What Makes a Firm Complex?

Indira Puri
Stern School of Business
New York University

Preliminary work in progress; please do not distribute.
Please click here for the latest version.

August 27, 2026

Abstract

We develop a revealed measure of complexity from professional forecasters' systematic mistakes. Using repeated forecasts by the same analysts across firms, we separate predictable errors associated with the forecaster from those associated with the firm being forecast. We define firms that generate larger absolute predictable errors as revealed complex. This measure, constructed without using stock returns, nevertheless organizes return predictability: complexity amplifies predictors that require investors to process changes in information, while leaving level-based predictors largely unaffected. The results indicate that revealed complexity can organize disparate patterns in real-world markets.

1 Introduction

Why do firm characteristics predict stock returns? A central debate in finance is whether these patterns reflect compensation for risk or systematic errors in how investors form beliefs. Distinguishing these explanations is difficult because both can generate predictable returns, and showing that an anomaly is stronger in some stocks rarely identifies why. We provide a new test based on a simple implication of imperfect information processing: if predictable returns arise partly because investors fail to fully incorporate information, they should be especially pronounced when that information concerns firms that are difficult to understand. Moreover, this prediction is selective. Complexity should amplify predictors that require

investors to process changes in information, but need not amplify characteristics that simply describe a firm's state. We find evidence for this pattern. Across the over one hundred predictable return factors documented in the literature, change-based return predictors become substantially stronger among complex firms, whereas level-based predictors do not.

This prediction is difficult to test because complexity itself is not directly observable. Existing approaches therefore typically begin by specifying observable characteristics that researchers believe ought to make a firm difficult to understand, such as its organizational structure, disclosures, or the nature of its assets. But observable complication need not coincide with actual difficulty. A firm with many segments or lengthy disclosures may nevertheless be easy to forecast, while a seemingly simple firm may repeatedly confound sophisticated forecasters. More fundamentally, complexity is task-dependent: what matters for belief formation is not whether a firm looks complicated to the researcher, but whether economic agents actually find it difficult to process information about the firm.

We therefore introduce a revealed approach to measuring firm complexity. Consider the same forecaster updating beliefs about two different firms. If her forecast revisions systematically leave larger predictable errors for one firm than for the other, we call that firm more complex. The idea is to infer difficulty from the mistakes an object induces rather than from characteristics chosen ex ante to proxy for difficulty. Because we observe the same analysts repeatedly forecasting different firms, we can distinguish an analyst's general tendency to make updating mistakes from the tendency of a particular firm to generate such mistakes.

We implement this idea using individual equity analysts' earnings forecasts and revisions. Within rolling windows, we regress subsequent forecast errors on forecast revisions while allowing both the intercept and the revision coefficient to vary across analysts and firms. Analyst effects absorb a forecaster's general tendency to forecast too high or too low and to over- or underreact to information. Firm effects capture predictable errors associated specifically with updating about a given firm. We define revealed complexity as the average absolute firm-specific predictable forecast error. A firm is therefore more complex when otherwise comparable professional forecasters systematically make larger mistakes when updating beliefs about it.

The measure is constructed entirely from forecasts and realized earnings; stock returns and anomaly characteristics play no role in its construction. This provides a natural out-of-sample test of its economic content. If revealed complexity captures difficulty in processing information, then it should matter particularly when investors must incorporate changes in firm information. We test this prediction using the universe of factors collected systematically by Jensen, Kelly and Pedersen (2023). We classify characteristics according to whether they

encode a change relative to a firm's own past or expected state, such as returns, earnings surprises, growth, accruals, or issuance, or instead measure a current level, such as valuation, profitability, size, liquidity, beta, or volatility. We then reconstruct each anomaly separately among firms with low and high revealed complexity.

The results exhibit a sharp asymmetry. An equal-weighted portfolio of change-based anomalies earns a high-minus-low-complexity capital asset pricing model (CAPM) alpha of 0.49 percent per month, whereas the corresponding alpha for level-based anomalies is only 0.05 percent and statistically indistinguishable from zero. The change-minus-level difference is 0.43 percent per month and statistically significant at the one-percent level. The result is equally strong after conditioning on size or idiosyncratic volatility, when the corresponding differences are 0.51 and 0.44 percent per month. Complexity therefore does not simply amplify all return predictability among stocks that may be harder to trade, value, or arbitrage. It amplifies precisely the class of predictors for which incorporating changing information is central.

The pattern also appears across economically distinct anomaly families. Complexity strongly amplifies momentum and profit-growth strategies and produces large effects for earnings and revenue surprises and changes in profitability. Conventional price momentum, for example, has a complexity-amplification CAPM alpha of 2.72 percent per month, which remains above 2.2 percent after conditioning separately on size and idiosyncratic volatility. By contrast, several prominent level-based patterns, including value, size, and low-risk characteristics, do not become stronger with complexity. Thus, revealed complexity does more than identify stocks with high average anomaly returns: it organizes which kinds of return predictability become stronger and which do not.

Our measurement approach builds on the forecast-error-on-revision regressions of Coibion and Gorodnichenko (2015). This literature uses the sign of predictable forecast errors to study under- and over-reaction to information. We build on this elegant framework by removing analyst-specific biases and updating tendencies and focusing on the *absolute* predictable error associated with each firm. Thus, the direction of the updating error is not relevant for our purpose: larger predictable errors for the same forecasters reveal greater firm complexity.

