}^{N_{jt}} \varpi_{it-1}q_{it}$, and $\overline{I/K}_{jt} \equiv \sum_{i=1}^{N_{jt}} \varpi_{it-1}(1-\tau_t) \times K_{it+1}/(K_{it+1} + W_{it+1}) \times (I_{it}/K_{it})$. In column (1), the dependent variable is the value-weighted operating profits-to-sales ratio for industry j at time t, the independent variable is the estimated value of γ for the same industry and time, and c_γ and b_γ are regression coefficients. Operating profits is measured by operating income before depreciation (item OIBDP). In column (2), the dependent variable is the value-weighted Tobin's q for industry j at time t, the independent variable is an interaction term between the estimated value of a for the same industry and time and the weighted industry average investment rate $\overline{I/K}_{jt}$, and c_a and b_a are regression coefficients. Tobin's q is measured by the market value divided by the book value of the firm. The market value of the firm is calculated as the book value of the firm minus the book value of equity plus the market value of equity. The weight ϖ_{it-1} is proportional to the market equity V_{it-1} as defined in equation (9). The t-values based on robust standard errors are reported in parentheses. The sample period is from fiscal year 1965 to 2017.

	$\frac{\Pi}{Y}$	q
γ_{jt}	0.13	
	(6.10)	
$a_{jt} \times \overline{I/K}_{jt}$		1.97
		(3.33)
Constant	0.18	2.16
	(34.83)	(34.49)
$Adj.R^2$	0.078	0.017
Observations	530	530

Table A.4: **Average firm characteristics of high and low decile portfolios**

This table reports the average I_{it}/K_{it}, I_{it+1}/K_{it+1}, Y_{it+1}/K_{it+1}, and W_{it+1}/K_{it+1} of the low (L) and high (H) decile portfolios for the 12 anomaly variables. The sample period is from June 1967 to December 2016 for all anomaly variables except for ROA, RD/M, and Ad/M. The sample starts at December 1972, December 1976, and December 1973 for ROA, RD/M, and Ad/M, respectively, due to data availability.

	BM		R11		I/A		ROE	
	L	H	L	H	L	H	L	H
I_{it}/K_{it}	0.55	0.18	0.39	0.49	0.25	0.53	0.41	0.46
I_{it+1}/K_{it+1}	0.45	0.19	0.25	0.48	0.28	0.40	0.30	0.43
Y_{it+1}/K_{it+1}	8.62	7.44	7.91	9.71	8.75	8.48	7.76	9.90
W_{it+1}/K_{it+1}	4.50	3.22	4.27	4.39	4.42	4.22	5.00	3.98
	Size		Accruals		NSI		$\Delta PI/A$	
	L	H	L	H	L	H	L	H
I_{it}/K_{it}	0.34	0.26	0.34	0.45	0.27	0.48	0.27	0.42
I_{it+1}/K_{it+1}	0.31	0.24	0.33	0.38	0.26	0.39	0.28	0.33
Y_{it+1}/K_{it+1}	9.43	4.27	7.51	11.88	7.90	7.94	9.46	5.85
W_{it+1}/K_{it+1}	4.34	1.79	3.45	5.73	3.21	4.35	4.40	2.38
	GP/A		ROA		RD/M		Ad/M	
	L	H	L	H	L	H	L	H
I_{it}/K_{it}	0.32	0.42	0.43	0.51	0.39	0.32	0.50	0.24
I_{it+1}/K_{it+1}	0.27	0.38	0.32	0.48	0.31	0.32	0.41	0.24
Y_{it+1}/K_{it+1}	3.54	12.40	7.90	10.06	6.74	9.09	7.58	10.45
W_{it+1}/K_{it+1}	3.30	4.40	5.23	4.44	2.98	6.06	4.37	3.90

References

Balakrishnan, Karthik, Eli Bartov and Lucile Faurel. 2010. "Post loss/profit announcement drift." Journal of Accounting and Economics 50(1):20–41.

Carhart, Mark M. 1997. "On persistence in mutual fund performance." Journal of Finance 52:57–82.

Chan, Louis KC, Josef Lakonishok and Theodore Sougiannis. 2001. "The stock market valuation of research and development expenditures." Journal of Finance 56(6):2431–2456.

Cooper, Michael J., Huseyin Gulen and Michael J. Schill. 2008. "Asset growth and the cross-section of stock returns." Journal of Finance 63:1609–1651.

Davis, James L., Eugene F. Fama and Kenneth R. French. 2000. "Characteristics, covariances, and average returns: 1929 to 1997." Journal of Finance 55(1):389–406.

Fama, Eugene F. and Kenneth R. French. 1992. "The cross-section of expected stock returns." Journal of Finance 47:427–465.

Fama, Eugene F. and Kenneth R. French. 1996. "Multifactor explanations of asset pricing anomalies." Journal of Finance 51(1):55–84.

Fama, Eugene F. and Kenneth R. French. 2008. "Average returns, B/M, and share issues." Journal of Finance 63:2971–2995.

Gonçalves, Andrei, Chen Xue and Lu Zhang. 2020. "Aggregation, capital heterogeneity, and the investment CAPM." Review of Financial Studies 33:2728–2771.

Hou, Kewei, Chen Xue and Lu Zhang. 2020. "Replicating anomalies." Review of Financial Studies 33(5):2019–2133.

Lyandres, Evgeny, Le Sun and Lu Zhang. 2008. "The new issues puzzle: Testing the investment-based explanation." Review of Financial Studies 21:2825–2855.

Novy-Marx, Robert. 2013. "The other side of value: The gross profitability premium." Journal of Financial Economics 108:1–28.

Robert, Christian and George Casella. 2013. Monte Carlo Statistical Methods. Springer Science & Business Media.

Sloan, Richard G. 1996. "Do stock prices fully reflect information in accruals and cash flows about future earnings?" The Accounting Review 71:289–315.