Fundamental Anomalies*

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

*We have benefited from helpful comments from Li An, Hang Bai, Frederico Belo, Jules Van Bingsbergen, Andrei Gonçalves, Feng Gao, Xiaodan Gao, Stefano Giglio, Bing Han, Xiaoji Lin, Clark Liu, Raymond Liaung, Xiaomeng Lu, Danqing Mei, Zhongzhi Song, Dacheng Xiu, Chen Xue, Lu Zhang, Miao Ben 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.

Fundamental Anomalies

July, 2022

Abstract

This paper proposes a portfolio-independent method to estimate q-theory models and examines whether an extensive set of stock market anomalies can be generated by a two-capital q-model. Model parameters are obtained using Bayesian Markov Chain Monte Carlo (MCMC) to match firm-level stock returns. 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 two-capital 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, Bayesian MCMC estimation, Anomalies, Investment, Profitability

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 assumption that a firm operates to maximize its market equity.[1] Liu, Whited and Zhang (2009) show that such a simple q-model fits value, earnings surprises, and investment anomalies well, when model parameters are chosen to match the average observed stock returns of decile portfolios sorted by these anomaly variables. Additional asset pricing anomalies can be explained in this framework as shown in subsequent studies, such as Liu and Zhang (2014) and Gonçalves, Xue and Zhang (2020) among others.

In this line of research, model parameters are estimated via the Generalized Method of Moments (GMM) with the average observed returns of testing portfolios as target moments. This practice is likely to miss information orthogonal to the sorting variables underlying these portfolios, and the resulting parameter estimates are generally portfolio-dependent. More importantly, this approach can not give a fair evaluation of the model's capability in explaining target anomalies because the decile portfolios of these same anomalies are used as the testing portfolios in parameter estimation. The set of parameters that are chosen to explain their target anomalies often fail to explain other anomalies. 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 Campbell (2017)'s critique, we propose a portfolio-independent methodology for estimating q-models, in which model parameters are chosen using the Bayesian Markov

[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.

Chain Monte Carlo (MCMC) method to match the model-implied fundamental returns with the observed stock returns at the firm level. Bayesian MCMC is widely used in the field of macroeconomics (Smets and Wouters, 2007, among others) and has also been used to study return dynamics (Li, Wells and Yu, 2008; Li et al., 2019, among others). Given its portfolio-independence, this methodology can be used to evaluate the capability of any given model in explaining stock market anomalies.

For comparison with prior literature, we estimate the same two-capital q-model as in Gonçalves, Xue and Zhang (2020) and examine its performance in explaining anomalies with portfolio-independent parameters. With the flexibility of Bayesian MCMC, we estimate the model under four specifications, each with constant, industry-specific, time-varying, or industry-specific and time-varying parameter values. Our simulation studies show that Bayesian MCMC is able to discover the true parameter values under this model framework. To examine the capability of the model in explaining stock market anomalies, we consider 12 well-documented anomalies covering six major categories classified by Hou, Xue and Zhang (2020):[2] 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

[2]We choose the 12 anomalies based on the following criteria: (1) in each of the six major categories, we select the anomaly variables that have been extensively documented to predict future returns and generate significant abnormal returns in our sample period (the only exception is size, which does not generate significant abnormal returns in our sample period but is kept due to its importance in the literature); (2) investment-based asset pricing models have been suggested by prior studies to explain these anomalies: value and size (Gomes, Kogan and Zhang, 2003; Carlson, Fisher and Giammarino, 2004; Zhang, 2005), momentum (Liu and Zhang, 2014), asset growth (Watanabe et al., 2013; Titman, Wei and Xie, 2013), investment-to-assets ratio and new stock issues (Lyandres, Sun and Zhang, 2008; Li, Livdan and Zhang, 2009), accruals (Wu, Zhang and Zhang, 2010), return-on-equity, return-on-assets, and gross profitability (Kogan, Li and Zhang, 2019; Ai, Li and Tong, 2021), research and development expenses (Li, 2011; Lin, 2012), and advertising expenses (Belo, Lin and Vitorino, 2014).

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 expenses-to-market ratio (RD/M) and advertising expenses-to-market ratio (Ad/M), and one trading frictions anomaly sorted on market capitalization (Size).

Several findings emerge from our study. First, the estimated q-model matches the mean, skewness, and kurtosis of firm-level returns well under all four specifications and is able to capture 30% to 56% of the volatility depending on the specification. The specification with industry-specific and time-varying parameter values leads to significantly lower mean absolute error (m.a.e.) than the other three specifications and is used as the baseline to examine fundamental return anomalies. We conduct an extensive set of tests to confirm that the industry and time variations in our estimates of the two model parameters, the investment adjustment cost parameter and production curvature parameter, are largely consistent with their economic interpretations.

Second, the fundamental returns exhibit large and significant size, momentum, investment (except the accruals), profitability, and intangibles premiums.[3] The differences between the realized and fundamental premiums, defined as alpha, are mostly insignificant at the 5% level. More importantly, the fundamental returns of these anomaly deciles match the dynamics of their counterparts in the data well. The fundamental and realized portfolio returns are all highly correlated, with an average correlation coefficient of 0.69 for decile portfolios and 0.43 for the high-minus-low deciles.

Third, comparative statics show that heterogeneity in firm characteristics contributes

[3]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).

more to anomaly premiums than the industry and time variations in parameter values. Among firm characteristics, sales-to-capital ratio, as a measure for profitability, is more important than lagged investment-to-capital ratio, which is opposite to what prior studies (Liu, Whited and Zhang, 2009; Liu and Zhang, 2014; Gonçalves, Xue and Zhang, 2020) find. The reason is that our estimates of the investment adjustment cost parameter are much smaller than their estimates, while our estimates of the production curvature parameters are in similar magnitude. For example, under the specification with constant parameter values, our estimate of the adjustment cost parameter is 0.14, while in contrast, the estimate is 2.84 in Gonçalves, Xue and Zhang (2020).[4] In the presence of quadratic investment adjustment costs, higher lagged investment rate means higher marginal costs of investment, while higher profitability means higher marginal benefits of investment. Therefore, fundamental stock return, which equals the levered investment return in the model, increases with profitability but decreases with lagged investment rate. The sensitivities of fundamental return to lagged investment rate and profitability increase with the adjustment cost parameter and the production curvature parameter, respectively. Smaller adjustment cost parameter thus largely decreases the importance of investment-to-capital ratio in explaining the cross-sectional return anomalies.

Despite the aforementioned success, the model fails to generate the value and accruals premiums. The fundamental value premium is 0.46% (t=0.26) per annum, compared to the realized one of 6.74% (t=2.57). This result contrasts drastically with prior studies (Liu, Whited and Zhang, 2009; Gonçalves, Xue and Zhang, 2020), in which estimated q-models can generate sizable value premium. The key reason behind this difference is the aforementioned

[4]Depending on methods and datasets, estimates of the investment adjustment cost parameter in the literature cover a wide range. Our estimates are in similar magnitude as those in Hall (2004) and Cooper and Haltiwanger (2006), among others.

small magnitude of our estimates of the adjustment costs parameter. Value firms have lower investment-to-capital and sales-to-capital ratios. To generate the value premium, the adjustment cost parameter has to be large enough so that the effect of investment dominates the effect of profitability. This failure highlights the importance of portfolio-independent parameter estimation in evaluating a model's capability to generate anomalies. Since returns on the value deciles are often among the target moments of the estimations in prior literature, their estimates of investment adjustment costs parameter are naturally larger. We explore two possible reasons why the model fails the value premium: the absence of asymmetric investment adjustment costs and intangible capitals. However, neither of them seems to explain the insignificant model-implied value premium.

The model also fails to generate the accruals anomaly. The fundamental accruals premium is 4.74% (t =4.45), while the realized one is -5.58% (t=-3.14). High accruals firms have higher profitability than low-accrual firms, resulting in higher fundamental returns. However, the profitability of high-accruals firms is likely overstated because numerous studies (Dechow and Dichev, 2002, among others) show that high-accruals firms are more likely to engage in earnings management activities and have more subsequent write-offs of account receivables than low-accruals firms do. Prior literature, such as Zhang (2007) and Wu, Zhang and Zhang (2010), argues that the accruals premium is driven by the cross-sectional spread in working capital investment. However, we show that adding adjustment costs in working capital investment leads to qualitatively similar results.

Furthermore, we compare the dynamics of the fundamental and realized factor premiums. We show that the model captures the short-lived dynamics of the momentum, ROE, and ROA premiums and the long-lived dynamics of the other anomalies. Fundamental factor premiums also exhibit largely consistent cyclicality as those of the realized ones, albeit with less cyclical variations. Overall, the model successfully captures the dynamics of these factor premiums.

At last, we explore the out-of-sample predictive power of the estimated model. 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 and portfolio-independent estimation is crucial for a fair evaluation of model performance.

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. (2022) 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. (2022) 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 contributes to the literature on cross sectional stock return prediction. Prior studies either run cross-sectional regressions of future stock returns on a few lagged stock characteristics (e.g., Fama and French, 2008b; Lewellen, 2015), or most recently use machine learning methods to harness a large collection of predictor variables, (e.g., Gu, Kelly and Xiu, 2020; Kozak, Nagel and Santosh, 2020). However, previous approaches do not impose economic structure on the data.[5] Our method is complementary to the existing literature by combining the Bayesian MCMC with a simple yet powerful q-theoretical model.

The rest of the paper is organized as follows. Section 2 outlines the model. Section 3 explains the data used in the estimation, describes the estimation procedure, and verifies the accuracy of Bayesian MCMC estimates under our model framework using simulation studies. Section 4 presents the estimation results and compares the performance of the four estimation specifications. Section 5 examines the 12 model-implied fundamental anomaly premiums and explores the economic mechanisms behind the capability of the estimated model in explaining anomalies and the limitations of the model. Section 6 discusses the recursive estimation with expanding window and out-of-sample forecasts. Section 7 concludes.

[5] Gu, Kelly and Xiu (2020) state that "Machine learning methods on their own do not identify deep fundamental associations among asset prices and conditioning variables" and call for future research in the direction that combines statistical methods with economic structures (e.g., Feng, Giglio and Xiu, 2020).

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 $\partial\Pi_{it}/\partial K_{it} = \gamma_{jt}^K Y_{it}/K_{it}$ and $\partial\Pi_{it}/\partial W_{it} = \gamma_{jt}^W Y_{it}/W_{it}$, respectively, in which γ_{jt}^K, $\gamma_{jt}^W > 0$ are the corresponding shares of capital in sales Y_{it} with $\gamma_{jt} \equiv \gamma_{jt}^K + \gamma_{jt}^W < 1$.[6] The model is estimated under four specifications, each with constant, industry-specific, time-varying, or industry-specific and time-varying parameter values. For generality, we formulate the model with industry-specific and time-varying parameter values in this section.

Note that the implications of the model hold regardless of whether model parameters are constant, industry-specific, or time-varying, as long as they are exogenous. We follow the convention in this line of research and assume rational expectation in the model, in which individuals know the true economic model, its parameters and shocks, and the nature of the stochastic processes that govern their evolution. The proofs and derivations in the Appendix do not rely on constant model parameters. The role of exogeneous and time varying parameters in the model is analogous to that of productivity shocks. The assumptions crucial to the model predictions are (1) profit function $\Pi(K_{it}, W_{it}, S_{it})$ is constant-return-to-scale;

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

(2) investment adjustment cost function is linearly homogeneous in investment and capital. In the estimations with time-varying parameters, we assume that these parameters follow random walks.[7]

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^B_{it+1} 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 $D_{it} \equiv (1-\tau_t)(\Pi_{it} - \Phi_{it}) - I_{it} - \Delta W_{it} + B_{it+1} - r^{Ba}_{it} B_{it} + \tau_t \delta_{it} K_{it}$, in which $r^{Ba}_{it} \equiv r^B_{it} - \tau_t(r^B_{it} - 1)$ 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 $E_t[M_{t+1} r^K_{it+1}] = 1$, in which r^K_{it+1} is the return on physical capital investment:

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

[7] We also try a specification that parameters evolve AR(1) process. The estimation yields close to one persistences and for simplicity, we assume random walk for our baseline estimation.

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:

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

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:

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

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, $w^K_{it} \equiv q^K_{it} K_{it+1}/(q^K_{it} K_{it+1} + W_{it+1})$ 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 physical capital. The marginal q of working capital is one in the absence of adjustment costs in working capital investment. 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

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

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

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

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, $r_{it+1}^S = r_{it+1}^F$, 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 and estimation methodology

3.1 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.

3.1.1 Anomalies We explore 12 anomalies covering all six categories defined in Hou, Xue and Zhang (2020): 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), and one trading frictions anomaly sorted on market capitalization (Size).[8] We choose these 12 anomalies based on the following criteria: (1) 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. (2) Investment-based asset pricing models have been suggested by prior studies to explain these anomalies, for example, value and size (Gomes, Kogan and Zhang, 2003; Carlson, Fisher and Giammarino, 2004; Zhang, 2005), momentum (Liu and Zhang, 2014), asset growth (Watanabe et al., 2013; Titman, Wei and Xie, 2013), investment-to-assets ratio and new stock issues (Lyandres, Sun and Zhang, 2008; Li, Livdan and Zhang, 2009), accruals (Wu, Zhang and Zhang, 2010), return-on-equity, return-on-assets, and gross profitability

[8]Hou, Xue and Zhang (2020) show that almost all well-known anomalies in the trading frictions category cannot be successfully replicated. For completeness, we include the Size anomaly from this category since it is one of the most widely studied anomalies in the literature.

(Kogan, Li and Zhang, 2019; Ai, Li and Tong, 2021), research and development expenses (Li, 2011; Lin, 2012), and advertising expenses (Belo, Lin and Vitorino, 2014). Section A 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.

3.1.2 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 $l+6$. The detailed description about the timing alignment is provided in Appendix C.

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 and portfolio compositions can also change monthly. 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 $l-5$ to month $l+6$ to match the fundamental returns constructed in month l. Applying this rolling procedure to the realized 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).

3.2 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. To completely avoid portfolio dependence in parameter estimates, we instead target the entire panel of firm-level stock returns using the Bayesian Markov Chain Monte Carlo (MCMC) method. The Bayesian MCMC method can efficiently extract large amount of information from firm-level stock returns and enable us to allow the parameters to vary across industries and over time. We consider four specifications in the estimation: constant parameters, industry variations only, parameters with time variations only, and parameters with industry and time variations. Next, we explain our methodology in details in terms of the specification that allows both industry and time variations in parameter values.

3.2.1 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.[9] For the specification with time variations only, the same random walk process is assumed for all industries, and for the specifications with no time variations, e^{γ}_{jt+1} and e^{a}_{jt+1} are set to zero.

Realized stock return of firm i (in industry j) at time $t+1$ is modeled as the corresponding

[9] 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.

fundamental return plus an estimation error:

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

where e_{it+1}^{r} 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.[10] 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).[11]

For the MCMC method, prior distributions of the model parameters need to be specified. We use inverse gamma distributions for the priors of variances: $\sigma_\gamma^2 \sim IG(\kappa_1^\gamma, \kappa_2^\gamma)$, $\sigma_a^2 \sim IG(\kappa_1^a, \kappa_2^a)$, 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

[10]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).

[11]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.

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 equations (7) and (8), the joint posterior distribution of $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$ can be written (in a proportional form) as:

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

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 $\mathcal{IG}(\cdot\,;\dot{\kappa_1},\dot{\kappa_2})$ 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.

3.2.2 Comparison of Bayesian MCMC with GMM and NLS 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.

Third, given the vast amount of information in firm-level stock returns exploited by our Bayesian MCMC estimation, we are able to accurately identify the true values of model parameters even for the specification where these parameters vary across industries and over time. This feature can be extremely important when the modeled economy is highly heterogeneous and changing over time. With the Fama-French 10-industry classification, we estimate 2×500 latent variables for the sample between 1967 to 2016. Although the number of latent variables is small compared to the number of observations used in our estimation $(136,598)$, it is extremely difficult to estimate them via either GMM, or a maximum likelihood based approach. Prior studies thus assume constant parameter values in general.

Lastly, the frequentist method Nonlinear Least Squares (NLS) can also be used to estimate industry-specific and time-varying parameters using firm-level stock returns as follows:

$$\hat{\theta}_{jt+1}^{NLS} = \arg\min_{\theta_{jt+1}} \sum_{i=1}^{N_{jt+1}} \varpi_{it} \left[f\left(X_{it}, X_{it+1} | \hat{\theta}_{jt}^{NLS}, \theta_{jt+1} \right) - r_{it+1}^{S} \right]^2 , \tag{11}$$

where 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). However, only the information of industry j at time t is used to identify $\hat{\theta}_{jt+1}^{NLS}$s. In contrast, the posterior of an industry-time specific parameter, θ_{jt}, in Bayesian MCMC utilizes the information of the entire data sample. 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.[12]

We use simulation studies to examine whether Bayesian MCMC can discover the true parameter values under our model framework, which is highly nonlinear, and also compare the performance of Bayesian MCMC and NLS.[13] The results show that Bayesian MCMC performs very well for our highly nonlinear model and is able to discover the true parameter values under all four specifications, while NLS often fails to discover the true parameter values that are time-varying. The details of the simulation studies are included in Section E of the Internet Appendix.

4 Parameter estimation and performance evaluation

We conduct estimations of four model specifications: constant parameters, industry-specific parameters, time-varying parameters, and industry-specific and time-varying parameters. We discuss whether the estimated parameter values make economic sense and compare the performance of these four specifications in terms of matching firm-level stock returns.

4.1 Parameter estimates

Our estimation generates the posterior distributions of the marginal product parameter γ and the physical investment adjustment costs parameter a. For each parameter estimate, we report its posterior mean and credible interval (CI), the latter of which refers to the interval

[12]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.

[13]Detailed discussion on the differences between Bayesian and NLS, and the comparison of their estimation accuracy under our model framework are provided in Section D of the Internet Appendix.

wherein a posterior draw of the parameter falls with the specified probability. A small CI indicates that the parameter is accurately estimated. [14]

Panel A of Table 3 shows that under the specification with constant parameters, the posterior mean of γ is 0.15 with a CI of $[0.15, 0.15]$ and the posterior mean of a is 0.14 with a CI of $[0.13, 0.14]$. These narrow CIs indicate that the parameter values are identified with high precision due to the simple model structure (only two structural parameters) and the large amount of information (136,598 firm-level observations).

Panels B to D present estimation results under the speciations with industry-specific parameters, time-varying (industry invariant) parameters, and industry-specific and time-varying parameters, respectively. For specifications with time-varying parameters (Panels C and D), we report the time-series averages of the posterior means and 95% credible intervals of γ and a and their time-series standard deviations under Columns $\sigma\gamma$ and σa. Several observations emerge. First, parameters are accurately estimated with narrow CI's even when parameters vary over time and across industries. This result highlights that the Bayesian MCMC is able to extract large amount of information from firm-level stock returns and precisely identify the parameter values even when they are industry-specific and time-varying.

Second, allowing time variation in parameters is crucial for getting unbiased estimates when the underlying model is nonlinear in these parameters. When time variation is shut down, the estimates of γ (in Panels A and B) are close to the time series average of their time-varying counterparts (in Panels C and D). In contrast, we see large changes in the estimates of a when time variation is shut down. For example, the posterior mean of γ for

[14]Formally, the posterior credible interval I_p of parameter $\boldsymbol{\theta}$ satisfies $P(\boldsymbol{\theta} \in I_p | \boldsymbol{X}, \boldsymbol{r^S}, \boldsymbol{r^{Ba}}) = p$, where p is the probability. A credible interval is the Bayesian equivalent of the confidence interval in frequentist statistics. A credible interval is the Bayesian counterpart of the confidence interval in frequentist statistics. Confidence intervals treat the estimated parameter as a fixed value and the bounds are random variables due to random sampling, whereas credible intervals treat the parameter as a random variable and the credible bounds are determined by the derived posterior distribution.

Business Equipment sector is 0.23 in Panel B, which is the same as its time-series average in Panel D. However, the posterior mean of a for Business Equipment sector is 0.78 in Panel B, less half of its time-series average of 1.78 in Panel D. The reason is that model-implied return is a linear function of γ but a nonlinear function of a as shown in equation (6). If the true value of the parameter is time-varying, when information is aggregated along the time dimension, a nonlinear relation used for the estimation often generates an estimate that is far away from its time-series average.

Third, the parameter estimates differ greatly across industries. For example, under the specification with industry and time variations (Panel D), γ is estimated to be 0.08 on average with a CI of $[0.07, 0.09]$ 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. The average posterior mean of a ranges from 0.25 for Utilities to 1.78 for Business Equipment. 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 value-weighted average of model-implied marginal q for physical capital, i.e., q^K (marginal q for working capital is one in our model). Table 3 shows that the business equipment sector has the highest q^K of 1.53 while the utilities sector has the lowest q^K of 1.02. Hence, these estimates of a 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. Similar cross-industry variations in parameter values can be observed from Panel B, which reports the estimation results under the specification with industry-specific but time invariant parameters.

Overall, Table 3 shows that there are large variations in parameter values across industries and over time. Estimations that fail to recognize heterogeneities in parameter values could lead to vastly different estimates from the sample average of their true values if the underlying model is nonlinear. Bayesian MCMC is a useful tool for such scenarios.

4.1.1 Comparison with parameter estimates in Gonçalves, Xue and Zhang (2020)

Gonçalves, Xue and Zhang (2020) argue that when portfolio fundamental returns are aggregated from firm-level fundamental returns, their parameter estimates become more stable and less dependent on the testing portfolios. In their baseline estimation targeting the average returns of the value, momentum, investment (I/A), and profitability (ROE) decile portfolios, γ is 0.18 (std = 0.019) and a is 2.84 (std = 0.47). Panel A of Table 3 shows that when matching firm-level stock returns directly using the same model and same data sample, the posterior means of a, 0.14 with a CI of $[0.13, 0.14]$, is almost 20 times smaller than that in Gonçalves, Xue and Zhang (2020), although our estimates of γ, 0.15 with a CI of $[0.15, 0.15]$, are quite close to theirs. The result echos our previous discussion that aggregation of information in estimation, which is along the cross section in Gonçalves, Xue and Zhang (2020)'s case, can lead to very different estimate of the parameter when the underlying model is a nonlinear function of the parameter.[15] Our results indicate that Campbell (2017)'s critique continues to be a serious concern even if portfolio returns are aggregated from model-implied firm-level stock returns as in Gonçalves, Xue and Zhang (2020).