Our findings speak to several literatures across finance and economics. First, they provide new evidence in the classical risk-versus-mispricing debate in asset pricing (Fama, 1998; Barberis, Shleifer and Vishny, 1998; Daniel, Hirshleifer and Subrahmanyam, 1998). The results do not imply that anomaly returns are unrelated to risk. But the selective amplification of change-based anomalies by an independently constructed measure of belief-formation difficulty is a natural implication of imperfect information processing and is harder to reconcile

with explanations in which complexity simply raises generic risk or arbitrage frictions.

Second, the paper contributes to the literature on expectations formation (Coibion and Gorodnichenko, 2015; Bordalo et al., 2020; Bouchaud et al., 2019). This literature shows that agents systematically under- or over-react to information. We ask a complementary question: what determines where these updating errors are largest? Our results identify persistent heterogeneity on the object side of belief formation. The same forecasters update more successfully about some firms than others.

Third, the paper contributes to work measuring firm complexity and the difficulty of processing firm information (Cohen and Lou, 2012; Loughran and Mcdonald, 2014; Peters and Taylor, 2017). Existing approaches tend to infer complexity from features of the firm or its disclosures. We instead infer it from the difficulty sophisticated agents actually have updating about the firm. This revealed approach does not require the researcher to specify in advance which observable attributes constitute complexity.

Fourth, the paper contributes to a growing literature on complexity in economic decision making (Puri (2024), Oprea (2024), Arrieta and Nielsen (2026); among many others). This literature develops theories of complexity and finds experimentally that harder problems generate greater cognitive costs, aversion, and mistakes. Some experimental papers also use behavior itself, such as decision time (Agranov, Schotter and Trevino, 2026) or errors (Enke and Shubatt, 2026), to measure complexity. In the laboratory, however, the researcher designs the task, controls the information set, and knows the correct solution. Measuring complexity in real-world settings requires a different approach. We use repeated forecasts by the same professionals across firms to separate forecaster-specific mistakes from mistakes systematically associated with the firm. This yields a revealed measure of firm complexity, which we then show is reflected in market prices. In this sense, the paper connects the emerging economics of complexity to expectation formation and real-world pricing.

More broadly, the results connect two central questions in economics: when departures from full-information belief formation arise, and whether those departures are reflected in equilibrium prices. Professional analysts operate in a setting with strong incentives, repeated opportunities to learn, and unusually rich information. Yet the same experts make systematically different updating mistakes across the firms they cover, and those differences predict where information-based return anomalies are strongest. Our results therefore suggest that predictable returns are partly organized by the difficulty of forming correct beliefs.

The paper proceeds as follows. Section 2 introduces the revealed complexity measure and its construction. Section 3 studies how complexity amplifies anomalies, and its organizing role. Section 4 concludes.

2 Revealed Complexity Measure

2.1 Thought Experiment

We use the following thought experiment. Consider the same person updating about two different firms. He makes more predictable mistakes for one firm than the other. We call the firm on which his forecast revision leaves a larger absolute predictable mistake, revealed complex.

2.2 Specification

To implement the thought experiment, we use analyst-level forecasts and revisions. Observing individual analysts allows us to separate systematic differences across analysts from systematic differences across firms. Let i index firms, y fiscal targets, j analysts, t revision dates, and q rolling-window quarters. Given the forecast revision $r_{i,y,j,t}$ and subsequent forecast errors $e_{i,y,j,t}$ (whose construction we detail in the next section), we estimate within each rolling window q,

$$e_{i,y,t,j} = \alpha_{i,q} + \gamma_{j,q} + (\beta_{i,q} + \beta_{j,q})\, r_{i,y,t,j} + \epsilon_{i,y,t,j}.$$

The analyst fixed effect $\gamma_{j,q}$ absorbs analyst j's general tendency to forecast too high or too low, and $\beta j, q$ captures analyst j's general tendency to over- or under-react to information. The firm intercept $\alpha_{i,q}$ captures persistent forecast bias for firm i, and the firm-slope $\beta_{i,q}$ captures how strongly revisions for firm i predict the forecast error that remains. Following the literature's practice for rolling analysis using analyst revisions (Dessaint, Foucault and Fresard, 2024; Bouchaud et al., 2019; De Silva and Thesmar, 2024), we estimate all coefficients jointly by ordinary least squares within each rolling window, using Newey-West heteroskedasticity robust standard errors.[1]

We now observe that – conditional on having enough variation for each firm and analyst – the above system contains, for every window q, an intercept and a slope for each of the F firms and each of the A analysts: $2F + 2A$ coefficients in total. Not all of them are separately identified. Every observation belongs to one firm and one analyst, so the data only ever reveal sums: the level $\alpha_{i,q} + \gamma_{j,q}$ and the loading $\beta_{i,q} + \beta_{j,q}$ for firm-analyst pairs that appear in the sample. Adding a constant to every firm intercept and subtracting it from every analyst intercept leaves every such sum, and hence every fitted value unchanged; the

[1]Note, however, that since the complexity measure relies on point estimates, the type of standard errors used do not affect the complexity measure.

same holds for the slopes. Because the firm-analyst network is connected,[2] the only way this can happen is through one common shift between all firm and analyst intercepts and one common shift between all firm and analyst slopes. Thus there are exactly two unidentified directions. Therefore the data provide $2F + 2A - 2$ independent restrictions on the $2F + 2A$ coefficients.

To add the two equations required for identification, we require that the average analyst fixed effect and the average analyst slope equals zero. As a result, the firm-level fixed effect and slope represent failures to update above and beyond analysts' typical failure to update in the sample, mapping back to the motivating thought experiment of Section 2.1. Formally, we require $\sum n_{j,q}\gamma_{j,q} = 0$ and $\sum n_{j,q}\beta j, q = 0$, where $n_{j,q}$ is the number of revisions by analyst j in rolling window q for any firm.

The complexity measure is then the fitted, predictable average absolute forecast error:

$$Complexity_{i,q} = \frac{1}{n_{i,q}} \sum_{(y,j,t)\in O_{i,q}} |\hat{\alpha}_{i,q} + \hat{\beta}_{i,q} r_{i,y,j,t}|, \tag{1}$$

where $O_{i,q}$ is the set of observations for firm i in window q, and $\hat{\alpha}_{i,q}$ and $\hat{\beta}_{i,q}$ are the regression-based estimates of $\alpha_{i,q}$ and $\beta_{i,q}$ after zeroing out average analyst fixed effects and slopes.

We take absolute values because complexity reflects the size, not the direction, of predictable forecast errors. Without them, offsetting positive and negative errors could make a difficult firm appear easy: in our thought experiment, any deviation from the objectively correct prediction, no matter its direction, reflects revealed failure to update. Zero therefore means no firm-specific predictable error, while larger values indicate greater revealed complexity.

2.3 Data and Cleaning

Data on analyst forecasts and realized earnings come from the I/B/E/S Unadjusted Detail and Unadjusted Actuals files, respectively. We pull all individual analysts' one- and two-year-ahead annual EPS forecasts[3] for ordinary common shares of U.S.-incorporated companies. The forecast sample covers announcements from January 1984 through December 2024. We match each forecast to its realization from the Unadjusted Actuals file. Data on stock prices,

[2]We check empirically that the firm-analyst network forms one connected component after imposing the standard data filters of Section 2.3. 99% of observations belong to one connected component.

[3]Bouchaud et al. (2019) and De Silva and Thesmar (2024) restrict forecasts to those issued within 45 days after an annual earnings announcement, which helps align analysts' information sets. Our object is instead the relation between an analyst's forecast revision and the subsequently realized error, regardless of the news prompting the revision, and we therefore retain all chronologically valid revisions.

security classifications, and cumulative share-adjustment factors come from the CRSP daily stock file. Data on firm characteristics and factors come from the Global Factor Data of Jensen, Kelly and Pedersen (2023).

We follow the literature in cleaning the I/B/E/S data and constructing forecast revisions (Clement and Tse, 2005; Bouchaud et al., 2019; De Silva and Thesmar, 2024; Jensen, Kelly and Pedersen, 2023). We define a revision pair as two consecutive EPS forecasts by the same analyst for the same firm and fiscal-period end date. Forecasts are ordered by announcement date and time, and each forecast is paired with the analyst's immediately preceding forecast. The two forecasts must occur on different calendar days, and each must strictly precede the report date of the realized annual EPS.[4]

We use CRSP cumulative share-adjustment factors to express the old forecast, the new forecast, realized EPS, and the normalization price on a common split-adjusted share basis, dividing each per-share quantity by the cumulative factor in effect on its own measurement date. Following Clement and Tse (2005), the normalization price is the CRSP price from the second trading day preceding the new forecast, adjusted using the share factor applying on that trading day.[5] For firm i, analyst j, fiscal-year target y, and new-forecast date t, we define the price-scaled revision and the subsequent realized forecast error as

$$r_{i,t,y,j} \;=\; \frac{F^{adj}_{new,i,t,y,j} - F^{adj}_{old,i,t,y,j}}{P^{adj}_{i,t-2}}, \qquad e_{i,t,y,j} \;=\; \frac{A^{adj}_{i,y} - F^{adj}_{new,i,t,y,j}}{P^{adj}_{i,t-2}}, \tag{2}$$

where $F^{adj}_{new,i,t,y,j}$ and $F^{adj}_{old,i,t,y,j}$ are the split-adjusted new and immediately preceding forecasts, $A^{adj}_{i,y}$ is split-adjusted realized annual EPS, and $P^{adj}_{i,t-2}$ is the split-adjusted stock price two trading days before the new forecast. Thus, $r_{i,t,y,j} > 0$ denotes an upward forecast revision, while $e_{i,t,y,j} > 0$ indicates that realized EPS exceeded the analyst's new forecast. Finally, following the prior literature, we remove observations for which the price-scaled new forecast, the price-scaled realized EPS, or the realized forecast error lies outside the interval $[Q_1 - 10\,IQR,\ Q_3 + 10\,IQR]$, where the quartiles are computed variable by variable over the pooled sample.