How reasonable are our estimates of adjustment cost parameter a? Regardless of the

[15] In fact, Panel B of Table 3 in Gonçalves, Xue and Zhang (2020) shows that their parameter estimates still show substantial variations with the testing portfolios in their baseline estimation, especially for the adjustment cost parameter a. For example, estimates of γ range from 0.1337 to 0.1762 when matching the average returns of the ten decile portfolios sorted on book-to-market (BM), prior 11-month returns (R11), asset growth (I/A), and return on equity (ROE), respectively. In contrast, estimates of a range between 1.63 to 8.11.

specification, our estimates of a reported in Table 3 are much smaller than the estimates in Gonçalves, Xue and Zhang (2020). Depending on methods and datasets, estimates of a in the literature range from over 20 (Hayashi, 1982) to essentially zero (Hall, 2004). Our estimates of adjustment cost parameter a are fairly close to the estimates in the influential work by Cooper and Haltiwanger (2006). By matching the serial correlation of investment rates, correlation of profit shocks and investment rates, and positive and negative investment spike rates,[16] Cooper and Haltiwanger (2006) estimate that the quadratic adjustment cost parameter is 0.455 when only quadratic adjustment cost is present and the value drops to 0.049 when non-convex adjustment costs are added. These estimates appear extremely low compared to prior estimates based on investment-q regressions (for example, 20 in Hayashi (1982) and 3 in Gilchrist and Himmelberg (1995)). Using simulated data, Cooper and Haltiwanger (2006) show that due to measurement errors in average q, the coefficient on average q in a regression of investment rates on a constant and average q implies an estimate of a 100 times as large as its true value.

In sum, using firm-level return data in estimation seems to generate estimates of a smaller than those estimated with portfolio-level return data (Liu, Whited and Zhang, 2009; Liu and Zhang, 2014; Gonçalves, Xue and Zhang, 2020, among others). When comparing to estimates in the literature without using return data, our estimates are close to the ones based on structural estimation of investment models (Cooper and Haltiwanger, 2006; Hall, 2004, among others) instead of the investment-q regression.

4.1.2 Do the time variations in parameter estimates make economic sense?

In this subsection, we investigate whether the estimates of the production function curvature parameter γ and the investment adjustment cost parameter a are consistent with their

[16]Cooper and Haltiwanger (2006) define episodes of investment rates in excess of 20% spikes.

economic underpinning. 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. In reality, 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.

If the estimated values of γ indeed capture the aforementioned economic, these estimates should be 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}:

$$\overline{\Pi/Y}_{jt} = c_\gamma + b_\gamma \, \gamma_{jt} + \epsilon^{\gamma}_{jt} \, , \tag{12}$$

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.

In terms of the investment adjustment cost parameter a, 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.

Similarly, variations in the estimated values of a should positively correlate with the variations in average Tobin's q across industries and over time since investment rate at the industry level is on average positive. Specifically, the following regression should yield a positive and significant coefficient on a_{jt}:

$$\overline{q}_{jt} = c_a + b_a\, a_{jt} + \epsilon^a_{jt}\,. \tag{13}$$

where the dependent variable is the value-weighted Tobin's q, defined as $\overline{q}_{jt} \equiv \sum_{i=1}^{N_{jt}} \varpi_{it-1} q_{it}$, for industry j at time t. We expect the coefficient on a_{jt} to be significantly positive.

We conduct regressions (12) and (13) using the firm-year panel between 1965 and 2017. Table 4 shows that b_γ is 0.13 with t-stat being 6.10 and b_a is 0.19 with t-stat being 3.60, both of which are positive and highly significant. The adjusted R-squareds are 0.078 and 0.023, respectively. These results confirm that our estimates of γ and a indeed capture the industry-time variations in firm's profitability and Tobin's q in the economically sensible way, both of which are not directly used in the estimation.

4.2 Overall fit of the estimation

In this subsection, we compare a set of key moments of the realized and fundamental firm-level stock returns to gauge the overall fit of the four estimation specifications. Table 5 reports the posterior means and 95% credible intervals of the mean, standard deviation, skewness, kurtosis, the time-series average of the cross-sectional correlations between fundamental and realized stock returns, and the mean absolute error (m.a.e.) for the specifications with constant, industry-specific, time-varying, and industry-specific and time-varying parameter values.[17] The m.a.e. is defined as

$$\text{m.a.e.} \equiv \frac{1}{T}\sum_{t=0}^{T-1}\frac{1}{N_{t+1}}\sum_{i=1}^{N_{t+1}}\left|r_{it}^{S}-r_{it}^{F}\right|, \tag{14}$$

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. The same moments of the realized stock returns are presented for comparison.

Several observations emerge from Table 5. First, 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. Second, standard deviations of the fundamental returns are much smaller than that of the realized one. The specification with industry-specific and time-varying parameters generates the highest standard deviation 34.17%, compared to 60.78% in the data, while the specification with industry-specific parameters generates the lowest standard deviation, 18.49%. Third, the time-series average of the cross-sectional correlations between realized and fundamental firm-level stock returns is highest for the specification with industry-specific and time-varying parameter values (0.20), and lowest

[17] 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 and any statistical moments of these returns.

for the specifications with constant parameter values (0.09). Lastly, the m.a.e. is lowest for the specification with industry-specific and time-varying parameter values (40.10%), followed in turn by the specifications with time-varying (40.85%), industry-specific (41.85%), and constant parameter values (42.45%).[18]

Figure 1 plots the histograms of realized (in blue) and fundamental (in orange) firm-level returns based on the posterior means of the parameter estimates under the four specifications. Consistent with what Table 5 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 the realized and the four fundamental distributions have longer right tails, resulting in positive skewness and kurtosis larger than 3. Next, we conduct a rigorous comparison in performance among the four specifications.

4.3 Performance comparison of the four specifications

We evaluate the performance of each specification based on the the mean absolute error (m.a.e.) of the fundamental firm-level stock returns. 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. Based on the posterior distribution of the m.a.e. under each specification, we test whether the differences in m.a.e. of the four specifications

[18]For comparison, we report the same statistics of the NLS estimation for the four specifications in Table A.1 in the Internet Appendix. Bayesian MCMC results in a smaller m.a.e. in every specification, echoing the superior performance of Bayesian approach compared to NLS documented in Sections E and D of the Internet Appendix.

are statistically significant using the following statistic:

$$d^a = \frac{1}{T}\sum_{t=0}^{T-1}\frac{1}{N_{t+1}}\sum_{i=1}^{N_{t+1}}\left(\left|r_{it+1}^{S} - r_{it+1}^{F(a)}\right| - \left|r_{it+1}^{S} - r_{it+1}^{F(b)}\right|\right)\,, \tag{15}$$

where $r^{F(b)}$ and $r^{F(a)}$ are the fundamental returns under the benchmark specification, defined as the one with the lowest m.a.e. (i.e., specification with industry-specific and time-varying parameters), and under an alternative specification, a, 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 (15), 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 benchmark specification performs significantly better than the alternative specification a in explaining firm-level stock returns.

Figure 2 plots the distribution of d^a for each of the three alternative specifications 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 ([1.80, 1.90], [1.22, 1.32], and [0.57, 0.67] for the specifications with constant, industry-specific, and time-varying parameter values) includes zero, indicating that the baseline performs significantly better in explaining the firm-level stock returns than these alternative specifications. Moreover, the location of these distributions in Panel (a) indicates that the performance of the model with time-varying parameters is closest to the benchmark, followed by the settings with industry-specific parameters and with constant parameters.

In sum, the specification with industry-specific and time-varying parameters has significantly better performance in matching firm-level stock returns than the other three specifications. In addition, time variation is more crucial to the superior performance than industry variation. Hereafter, we use the specification with industry-specific and time-varying parameter values as the baseline for the analysis of stock market anomalies.[19]

5 Fundamental anomalies

In this section, we construct the fundamental returns for a set of well-documented and robust anomalies. We choose these anomalies based on the replication study by Hou, Xue and Zhang (2020), who replicate 452 anomalies, classified in six categories: the momentum, value versus growth, investment, profitability, intangibles, and trading frictions. They show that for the sample period as ours, from June 1967 to December 2017, only 153 anomalies can be successfully replicated.[20] We then choose 12 anomalies, covering all six categories, that are most widely studied in the literature and can be successfully replicated in our sample period. More importantly, investment-based asset pricing models have been suggested by prior studies to explain these anomalies: 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

[19]The analysis of stock market anomalies under the specifications with constant, industry-specific, and time-varying parameters are presented in the Internet Appendix.

[20]The replication of an anomaly is successful if the average return of its high-minus-low decile is significant at the 5% threshold based on portfolio sorts with NYSE breakpoints and value-weighted returns.

anomalies sorted on return-on-equity (ROE), return-on-assets (ROA), and gross profitability (GP/A); and two intangibles anomalies sorted on R&D expenses-to-market ratio (RD/M) and advertising expenses-to-market ratio (Ad/M), and one trading frictions anomaly sorted on market capitalization (Size).[21] In addition, prior literature has proposed investment-based asset pricing models as potential explanations for these anomalies. We ask whether the model-implied fundamental returns exhibit the same return regularities when the model parameters are estimated to match firm-level stock returns.

5.1 Anomaly premiums

Figure 3 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

[21] Hou, Xue and Zhang (2020) show that almost all well-known anomalies in the trading frictions category cannot be successfully replicated. For completeness, we include the Size anomaly from this category since it is one of the most widely studied anomalies in the literature.

information confine the parameter estimates and the fundamental factor premiums to a small set of possible values.

Table 6 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, $\alpha \equiv r^S - r^F$.[22] 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, the 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, which reflects how significant the corresponding fundamental value premium is. In contrast, the credible interval of the fundamental value premium indicates the range of model-implied values given the observed firm fundamentals but allows all possible posterior draws of parameter values. We conclude that a given anomaly exists in fundamental returns if its fundamental anomaly premium is significant at the 5% level (the absolute t-value, $|t| > 1.96$) with a posterior probability higher than 95%. That is, the 95% credible interval of $t(r^F)$ is on the right of 1.96 if $t(r^F) > 0$ and on the left of -1.96 if $t(r^F) < 0$.

Table 6 shows that the model is able to generate significant momentum (R11), investment (I/A, NSI, and ΔPI/A), profitability (ROE, ROA, and GP/A), intangibles (R&D and advertising), and size premiums. The t-values of these premiums all have credible intervals larger than 1.96 (indicating 5% significance level) in absolute value. In

[22]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. For completeness, we report the fundamental anomaly premiums and their corresponding alphas in Table A.2. We analyze the importance of industry and time variations in parameter estimates in term of generating anomaly premiums in Section 5.3.

general, the t-values of the fundamental anomaly premiums are larger than their counterparts in the data due to the low variability of fundamental returns. For example, in terms of the posterior means, the fundamental momentum premium is 11.82% per annum with a t-value of 12.51, while its value is 13.75% (t = 4.15) in the data. In terms of matching the magnitude of the realized anomaly premiums, the alphas of seven out of ten anomalies are insignificantly different from zero, while the alphas of three anomalies, I/A, NSI, and GP/A, are significant at the 5% level.

However, the model fails to generate significant premiums for value and accruals premiums that are consistent with the data. The model generates positive but statistically insignificant value premium. The posterior means of the fundamental value premium is 0.46% per annum (CI=$[0.31\%, 0.60\%]$) with a t-value of 0.26 (CI=$[0.18, 0.35]$). The fundamental accruals premium is 4.74% (t = 4.45) in contrast to -5.58% ($t = -3.14$) in the data. We explore possible explanations for this failure later in Section 5.3.

Note that our results contrast sharply with the findings of Gonçalves, Xue and Zhang (2020) who show that the same model can generate fundamental value and I/A premiums with insignificant alphas (t-values being 1.37 and -0.04) in the same sample period. The difference is that Gonçalves, Xue and Zhang (2020) estimate model parameters using the value, I/A, R11, and ROE deciles as testing portfolios, which improves the model's performance in terms of matching the value and I/A premiums. [23] This comparison validates Campbell (2017)'s critique and highlights the importance of using portfolio-independent parameters to evaluate the capability of a model in generating stock market anomalies.

[23]Our baseline estimation allows parameters to vary across industry and time. When the model with constant parameters, as the one in Gonçalves, Xue and Zhang (2020), are estimated to match firm-level stock return, the model performs even worse in terms of generating value, I/A, R11, and ROE premiums: only the alpha of the ROE premium is insignificantly different from zero. Results are available upon request.

5.2 Dynamics of factor premiums

In this section, we use the fundamental returns implied from the baseline estimation to study the dynamics of factor premiums, including the correlation between realized and fundamental portfolio returns, the persistence of the factor premiums, and the relation between the factor premiums and market states. Fundamental stock returns in this subsection 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 subsection can be largely carried over to any set of parameter values within the 95% credible interval of the posterior distribution.

5.2.1 Correlation between realized and fundamental portfolio returns In this subsection we examine how well the fundamental returns of these anomaly deciles match the dynamics of their counterparts in the data. Table 7 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.

An interesting observation is that although the model implies a tiny value premium and an accruals premium with the wrong sign, the fundamental returns on the BM and accruals deciles are highly correlated with the realized ones. The correlations of these 20 deciles range between 0.63 and 0.78, and the correlations of the high-minus-low deciles are 0.53 for BM and 0.41 for accruals. This contrasts sharply with the corresponding correlations reported in Panel B of Table 6 in Gonçalves, Xue and Zhang (2020). Their model generates much higher correlations for the high, low, and high-minus-low portfolios than the deciles in the middle. For example, the correlations between the fundamental and realized returns on the low, high, and high-minus-low I/A deciles are 0.19, 0.30, and 0.42, respectively, all of which are highly significant. In contrast, the correlations for deciles two to nine range from −0.03 to 0.12, none of which are significant.

Figure 4 plots the time-series returns of the 12 high-minus-low deciles. The fundamental and realized decile returns show strong comovements, consistent with the reported high correlations in Table 7. Overall, our model with the baseline estimation matches the dynamics of these 12 anomaly portfolios very well. It also generates significant premiums for a large set of anomalies, but fails to generate value and accruals premiums. Next, we explore the reasons behind the successes and failures of the model.

5.2.2 Persistence of factor premiums One important aspect of a factor premium is its persistence, which varies greatly across anomalies. Figure 5 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.

5.2.3 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.[24]

Table 8 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

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

both up and down markets.

5.3 Inspecting the economic mechanism

In this section, we first conduct comparative statics to quantify the importance of firm characteristics and industry-time variations in parameter values in generating fundamental return anomalies. We then investigate possible explanations why the model fails to generate the value and accruals premiums.

5.3.1 Comparative statics The set of firm characteristics includes lagged invest-to-physical capital ratio I_{it}/K_{it}, current invest-to-physical capital ratio I_{it+1}/K_{it+1}, current sales-to-physical capital ratio Y_{it+1}/K_{it+1}, and current working-to-physical capital ratio W_{it+1}/K_{it+1}, all of which affect stock returns r^F_{it+1} as indicated by equation (6). We then explore possible explanations why the model fails to generate the value and accruals anomalies.

Intuitively, the importance of a firm characteristic to a specific anomaly premium depends on (1) the spread in this characteristic among high and low deciles; (2) sensitivity of firm-level return to this characteristic. For example, for the I/A premium, firms in its high and low deciles differ most in lagged investment rate I_{it}/K_{it}, because the I/A deciles are sorted on asset growth rate, which is highly correlated with the firm characteristic I_{it}/K_{it}.[25] Naturally, we would expect that I_{it}/K_{it} is most crucial for generating the I/A premium. Based on equation (6), firm-level stock return is more sensitive to investment rate if the adjustment cost parameter a is large, and more sensitive to profitability (Y/K) if the production curvature

[25]Table A.3 in the Internet Appendix shows the average firm characteristics for decile portfolios of the 12 anomaly variables. For example, with respect to the I/A low and high deciles, I_{it}/K_{it} is 0.26 vs. 0.53, I_{it+1}/K_{it+1} is 0.28 vs. 0.40, Y_{it+1}/K_{it+1} is 8.75 vs. 8.43, and W_{it+1}/K_{it+1} is 4.40 vs. 4.19.

parameter γ is large.[26] Thus, the I/A premium is likely to increase with the magnitude of a. From the same reasoning, Y_{it+1}/K_{it+1} is likely to be more important for anomalies sorted on profitability measures if the magnitude of γ is larger.

To conduct comparative statics on a firm characteristic, for example, 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 anomaly 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. To quantify the importance of industry and time variations in parameter values, we shut down the time and industry variations, separately, in parameter values in estimation (namely, specifications $\boldsymbol{\theta}_j$ and $\boldsymbol{\theta}_t$), and compare the resulting alphas with those from the baseline estimation. The results are presented in Table 9 and we summarize the main findings below.

Firstly, profitability Y_{it+1}/K_{it+1} is more important than lagged investment rate I_{it}/K_{it} for all 12 anomalies, even for anomalies sorted on investment measures (I/A, NSI, and ΔPI/A). For example, when firm heterogeneity in I_{it}/K_{it} is turned off, alpha of the fundamental I/A premium becomes -9.88 percent per annum, which is -3.16 in the baseline. In comparison, the I/A alpha becomes 5.53 when heterogeneity in Y_{it+1}/K_{it+1} is turned off. Previous literature often finds that heterogeneity in lagged investment rate is the key driver of anomalies sorted on investment measures (for example, Liu, Whited and Zhang 2009 and Gonçalves, Xue and Zhang 2020). The reason behind this difference is

[26]Section 5.3 in the Internet Appendix provides the proof on the relation between fundamental return and firm characteristics, and the sensitivity of this relation to the parameter values.

that our estimates of a are generally smaller than previous estimates as discussed in subsection 4.1.1, which decreases the importance of investment rate to stock returns.

The importance of Y_{it+1}/K_{it+1} relative to I_{it}/K_{it} is more pronounced for anomalies sorted on profitability measures (ROE, ROA, and GP/A) as expected. Y_{it+1}/K_{it+1} is also the most important factor to generate anomalies sorted on intangibles, sorted on R&D expenses-to-market ratio (RD/M) and advertising expenses-to-market ratio (Ad/M). Neither RD/M nor Ad/M appears in the model-implied relation between stock return and firm fundamentals in equation 6. However, Table A.3 in the Internet Appendix shows that high RD/M (Ad/M) firms have higher profitability than low RD/M (Ad/M) firms, with Y_{it+1}/K_{it+1} being 9.14 (10.47) compared to 6.92 (7.54), which implies positive fundamental intangible premiums.

Secondly, for all 12 anomalies, heterogeneity in firm characteristics is more important for generating anomaly premiums than heterogeneity in parameter values. Shutting down industry or time variations in parameter value never leads to larger deviation from the baseline alphas compared to shutting down variations in firm characteristics. In addition, shutting down industry or time variations generate smaller alphas than the baseline specification for some anomalies. For example, 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.

In sum, our results show that for generating the 12 anomaly premiums, firm fundamental characteristics are more important than variations in parameter values. And among these fundamental characteristics, sales-to-capital ratio Y_{it+1}/K_{it+1} is more important than the investment rates I_{it}/K_{it} and I_{it+1}/K_{it+1}. Next, we investigate possible reasons why the model fails the value and accruals anomalies.

5.3.2 The value premium

Prior studies with similar or same models (Liu, Whited and Zhang, 2009; Gonçalves, Xue and Zhang, 2020) find that the differences in lagged investment ratio I_{it}/K_{it} between value and growth firms contribute the most to the value premium. Table A.3 in the Internet Appendix shows that the largest difference between value and growth firms is that growth firms have higher lagged and current investment rate. However, since our estimated adjustment cost parameter is small, this difference in investment rate is not able to generate large enough return spread. We next discuss possible reasons why the model fails to generate the value premium.

Asymmetric adjustment costs — In our baseline model, we assume adjustment cost of investments to be symmetric and quadratic to be consistent with Gonçalves, Xue and Zhang (2020) for comparison. However, there is a large literature that argues that asymmetric adjustment costs are critical for explaining the observed investment dynamics (Abel and Eberly, 1994; Cooper and Haltiwanger, 2006) and could be a driver for the value premium (Zhang, 2005). We thus consider two versions of asymmetric adjustment costs. The first version is to keep the quadratic form but allow parameter a_{jt} to have different values for positive and negative investments, denoted as a_{jt}^{+} and a_{jt}^{-},[27]

$$\Phi(I_{it}, K_{it}) = \frac{a_{jt}^{+}\mathbb{I}_{I_{it}>=0} + a_{jt}^{-}(1 - \mathbb{I}_{I_{it}>=0})}{2} \left(\frac{I_{it}}{K_{it}}\right)^{2} K_{it}$$

where $\mathbb{I}_{I_{it}>=0}$ is an indicator equals to one if investment is positive and to zero otherwise. Following Belo et al. (2022), our second version of asymmetric adjustment costs function is

[27]Our estimation is based on equation (6), which is derived from the first order condition of firm's optimal investments and holds only if investment is not zero. When investment is zero, firm's optimization problem does not have an interior solution and the first order condition fails to hold. However, in the data the observations with zero investment are rare, less than 0.1% of the sample. We thus ignore these observations in the estimation.

smooth and homogeneous of degree one in investment and capital (Belo et al., 2022):

$$\Phi_{it} \equiv \frac{\theta_{jt}}{\nu_{jt}^2}\left[\exp\left(-\nu_{jt}\frac{I_{it}}{K_{it}}\right) + \nu_{jt}\frac{I_{it}}{K_{it}} - 1\right],$$

where θ_{jt} for industry j at time to t is positive and disinvestment is more costly than investment if $\nu_{jt} > 0$. The relation between fundamental stock return r_{it}^F and firm characteristics under these two versions of asymmetric adjustment costs are provided in Section A.3 of the Internet Appendix.

As in our baseline estimation, we allow the adjustment cost parameters to be industry-specific and time-varying. Table A.4 in the Internet Appendix presents the time-series averages of the posterior means and 95% credible intervals (CI) for each parameter and for each industry. In general, adjustment costs parameters for negative investments are larger than those for positive investments: a^- is larger than a^+ for all industries except for the business equipment sector and the estimates of ν are significantly positive for all sectors. However, asymmetric adjustment costs do not help to generate the value premium. Table A.5 in the Internet Appendix presents the fundamental anomaly premiums and their alphas under these two versions of asymmetric adjustment costs. The results are qualitatively close to the baseline results. Specifically, the fundamental value premium is still insignificant, 0.42% per annum with t-stat of 0.24 under the quadratic version and −0.18% per annum with t-stat of −0.10 under the exponential version of asymmetric adjustment costs. Therefore, the absence of asymmetric investment adjustment costs in the baseline model does not seem to explain its failure to generate the value premium.