To implement the complexity measure of Section 2.2, we use a rolling window length of 12 months. To help ensure enough variation to estimate intercepts and slopes, we follow prior literature on rolling analyst revisions Dessaint, Foucault and Fresard (2024) in requiring a

[4]This procedure follows Ivkovic and Jegadeesh (2004) and Clement and Tse (2005) in retaining all chronologically valid forecast revisions; useful for our exercise because our object of interest is the analyst- and firm-level relation between the forecast revision and the subsequently realized error, retaining all valid revisions helps maximize power.

[5]They, and we, use the price two trading days before the new forecast so that the normalization price is not affected by information released with the forecast revision.

minimum of 20 observations, in our case, both for each firm as well as for each analyst within the window; this requirement retains 87% percent of the data.

3 Revealed Complexity Organizes Anomalies

We now ask whether revealed complexity organizes empirical patterns in finance. This constitutes an out-of-sample test of the measure's economic content, because stock returns and anomaly characteristics do not enter the construction of revealed complexity (Section 2.2).

Our hypothesis is that complexity should matter most when investors must process changing firm information and update beliefs about future fundamentals. This is because the revealed complexity measure shows which firms the same experts make more updating mistakes on.

As a result, complexity should organize the factor zoo in finance, amplifying predictable patterns that are based on changes, but having little effect on predictable patterns based on more stable firm characteristics. We find economically and statistically significant evidence for this hypothesis below.

3.1 Data and Portfolio Construction

We use the 153 global factor database of Jensen, Kelly and Pedersen (2023) (henceforth JKP), who collect all factors found in the finance literature into a readily accessible database. We match each stock-month to the most recent quarterly observation of revealed complexity. The matched sample contains 788,966 stock-months from 8,040 firms. The median month contains 1,786 firms.

Each month, we sort firms into terciles of complexity. We then construct each anomaly separately within each complexity tercile. For anomaly k, complexity tercile c, ad month t, let $R^{H,c}_{k,t}$ and $R^{L,c}_{k,t}$ denote the value-weighted one-month-ahead excess returns of the high and low characteristic portfolios. By high characteristic, we mean 70th percentile of the characteristic within the complexity tercile; by low, 30th percentile.[6]

We orient each anomaly in the direction reported by JKP. Let $d_k \in \{-1, 1\}$ denot the direction indicator. If $d_k = 1$, the documented anomaly return is high minus low; if -1, low minus high. We define $A^c_{k,t} \coloneqq d_k \left(R^{H,c}_{k,t} - R^{L,c}_{k,t} \right)$. This normalization gives every anomaly sort the same directional interpretation: a positive $A^c_{k,t}$ means the anomaly earns a positive return in its documented direction.

[6]These percentiles breaks follow Fama and French (1993) and subsequent work.

Our object of interest is:

$$\text{Complexity Amplification for Anomaly } k \text{ in month } t \coloneqq \Delta A_{k,t} \coloneqq A^3_{k,t} - A^1_{k,t},$$

where terciles 1 and 3 contain the least and most complex firms, respectively. A positive $\Delta A_{k,t}$ means that the anomaly is stronger among high-complexity firms. We call this amplification.

We next check, for robustness, whether this relation survives conditioning on size or idiosyncratic volatility, characteristics known to amplify anomalies in general (Fama and French, 2008). To do so, we construct triple sorts. For each control variable (size or idiosyncratic volatility) we follow standard practice in first sorting firms into monthly terciles using breakpoints estimated from NYSE firms (Fama and French, 1993). We measure size as market equity, and idiosyncratic volatility as the 21-day Fama-French three factor idiosyncratic volatility measure (Ang et al., 2006). We then form complexity terciles within each control tercile. Finally, we estimate the anomaly's 30/70 characteristic cutoffs within each control-by-complexity cell and form the same direction-normalized, value-weighted anomaly return. For a given complexity tercile, we equal-weight the anomaly returns across the three control terciles.

For each of: the complexity sort; the size-then-complexity sort; and the idiosyncratic volatility-then-complexity sort, we report the mean excess return amplification $\Delta A_{k,t}$ and its corresponding CAPM alpha, estimated the usual way as $\Delta A_{k,t} = \alpha^{\Delta}_k + \beta_k R^{MKT}_t + u_{k,t}$, with R^{MKT}_t the market return and $u_{k,t}$ a noise term. Inference throughout uses Newey-West heteroskedasticity and autocorrelation robust standard errors.

3.2 Complexity Amplification by Factor

Complexity amplifies many anomalies. The pattern survives both robustness tests. In the baseline specification, 89 of the 163 anomalies have positive CAPM differences, with 51 statistically significant at the 5% level. After conditioning on size, the corresponding counts increase to 99 and 58; and after conditioning on idiosyncratic volatility, 94 and 46.

The portfolio $\Delta A_{k,t}$ has a direct trading interpretation. It is long the documented anomaly strategy among high-complexity firms and short the same anomaly strategy among low-complexity firms. Thus, the raw mean of $\Delta A_{k,t}$ is the spread return to the complexity-amplification strategy before implementation costs, while α^{Δ}_k is its CAPM-adjusted spread return.

The magnitude of these differences is economically large. Consider price momentum. In the baseline specification, the complexity-amplification portfolio earns 2.36 percent per

month before risk adjustment and has a CAPM alpha of 2.72 percent per month. The alpha remains 2.27 percent after conditioning on size and 2.26 percent after conditioning on idiosyncratic volatility. These figures are before implementation costs and do not account for transaction costs, short-sale costs, or other trading frictions.

Table 3.1 shows complexity amplification results across the factor themes statistically defined by JKP. Momentum is the strongest theme in every specification. Profit growth is the next strongest theme. Profitability, quality, and debt issuance are also positive across all three specifications. By contrast, value is negative, while short-term reversal and size become weaker rather than stronger among high-complexity firms. Complexity therefore does not mechanically amplify every anomaly.

Table 3.2 reports the individual anomalies with the largest complexity-amplification alphas. The largest effects are concentrated in momentum, earnings changes, and surprises. Several exceed one percent per month even after conditioning on size and idiosyncratic volatility. The largest baseline estimate is 3.15 percent per month for the year-one lagged return strategy. The conventional momentum strategy has a complexity amplification alpha of 2.72 percent per month. Changes in quarterly return on assets and changes in quarterly return on equity have alphas of 1.38 percent per month, and standardized earnings surprise has an alpha of 1.29 percent. These effects remain economically large in both controlled specifications. Appendix A.1 reports results for all 153 characteristics.

3.3 Complexity Organizes Anomalies

The theme-level ordering is consistent with the updating-based hypothesis: the amplified themes condition on news, while the unaffected themes condition on states. To test this organizing hypothesis directly, we classify all 153 characteristics by their construction into whether the characteristic is change- or level-based.

A change-based characteristic encodes movement relative to another state of the same firm or price: explicit changes and growth rates, earnings and revenue surprises, abnormal values relative to a firm's own benchmark, accruals, issuance, and return-based measures, because a return is itself a price change over an interval.

A level-based characteristic measures the current state of a firm, or a rolling statistic of its data, without comparison to the firm's own past or expected value: valuation ratios, profitability ratios, size, liquidity, betas, volatilities, and skewness. The other category collects composite scores that mix factors, for example, the Altman Z-score. The rule yields 67 Change, 79 Level, and 7 Other signals.

We then re-perform the same complexity portfolio sorts on changes vs. levels. Table 3.3

TABLE 3.1: Complexity Amplification Across Themes

Theme	Complexity	Size → Complexity	IVOL → Complexity
Momentum	2.37***	1.88***	1.96***
	(0.24)	(0.14)	(0.16)
Profit Growth	0.96***	0.78***	0.77***
	(0.14)	(0.07)	(0.10)
Profitability	0.64***	0.25**	0.32***
	(0.17)	(0.10)	(0.12)
Quality	0.55***	0.27**	0.41***
	(0.18)	(0.11)	(0.13)
Debt Issuance	0.33***	0.45***	0.26***
	(0.11)	(0.07)	(0.08)
Low Leverage	0.22	0.28***	0.30***
	(0.16)	(0.10)	(0.11)
Seasonality	0.13*	0.08*	0.02
	(0.07)	(0.05)	(0.06)
Accruals	0.11	0.16*	0.25**
	(0.16)	(0.09)	(0.10)
Low Risk	−0.06	−0.01	−0.07
	(0.16)	(0.07)	(0.09)
Investment	−0.15	0.22**	−0.02
	(0.17)	(0.10)	(0.11)
Value	−0.19	−0.33***	−0.39***
	(0.14)	(0.11)	(0.12)
Short-Term Reversal	−0.42***	−0.31***	−0.16
	(0.11)	(0.07)	(0.10)
Size	−0.67***	−0.18*	−0.60***
	(0.17)	(0.09)	(0.13)

Notes: Entries are CAPM alphas of the high-complexity anomaly return minus the low-complexity anomaly return, equal-weighted within theme, in percent per month. In the baseline, anomaly legs use 30/70 NYSE breakpoints estimated within each complexity tercile. The controlled specifications first form NYSE terciles of the control, then complexity terciles within each control tercile, then anomaly legs within each control-by-complexity cell, and average across control terciles. Newey-West standard errors are in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels.

TABLE 3.2: Largest Complexity-Amplified Anomalies

Rank	Anomaly	Complexity	Size → Complexity	IVOL → Complexity
1	Year 1-lagged return, nonannual	3.15***	2.51***	2.59***
		(0.32)	(0.19)	(0.23)
2	Price momentum $t-9$ to $t-1$	2.80***	2.19***	2.18***
		(0.33)	(0.18)	(0.21)
3	Price momentum $t-12$ to $t-1$	2.72***	2.27***	2.26***
		(0.32)	(0.18)	(0.20)
4	Current price to high price over last year	2.53***	2.06***	2.21***
		(0.27)	(0.18)	(0.20)
5	Price momentum $t-6$ to $t-1$	2.45***	1.89***	1.99***
		(0.29)	(0.16)	(0.19)
6	Residual momentum $t-12$ to $t-1$	2.05***	1.62***	1.74***
		(0.24)	(0.14)	(0.18)
7	Price momentum $t-12$ to $t-7$	1.76***	1.21***	1.43***
		(0.26)	(0.16)	(0.20)
8	Price momentum $t-3$ to $t-1$	1.73***	1.28***	1.35***
		(0.24)	(0.13)	(0.18)
9	Residual momentum $t-6$ to $t-1$	1.49***	1.22***	1.33***
		(0.26)	(0.14)	(0.20)
10	Change in quarterly return on assets	1.38***	1.36***	1.11***
		(0.23)	(0.14)	(0.17)
11	Change in quarterly return on equity	1.38***	1.31***	1.14***
		(0.24)	(0.14)	(0.18)
12	Standardized earnings surprise	1.29***	1.26***	1.15***
		(0.24)	(0.14)	(0.17)
13	Mispricing factor: performance	1.16***	0.81***	0.93***
		(0.28)	(0.17)	(0.21)
14	Piotroski F-score	1.15***	0.62***	0.84***
		(0.21)	(0.11)	(0.16)
15	Change sales minus change SG&A	1.11***	0.57***	0.69***
		(0.29)	(0.16)	(0.20)
16	Tax expense surprise	1.09***	0.83***	0.79***
		(0.24)	(0.14)	(0.17)
17	Quarterly return on equity	1.05***	0.51***	0.55***
		(0.23)	(0.15)	(0.21)
18	Quarterly return on assets	1.01***	0.55***	0.44**
		(0.26)	(0.15)	(0.19)
19	Asset tangibility	0.95***	0.83***	0.90***
		(0.27)	(0.17)	(0.22)
20	Standardized revenue surprise	0.83***	0.90***	0.92***
		(0.23)	(0.14)	(0.18)