Intangibles — Several recent papers (for example, Eisfeldt, Kim and Papanikolaou, 2022; Belo et al., 2022) show that intangible capitals become increasingly important for cross-sectional return and valuation differences. If intangibles are the driving forces for the value premium, and if our model cannot capture the effects of intangibles, the model will naturally fail to generate the value premium. There are mainly three types of intangibles studied in the asset pricing literature: knowledge capital proxied by R&D expenses (Chan, Lakonishok and Sougiannis, 2001), brand capital proxied by advertising expenses (Chan, Lakonishok and Sougiannis, 2001; Belo, Lin and Vitorino, 2014), and organizational capital proxied by organizational capital-to-book assets (Eisfeldt and Papanikolaou, 2013). Table 6 shows that our model is able to capture return spreads sorted on intangibles proxied by R&D and advertising expenses-to-market ratios. We also construct 10 decile portfolios sorted on the industry-adjusted organizational capital-to-book assets (Ioca) following (Eisfeldt and Papanikolaou, 2013) and compute the high-minus-low portfolio returns.[28] The fundamental organizational capital premium is 6.66% (t = 5.12) per annum, compared to 5.28% (t = 3.26) in the data. These results suggest that although our model does not explicitly model intangible capitals, differences in physical investment rate and profitability between firms with high and low intangibles are able to capture their return spreads. Therefore, the lack of intangibles in the model does not seem to explain its failure to generate the value premium. Explicitly modeling intangible capital may yield different results and is an interesting research topic in its own right, which we leave for future work.

5.3.3 The accruals anomaly Table 9 shows that Y_{it+1}/K_{it+1} is the most important driver of the return spreads across firms sorted on accruals. According to Table A.3 in the Internet Appendix, high-accruals firms have higher Y_{it+1}/K_{it+1} than low-accruals firms do, which leads to a positive high-minus-low accruals spread, opposite of the one in the data.

[28]The details on how to construct Ioca is included in the Appendix A.

However, differences in Y_{it+1}/K_{it+1} between high- and low-accruals firms are likely overstated. The concept of accruals is absent in our model and cash and accruals basis accountings are treated the same. The earnings of the high- and low-accruals firms are assumed to have the same quality. High accruals mean high profitability in the model, which however is not necessarily true in the data. Data shows that subsequent write-offs of account receivables happen more often to high-accruals firms (Dechow and Dichev, 2002, among others). Therefore, the true difference in profitability between high- and low-accruals firms is likely much smaller, which may lead to a smaller or even negative fundamental accruals premium.

Prior literature, such as Wu, Zhang and Zhang (2010) and Zhang (2007) among others, argues that high-accruals firms have lower average returns due to their higher investment rate on working capital. The intuition is analogous to that of the physical investment premium. Assuming that investment in working capital incurs adjustment cost, higher investment rate on working capital leads to higher marginal cost of working capital investment and thus lower fundamental stock return. Our baseline model assumes zero adjustment costs on working capital investment for comparison with Gonçalves, Xue and Zhang (2020). We thus estimate an extended model with quadratic adjustment costs on working capital investment and details are explained in Section A.4 of the Internet Appendix. Parameter estimates and the corresponding fundamental anomaly premiums are reported in Table A.6 and Table A.7, respectively. The results are quantitatively similar to the baseline ones and specifically, the accruals premium stays positive and significant. These results indicate that the failure to generate accruals premium is not due to the absence of adjustment costs on working capital investment.[29] Overall, our evidence suggests that explicitly modeling earnings quality is

[29] Our estimates of the adjustment cost parameter on working capital investment is close to the estimates in Gonçalves, Xue and Zhang (2020). When using Gonçalves, Xue and Zhang (2020)'s baseline parameters

a more promising direction to generate the accruals anomaly, which we leave for future research.

6 Out-of-sample performance

In this section, we discuss whether the in-sample feature of the baseline estimation is critical for the model's ability to explain anomalies by recursively estimating model parameters. In addition, we examine the model's out-of-sample predictive power on stock returns at the cross section.

6.1 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 2016. 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

estimates for the extended model with adjustment costs on working capital investment (reported in Section E of Internet Appendix), we also find positive implied accruals premiums as well.

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 2016. 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.

6.2 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.[30] 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 6.1.[31] To reduce 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

[30]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.

[31]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 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 6.1, 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 A.4 in the Internet Appendix.

at the 2.5%-97.5% level.

At the beginning of each month t from October 1980 to December 2016, 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 6.1.

At the beginning of each month t from October 1980 to December 2016, 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-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.

7 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 the same set of anomaly returns that these studies aim to explain. In this paper, we propose a portfolio-independent estimation methodology based on Bayesian MCMC, which can be used to fairly evaluate the performance of any given model in explaining anomalies. Applying this method on a two-capital q-model, we show that the estimated model generates large and significant size, momentum, investment, profitability, and intangibles premiums, but fails to generate the value and accruals premiums. Our results call for future studies on the economic mechanism behind the value and accruals anomalies.

In addition, the estimated 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. Combining statistical methods with economic structures is a promising direction for future research in cross-sectional return predictability.

References

Abel, Andrew and Janice Eberly. 1994. “A United model of investment under uncertainty.” American Economic Review 84(5):1369–1384.

Ai, Hengjie, Jun Li and Jincheng Tong. 2021. “Equilibrium value and profitability premiums.” the Carlson School of Management, University of Minnesota.

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

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

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.

Belo, Frederico, Chen Xue and Lu Zhang. 2013. “A supply approach to valuation.” Review of Financial Studies 26(12):3029–3067.

Belo, Frederico, Vito Gala, Juliana Salomao and Maria Ana Vitorino. 2022. “Decomposing firm value.” Journal of Financial Economics 143:619–639.

Belo, Frederico, Xiaoji Lin and Maria Ana Vitorino. 2014. “Brand capital and firm value.” Review of Economic Dynamics 17(1):150–169.

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.

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

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.

Chan, Lious K., Josef Lakonishok and Theodore Sougiannis. 2001. "The stock market valuation of research and development expenditures." Journal of Finance 56:2431–56.

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

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.

Cooper, Russell W. and John C. Haltiwanger. 2006. "On the nature of capital adjustment costs." Review of Economic Studies 73:611–633.

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.

Dechow, Patricia M. and Ilia D. Dichev. 2002. "The quality of accruals and earnings: The role of accrual estimation 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. and Dimitris Papanikolaou. 2013. "Organizational capital and the cross-section of expected returns." Journal of Finance 68:1365–406.

Eisfeldt, Andrea L., Edward Kim and Dimitris Papanikolaou. 2022. "Intangible value." Critical Finance Review 11:299–332.

Erickson, Timothy and Toni M. 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. 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. 1997. "Industry costs of equity." Journal of Finance 43:153–193.

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

Fama, Eugene F. and Kenneth R. French. 2008b. "Dissecting anomalies." Journal of Finance 63:1653–1678.

Feng, Guanhao, Stefano Giglio and Dacheng Xiu. 2020. "Taming the factor zoo: A test of new factors." Journal of Finance 75(3):1327–1370.

Gilchrist, Simon. and Charles P. Himmelberg. 1995. "Evidence on the role of cash flow for investment." Journal of Monetary Economics 36:541–572.

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

Gonçalves, Andrei S., 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.

Hall, Robert E. 2004. "Measuring factor adjustment costs." Quarterly Journal of Economics 86:971–987.

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.

Kogan, Leonid, Jun Li and Harold H. Zhang. 2019. "Operating hedge and gross profitability premium." The Sloan School of Management, Massachusetts Institute of Technology.

Kozak, Serhiy, Stefan Nagel and Shrihari Santosh. 2020. "Shrinking the cross-section." Journal of Financial Economics 135(2):271–292.

Lewellen, Jonathan. 2015. "The cross-section of expected stock returns." Critical Finance Review 4:1–44.

Li, Dongmei. 2011. "Financial constraints, R&D investment, and stock returns." Review of Financial Studies 24(9):2974–3007.

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.

Lin, Xiaoji. 2012. "Endogenous technological progress and the cross-section of stock returns." Journal of Financial Economics 103:411–427.

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 M. Whited and Lu Zhang. 2009. "Investment-based expected stock returns." Journal of Political Economy 117(6):1105–1139.

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.

Papanikolaou, Dimitris. 2011. "Investment shocks and asset prices." Journal of Political Economy 119(4):639–685.

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

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

Titman, Sheridan, KC John Wei and Feixue Xie. 2013. "Market development and the asset growth effect: International evidence." Journal of Financial and Quantitative Analysis 48(5):1405–1432.

Watanabe, Akiko, Yan Xu, Tong Yao and Tong Yu. 2013. "The asset growth effect: Insights from international equity markets." Journal of Financial Economics 108(2):529–563.

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. 2007. “Accruals, investment, and the accrual anomaly.” The Accounting Review 82(5):1333–1363.

Appendix

A 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, 2008a) 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.

OCA and IOCA (Eisfeldt and Papanikolaou, 2013; Hou, Xue and Zhang, 2020) We measure OCA and IOCA as organizational capital-to-assets ratio and industry-adjusted organizational capital-to-assets ratio. We construct the stock of organization capital using the perpetual inventory method:

$$OC_{it} = (1-\delta)OC_{it-1} + SG\&A_{it}/CPI_t, \tag{A.1}$$

where OC_{it} is the organization capital of firm i at the end of year t, $SG\&A_{it}$ is selling, general, and administrative ($SG\&A$) expenses (Compustat annual item XSGA), CPI_t denotes the consumer price index, and δ is the annual depreciation rate of OC. The initial stock of OC

is defined as:

$$OC_{i0} = SG\&A_{i0}/(g + \delta), \tag{A.2}$$

where $SG\&A_{i0}$ is the first valid $SG\&A$ observation (zero or positive) for firm i. g is the long-term growth rate of $SG\&A$ and is assumed to be 10% for $SG\&A$. δ is the depreciation rate for OC and is assumed to be 15%. Missing $SG\&A$ values after the starting date are treated as zero. OCA is defined as OC scaled by total assets. To calculate IOCA, we demean a firm's OCA by its industry mean and then divide the demeaned OCA by the standard deviation of OCA within its industry. We use the Fama and French (1997) 17-industry classification. We winsorize OCA at the 1% and 99% levels of all firms each year before the industry standardization to alleviate the impact of outliers. At the end of June of each year t, we use NYSE breakpoints to sort stocks into deciles based on OCA and IOCA 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. We require $SG\&A$ to be nonmissing in calendar year t-1 because this $SG\&A$ receives the highest weight in OC. We also exclude firms with zero OC.

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 γ reports the the posterior means of the marginal product parameter γ; column CI_γ reports the 95% credible intervals of γ; and column $\sigma(\gamma)$ in Panels C and D report the time-series standard deviation of the posterior means of γ when the parameters are time-varying. Similar definitions apply to columns a, CI_a, and $\sigma(a)$ for the adjustment costs parameter a. Column q^K reports the value-weighted average of model-implied marginal q for physical capital, defined as $q_{it}^K \equiv 1 + a_{jt}(1-\tau_t)I_{it}/K_{it}$ for firm i in industry j at time t. For specifications with time-varying parameters, we report the time-series averages of these statistics due to space constraint.

Industry	γ	CI_γ	$\sigma(\gamma)$	a	CI_a	$\sigma(a)$	q^K
	Panel A: Constant parameters $\boldsymbol{\theta}$						
All industries	0.15	$[0.15, 0.15]$		0.14	$[0.13, 0.14]$		1.02
	Panel B: Industry-specific parameters $\boldsymbol{\theta}_j$						
Consumer Nondurables	0.13	[0.13, 0.13]		0.07	[0.05, 0.09]		1.01
Consumer Durables	0.16	[0.15, 0.16]		0.29	[0.25, 0.34]		1.06
Manufacturing	0.16	[0.16, 0.16]		0.13	[0.11, 0.14]		1.02
Energy	0.20	[0.20, 0.21]		0.10	[0.09, 0.12]		1.01
Business Equipment	0.23	[0.23, 0.24]		0.78	[0.75, 0.81]		1.23
Telecom	0.24	[0.24, 0.25]		0.07	[0.06, 0.08]		1.01
Wholesale & Retail	0.09	[0.08, 0.09]		0.18	[0.16, 0.20]		1.03
Healthcare	0.17	[0.17, 0.18]		0.12	[0.10, 0.15]		1.02
Utilities	0.29	[0.29, 0.30]		0.01	[0.00, 0.02]		1.00
Others	0.15	[0.15, 0.15]		0.15	[0.14, 0.16]		1.03
	Panel C: Time-series average of time-varying parameters $\boldsymbol{\theta}_t$						
All industries	0.15	[0.14, 0.15]	0.09	0.29	[0.26, 0.32]	0.30	1.02
	Panel D: Time-series average of industry-specific and time-varying parameters $\boldsymbol{\theta}_{jt}$						
Consumer Nondurables	0.13	[0.11, 0.14]	0.09	0.42	[0.29, 0.55]	0.43	1.05
Consumer Durables	0.17	[0.14, 0.19]	0.18	1.15	[0.83, 1.47]	1.13	1.21
Manufacturing	0.16	[0.15, 0.17]	0.11	0.57	[0.51, 0.65]	0.98	1.08
Energy	0.20	[0.18, 0.22]	0.13	0.45	[0.40, 0.48]	0.54	1.06
Business Equipment	0.23	[0.21, 0.24]	0.19	1.78	[1.66, 1.83]	2.00	1.53
Telecom	0.28	[0.25, 0.30]	0.21	0.71	[0.65, 0.76]	0.66	1.09
Wholesale & Retail	0.08	[0.07, 0.09]	0.06	0.87	[0.77, 0.96]	0.98	1.13
Healthcare	0.19	[0.17, 0.21]	0.16	0.59	[0.44, 0.73]	0.68	1.09
Utilities	0.29	[0.25, 0.32]	0.17	0.25	[0.20, 0.32]	0.32	1.02
Others	0.17	[0.15, 0.18]	0.13	0.48	[0.47, 0.54]	0.52	1.09

Table 4: **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:

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

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

where $\overline{\Pi/Y}_{jt} \equiv \sum_{i=1}^{N_{jt}} \varpi_{it-1}\Pi_{it}/Y_{it}$, and $\overline{q}_{jt} \equiv \sum_{i=1}^{N_{jt}} \varpi_{it-1}q_{it}$.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 the estimated value of a for the same industry and time, 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}		0.19
		(3.60)
Constant	0.18	2.16
	(34.83)	(34.49)
$Adj.R^2$	0.078	0.023
Observations	530	530

Table 5: **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 setup ($\boldsymbol{\theta}$) in which the estimated parameters are constant over time and across industries; 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 (under column $\boldsymbol{\theta}_{jt}$) in which the estimated parameters are industry-specific and time-varying. The sample period is from June 1967 to December 2016.

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

Table 6: **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 7: **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. 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. All correlation coefficients are significant at the 1% level.

	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 8: **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 9: **Comparative statics of fundamental anomaly premiums under the baseline estimation**

This table reports the alpha of the fundamental factor premium, defined as $\alpha \equiv r^S - r^F$, from the baseline estimation and six 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. In the comparative static analysis denoted $\boldsymbol{\theta}_j$, time variation in parameter values is shut down, while in the comparative static analysis denoted $\boldsymbol{\theta}_t$, industry variation in parameter values is shut down. The t-values reported in parentheses are adjusted for heteroscedasticity and autocorrelations of up to 24 lags.

	BM	R11	I/A	ROE	Size	Accruals	NSI	$\Delta PI/A$	GP/A	ROA	RD/M	Ad/M
Baseline	6.27	1.87	−3.16	3.05	1.14	−10.33	−4.56	0.04	−3.39	2.63	3.47	−1.39
	(3.33)	(0.57)	(−2.06)	(1.33)	(0.34)	(−6.29)	(−2.90)	(0.03)	(−2.64)	(1.11)	(1.42)	(−0.59)
$\overline{I_{it}/K_{it}}$	13.36	−0.63	−9.88	2.35	1.42	−9.54	−6.91	−5.73	−7.07	1.79	3.90	3.46
	(5.90)	(−0.19)	(−5.47)	(0.96)	(0.40)	(−5.29)	(−3.64)	(−3.92)	(−4.69)	(0.73)	(1.17)	(1.38)
$\overline{I_{it+1}/K_{it+1}}$	−0.12	8.29	0.92	5.36	0.51	−11.67	−3.26	3.75	0.40	4.36	5.13	−5.42
	(−0.06)	(2.24)	(0.55)	(1.93)	(0.15)	(−7.01)	(−2.15)	(2.31)	(0.29)	(1.59)	(2.07)	(−2.16)
$\overline{Y_{it+1}/K_{it+1}}$	−62.26	14.88	5.53	25.52	−14.71	11.17	−36.32	−8.33	109.05	31.19	27.80	6.91
	(−7.47)	(3.69)	(1.72)	(7.62)	(−4.23)	(4.08)	(−5.14)	(−2.36)	(11.25)	(6.67)	(8.17)	(2.54)
$\overline{W_{it+1}/K_{it+1}}$	12.25	1.19	−1.79	2.52	15.83	−19.60	−4.04	9.41	−7.58	3.03	−2.81	−0.44
	(6.03)	(0.32)	(−0.48)	(1.06)	(3.95)	(−7.90)	(−2.00)	(2.48)	(−3.68)	(1.35)	(−0.83)	(−0.13)
$\boldsymbol{\theta}_j$	7.38	10.16	−6.24	3.80	0.98	−10.48	−4.85	−1.58	−2.90	3.40	6.18	−0.29
	(3.25)	(4.06)	(−3.42)	(2.10)	(0.28)	(−5.96)	(−2.44)	(−0.87)	(−1.47)	(1.77)	(1.49)	(−0.10)
$\boldsymbol{\theta}_t$	10.52	8.00	−5.62	2.36	2.73	−14.38	−3.07	−2.67	−11.61	2.35	8.22	−7.79
	(4.20)	(2.93)	(−3.18)	(1.29)	(0.73)	(−7.68)	(−1.49)	(−1.48)	(−5.41)	(1.24)	(2.07)	(−2.42)

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)		

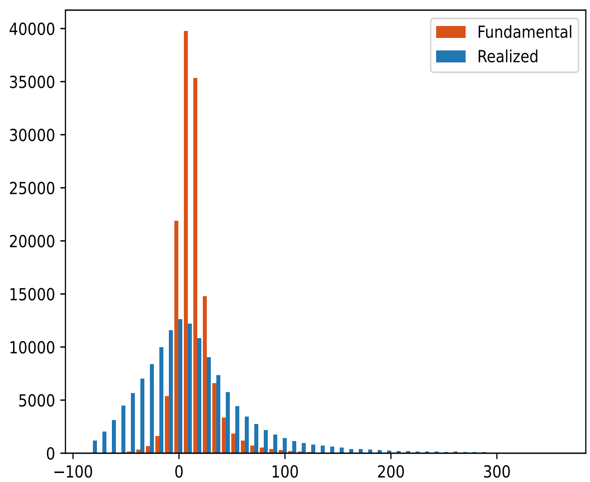
Table 11: **Realized returns on decile portfolios formed on the predicted returns, October 1980 - December 2016**

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.

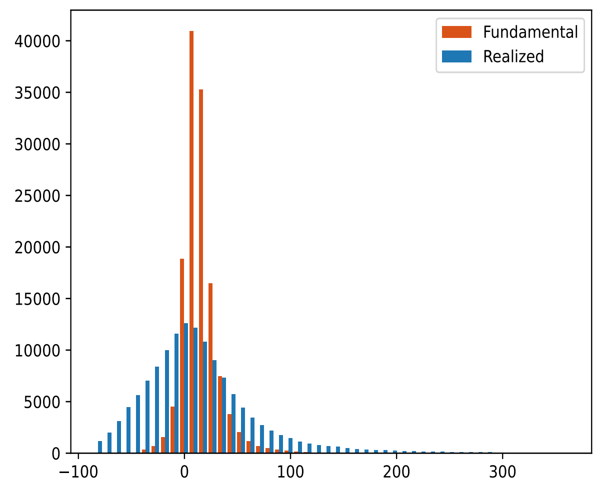
	L	2	3	4	5	6	7	8	9	H	H-L
R-rf	0.36	0.69	0.69	0.76	0.62	0.76	0.72	0.77	0.54	0.80	0.45
	(1.26)	(3.16)	(3.54)	(3.70)	(3.14)	(3.81)	(3.39)	(3.63)	(2.27)	(3.20)	(2.45)
α_{CAPM}	-0.36	0.09	0.10	0.15	0.05	0.17	0.09	0.11	-0.17	0.07	0.43
	(-2.40)	(1.04)	(1.47)	(1.79)	(0.72)	(1.94)	(1.12)	(1.31)	(-1.72)	(0.61)	(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)

Figure 1: **Distribution of firm-level returns: realized vs. fundamental**

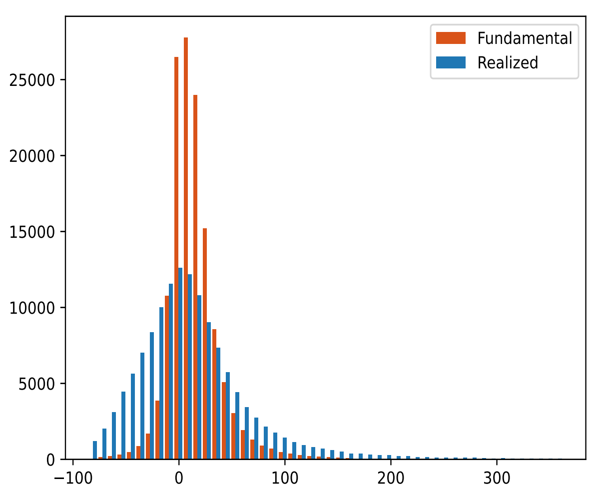
This table presents the histograms of the realized (in blue) and fundamental (in orange) firm-level stock returns from June 1967 to December 2016 under the specifications with constant, industry-specific, time-varying, and industry-specific and time-varying parameter values in Panel (a) - (d), respectively. Returns are in percentage per annum. The number of observations is 136,598. Observations are trimmed at 0.5 and 99.5 percentiles for illustration purposes.



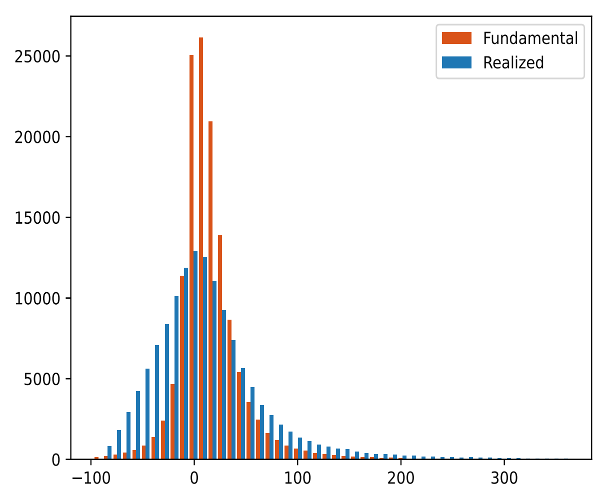
(a) Constant parameters



(b) Industry-specific parameters



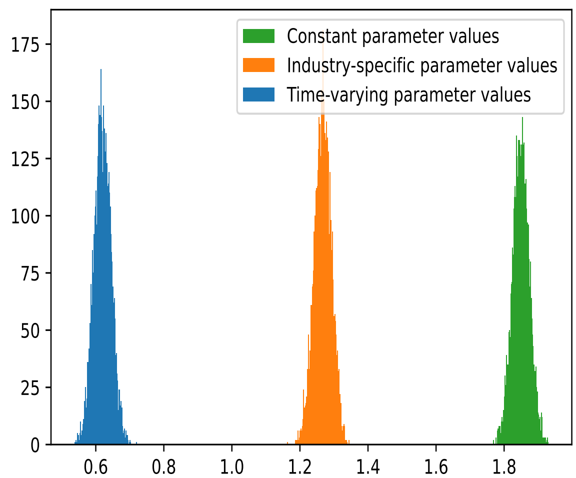
(c) Time-varying parameter values



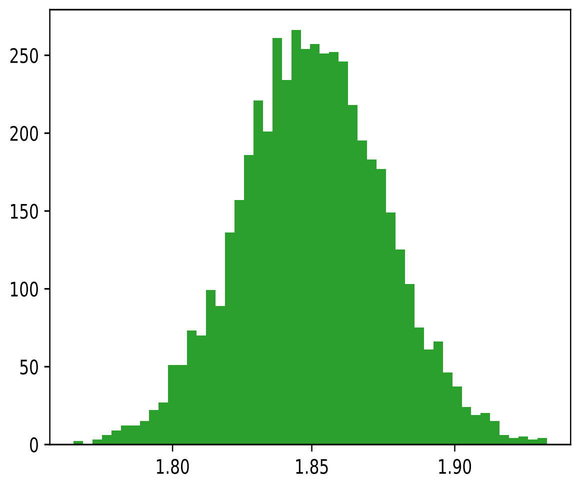
(d) Industry-specific and time-varying parameters

Figure 2: **Performance comparison**

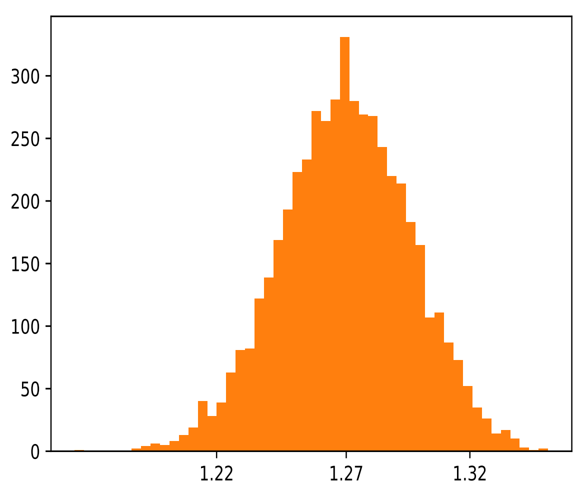
This figure plots the posterior distributions of the differences in m.a.e. among the four specifications. The differences in m.a.e. between the benchmark specification, defined as the one with the lowest m.a.e. (i.e., specification with industry-specific and time-varying parameters), and an alternative specification is given by $d^a = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N_t}\sum_{i=1}^{N_t}\left(\left|r^S_{it+1} - r^{F(a)}_{it+1}\right| - \left|r^S_{it+1} - r^{F(b)}_{it+1}\right|\right)$, where $r^{F(b)}$ and $r^{F(a)}$ are the fundamental returns under the benchmark specification b and under an alternative specification a, respectively, and d^a is in percentage per annum. The three alternative specifications are the ones with constant, industry-specific, and time-varying parameter values, respectively.



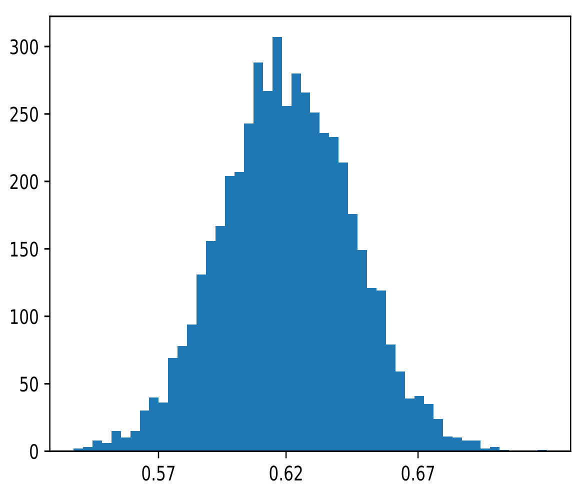
(a) Three alternative estimations



(b) Constant parameters



(c) Industry-specific parameters



(d) Time-varying parameters

Figure 3: **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.

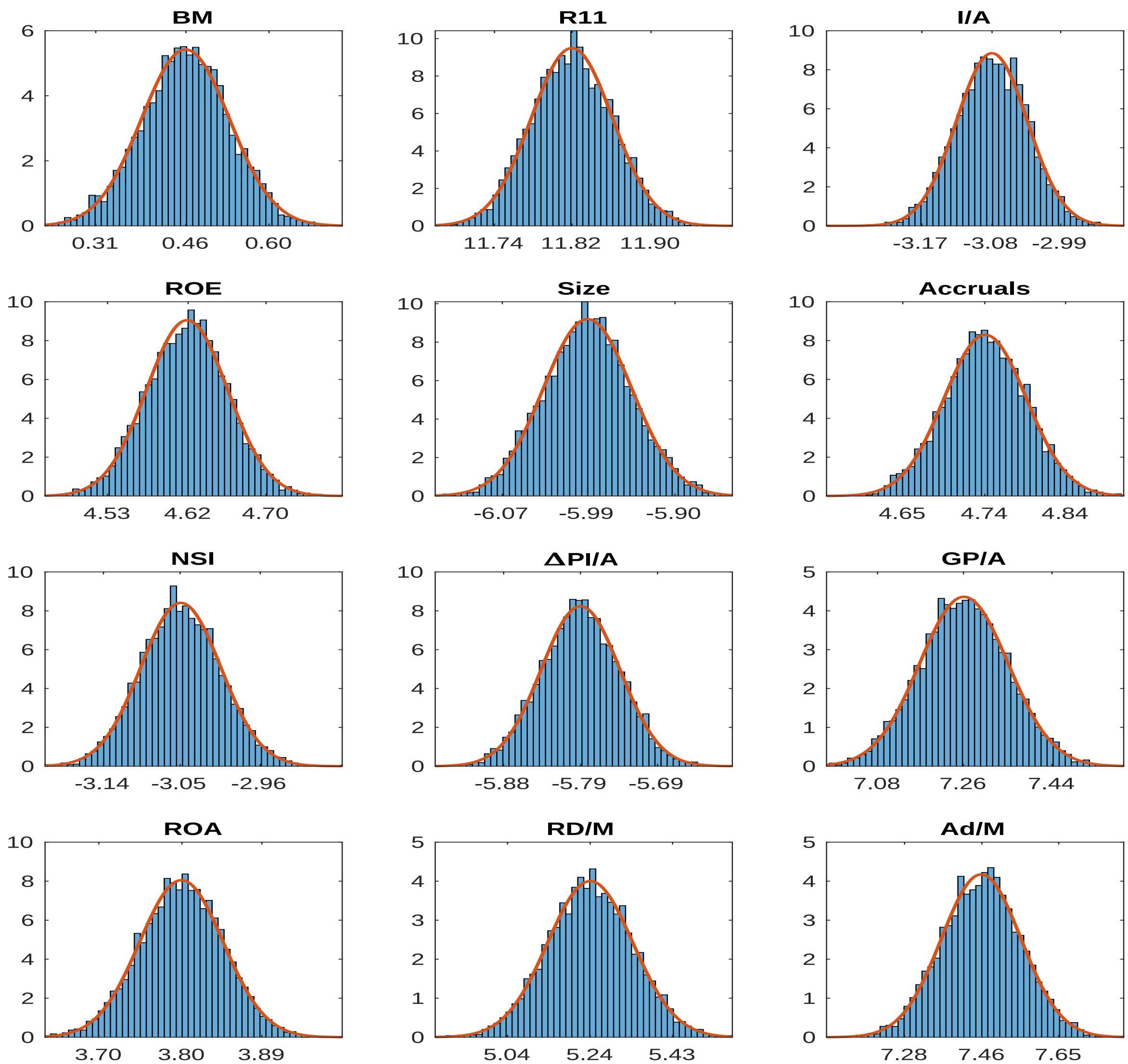


Figure 4: **Time series of factor premiums: realized vs. fundamental**

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.

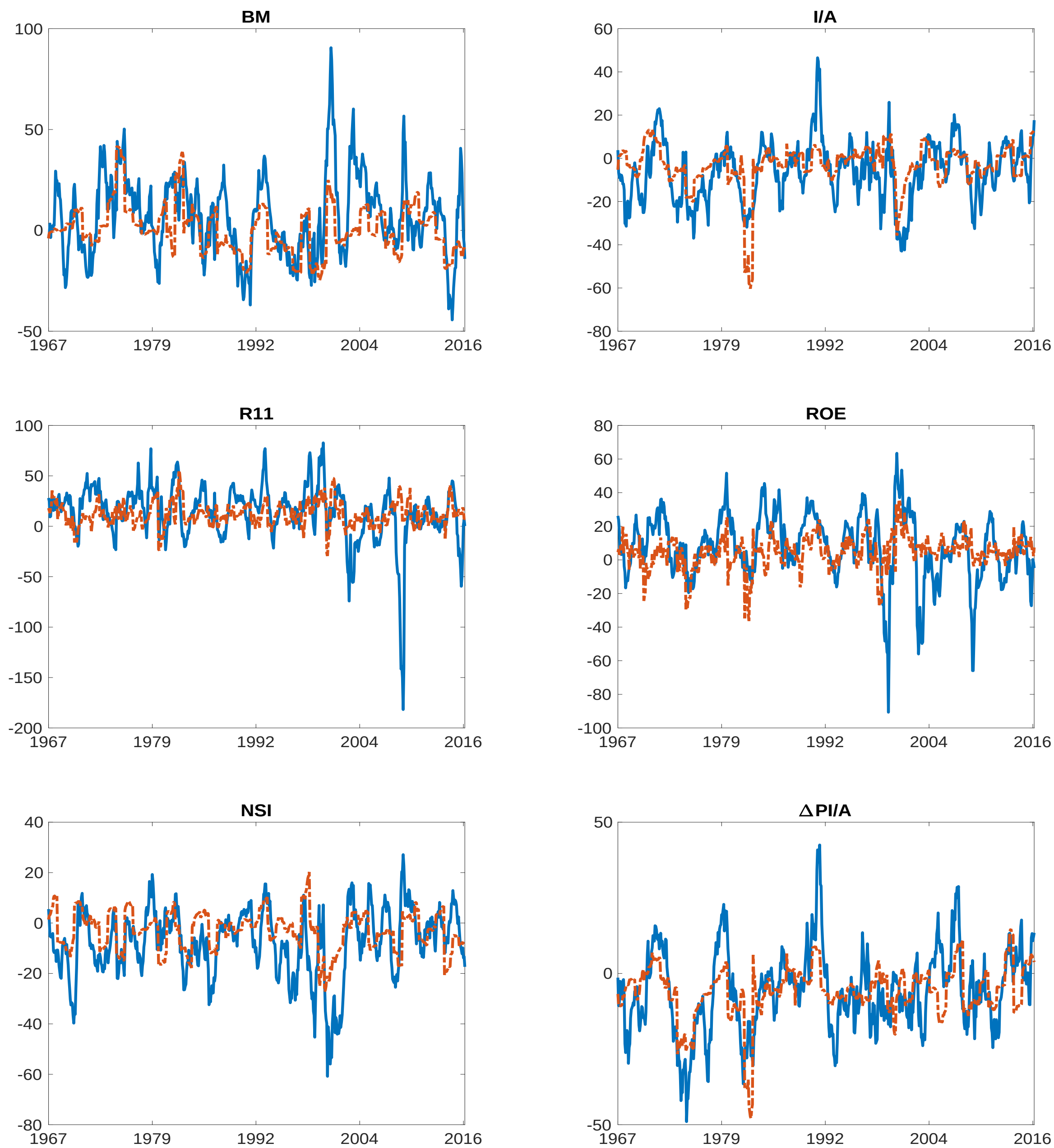


Figure 4: **Time series of factor premiums: realized vs. fundamental (continued)**

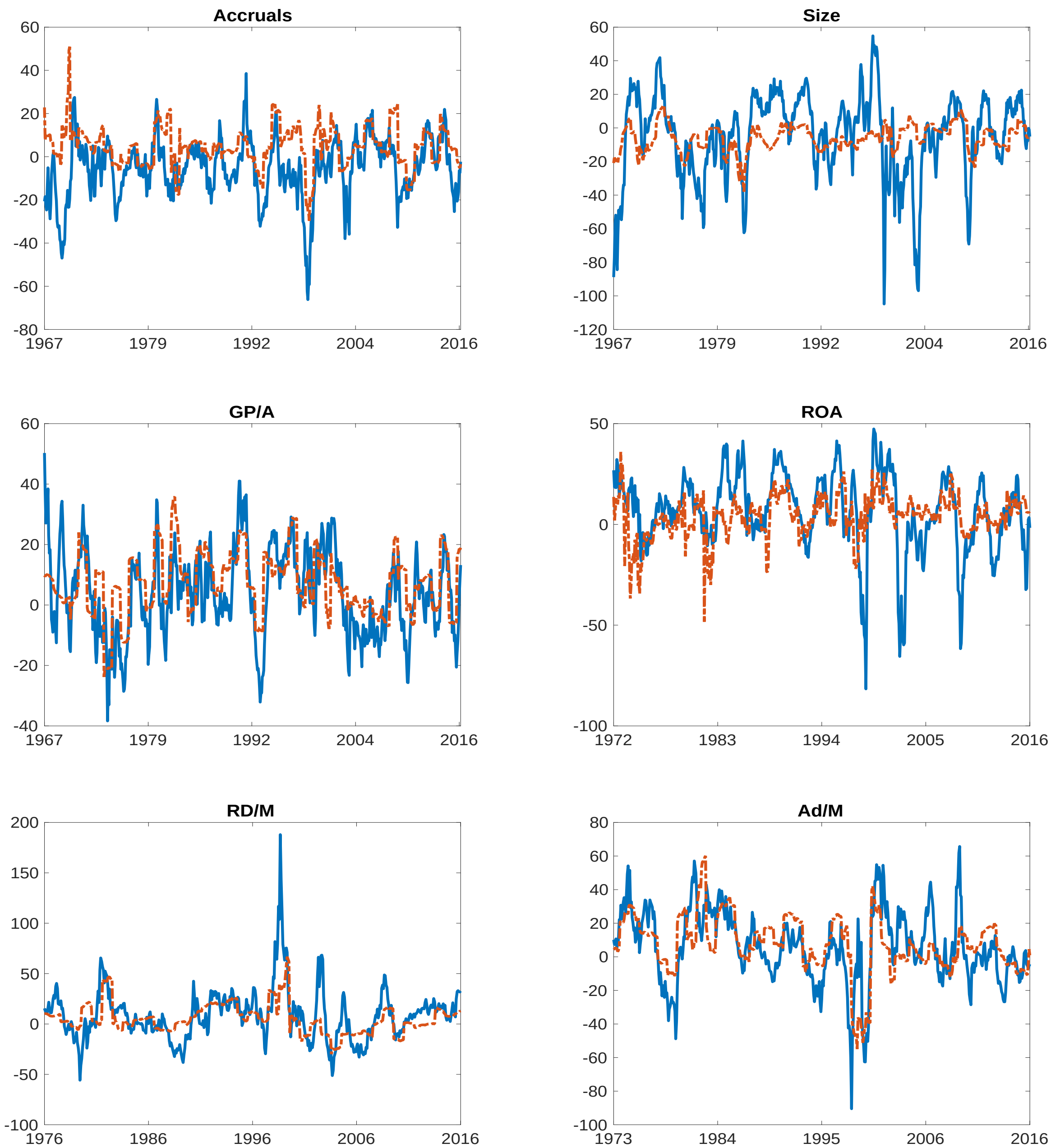
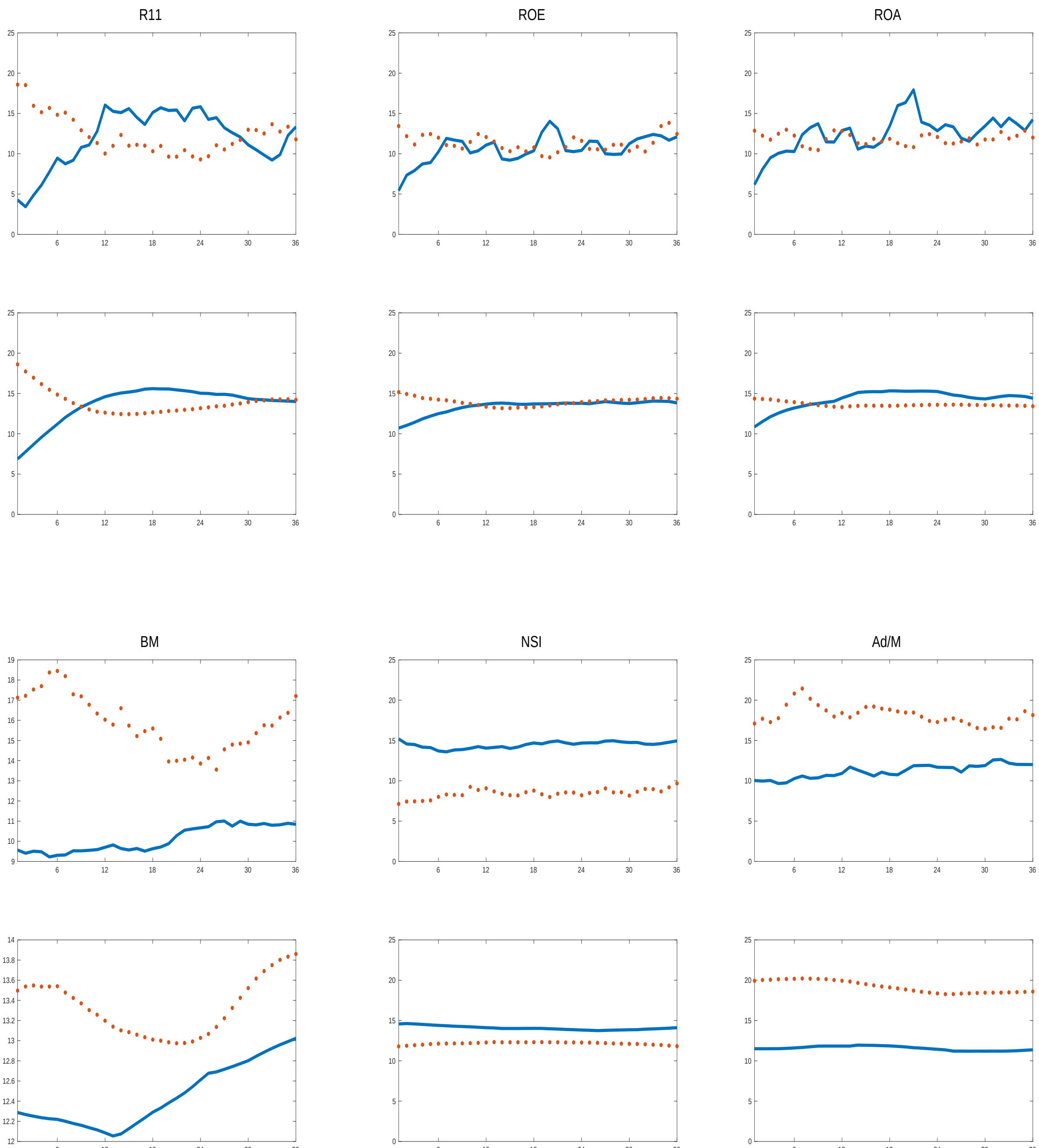
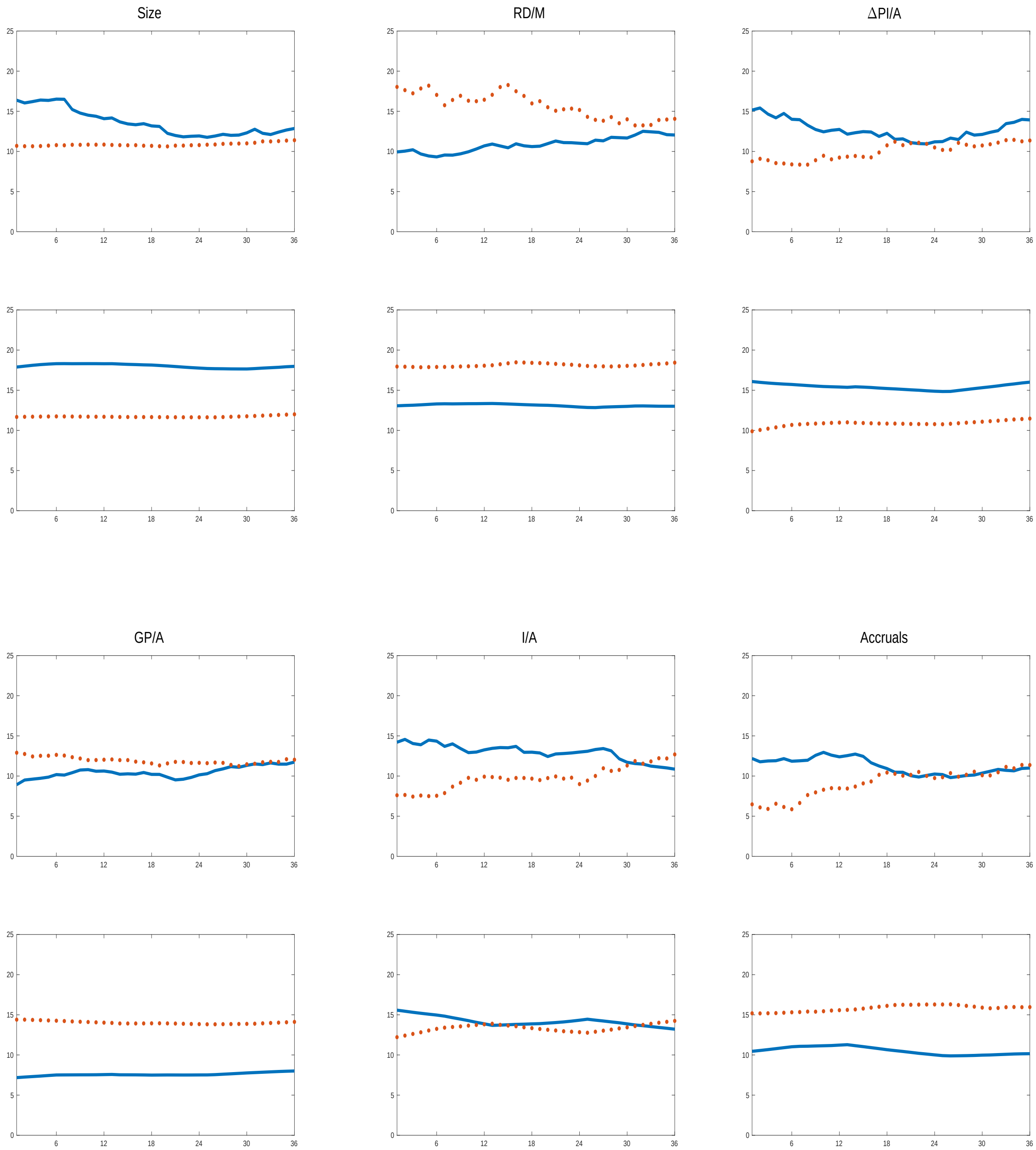


Figure 5: **Persistence of factor premiums: realized vs. fundamental**

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

Internet Appendix: “Fundamental Anomalies”

(for Online Publication Only)

July, 2022

Abstract

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

A Derivations in the two-capital Model

We show that firm's profit is constant-returns-to-scale in physical and working capital when the adjustment intermediate input is optimally chosen in Section A.1, and that stock return equals levered investment return in Section A.2. We also derive the fundamental stock return in the extended models with asymmetric adjustment costs in Section A.3 and with adjustment costs in the working capital investment in Section A.4.