Notes: The table reports the twenty largest positive complexity-amplification effects, ranked by the baseline CAPM alpha. Entries are CAPM alphas of the high-complexity anomaly return minus the low-complexity anomaly return, in percent per month. The controlled specifications condition first on size or idiosyncratic volatility as described in the text. Newey-West standard errors are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

reports the results. Panel A equal-weights anomaly returns within each category and reports the CAPM amplification of the resulting portfolios. The Change portfolio earns 0.49 percent per month in the baseline, 0.52 percent after conditioning on size, and 0.43 percent after conditioning on idiosyncratic volatility, all statistically significant at the 1 percent level. In contrast, the Level portfolio earns 0.05, 0.00, and −0.01 percent; none is statistically distinguishable from zero.

Panel B tests the difference directly by forming the Change-minus-Level amplification portfolio; the difference is 0.43 to 0.51 percent per month and statistically significant at the 1 percent level in every specification.

These results indicate that complexity organizes anomalies, tending to exacerbate change-based, but not level-based, factors in the literature.

TABLE 3.3: Complexity Amplification by Signal Type: Change vs. Level

	Complexity	Size → Complexity	IVOL → Complexity
Panel A. Category Portfolios			
Change	0.49***	0.52***	0.43***
	(0.06)	(0.04)	(0.05)
Level	0.05	0.00	−0.01
	(0.06)	(0.04)	(0.04)
Panel B. Change − Level			
Raw return	0.44***	0.53***	0.45***
	(0.08)	(0.05)	(0.06)
CAPM alpha	0.43***	0.51***	0.44***
	(0.08)	(0.06)	(0.06)

Notes: Panel A reports CAPM alphas of the high-minus-low complexity return on portfolios that equal-weight anomaly returns within each signal category, in percent per month. Panel B reports the raw excess return and CAPM alpha of the Change amplification portfolio minus the Level amplification portfolio. Newey-West standard errors are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4 Conclusion

This paper develops a revealed measure of firm complexity from the systematic mistakes professional forecasters make when updating beliefs. By comparing repeated forecasts by the same analysts across firms, we separate predictable errors associated with the forecaster from those associated with the firm being forecast. Complexity amplifies factors based on changes in information, while leaving level-based predictors largely unaffected. The pattern suggests that the difficulty of forming and updating beliefs is an important dimension of

when predictable returns emerge in markets.

More broadly, the results point to a useful way of studying complexity outside controlled settings. In many economic environments, researchers do not observe task difficulty directly and cannot define the correct benchmark ex ante. Repeated behavior can nevertheless reveal which objects systematically generate mistakes. This perspective shifts attention from the characteristics an object appears to have toward the difficulty agents experience when reasoning about it. The revealed complexity measure introduced here may therefore provide a useful framework for studying belief formation, decision making, and outcomes in market settings.

References

Agranov, Marina, Andrew Schotter, and Isabel Trevino. 2026. "Complex for Whom? An Experimental Approach to Subjective Complexity." *Working Paper*.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2006. "The Cross-Section of Volatility and Expected Returns." *The Journal of Finance*, 61(1): 259–299.

Arrieta, Gonzalo, and Kirby Nielsen. 2026. "Procedural Decision-Making in the Face of Complexity." *Working Paper*.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. 1998. "A model of investor sentiment." *Journal of Financial Economics*, 49(3): 307–343.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2020. "Overreaction in Macroeconomic Expectations." *American Economic Review*, 110(9): 2748–2782.

Bouchaud, Jean-Philippe, Philipp Kruger, Augustin Landier, and David Thesmar. 2019. "Sticky Expectations and the Profitability Anomaly." *The Journal of Finance*, 74(2): 639–674.

Clement, Michael B., and Senyo Y. Tse. 2005. "Financial Analyst Characteristics and Herding Behavior in Forecasting." *The Journal of Finance*, 60(1): 307–341.

Cohen, Lauren, and Dong Lou. 2012. "Complicated firms." *Journal of Financial Economics*, 104(2): 383–400.

Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review*, 105(8): 2644–2678.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam. 1998. "Investor Psychology and Security Market Under- and Overreactions." *The Journal of Finance*, 53(6): 1839–1885.