A.1

Let the production function be $Y_{it} \equiv Y(K_{it}, W_{it}, S_{it}, X_{it}) = X_{it} K_{it}^{\gamma_{jt}^K} W_{it}^{\gamma_{jt}^W} S_{it}^{1-\gamma_{jt}^K-\gamma_{jt}^W}$ for firm i in industry j at time t, 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_{jt}^K, γ_{jt}^W, and $1-\gamma_{jt}^K-\gamma_{jt}^W$, respectively. The operating profits function solves the static optimization problem:

$$\Pi(K_{it}, W_{it}, X_{it}) = \max_{\{S_{it}\}} X_{it} K_{it}^{\gamma_{jt}^K} W_{it}^{\gamma_{jt}^W} S_{it}^{1-\gamma_{jt}^K-\gamma_{jt}^W} - p_t^S S_{it}.$$

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

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

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

$$\Pi(K_{it}, W_{it}, X_{it}) = (\gamma_{jt}^K + \gamma_{jt}^W) X_{it}^{\frac{1}{\gamma_{jt}^K+\gamma_{jt}^W}} \left(\frac{1-\gamma_{jt}^K-\gamma_{jt}^W}{p_t^S}\right)^{\frac{1-\gamma_{jt}^K-\gamma_{jt}^W}{\gamma_{jt}^K+\gamma_{jt}^W}} K_{it}^{\frac{\gamma_{jt}^K}{\gamma_{jt}^K+\gamma_{jt}^W}} W_{it}^{\frac{\gamma_{jt}^K}{\gamma_{jt}^K+\gamma_{jt}^W}}.$$

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 $\gamma^K_{jt}/(\gamma^K_{jt} + \gamma^W_{jt})$ and $\gamma^W_{jt}/(\gamma^K_{jt} + \gamma^W_{jt})$, respectively, sum to one. In particular, $\partial\Pi_{it}/\partial K_{it} = [\gamma^K_{jt}/(\gamma^K_{jt} + \gamma^W_{jt})](\Pi_{it}/K_{it}) = \gamma^K_{jt} Y_{it}/K_{it}$. Similarly, $\partial\Pi_{it}/\partial W_{it} = \gamma^W_{jt} Y_{it}/W_{it}$.

A.2

The optimization problem of firm i can be written as:

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

where $D_{it} = (1-\tau_t)(\Pi_{it} - \Phi_{it}) - I_{it} - \Delta W_{it} + B_{it+1} - r^B_{it} B_{it} + \tau_t \delta_{it} K_{it} + \tau_t (r^B_{it} - 1) B_{it}$. Let q^K_{it} and q^W_{it} 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:

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

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

$$q^K_{it} = 1 + (1 - \tau_t) \frac{\partial \Phi_{it}}{\partial I_{it}} \tag{A.3}$$

$$q^W_{it} = 1 \tag{A.4}$$

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

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

$$1 = E_t[M_{t+1}(r^B_{it+1} - (r^B_{it+1} - 1)\tau_{t+1})] = E_t[M_{t+1} r^{Ba}_{it+1}] \tag{A.7}$$

Combining equations (A.3) and (A.5) leads to

$$E_t[M_{t+1} r^K_{it+1}] = 1$$

where

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

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

$$E_t[M_{t+1} r^W_{it+1}] = 1$$

where

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

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

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

we proceed in three steps:

1. Show that firm asset value V^a_{it} can be written as

$$V^a_{it} = P_{it} + B_{it+1} = \mathbb{E}_t\left[\sum_{s=1}^{\infty} M_{t+s} D^a_{it+s}\right]$$

where

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

Proof:

$$V^a_t \;\; = \;\; P_t + B_{t+1} = \mathbb{E}_t\left[M_{t+1}\left(D_{t+1} + P_{t+1}\right)\right] + B_{t+1}$$

$$
\begin{aligned}
&= \mathbb{E}_t\left[M_{t+1}\left[(1-\tau_{t+1})(\Pi_{t+1}-\Phi_{t+1})+\tau_{t+1}\delta_{t+1}K_{t+1}-I_{t+1}-\Delta W_{it+1}+B_{t+2}-r^B_{t+1}B_{t+1}\right.\right.\\
&\quad \left.\left.+\tau_{t+1}(r^B_{t+1}-1)B_{t+1}+P_{t+1}\right]\right]+B_{t+1} \qquad \text{(A.8)}
\end{aligned}
$$

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

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

Substitute the above equation into equation (A.8) and get

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

Q.E.D.

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

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

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

where the second equality is derived using equations (A.3), (A.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^K_{it+1} and r^W_{it+1} and the proof in Step 2, it is straightforward to show that

$$
\begin{aligned}
w^K_{it} r^K_{it+1} + (1 - w^K_{it}) r^W_{it+1} &= \frac{q^K_{it} K_{it+1} r^K_{it+1} + W_{it+1} r^W_{it+1}}{q^K_{it} K_{it+1} + W_{it+1}} \\
&= \frac{D^a_{it+1} + q^K_{it+1} K_{it+2} + W_{t+2}}{q^K_{it} K_{it+1} + W_{it+1}} \\
&= \frac{D^a_{it+1} + P_{t+1} + B_{t+2}}{P_t + B_{t+1}} \\
&= w^B_{it} r^{Ba}_{it+1} + (1 - w^B_{it}) r^S_{it+1}\,.
\end{aligned}
$$

Q.E.D.

It is thus easy to show that the fundamental stock return, r^F_{it+1}, is given by

$$
\begin{aligned}
r^F_{it+1} &\equiv f\left(X_{it}, X_{it+1} | \theta_t, \theta_{t+1}\right) \\
&= \left\{(1 - \tau_{t+1}) \left[\left(K_{it+1}\frac{\partial \Pi_{it+1}}{\partial K_{it+1}} + W_{it+1}\frac{\partial \Pi_{it+1}}{\partial W_{it+1}}\right) - \frac{\partial \Phi_{it+1}}{\partial K_{it+1}}\right] + \tau_{t+1}\delta_{it+1} + \frac{W_{it+1}}{K_{it+1}}\right. \\
&\quad \left. +(1 - \delta_{it+1}) \left[1 + (1 - \tau_{t+1})\frac{\partial \Phi_{it+1}}{\partial I_{it+1}}\right]\right\} \Big/ \left\{\left(1 - w^B_{it}\right) \left[1 + (1 - \tau_t)\frac{\partial \Phi_{it}}{\partial I_{it}} + \frac{W_{it+1}}{K_{it+1}}\right]\right\} \\
&\quad - \frac{w^B_{it} r^{Ba}_{it+1}}{1 - w^B_{it}}\,, \qquad \text{(A.9)}
\end{aligned}
$$

and Section A.1 shows that $K_{it+1}\partial\Pi_{it+1}/\partial K_{it+1} + W_{it+1}\partial\Pi_{it+1}/\partial W_{it+1} = \gamma_{jt+1} Y_{it+1}$, where $\gamma_{jt+1} = \gamma^K_{jt+1} + \gamma^W_{jt+1}$.

Note that the derivation in this Section holds for any model with production and adjustment cost functions that are homogeneous of degree one in capital and investment. It also holds if model parameters are time varying as long as parameter values are exogenously given. The role of time-varying parameters in the model is analogous to that of time-varying productivity shocks.

A.3 Extended model with Asymmetric adjustment costs

For the quadratic adjustment cost function, we ignore the corner solution with investment being zero since it happens in less than 0.1% of the observations in our sample. Equation (6) still holds if we replace the adjustment costs parameter a_{jt} for industry j at time t with $a_{jt} = a^{+}_{jt}\mathbb{I}_{I_{it}>=0} + a^{-}_{jt}(1 - \mathbb{I}_{I_{it}>=0})$ where $\mathbb{I}_{I_{it}>=0}$ is an indicator equal to one if investment is positive and zero otherwise. For the exponential adjustment cost function,

$$\Phi_{it} \equiv \frac{\theta_{jt}}{\nu^2_{jt}} \left[\exp\left(-\nu_{jt} \frac{I_{it}}{K_{it}} \right) + \nu_{jt} \frac{I_{it}}{K_{it}} - 1 \right] ,$$

equation (A.9) implies that the fundamental return is given by

$$\begin{aligned} r^{a}_{it+1} = \Bigg\{ & (1-\tau_{t+1}) \left[\gamma_{jt+1} \frac{Y_{it+1}}{K_{it+1}} - \frac{\theta_{jt+1}}{v^2_{jt+1}} \left[\exp\left\{ -\frac{I_{it+1}}{K_{it+1}} v_{jt+1} \right\} \left(1 + \frac{I_{it+1}}{K_{it+1}} v_{jt+1} \right) - 1 \right] \right] \\ & + \tau_{t1+} \delta_{it+1} + (1-\delta_{it+1}) \left[1 + (1-\tau_{t+1}) \frac{\theta_{jt+1}}{v_{jt+1}} \left(1 - \exp\left\{ -v_{jt+1} \frac{I_{it+1}}{K_{it+1}} \right\} \right) \right] \\ & + \frac{W_{it+1}}{K_{it+1}} \Bigg\} \Bigg/ \left\{ \left(1 - w^B_{it}\right) \left[1 + (1-\tau_t) \frac{\theta_{jt}}{v_{jt}} \left(1 - \exp\left\{ -\frac{I_{it}}{K_{it}} v_{it} \right\} \right) + \frac{W_{it+1}}{K_{it+1}} \right] \right\} + \frac{w^B_{it} r^{Ba}_{it+1}}{1 - w^B_{it}} . \end{aligned} \tag{A.10}$$

The estimation results are presented in Table A.4 and the implied fundamental anomalies are presented in Table A.5.

A.4 Extended model with adjustment costs on the working capital investment

We extend the baseline model and add quadratic adjustment costs in working capital investment: $\Phi(\Delta W_{it}, W_{it}) = \frac{b_{jt}}{2}\left(\frac{\Delta W_{it}}{W_{it}}\right)^2 W_{it}$, where b_{jt} is the adjustment costs parameter for industry j at time t. We present the fundamental stock return in equation (A.11). Detailed derivations can be found in Section E of Gonçalves, Xue and Zhang (2020) Online Appendix.

$$\begin{aligned} r^F_{it+1} = & \left\{(1-\tau_{t+1})\left[\gamma_{jt+1}\left(\frac{Y_{it+1}}{K_{it+1}}\right) + \frac{a_{jt+1}}{2}\left(\frac{I_{it+1}}{K_{it+1}}\right)^2 + \frac{b_{jt+1}}{2}\left(\frac{\Delta W_{it+1}}{W_{it+1}}\right)^2 \frac{W_{it+1}}{K_{it+1}}\right.\right. \\ & \left.\left. + b_{jt+1}\left(\frac{\Delta W_{it+1}}{W_{it+1}}\right)\frac{W_{it+1}}{K_{it+1}}\right] + \tau_{t+1}\delta_{it+1} + (1-\delta_{it+1})\left[1 + (1-\tau_{t+1})a_{jt+1}\left(\frac{I_{it+1}}{K_{it+1}}\right)\right] + \frac{W_{it+1}}{K_{it+1}}\right\} \\ & \Bigg/ \left\{(1-w^B_{it})\left[1 + (1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right) + \frac{W_{it+1}}{K_{it+1}}\left(1 + (1-\tau_t)b_{jt}\frac{\Delta W_{it}}{W_{it}}\right)\right]\right\} - \frac{w^B_{it} r^{Ba}_{it+1}}{1-w^B_{it}}, \end{aligned} \tag{A.11}$$

The estimation results are presented in Table A.6 and the implied fundamental anomalies are presented in Table A.7.

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

$$\begin{aligned}
&\mathcal{P}(\boldsymbol{\theta}, \boldsymbol{\sigma} | \boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba}) \\
=&\frac{\mathcal{P}(\boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba})}{\mathcal{P}(\boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba})} \\
\propto&\mathcal{P}(\boldsymbol{r}^S | \boldsymbol{X}, \boldsymbol{r}^{Ba}; \boldsymbol{\theta}, \boldsymbol{\sigma}) \mathcal{P}(\boldsymbol{\theta}, \boldsymbol{\sigma}) \\
=&\mathcal{P}(\boldsymbol{r}^S | \boldsymbol{X}, \boldsymbol{r}^{Ba}; \boldsymbol{\theta}, \boldsymbol{\sigma}) \mathcal{P}(\boldsymbol{\theta} | \boldsymbol{\sigma}) \pi(\boldsymbol{\sigma}),
\end{aligned} \tag{B.1}$$

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]

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

The newly defined ret can be seen as a function of latent variables, which we denote function-

[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.

ally, for firm i that belongs to industry j at time $t+1$ but to industry j' at time t, as

$$ret_{it+1} \equiv \Lambda_{it+1}\Big(\gamma_{jt+1}, a_{jt+1}, a_{j't}\Big) + \sigma_r e^r_{it+1}, \tag{B.3}$$

where $\Lambda_{it+1}(\gamma_{jt+1}, a_{jt+1}, a_{j't}) = \frac{\varpi_{it}^{1/2}}{1-w_{it}^B} r^K_{it+1}$ and r^K_{it+1} 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 prior specifications, equation (B.1) can be written in proportion as:

$$\begin{aligned}\mathcal{P}(\boldsymbol{\theta}, \boldsymbol{\sigma} | \boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba}) \propto & \prod_{t=0}^{T-1} \prod_{i=1}^{N_{t+1}} \mathcal{N}\Big(ret_{it+1}; \Lambda_{it+1}, \sigma_r^2\Big) \\ & \cdot \prod_{t=0}^{T-1} \prod_{j=1}^{N_d} \mathcal{N}\Big(\gamma_{jt+1}; \gamma_{jt}, \sigma_\gamma^2\Big) \cdot \prod_{t=0}^{T-1} \prod_{j=1}^{N_d} \mathcal{N}\Big(a_{jt+1}; a_{jt}, \sigma_a^2\Big) \\ & \cdot \mathcal{IG}\Big(\sigma_r^2; \kappa_1^r, \kappa_2^r\Big) \cdot \mathcal{IG}\Big(\sigma_\gamma^2; \kappa_1^\gamma, \kappa_2^\gamma\Big) \cdot \mathcal{IG}\Big(\sigma_a^2; \kappa_1^a, \kappa_2^a\Big),\end{aligned} \tag{B.4}$$

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

[2] ϖ is not present in the formula to be consistent with equation (B.1). Besides, it has no influence in the following derivation. Also note that in equation (B.4), 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.

conditional posteriors:

$$\mathcal{P}(\boldsymbol{\theta}, \boldsymbol{\sigma}|\boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba}) \Longleftrightarrow \mathcal{P}(\boldsymbol{\theta}|\boldsymbol{\sigma}, \boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba}) \text{ and } \mathcal{P}(\boldsymbol{\sigma}|\boldsymbol{\theta}, \boldsymbol{X}, \boldsymbol{r}^S, \boldsymbol{r}^{Ba}). \tag{B.5}$$

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 posterior 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, $\mathbb{1} = 0$:

$$p\left(\gamma_{jt+1} \middle| \{\gamma_{jt}\}, \{a_{jt}\}, \sigma_r^2, \sigma_\gamma^2, \sigma_a^2\right) \propto \mathcal{N}\left(\gamma_{jt+1}; \frac{v_1}{u_1}, \frac{1}{u_1}\right), \tag{B.6}$$

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:

$$p\left(a_{jt+1}\middle|\left\{\gamma_{jt+1}\right\},\left\{a_{jt+1}\right\},\sigma_r^2,\sigma_\gamma^2,\sigma_a^2\right)\propto\prod_{t=0}^{T-1}\prod_{i=1}^{N}\mathcal{N}\left(ret_{it+1};\Lambda_{it+1},\sigma_r^2\right)\prod_{t=0}^{T-1}\mathcal{N}\left(a_{jt+1};a_{jt},\sigma_a^2\right). \tag{B.7}$$

We propose from

$$\mathcal{N}\left(a_{jt+1};\frac{v_2}{u_2},\frac{1}{u_2}\right) \tag{B.8}$$

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

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

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:

$$p\left(\sigma_r^2 \middle| \{\gamma_{jt}\}, \{a_{jt}\}, \sigma_\gamma^2, \sigma_a^2\right) \sim \mathcal{IG}\left(\kappa_1^r + \frac{\sum_{t=0}^{T-1} N_{t+1}}{2}, \kappa_2^r + \frac{1}{2} \sum_{t=0}^{T-1} \sum_{i=1}^{N} \left(ret_{it+1} - \Lambda_{it+1}\right)^2\right). \quad \text{(B.9)}$$

$$p\left(\sigma_\gamma^2 \middle| \{\gamma_{jt}\}, \{a_{jt}\}, \sigma_r^2, \sigma_a^2\right) \sim \mathcal{IG}\left(\kappa_1^\gamma + \frac{N_d T}{2}, \kappa_2^\gamma + \frac{1}{2} \sum_{t=0}^{T-1} \sum_{k=1}^{N_d} \left(\gamma_{jt+1} - \gamma_{jt}\right)^2\right). \quad \text{(B.10)}$$

$$p\left(\sigma_a^2 \middle| \{\gamma_{jt}\}, \{a_{jt}\}, \sigma_r^2, \sigma_\gamma^2\right) \sim \mathcal{IG}\left(\kappa_1^a + \frac{N_d T}{2}, \kappa_2^a + \frac{1}{2} \sum_{t=0}^{T-1} \sum_{k=1}^{N_d} \left(a_{jt+1} - a_{jt}\right)^2\right). \quad \text{(B.11)}$$

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 Timing alignment

In this section, we explain the timing of variables used to construct fundamental return and how we align the timing of fundamental return with that of actual return. As we explain in Section 3.1, model-implied fundamental returns are constructed in annual frequency. In the model, time-t stock variables (such as capital K and debt B) are at the beginning of year t, and time-t flow variables (such as sales Y and depreciations) 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.

To construct the fundamental return for firm i from t to $t+1$, r^F_{it+1}, we need accounting information from fiscal year $t-1$ to construct K_{it}, accounting information from fiscal year t to construct K_{t+1} and leverage ratio w_{it} (defined as the ratio of total debt B_{it+1} to the sum of total debt and market equity $V_{it} - D_{it}$), and accounting information from fiscal year $t+1$ to construct Y_{it+1}, δ_{it+1}, and K_{it+2}. Finally, I_{it} and I_{it+1} is constructed based on the law of motion: $I_{it} = K_{it+1} - (1-\delta_{it})K_{it}$ and $I_{it+1} = K_{it+2} - (1-\delta_{it+1})K_{it+1}$, respectively.

The fundamental return r^F_{it+1} computed based on equation (6) corresponds to return of the 12-month period between the 5th month prior the fiscal year t ending month till the 6th month after, namely r^S_{it+1}. Parameters are estimated by matching r^F_{it+1} with r^S_{it+1} for the entire sample.

Our timing alignment is consistent with that in GXZ, who construct monthly fundamental returns from annual accounting variables to match with monthly stock returns. For each month, they take firm-level accounting variables from the fiscal year end that is closest to the month in question to measure (flow) variables dated t in the model and take accounting variables from the subsequent fiscal year end to measure (flow) variables dated $t+1$ in the model.

For comparison with GXZ, we construct fundamental returns in monthly frequency for anomaly portfolios between June 1967 and December 2016. However, it is important to note that the accounting variables underlying the fundamental returns for June 1967 can come from the fiscal year ending in calendar time as early as December 1965, and the accounting variables underlying the fundamental returns for December 2016 can come from as late as May 2018. It also means that in terms of the fundamental returns in annual frequency (in fiscal year), we need to have returns of as early as fiscal year 1967 and as late as fiscal year 2018. For example, if a firm i's fiscal ending month is May and its return on December 2016 is aligned with the fundamental return of fiscal year 2018 (from year 2017 to year 2018). On the other hand, if a firm i's fiscal ending month is December and its return on June 1967 is aligned with the fundamental return of fiscal year 1967.

Equation (6) shows that the year t value of fundamental return depends on the adjustment cost parameters in $t-1$ and t. To compute the fundamental returns in the starting year of our sample, we assume that the t and $t-1$ values of the adjustment cost parameter for the starting year t are the same. For accuracy, returns of the starting year in the estimation are not used in our analysis of portfolio returns. Therefore, parameters are estimated by matching fundamental and actual returns in annual frequency between 1966 and 2018, and fundamental returns between 1967 and 2018 are used to construct monthly returns of anomaly portfolios between June 1967 and December 2016.

D 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]

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

where 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 A.2 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%

[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

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

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.

Figure A.2 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.8 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

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

assuming $\theta_{j0} = \theta_{j1}$, where $\mathbf{0}$ are vectors of zeros of corresponding dimensions.[5] 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.

[5]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

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

E Simulation study

We use simulation studies to examine whether Bayesian MCMC can discover the true parameter values under our model framework, which is highly nonlinear. We 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. 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 constant values, with industry variations only, with time variations only, and with industry and time variations. 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 3.2, we draw from the posterior distributions of the latent variables given the financial and accounting information $\boldsymbol{X}$ and the simulated stock return r^S of this subgroup of firms. The initial guesses, $\boldsymbol{\theta}^{(0)}_{jt+1}$, 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

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

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.

Figure A.2 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.

Figure A.2 shows that the true values of the model parameters are almost always confined in 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. 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.8. in the Internet Appendix.

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 specifications.

F Comparative statistics

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

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

$$\frac{\partial r^F_{it+1}}{\partial(I_{it+1}/K_{it+1})} = \frac{(1-\tau_{t+1})\left(1+\frac{I_{it+1}}{K_{it+1}}-\delta_{it+1}\right)a_{jt+1}}{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)+\frac{W_{it+1}}{K_{it+1}}\right]} > 0 \tag{F.2}$$

$$\frac{\partial r^F_{it+1}}{\partial(Y_{it+1}/K_{it+1})} = \frac{(1-\tau_{t+1})\gamma_{jt+1}}{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)+\frac{W_{it+1}}{K_{it+1}}\right]} > 0 \tag{F.3}$$

$$\frac{\partial r^F_{it+1}}{\partial(W_{it+1}/K_{it+1})} = \frac{1-r^{Fwacc}_{it+1}}{(1-w^B_{it})\left[1+(1-\tau_t)a_{jt}\left(\frac{I_{it}}{K_{it}}\right)+\frac{W_{it+1}}{K_{it+1}}\right]}, \tag{F.4}$$

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} with respect to I_{it}/K_{it}, I_{it+1}/K_{it+1} and Y_{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

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

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 $\tau = 0.39$.

References

Efron, Bradley and Robert J Tibshirani. 1994. An Introduction to the Bootstrap. CRC press.

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

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

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

Table A.1: **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 $t+1$. 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 setup (under column $\boldsymbol{\theta}$) in which the estimated parameters are constant over time and across industries; 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 baseline setup ($\boldsymbol{\theta}_{jt}$) in which the estimated parameters are industry specific and time varying. The sample period is from June 1967 to December 2016.

	Data	$\boldsymbol{\theta}$	$\boldsymbol{\theta}_j$	$\boldsymbol{\theta}_t$	$\boldsymbol{\theta}_{jt}$
mean	14.45	14.89	20.77	14.09	14.71
StdDev	60.78	19.75	23.69	30.12	32.82
Skewness	2.15	2.13	2.07	1.05	1.34
Kurtosis	11.05	13.36	10.98	11.40	11.84
Correlation	na	0.09	0.09	0.15	0.17
m.a.e.	na	42.45	44.25	43.66	43.41

Table A.2: **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 specification (under column $\boldsymbol{\theta}$) with constant 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 baseline specification ($\boldsymbol{\theta}_{jt}$) with industry-specific and time-varying 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			
	$\boldsymbol{\theta}$	$\boldsymbol{\theta}_j$	$\boldsymbol{\theta}_t$	$\boldsymbol{\theta}_{jt}$	$\boldsymbol{\theta}$	$\boldsymbol{\theta}_j$	$\boldsymbol{\theta}_t$	$\boldsymbol{\theta}_{jt}$
BM	-3.25^{*}	-0.64	-3.78^{**}	0.46	9.99^{***}	7.38^{***}	10.52^{***}	6.28^{***}
	(-1.95)	(-0.52)	(-2.17)	(0.26)	(4.23)	(3.25)	(4.20)	(3.33)
R11	3.82^{***}	3.59^{***}	5.75^{***}	11.82^{***}	9.92^{***}	10.16^{***}	8.00^{***}	1.93
	(6.38)	(7.75)	(6.88)	(12.51)	(3.82)	(4.06)	(2.93)	(0.78)
I/A	0.48	-0.06	-0.68	-3.08^{**}	-6.78^{***}	-6.24^{***}	-5.62^{***}	-3.22^{**}
	(0.48)	(-0.09)	(-0.43)	(-2.25)	(-4.03)	(-3.42)	(-3.18)	(-2.10)
ROE	4.72^{***}	3.89^{***}	5.32^{***}	4.62^{***}	2.97	3.80^{**}	2.36	3.07^{*}
	(9.30)	(9.65)	(9.26)	(5.72)	(1.60)	(2.10)	(1.29)	(1.81)
Size	-8.10^{***}	-5.82^{***}	-7.57^{***}	-5.99^{***}	3.26	0.98	2.73	1.15
	(-8.84)	(-8.62)	(-6.05)	(-5.63)	(0.93)	(0.28)	(0.73)	(0.34)
Accruals	9.01^{***}	4.90^{***}	8.80^{***}	4.74^{***}	-14.59^{***}	-10.48^{***}	-14.38^{***}	-10.32^{***}
	(16.76)	(7.96)	(14.55)	(4.45)	(-8.08)	(-5.96)	(-7.68)	(-6.28)
NSI	-4.70^{***}	-2.80^{***}	-4.59^{***}	-3.05^{***}	-2.96	-4.85^{**}	-3.07	-4.60^{***}
	(-7.18)	(-4.28)	(-6.70)	(-3.36)	(-1.51)	(-2.44)	(-1.49)	(-2.93)
ΔPI/A	-2.29^{**}	-4.21^{***}	-3.12^{**}	-5.79^{***}	-3.50^{*}	-1.58	-2.67	-0.01
	(-2.26)	(-6.99)	(-2.07)	(-4.81)	(-1.90)	(-0.87)	(-1.48)	(-0.00)
GP/A	15.28^{***}	6.77^{***}	15.48^{***}	7.26^{***}	-11.42^{***}	-2.90	-11.61^{***}	-3.39^{***}
	(19.27)	(13.05)	(14.63)	(5.84)	(-5.20)	(-1.47)	(-5.41)	(-2.63)
ROA	3.84^{***}	3.06^{***}	4.10^{***}	3.80^{***}	2.62	3.40^{*}	2.35	2.66
	(6.85)	(6.25)	(5.80)	(3.99)	(1.34)	(1.77)	(1.24)	(1.48)
RD/M	0.52	2.52^{*}	0.47	5.24^{**}	8.18^{**}	6.18	8.22^{**}	3.46
	(0.49)	(1.92)	(0.50)	(2.12)	(1.97)	(1.49)	(2.07)	(1.42)
Ad/M	14.14^{***}	6.39^{***}	13.89^{***}	7.46^{***}	-8.04^{***}	-0.29	-7.79^{**}	-1.36
	(8.96)	(5.75)	(7.32)	(2.82)	(-2.71)	(-0.10)	(-2.42)	(-0.58)

Table A.3: **Average firm characteristics of 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 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.

	L	2	3	4	5	6	7	8	9	H
					BM					
I_{it}/K_{it}	0.56	0.44	0.38	0.34	0.33	0.29	0.27	0.24	0.22	0.18
I_{it+1}/K_{it+1}	0.45	0.37	0.34	0.31	0.29	0.27	0.25	0.23	0.22	0.19
Y_{it+1}/K_{it+1}	8.58	8.09	7.93	7.65	7.44	7.32	7.09	6.77	7.04	7.46
W_{it+1}/K_{it+1}	4.48	3.78	3.51	3.29	3.17	3.07	2.95	2.85	3.05	3.22
					R11					
I_{it}/K_{it}	0.39	0.35	0.33	0.32	0.31	0.31	0.32	0.34	0.37	0.49
I_{it+1}/K_{it+1}	0.25	0.27	0.27	0.27	0.28	0.28	0.30	0.32	0.36	0.48
Y_{it+1}/K_{it+1}	7.88	7.63	7.31	7.13	6.99	6.97	7.14	7.47	8.00	9.68
W_{it+1}/K_{it+1}	4.25	3.62	3.29	3.13	3.02	2.97	3.03	3.15	3.39	4.36
					I/A					
I_{it}/K_{it}	0.26	0.22	0.23	0.24	0.25	0.28	0.30	0.34	0.40	0.53
I_{it+1}/K_{it+1}	0.28	0.24	0.23	0.23	0.25	0.26	0.28	0.30	0.34	0.40
Y_{it+1}/K_{it+1}	8.75	7.12	6.69	6.40	6.56	6.59	7.10	7.51	8.19	8.43
W_{it+1}/K_{it+1}	4.40	3.06	2.79	2.58	2.66	2.70	2.95	3.16	3.57	4.19
					ROE					
I_{it}/K_{it}	0.41	0.31	0.30	0.31	0.32	0.34	0.35	0.38	0.41	0.46
I_{it+1}/K_{it+1}	0.30	0.25	0.26	0.27	0.29	0.31	0.33	0.35	0.38	0.44
Y_{it+1}/K_{it+1}	7.70	6.97	6.94	6.99	7.25	7.58	7.81	8.24	8.71	9.92
W_{it+1}/K_{it+1}	4.92	3.37	3.10	3.00	3.00	3.05	3.13	3.28	3.46	3.97
					Size					
I_{it}/K_{it}	0.34	0.37	0.37	0.36	0.33	0.31	0.30	0.27	0.26	0.26
I_{it+1}/K_{it+1}	0.31	0.32	0.33	0.31	0.30	0.28	0.27	0.25	0.24	0.24
Y_{it+1}/K_{it+1}	9.40	7.95	7.14	6.79	6.29	5.62	5.43	5.00	4.60	4.22
W_{it+1}/K_{it+1}	4.32	3.65	3.20	2.95	2.70	2.31	2.21	2.07	1.84	1.74
					Accruals					
I_{it}/K_{it}	0.34	0.29	0.28	0.27	0.26	0.27	0.28	0.31	0.34	0.45
I_{it+1}/K_{it+1}	0.33	0.27	0.27	0.25	0.25	0.25	0.26	0.28	0.31	0.38
Y_{it+1}/K_{it+1}	7.50	6.04	5.87	5.67	5.82	6.13	6.58	7.59	8.92	11.84
W_{it+1}/K_{it+1}	3.43	2.53	2.39	2.38	2.41	2.65	2.95	3.43	4.05	5.69

Continued from previous page

	L	2	3	4	5	6	7	8	9	H
					NSI					
I_{it}/K_{it}	0.27	0.25	0.26	0.26	0.27	0.29	0.31	0.34	0.40	0.48
I_{it+1}/K_{it+1}	0.27	0.25	0.25	0.25	0.26	0.27	0.29	0.31	0.35	0.39
Y_{it+1}/K_{it+1}	7.88	7.42	7.49	7.01	6.89	7.02	7.22	7.70	8.29	7.91
W_{it+1}/K_{it+1}	3.20	2.94	3.11	2.95	2.91	2.98	3.16	3.51	4.01	4.31
					$\Delta PI/A$					
I_{it}/K_{it}	0.27	0.27	0.30	0.31	0.30	0.31	0.31	0.33	0.36	0.42
I_{it+1}/K_{it+1}	0.29	0.28	0.29	0.29	0.28	0.28	0.28	0.29	0.30	0.33
Y_{it+1}/K_{it+1}	9.48	9.59	9.07	8.26	7.49	6.96	6.64	6.61	6.36	5.85
W_{it+1}/K_{it+1}	4.42	4.57	4.37	3.86	3.36	3.08	2.79	2.75	2.58	2.38
					GP/A					
I_{it}/K_{it}	0.32	0.22	0.27	0.30	0.30	0.32	0.33	0.35	0.38	0.42
I_{it+1}/K_{it+1}	0.27	0.21	0.24	0.26	0.28	0.29	0.30	0.32	0.34	0.38
Y_{it+1}/K_{it+1}	3.53	3.39	4.67	5.80	6.75	7.35	8.21	9.22	10.68	12.39
W_{it+1}/K_{it+1}	3.27	1.59	2.14	2.63	3.00	3.30	3.65	4.02	4.53	4.40
					ROA					
I_{it}/K_{it}	0.43	0.29	0.29	0.29	0.31	0.33	0.35	0.37	0.41	0.52
I_{it+1}/K_{it+1}	0.32	0.23	0.24	0.26	0.28	0.30	0.32	0.35	0.38	0.48
Y_{it+1}/K_{it+1}	7.83	6.93	7.03	7.14	7.41	7.66	7.89	8.17	8.46	10.06
W_{it+1}/K_{it+1}	5.15	3.15	3.03	2.95	3.03	3.09	3.15	3.27	3.45	4.43
					RD/M					
I_{it}/K_{it}	0.41	0.37	0.38	0.39	0.38	0.36	0.37	0.36	0.37	0.32
I_{it+1}/K_{it+1}	0.32	0.31	0.33	0.33	0.32	0.34	0.33	0.33	0.34	0.33
Y_{it+1}/K_{it+1}	6.92	6.70	6.84	7.12	7.07	7.31	7.36	7.53	8.12	9.14
W_{it+1}/K_{it+1}	3.12	3.11	3.36	3.68	3.68	3.89	4.03	4.39	5.00	6.15
					Ad/M					
I_{it}/K_{it}	0.50	0.39	0.36	0.33	0.31	0.29	0.28	0.27	0.27	0.24
I_{it+1}/K_{it+1}	0.41	0.34	0.32	0.30	0.28	0.28	0.27	0.26	0.25	0.24
Y_{it+1}/K_{it+1}	7.54	8.12	8.41	8.36	8.02	8.23	8.37	8.96	9.70	10.47
W_{it+1}/K_{it+1}	4.36	3.97	3.91	3.66	3.45	3.46	3.36	3.55	3.74	3.89

Table A.4: **Parameter estimates for models with asymmetric adjustment costs**

The model with quadratic asymmetric investment adjustment costs function, $\Phi(I_{it}, K_{it}) = \frac{a_{jt}^{+}\mathbb{I}_{I_{it}>=0} + a_{jt}^{-}(1-\mathbb{I}_{I_{it}>=0})}{2} \left(\frac{I_{it}}{K_{it}}\right)^2 K_{it}$ and the one with exponential asymmetric investment adjustment costs function, $\Phi_{it} \equiv \frac{\theta_{jt}}{\nu_{jt}^2}\left[\exp\left(-\nu_{jt}\frac{I_{it}}{K_{it}}\right) + \nu_{jt}\frac{I_{it}}{K_{it}} - 1\right]$ are estimated. This table reports the estimation results for these two models in Panels A and B, respectively. Column γ reports the the posterior means of the marginal product parameter γ; column CI_γ reports the 95% credible intervals of γ; and column $\sigma(\gamma)$ reports the time-series standard deviation of the posterior means of γ. Similar definitions apply to other columns.

Industry	Panel A: Quadratic asymmetric adjustment cost								
	γ	CI_γ	$\sigma(\gamma)$	a^+	CI_{a_+}	$\sigma(a^+)$	a^-	CI_{a^-}	$\sigma(a^-)$
Consumer Nondurables	0.13	[0.11, 0.14]	0.09	0.44	[0.29, 0.62]	0.46	0.96	[0.21, 2.06]	0.77
Consumer Durables	0.17	[0.14, 0.20]	0.18	1.24	[0.87, 1.65]	1.15	2.45	[0.95, 4.43]	2.35
Manufacturing	0.16	[0.15, 0.17]	0.11	0.54	[0.48, 0.62]	0.88	0.92	[0.32, 1.71]	1.38
Energy	0.20	[0.18, 0.22]	0.14	0.45	[0.40, 0.51]	0.54	0.66	[0.28, 1.21]	0.79
Business Equipment	0.23	[0.21, 0.25]	0.19	1.69	[1.60, 1.79]	1.95	1.18	[0.33, 2.49]	1.09
Telecom	0.28	[0.25, 0.32]	0.21	0.71	[0.63, 0.80]	0.67	1.63	[1.06, 2.50]	2.50
Wholesale & Retail	0.08	[0.07, 0.09]	0.06	0.83	[0.73, 0.96]	1.00	1.01	[0.28, 2.05]	0.86
Healthcare	0.19	[0.17, 0.21]	0.16	0.64	[0.48, 0.83]	0.82	1.72	[0.69, 3.00]	1.57
Utilities	0.29	[0.25, 0.33]	0.17	0.26	[0.19, 0.36]	0.32	0.40	[0.13, 0.85]	0.46
Others	0.16	[0.15, 0.18]	0.12	0.45	[0.42, 0.48]	0.50	0.52	[0.08, 1.18]	0.48
	Panel B: Exponential adjustment cost								
	γ	CI_γ	$\sigma(\gamma)$	θ	CI_θ	$\sigma(\theta)$	ν	CI_ν	$\sigma(\nu)$
Consumer Nondurables	0.12	[0.08, 0.15]	0.09	3.64	[2.73, 4.52]	4.34	6.56	[2.78, 11.55]	4.01
Consumer Durables	0.18	[0.15, 0.23]	0.20	19.77	[18.38, 21.05]	35.48	9.35	[4.82, 13.55]	7.99
Manufacturing	0.14	[0.10, 0.18]	0.10	3.05	[2.53, 3.79]	2.99	7.39	[4.51, 11.85]	5.95
Energy	0.18	[0.13, 0.24]	0.13	3.45	[2.97, 4.11]	7.02	15.16	[11.62, 20.19]	28.69
Business Equipment	0.20	[0.18, 0.21]	0.12	15.29	[14.98, 15.75]	10.36	6.03	[4.95, 7.36]	6.38
Telecom	0.23	[0.18, 0.30]	0.19	15.69	[15.13, 17.09]	22.44	15.00	[11.52, 19.04]	14.18
Wholesale & Retail	0.07	[0.05, 0.09]	0.05	3.69	[2.68, 4.97]	3.08	8.17	[3.97, 13.44]	7.70
Healthcare	0.21	[0.17, 0.25]	0.20	21.82	[18.20, 24.07]	34.78	9.02	[6.39, 12.67]	6.54
Utilities	0.24	[0.15, 0.33]	0.15	8.12	[7.43, 8.93]	19.53	29.23	[25.81, 33.12]	44.77
Others	0.14	[0.09, 0.20]	0.10	2.22	[1.65, 2.92]	2.05	6.47	[2.61, 11.58]	4.05

Table A.5: **Anomaly premiums and alphas under asymmetric adjustment costs**

The fundamental anomaly premiums, r^F, and the corresponding alphas, $\alpha \equiv r^S - r^F$, of the 12 anomalies are reported for two estimated models with quadratic and exponential, respectively, asymmetric adjustment costs functions under Columns 'Quadratic' and 'Exponential'. Fundamental returns are computed using the posterior means of the parameter estimates. The t-values are adjusted for heteroscedasticity and autocorrelations with lags up to 24 months. 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.

	r^F		α_{H-L}	
	Quadratic	Exponential	Quadratic	Exponential
BM	0.42 (0.24)	−0.18 (−0.10)	6.32 (3.36)	6.92 (3.72)
R11	11.91 (12.43)	14.27 (12.51)	1.84 (0.75)	−0.52 (−0.21)
I/A	−3.09 (−2.28)	−1.17 (−0.94)	−3.21 (−2.09)	−5.13 (−2.84)
ROE	4.71 (5.82)	7.06 (7.66)	2.97 (1.76)	0.63 (0.37)
Size	−5.96 (−5.53)	−2.68 (−1.77)	1.12 (0.33)	−2.16 (−0.55)
Accruals	4.79 (4.49)	5.23 (3.71)	−10.37 (−6.37)	−10.81 (−6.28)
NSI	−3.02 (−3.34)	−2.86 (−2.32)	−4.63 (−2.98)	−4.79 (−2.96)
ΔP/A	−5.78 (−4.82)	−3.76 (−3.67)	−0.01 (−0.01)	−2.03 (−1.02)
GP/A	7.25 (5.89)	5.95 (3.06)	−3.38 (−2.65)	−2.09 (−1.38)
ROA	3.92 (4.07)	5.31 (4.35)	2.54 (1.42)	1.15 (0.66)
RD/M	5.35 (2.14)	2.43 (0.95)	3.35 (1.39)	6.27 (1.76)
Ad/M	7.41 (2.81)	5.28 (2.79)	−1.31 (−0.56)	0.82 (0.33)

Table A.6: **Parameter estimates for the extended model with adjustment costs in working capital investment**

The table reports parameter estimates of an extended model with quadratic adjustment costs in working capital investment: $\Phi(\Delta W_{it}, W_{it}) = \frac{b_{jt}}{2}\left(\frac{\Delta W_{it}}{W_{it}}\right)^2 W_{it}$, where b_{jt} is the adjustment costs parameter for industry j at time t. Parameter values are industry-specific and time-varying. Column γ reports the time-series averages of the posterior means of the marginal product parameter γ; column CI$_\gamma$ reports the time-series averages of the 95% credible intervals of γ; and column $\sigma(\gamma)$ reports the time-series standard deviation of the posterior means of γ. Similar definitions apply to other columns.

	Old approach: Extended model with adjusment costs on working capital								
	γ	CI_γ	$\sigma(\gamma)$	a	CI_a	$\sigma(a)$	b	CI_b	$\sigma(b)$
Consumer Nondurables	0.13	[0.11, 0.14]	0.09	0.30	[0.16, 0.46]	0.34	0.23	[0.15, 0.34]	0.34
Consumer Durables	0.16	[0.13, 0.18]	0.17	1.98	[1.79, 2.21]	2.33	0.41	[0.31, 0.55]	0.56
Manufacturing	0.16	[0.15, 0.17]	0.11	0.30	[0.23, 0.38]	0.49	0.28	[0.22, 0.34]	0.39
Energy	0.20	[0.18, 0.22]	0.14	0.37	[0.31, 0.44]	0.57	0.38	[0.28, 0.50]	0.57
Business Equipment	0.23	[0.21, 0.25]	0.20	2.39	[2.29, 2.49]	2.57	0.77	[0.72, 0.83]	1.35
Telecom	0.28	[0.25, 0.31]	0.33	1.29	[1.20, 1.39]	2.59	0.94	[0.77, 1.12]	1.71
Wholesale & Retail	0.08	[0.07, 0.09]	0.06	0.81	[0.67, 0.96]	1.06	0.31	[0.22, 0.42]	0.38
Healthcare	0.19	[0.17, 0.21]	0.16	0.44	[0.31, 0.60]	0.62	0.32	[0.24, 0.42]	0.43
Utilities	0.29	[0.25, 0.33]	0.17	0.23	[0.15, 0.33]	0.27	0.20	[0.06, 0.41]	0.20
Others	0.16	[0.14, 0.18]	0.12	0.29	[0.22, 0.35]	0.43	0.23	[0.16, 0.30]	0.31

Table A.7: **Anomaly premiums and alphas under the extended model with adjustment costs in working capital investment**

The fundamental anomaly premiums, r^F, and the corresponding alphas, $\alpha \equiv r^S - r^F$, of the 12 anomalies are reported for the estimated extended model with adjustment costs in working capital investment. Fundamental returns are computed using the posterior means of the parameter estimates. The t-values are adjusted for heteroscedasticity and autocorrelations with lags up to 24 months. 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.

	r_F	α_{H-L}
BM	1.58	5.19
	(0.87)	(2.72)
R11	12.12	1.55
	(11.40)	(0.66)
I/A	−3.17	−3.07
	(−2.27)	(−2.19)
ROE	3.74	3.89
	(5.01)	(2.27)
Size	−7.06	2.22
	(−5.76)	(0.68)
Accruals	4.65	−10.22
	(3.86)	(−5.98)
NSI	−3.81	−3.84
	(−4.03)	(−2.61)
ΔP/A	−6.20	0.47
	(−4.94)	(0.37)
GP/A	6.44	−2.50
	(4.80)	(−1.65)
ROA	3.14	3.23
	(3.50)	(1.74)
RD/M	5.38	3.23
	(2.17)	(1.27)
Ad/M	7.84	−1.70
	(2.80)	(−0.71)

Table A.8: **Simulation study: Bayesian MCMC vs. NLS**

This table reports the true values, the NLS estimates and its confidence interval (in square brackets), 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. Bootstrap method (Efron and Tibshirani, 1994) is implemented to calculate the confidence intervals for NLS because the parameters are estimated with constraints. We sample the simulation data sample (13,038 observations) with replacement from the original data set and estimate the parameters. This with-replacement-sampling-estimating procedure is repeated for 1,000 times to obtain the confidence intervals of the NLS estimates.