De Silva, Tim, and David Thesmar. 2024. "Noise in Expectations: Evidence from Analyst Forecasts." *The Review of Financial Studies*, 37(5): 1494–1537.

Dessaint, Olivier, Thierry Foucault, and Laurent Fresard. 2024. "Does Alternative Data Improve Financial Forecasting? The Horizon Effect." *The Journal of Finance*, 79(3): 2237–2287.

Enke, Benjamin, and Cassidy Shubatt. 2026. "Quantifying Lottery Choice Complexity." *Working Paper*.

Fama, Eugene F. 1998. "Market efficiency, long-term returns, and behavioral finance1The comments of Brad Barber, David Hirshleifer, S.P. Kothari, Owen Lamont, Mark Mitchell, Hersh Shefrin, Robert Shiller, Rex Sinquefield, Richard Thaler, Theo Vermaelen, Robert Vishny, Ivo Welch, and a referee have been helpful. Kenneth French and Jay Ritter get special thanks.1." *Journal of Financial Economics*, 49(3): 283–306.

Fama, Eugene F., and Kenneth R. French. 1993. "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics*, 33(1): 3–56.

Fama, Eugene F., and Kenneth R. French. 2008. "Dissecting Anomalies." *The Journal of Finance*, 63(4): 1653–1678.

Ivkovic, Zoran, and Narasimhan Jegadeesh. 2004. "The timing and value of forecast and recommendation revisions." *Journal of Financial Economics*, 73(3): 433–463.

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen. 2023. "Is There a Replication Crisis in Finance?" *The Journal of Finance*, 78(5): 2465–2518.

Loughran, Tim, and Bill Mcdonald. 2014. "Measuring Readability in Financial Disclosures." *The Journal of Finance*, 69(4): 1643–1671.

Oprea, Ryan. 2024. "Decisions under Risk Are Decisions under Complexity." *American Economic Review*, 114(12): 3789–3811.

Peters, Ryan H., and Lucian A. Taylor. 2017. "Intangible capital and the investment-q relation." *Journal of Financial Economics*, 123(2): 251–272.

Puri, Indira. 2024. "Simplicity and Risk." *The Journal of Finance*, jofi.13417.

A Appendix: Additional Factor Tests

A.1 All Factors

TABLE A.1: Factor-Level Complexity Amplification

Anomaly	JKP Variable	JKP Theme	Change or Level	Complexity	Size → Complexity	IVOL → Complexity
Change in current operating working capital	cowc_gr1a	Accruals	Change	0.52**	0.59***	0.80***
				(0.24)	(0.12)	(0.19)
Operating accruals	oaccruals_at	Accruals	Change	0.59**	0.59***	0.71***
				(0.25)	(0.14)	(0.16)
Percent operating accruals	oaccruals_ni	Accruals	Change	0.10	0.16	0.26
				(0.30)	(0.15)	(0.22)
Year 16-20 lagged returns, non annual	seas_16_20na	Accruals	Change	−0.02	0.05	0.03
				(0.30)	(0.16)	(0.25)
Total accruals	taccruals_at	Accruals	Change	−0.19	−0.16	−0.04
				(0.20)	(0.13)	(0.15)
Percent total accruals	taccruals_ni	Accruals	Change	−0.33	−0.25**	−0.25
				(0.20)	(0.12)	(0.16)
Abnormal corporate investment	capex_abn	Debt Issuance	Change	0.42*	0.61***	0.29*
				(0.23)	(0.14)	(0.15)
Growth in book debt (3 years)	debt_gr3	Debt Issuance	Change	0.46**	0.54***	0.32*
				(0.21)	(0.12)	(0.17)
Change in financial liabilities	fnl_gr1a	Debt Issuance	Change	0.73***	0.66***	0.58***
				(0.19)	(0.12)	(0.15)
Change in concurrent operating liabilities	ncol_gr1a	Debt Issuance	Change	−0.35	0.12	−0.27
				(0.23)	(0.11)	(0.17)
Change in net financial assets	nfna_gr1a	Debt Issuance	Change	0.45**	0.61***	0.40***
				(0.20)	(0.11)	(0.15)
Earnings persistence	ni_ar1	Debt Issuance	Level	−0.04	−0.11	−0.02
				(0.22)	(0.11)	(0.16)
Net operating assets	noa_at	Debt Issuance	Level	0.67***	0.71***	0.49***
				(0.23)	(0.15)	(0.19)
Liquidity of book assets	aliq_at	Investment	Level	−0.31	−0.15	0.04
				(0.22)	(0.14)	(0.18)
Asset growth	at_gr1	Investment	Change	−0.25	−0.03	−0.22
				(0.22)	(0.14)	(0.18)