	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
		[0.3237, 0.3344]	[0.3241, 0.3347]		[5.4704, 5.6489]	[5.4930, 5.6502]
Manufacturing	0.7138	0.7107	0.7112	3.9548	3.9581	3.9596
		[0.7051, 0.7157]	[0.7065, 0.7158]		[3.8781, 4.0064]	[3.9129, 4.0013]
Business Equipment	0.6348	0.6378	0.6356	6.0973	6.0769	6.0235
		[0.6284, 0.6472]	[0.6281, 0.6421]		[5.8760, 6.2607]	[5.8258, 6.1722]
Wholesale & Retail	0.5154	0.5178	0.5180	5.0573	5.0686	5.0888
		[0.5135, 0.5224]	[0.5138, 0.5219]		[4.9370, 5.1793]	[4.9892, 5.1812]
Healthcare	0.3791	0.3813	0.3800	5.5589	5.5051	5.5078
		[0.3741, 0.3897]	[0.3723, 0.3875]		[5.3258, 5.6259]	[5.4267, 5.5977]
Utilities	0.3068	0.2902	0.2901	4.8656	4.8718	4.8729
		[0.2751, 0.3049]	[0.2769, 0.3030]		[4.8388, 4.9055]	[4.8470, 4.8986]
Other	0.4607	0.4666	0.4661	6.1648	6.1627	6.1626
		[0.4549, 0.4764]	[0.4559, 0.4763]		[6.0917, 6.2243]	[6.1346, 6.1926]
			Panel B: Parameters with constant values			
	0.1500	0.1501	0.1500	0.1300	0.1280	0.1300
		[0.1481, 0.1523]	[0.1487,0.1520]		[0.1025, 0.1511]	[0.1297,0.1407]

Table A.9: **Alphas for decile portfolios based on the posterior means of the baseline parameter estimates**

This table reports the annualized alphas, α, of decile portfolios and the average absolute alpha over decile portfolios, $\overline{|\alpha|}$ for each of the 12 anomalies. 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.

	L	2	3	4	5	6	7	8	9	H	H-L	$\overline{\lvert\alpha\rvert}$
BM	-2.60**	-0.84	-1.49	-1.93	-1.48	0.22	1.04	2.40**	2.38**	3.67**	6.27***	1.81
	(-2.30)	(-0.96)	(-1.91)	(-1.48)	(-1.49)	(0.24)	(0.90)	(2.30)	(2.17)	(2.52)	(3.33)	
R11	-2.09	0.67	1.36	0.80	-0.53	-1.25	-1.80	-1.19	-2.26	-0.22	1.87	1.22
	(-0.72)	(0.33)	(1.14)	(0.73)	(-0.50)	(-1.22)	(-1.86)	(-1.18)	(-1.66)	(-0.13)	(0.57)	
I/A	-1.22	0.10	0.17	-0.23	-0.82	-0.35	-1.09	-0.76	-1.08	-4.37***	-3.16**	1.02
	(-0.80)	(0.10)	(0.18)	(-0.19)	(-0.95)	(-0.41)	(-1.44)	(-0.81)	(-0.78)	(-2.60)	(-2.06)	
ROE	-4.77**	-1.73	-0.40	-0.97	0.74	-1.02	-0.97	-2.45***	-2.17**	-1.72	3.05	1.70
	(-2.23)	(-1.05)	(-0.32)	(-0.87)	(0.52)	(-0.78)	(-0.95)	(-2.64)	(-2.23)	(-1.33)	(1.33)	
Size	-2.00	-3.11	-3.15	-2.29	-1.37	-1.56	-0.74	-0.36	-0.47	-0.86	1.14	1.59
	(-0.69)	(-1.55)	(-1.72)	(-1.53)	(-0.87)	(-1.24)	(-0.58)	(-0.34)	(-0.50)	(-0.90)	(0.34)	
Accruals	1.77	0.79	1.41	-1.43	0.21	-0.04	-1.24	-3.75***	-5.34***	-8.56***	-10.33***	2.45
	(1.38)	(0.95)	(1.62)	(-1.77)	(0.19)	(-0.04)	(-1.18)	(-3.09)	(-3.89)	(-4.44)	(-6.29)	
NSI	0.20	-0.81	-0.82	-0.12	-0.84	-1.10	0.41	-0.34	-2.42**	-4.36***	-4.56***	1.14
	(0.18)	(-0.76)	(-0.92)	(-0.12)	(-0.96)	(-1.07)	(0.57)	(-0.28)	(-2.01)	(-2.68)	(-2.90)	
$\Delta PI/A$	-0.98	-2.15**	-1.59	-1.45	-1.73	-0.62	-0.42	-0.24	-1.28	-0.94	0.04	1.14
	(-0.65)	(-2.43)	(-1.80)	(-1.34)	(-1.59)	(-0.63)	(-0.43)	(-0.25)	(-0.99)	(-0.56)	(0.03)	
GP/A	1.72	0.20	-0.06	-1.31	0.12	-1.73	-2.89**	-1.64	-1.23	-1.68	-3.39***	1.26
	(1.22)	(0.21)	(-0.06)	(-1.27)	(0.11)	(-1.58)	(-2.57)	(-1.69)	(-1.18)	(-1.77)	(-2.64)	
ROA	-4.04	-2.26	0.07	-0.74	-1.85	0.98	-0.98	-2.16**	-1.24	-1.41	2.63	1.57
	(-1.72)	(-1.29)	(0.05)	(-0.58)	(-1.20)	(0.83)	(-0.75)	(-2.16)	(-1.20)	(-1.12)	(1.11)	
RD/M	-2.81	-1.90	-4.43***	-2.10	-1.44	-1.34	-0.53	1.01	-0.10	0.66	3.47	1.63
	(-1.94)	(-1.51)	(-3.03)	(-1.48)	(-1.32)	(-1.04)	(-0.46)	(0.88)	(-0.07)	(0.32)	(1.42)	
Ad/M	-1.16	0.06	-1.09	-1.23	-2.12	1.30	0.12	-1.22	2.16	-2.55	-1.39	1.30
	(-0.70)	(0.04)	(-0.69)	(-1.12)	(-1.70)	(0.93)	(0.10)	(-0.88)	(1.23)	(-1.47)	(-0.59)	

Table A.10: **Realized stock returns, fundamental returns, and alphas for portfolio deciles – Parameters with no time variation**

Panel A reports the annualized alpha of decile portfolios and the average absolute alpha over decile portfolios for each of the 12 anomaly variables. The alpha is the average difference between portfolio stock returns and constructed fundamental returns. Panel B reports the ratio of fundamental return to realized return for decile portfolios and high-minus-low decile portfolios. The sample period is from June 1967 to December 2016.

	L	2	3	4	5	6	7	8	9	H	H-L
Bm	-3.08	-0.80	-0.86	-1.22	-1.30	0.10	2.12	2.55	2.61	4.52**	7.59***
	(-1.75)	(-0.49)	(-0.56)	(-0.72)	(-0.86)	(0.07)	(1.33)	(1.67)	(1.70)	(2.39)	(4.48)
R11	-6.96**	-2.32	-0.52	-0.81	-1.11	-1.44	-1.37	-0.16	-0.06	3.87	10.83***
	(-2.50)	(-1.03)	(-0.30)	(-0.50)	(-0.74)	(-0.92)	(-0.97)	(-0.11)	(-0.04)	(1.78)	(4.31)
I/A	0.89	0.97	0.74	-0.05	0.18	0.02	-0.02	-1.26	-1.24	-5.71***	-6.60***
	(0.47)	(0.59)	(0.52)	(-0.04)	(0.12)	(0.01)	(-0.01)	(-0.83)	(-0.69)	(-2.89)	(-5.45)
ROE	-5.24**	-2.39	-0.51	-0.76	0.51	-1.30	-1.19	-2.38	-1.61	-1.09	4.16**
	(-2.16)	(-1.23)	(-0.30)	(-0.54)	(0.32)	(-0.81)	(-0.79)	(-1.50)	(-1.03)	(-0.61)	(2.26)
Size	-1.75	-2.84	-2.79	-2.22	-0.67	-1.37	-0.55	0.08	-0.37	-0.74	1.01
	(-0.65)	(-1.27)	(-1.41)	(-1.15)	(-0.35)	(-0.79)	(-0.31)	(0.05)	(-0.24)	(-0.49)	(0.44)
Accruals	1.69	1.41	1.13	-1.43	0.66	-0.33	-0.89	-3.88**	-5.78***	-8.82***	-10.52***
	(0.84)	(0.87)	(0.79)	(-0.98)	(0.39)	(-0.22)	(-0.52)	(-2.25)	(-3.26)	(-4.36)	(-8.42)
NSI	0.56	0.08	-0.75	0.89	-0.46	-0.51	0.82	-0.11	-3.23	-4.46**	-5.02***
	(0.35)	(0.05)	(-0.46)	(0.56)	(-0.31)	(-0.35)	(0.49)	(-0.06)	(-1.84)	(-2.37)	(-3.78)
ΔPI/A	0.27	-1.50	-1.61	-1.29	-1.12	-0.34	-0.02	0.18	-2.31	-1.73	-2.01
	(0.14)	(-0.92)	(-1.10)	(-0.81)	(-0.70)	(-0.22)	(-0.01)	(0.11)	(-1.30)	(-0.89)	(-1.63)
GP/A	1.70	0.49	0.46	-1.41	0.66	-1.83	-2.60	-2.25	-0.78	-1.30	-3.00**
	(0.99)	(0.32)	(0.28)	(-0.89)	(0.41)	(-1.15)	(-1.50)	(-1.32)	(-0.46)	(-0.79)	(-2.28)
ROA	-4.31	-2.88	0.38	-0.46	-2.13	0.40	-1.33	-2.32	-1.19	-0.62	3.70
	(-1.54)	(-1.25)	(0.21)	(-0.29)	(-1.18)	(0.22)	(-0.75)	(-1.40)	(-0.68)	(-0.32)	(1.90)
RD/M	-3.29	-1.46	-3.90**	-1.37	-0.77	-0.40	-0.31	1.95	0.41	2.92	6.21**
	(-1.80)	(-0.79)	(-1.96)	(-0.73)	(-0.42)	(-0.22)	(-0.16)	(0.94)	(0.19)	(0.99)	(2.19)
Ad/M	-1.26	-0.48	-2.05	0.34	-0.53	2.62	0.08	-0.05	2.86	-1.30	-0.05
	(-0.62)	(-0.24)	(-1.05)	(0.18)	(-0.30)	(1.40)	(0.05)	(-0.03)	(1.20)	(-0.60)	(-0.02)

Table A.11: **Realized stock returns, fundamental returns, and alphas for portfolio deciles – Parameters with no industry variation**

This table reports the annualized alpha of decile portfolios and the average absolute alpha over decile portfolios for each of the 12 anomaly variables. The alpha is the average difference between portfolio stock returns and constructed fundamental returns. The sample period is from June 1967 to December 2016.

	L	2	3	4	5	6	7	8	9	H	H-L
Bm	-1.67	0.08	0.56	0.30	0.94	2.79***	5.77***	6.62***	7.26***	8.18***	9.85***
	(-1.42)	(0.08)	(0.61)	(0.27)	(0.89)	(2.80)	(4.85)	(5.59)	(6.42)	(5.37)	(5.39)
R11	-1.43	1.00	2.19	1.83	1.00	0.76	0.57	1.64	1.30	4.19***	5.61**
	(-0.56)	(0.53)	(1.67)	(1.57)	(0.92)	(0.74)	(0.60)	(1.74)	(1.09)	(2.63)	(2.06)
I/A	1.24	2.99***	3.03***	2.22**	2.05**	1.99**	1.30	0.57	1.18	-2.89**	-4.13***
	(0.92)	(2.85)	(3.32)	(2.37)	(2.39)	(2.11)	(1.47)	(0.59)	(0.97)	(-2.09)	(-3.17)
ROE	-1.70	1.42	2.92**	2.24**	3.40***	1.28	0.61	-1.15	-0.21	-0.32	1.38
	(-0.88)	(0.96)	(2.50)	(2.28)	(3.06)	(1.12)	(0.67)	(-1.19)	(-0.22)	(-0.28)	(0.75)
Size	-0.89	-1.53	-2.02	-0.71	0.56	0.54	1.59	1.69	1.22	1.80**	2.69
	(-0.41)	(-0.94)	(-1.48)	(-0.56)	(0.41)	(0.47)	(1.38)	(1.57)	(1.25)	(2.00)	(1.18)
Accruals	5.35***	3.79***	3.55***	1.15	2.19	2.19**	0.88	-1.96	-5.03***	-8.61***	-13.96***
	(3.60)	(3.27)	(3.62)	(1.25)	(1.96)	(2.10)	(0.83)	(-1.87)	(-4.20)	(-6.05)	(-10.52)
NSI	2.51**	1.11	1.11	2.15**	0.47	0.57	2.50**	3.30***	0.61	0.22	-2.29
	(2.50)	(1.18)	(1.09)	(2.23)	(0.48)	(0.63)	(2.30)	(2.61)	(0.50)	(0.16)	(-1.69)
ΔPI/A	1.24	-0.09	0.07	1.06	1.07	2.25**	2.05	2.30**	-1.03	0.01	-1.23
	(0.89)	(-0.10)	(0.08)	(1.16)	(0.97)	(2.21)	(1.80)	(2.20)	(-0.84)	(0.00)	(-0.94)
GP/A	7.83***	5.83***	4.19***	1.70	3.62***	0.83	-0.51	-0.60	-0.84	-3.58***	-11.41***
	(5.99)	(5.11)	(3.76)	(1.53)	(3.33)	(0.86)	(-0.46)	(-0.56)	(-0.81)	(-3.16)	(-8.21)
ROA	-0.38	1.70	4.15***	2.94**	1.00	3.21***	0.58	-1.19	-0.24	1.13	1.52
	(-0.18)	(1.00)	(3.03)	(2.46)	(0.76)	(2.77)	(0.49)	(-1.11)	(-0.23)	(0.92)	(0.80)
RD/M	-0.55	1.28	-1.12	1.34	1.61	2.63**	3.31**	5.38***	4.52***	7.50***	8.05***
	(-0.46)	(0.98)	(-0.79)	(0.94)	(1.28)	(2.01)	(2.48)	(3.60)	(3.02)	(2.98)	(2.98)
Ad/M	1.33	1.73	-0.38	0.83	-0.28	1.27	-1.71	-3.31**	-1.47	-7.04***	-8.37***
	(0.88)	(1.23)	(-0.27)	(0.60)	(-0.22)	(0.96)	(-1.19)	(-2.19)	(-0.78)	(-3.81)	(-3.74)

Table A.12: **Realized stock returns, fundamental returns, and alphas for portfolio deciles – Constant parameters**

This table reports the annualized alpha of decile portfolios and the average absolute alpha over decile portfolios for each of the 12 anomaly variables. The alpha is the average difference between portfolio stock returns and constructed fundamental returns. The sample period is from June 1967 to December 2016.

	L	2	3	4	5	6	7	8	9	H	H-L
Bm	-2.07	-0.38	0.33	0.46	1.23	3.14**	5.63***	6.45***	6.69***	7.92***	9.99***
	(-1.17)	(-0.23)	(0.22)	(0.27)	(0.81)	(2.20)	(3.56)	(4.20)	(4.30)	(4.07)	(5.87)
R11	-4.63	-0.61	1.18	1.11	0.74	0.60	0.59	1.71	1.54	5.31**	9.94***
	(-1.65)	(-0.26)	(0.66)	(0.67)	(0.49)	(0.38)	(0.42)	(1.09)	(0.88)	(2.42)	(3.83)
I/A	2.85	3.33**	3.20**	2.55	2.08	1.64	1.29	0.21	0.22	-3.93**	-6.78***
	(1.54)	(2.01)	(2.21)	(1.78)	(1.47)	(1.05)	(0.84)	(0.14)	(0.12)	(-2.01)	(-5.89)
ROE	-2.91	0.36	2.49	2.02	3.11**	0.79	0.43	-1.29	-0.46	0.07	2.98
	(-1.19)	(0.19)	(1.48)	(1.41)	(1.97)	(0.49)	(0.28)	(-0.81)	(-0.29)	(0.04)	(1.60)
Size	-1.68	-2.54	-2.18	-1.44	0.50	0.05	1.03	1.67	1.26	1.58	3.26
	(-0.62)	(-1.14)	(-1.09)	(-0.74)	(0.26)	(0.03)	(0.57)	(1.02)	(0.82)	(1.05)	(1.42)
Accruals	5.26***	4.06**	3.79***	0.65	2.23	1.71	0.72	-2.81	-5.51***	-9.33***	-14.59***
	(2.62)	(2.48)	(2.69)	(0.44)	(1.31)	(1.12)	(0.42)	(-1.62)	(-3.12)	(-4.58)	(-11.20)
NSI	2.17	1.52	0.86	2.58	0.59	0.62	2.51	2.73	0.06	-0.79	-2.96**
	(1.39)	(1.02)	(0.53)	(1.59)	(0.40)	(0.43)	(1.47)	(1.46)	(0.03)	(-0.42)	(-2.25)
ΔPI/A	2.22	0.67	0.59	1.09	0.97	2.07	1.87	1.63	-1.97	-1.29	-3.51***
	(1.18)	(0.42)	(0.41)	(0.70)	(0.59)	(1.33)	(1.10)	(0.98)	(-1.09)	(-0.66)	(-2.84)
GP/A	7.78***	6.10***	4.14**	1.51	3.27**	0.59	-0.73	-0.77	-1.29	-3.64**	-11.42***
	(4.56)	(3.96)	(2.49)	(0.95)	(2.03)	(0.37)	(-0.43)	(-0.46)	(-0.74)	(-2.17)	(-8.12)
ROA	-1.50	0.43	3.97**	2.68	0.57	2.79	0.22	-1.36	-0.42	1.13	2.63
	(-0.53)	(0.19)	(2.20)	(1.65)	(0.31)	(1.51)	(0.12)	(-0.81)	(-0.24)	(0.59)	(1.34)
RD/M	-0.53	1.22	-0.76	1.66	2.04	2.89	3.48	5.78***	4.56**	7.65**	8.18***
	(-0.30)	(0.65)	(-0.38)	(0.86)	(1.10)	(1.59)	(1.84)	(2.84)	(2.11)	(2.56)	(2.92)
Ad/M	1.05	1.78	-0.52	1.05	-0.44	1.91	-2.13	-3.31	-1.46	-6.99***	-8.04***
	(0.50)	(0.89)	(-0.27)	(0.55)	(-0.24)	(1.02)	(-1.25)	(-1.74)	(-0.62)	(-3.11)	(-4.04)

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".

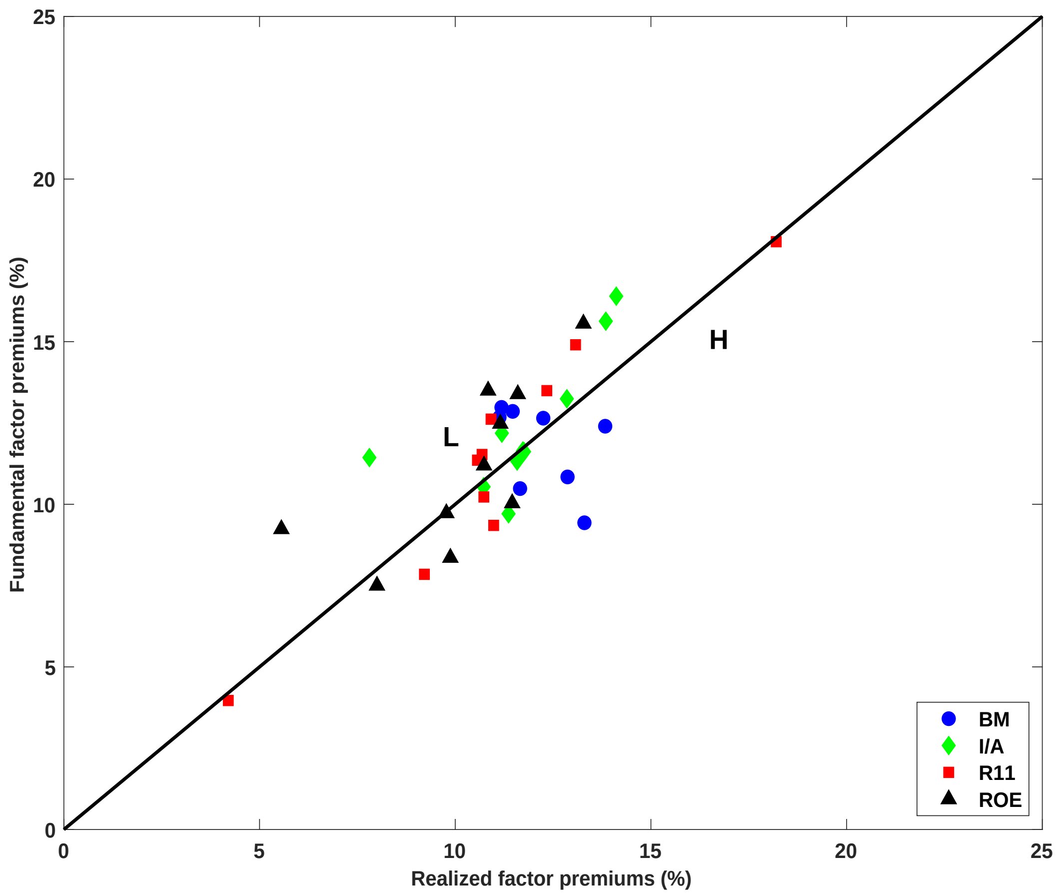
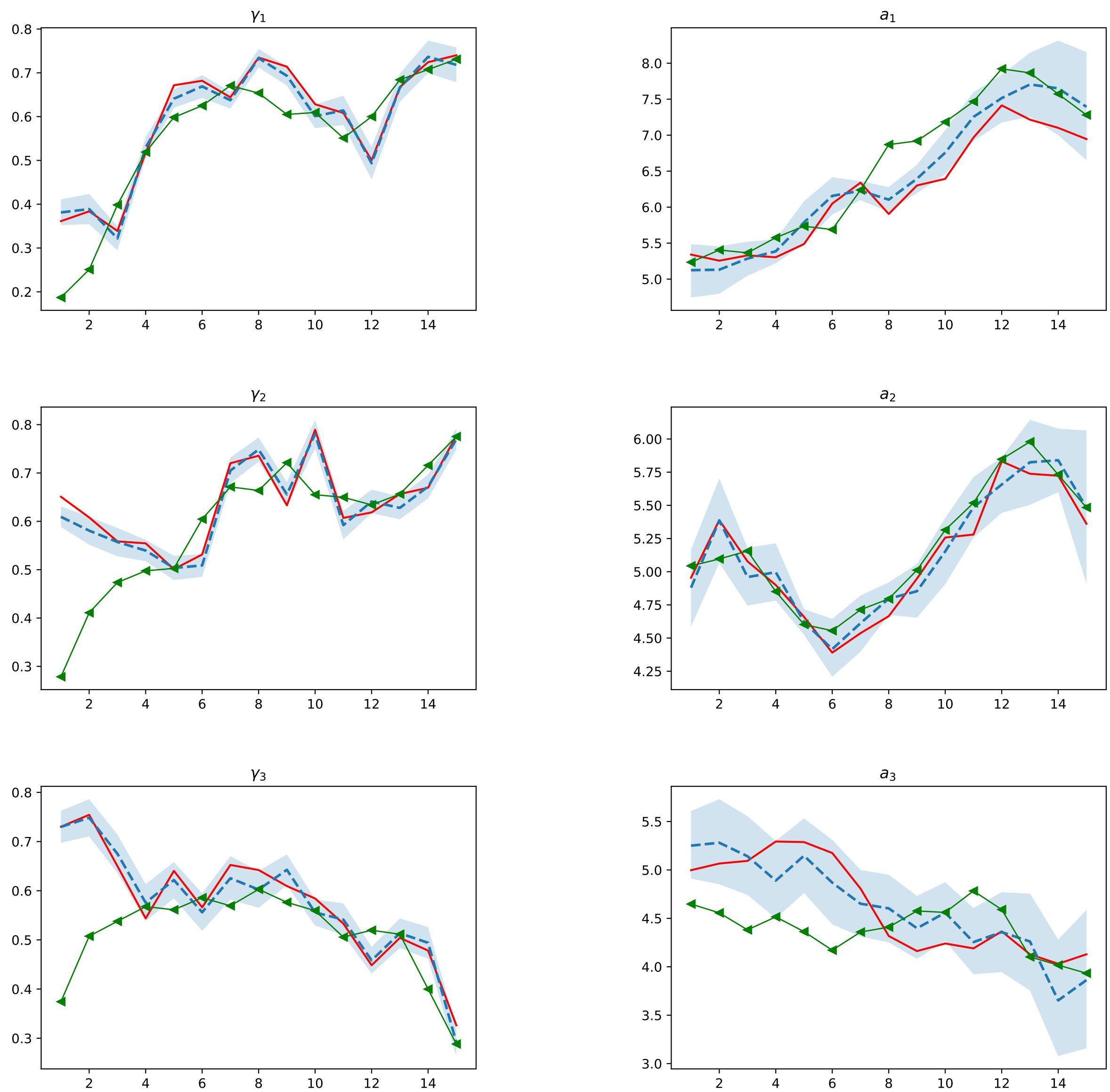


Figure A.2: **Simulation study**

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% credible intervals (in shaded areas) of the model parameters estimated from the simulated data. The marginal product and adjustment costs parameters of Consumer Nondurables, Manufacturing, Business Equipment industries, Wholesale & Retail, Healthcare, Utilities, and Others industries are denoted as γ_j and a_j for $j = 1, 2, 3, 4, 5, 6, 7$, respectively..



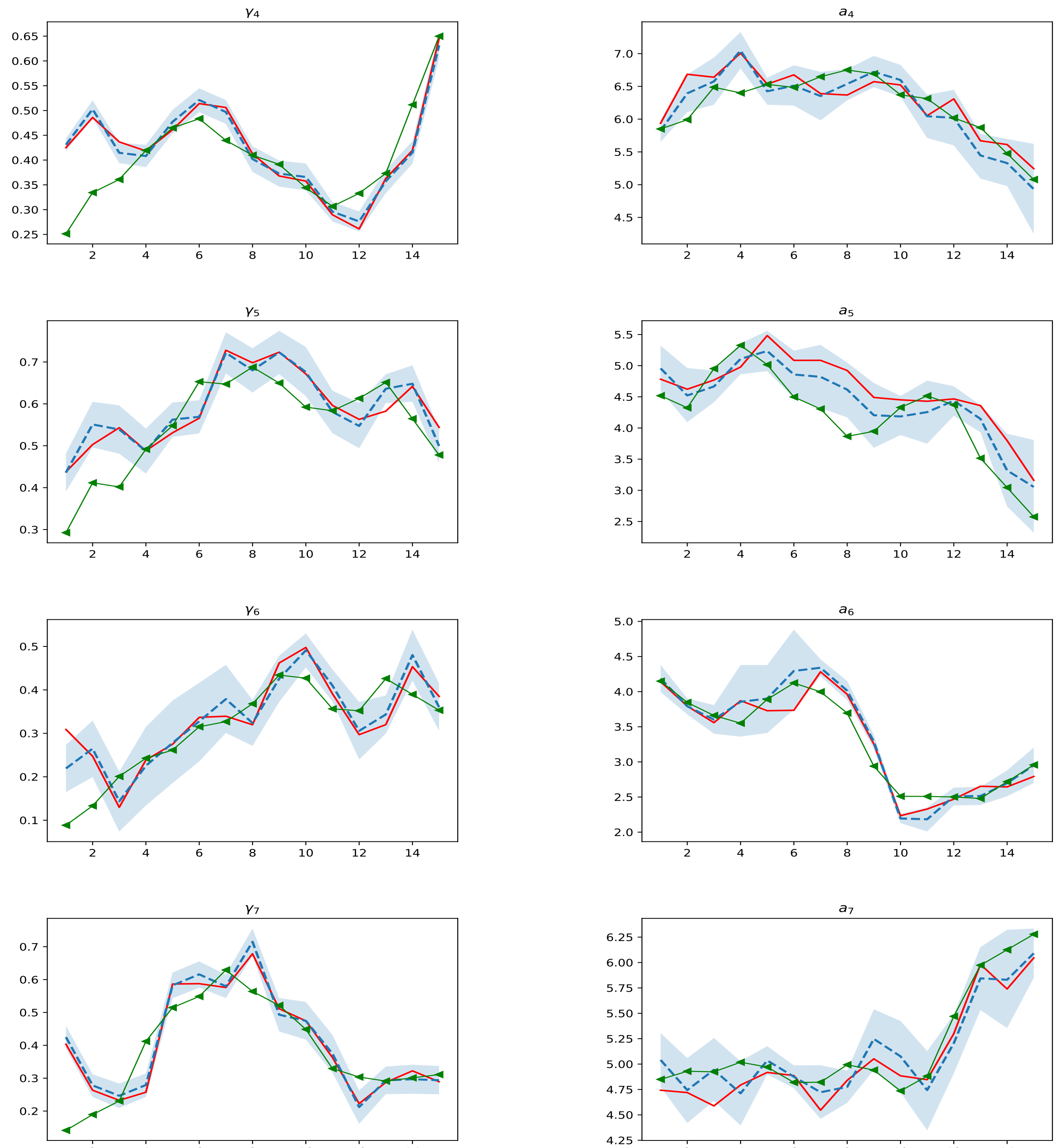

γ4
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
2 4 6 8 10 12 14
a4
7.0
6.5
6.0
5.5
5.0
4.5
2 4 6 8 10 12 14
γ5
0.7
0.6
0.5
0.4
0.3
2 4 6 8 10 12 14
a5
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2 4 6 8 10 12 14
γ6
0.5
0.4
0.3
0.2
0.1
2 4 6 8 10 12 14
a6
5.0
4.5
4.0
3.5
3.0
2.5
2.0
2 4 6 8 10 12 14
γ7
0.7
0.6
0.5
0.4
0.3
0.2
2 4 6 8 10 12 14
a7
6.25
6.00
5.75
5.50
5.25
5.00
4.75
4.50
4.25
2 4 6 8 10 12 14

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.

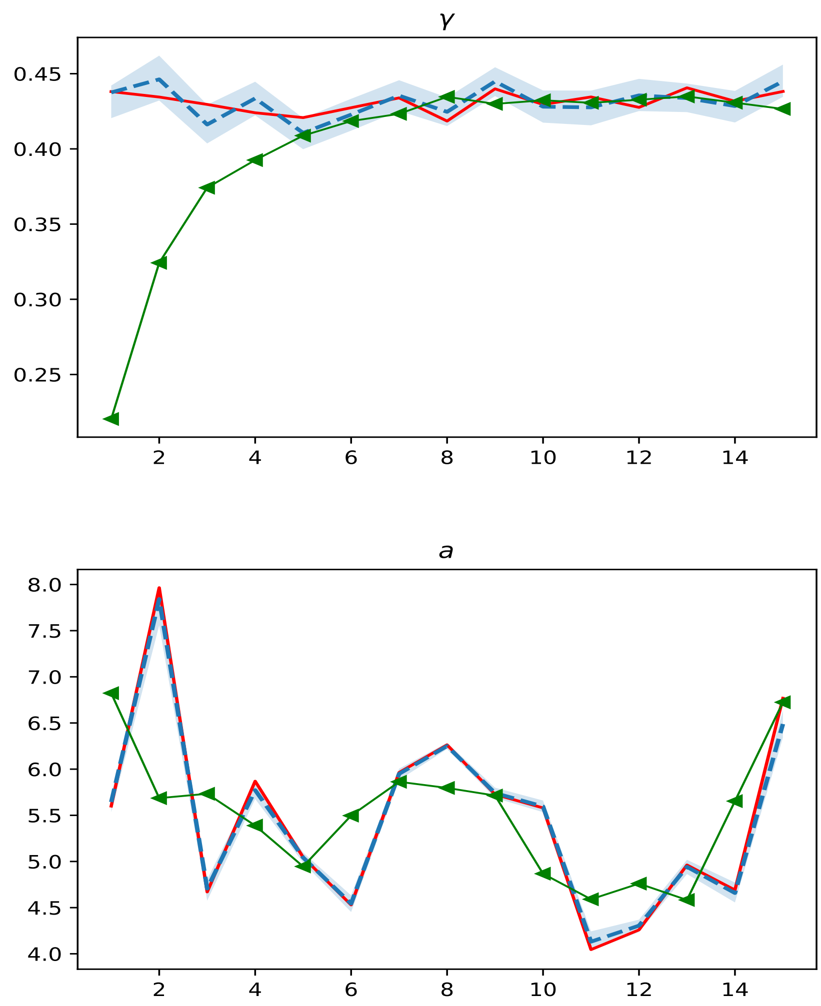
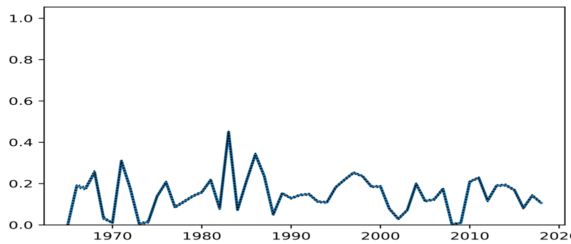


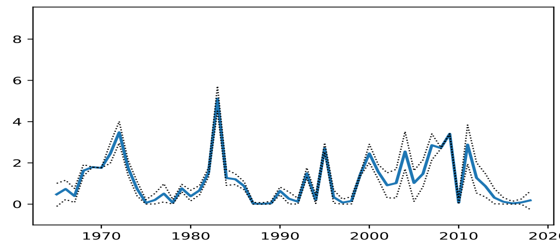
Figure A.4: **Time series of parameter estimates**

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 under the specification with time-varying parameter values in Panel A, and for Fama-French 10 industries under the specification with industry-specific and time-varying parameter values in Panel B.

Panel A: Time-varying parameter values ($\boldsymbol{\theta}_t$)

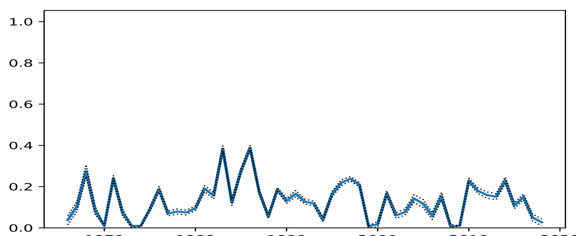


(a) Consumer Nondurables: γ

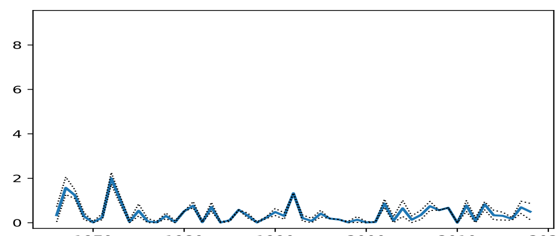


(b) Consumer Nondurables: a

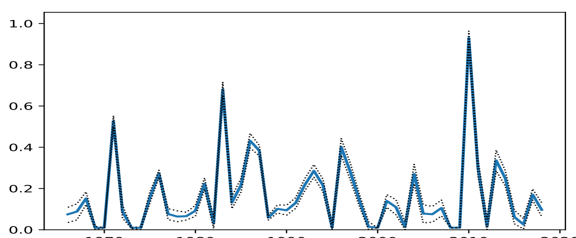
Panel B: Industry-specific and time-varying parameter values ($\boldsymbol{\theta}_{jt}$)



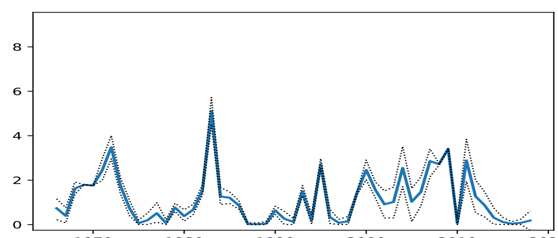
(c) Consumer Nondurables: γ



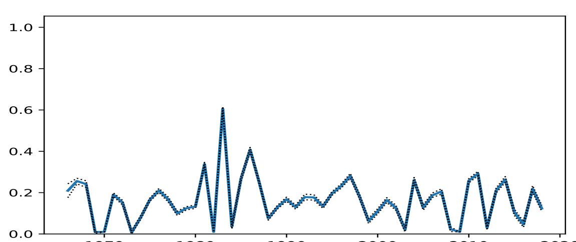
(d) Consumer Nondurables: a



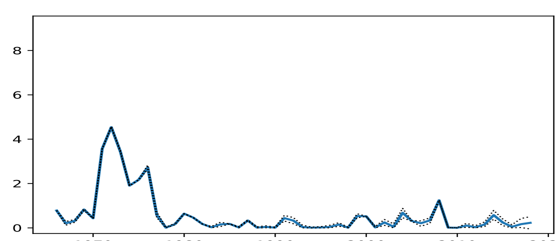
(e) Consumer Durables: γ



(f) Consumer Durables: a

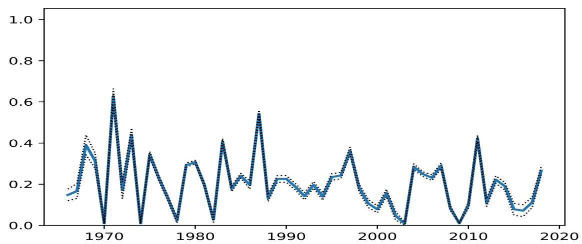


(g) Manufacturing: γ

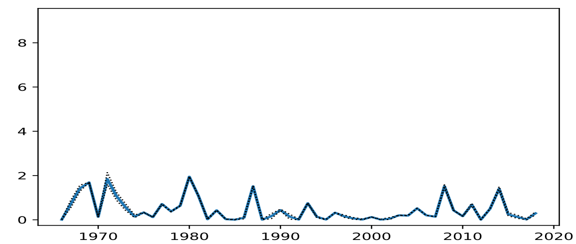


(h) Manufacturing: a

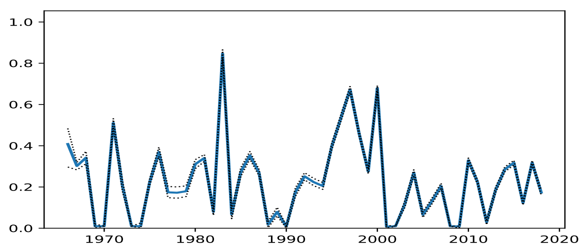
Figure A.4: **Time series of parameter estimates (continued)**



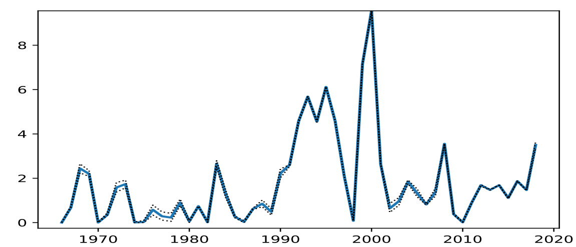
(i) Energy: γ



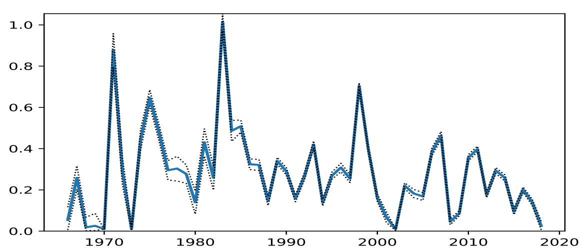
(j) Energy: a



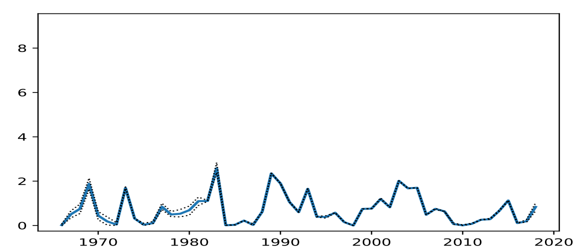
(k) Business Equipment: γ



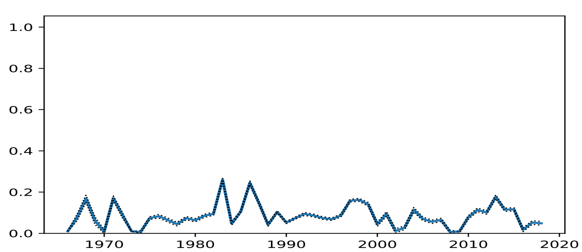
(l) Business Equipment: a



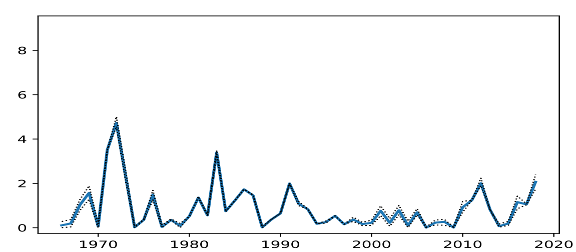
(m) Telecom: γ



(n) Telecom: a

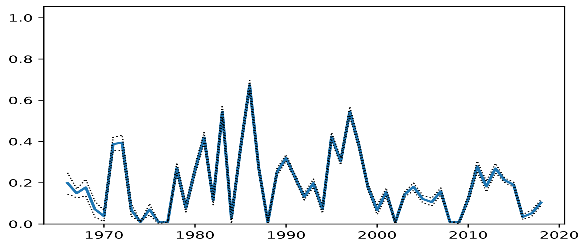


(o) Wholesale & Retail: γ

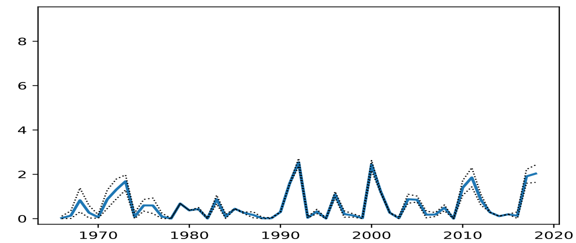


(p) Wholesale & Retail: a

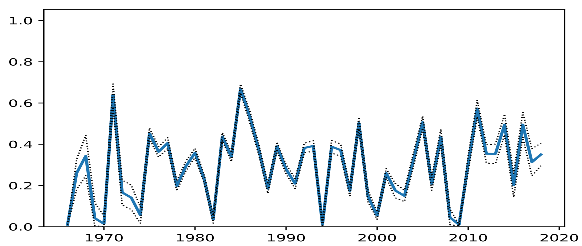
Figure A.4: **Time series of parameter estimates (continued)**



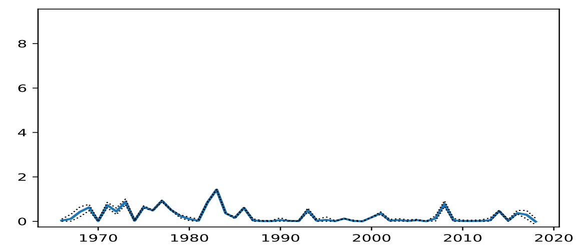
(q) Healthcare: γ



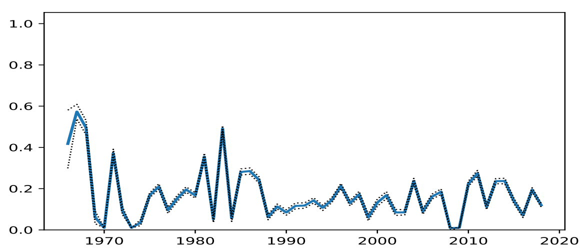
(r) Healthcare: a



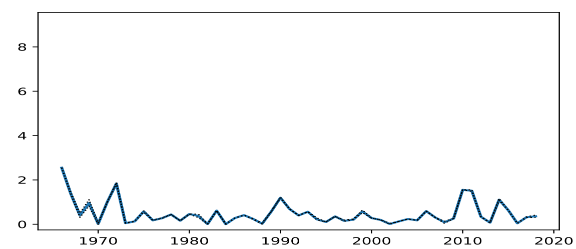
(s) Utilities: γ



(t) Utilities: a



(u) Others: γ



(v) Others: a