Continued on next page

Table A.1 – *continued from previous page*

Anomaly	JKP Variable	JKP Theme	Type	Complexity	Size → Complexity	IVOL → Complexity
Change in common equity	be_gr1a	Investment	Change	−0.35	−0.15	−0.31
				(0.26)	(0.15)	(0.19)
CAPEX growth (1 year)	capx_gr1	Investment	Change	−0.11	0.19	0.08
				(0.25)	(0.14)	(0.20)
CAPEX growth (2 year)	capx_gr2	Investment	Change	0.19	0.48***	0.16
				(0.26)	(0.15)	(0.19)
CAPEX growth (3 years)	capx_gr3	Investment	Change	0.20	0.46***	0.11
				(0.26)	(0.15)	(0.20)
Change in current operating assets	coa_gr1a	Investment	Change	−0.30	0.17	−0.11
				(0.23)	(0.13)	(0.17)
Change in current operating liabilities	col_gr1a	Investment	Change	−0.86***	−0.34**	−0.65***
				(0.27)	(0.14)	(0.19)
Hiring rate	emp_gr1	Investment	Change	0.29	0.40***	0.36**
				(0.21)	(0.14)	(0.18)
Inventory growth	inv_gr1	Investment	Change	−0.02	0.34***	0.23
				(0.22)	(0.13)	(0.17)
Inventory change	inv_gr1a	Investment	Change	−0.03	0.53***	0.16
				(0.23)	(0.12)	(0.16)
Change in long-term net operating assets	lnoa_gr1a	Investment	Change	0.11	0.56***	0.17
				(0.20)	(0.11)	(0.16)
Mispricing factor: management	mispricing_mgmt	Investment	Other	0.29	0.36***	0.17
				(0.21)	(0.14)	(0.18)
Change in concurrent operating assets	ncoa_gr1a	Investment	Change	0.30	0.63***	0.26*
				(0.21)	(0.12)	(0.16)
Change in net concurrent operating assets	nncoa_gr1a	Investment	Change	0.44**	0.63***	0.35**
				(0.21)	(0.12)	(0.17)
Change in net operating assets	noa_gr1a	Investment	Change	0.51**	0.73***	0.52***
				(0.21)	(0.13)	(0.17)
Change PPE and Inventory	ppeinv_gr1a	Investment	Change	−0.01	0.58***	0.06
				(0.24)	(0.14)	(0.19)
Long-term reversal	ret_60_12	Investment	Change	−0.42	0.27*	−0.09
				(0.30)	(0.16)	(0.20)
Sales Growth (1 year)	sale_gr1	Investment	Change	−0.86***	−0.35*	−0.68***
				(0.28)	(0.18)	(0.22)
Sales Growth (3 years)	sale_gr3	Investment	Change	−0.45*	0.18	−0.01
				(0.26)	(0.15)	(0.21)

Continued on next page

Table A.1 – *continued from previous page*

Anomaly	JKP Variable	JKP Theme	Type	Complexity	Size → Complexity	IVOL → Complexity
Sales growth (1 quarter)	saleq_gr1	Investment	Change	−1.12***	−0.89***	−0.99***
				(0.28)	(0.17)	(0.20)
Years 2-5 lagged returns, nonannual	seas_2_5na	Investment	Change	−0.47	0.25	−0.15
				(0.30)	(0.16)	(0.21)
Firm age	age	Low Leverage	Level	−0.28	0.14	0.13
				(0.24)	(0.12)	(0.17)
Liquidity of market assets	aliq_mat	Low Leverage	Level	−0.23	−0.41***	−0.29*
				(0.23)	(0.15)	(0.17)
Book leverage	at_be	Low Leverage	Level	0.67**	0.19	0.44**
				(0.26)	(0.15)	(0.19)
High-low bid-ask spread	bidaskhl_21d	Low Leverage	Level	−0.30	0.02	−0.18
				(0.26)	(0.15)	(0.17)
Cash-to-assets	cash_at	Low Leverage	Level	0.70**	0.67***	0.68***
				(0.28)	(0.16)	(0.18)
Net debt to price	netdebt_me	Low Leverage	Level	0.59***	0.65***	0.54***
				(0.20)	(0.16)	(0.18)
Earnings volatility	ni_ivol	Low Leverage	Level	−0.01	0.13	0.24
				(0.23)	(0.15)	(0.20)
RD to sales	rd_sale	Low Leverage	Level	−0.15	0.31	0.12
				(0.37)	(0.22)	(0.22)
RD capital to book assets	rd5_at	Low Leverage	Level	0.10	0.26	0.19
				(0.37)	(0.24)	(0.22)
Asset tangibility	tangibility	Low Leverage	Level	0.95***	0.83***	0.90***
				(0.27)	(0.17)	(0.22)
Altman Z-score	z_score	Low Leverage	Other	0.38	0.27*	0.52**
				(0.26)	(0.16)	(0.21)
Market Beta	beta_60m	Low Risk	Level	−0.17	0.05	−0.12
				(0.24)	(0.16)	(0.17)
Dimson beta	beta_dimson_21d	Low Risk	Level	−0.12	0.21	−0.05
				(0.21)	(0.15)	(0.16)
Frazzini-Pederson market beta	betabab_1260d	Low Risk	Level	−0.18	0.24	−0.07
				(0.28)	(0.17)	(0.19)
Downside beta	betadown_252d	Low Risk	Level	−0.57**	−0.18	−0.48***
				(0.23)	(0.15)	(0.17)
Earnings variability	earnings_variability	Low Risk	Level	0.14	−0.08	−0.20
				(0.21)	(0.13)	(0.17)

Continued on next page

Table A.1 – *continued from previous page*

Anomaly	JKP Variable	JKP Theme	Type	Complexity	Size → Complexity	IVOL → Complexity
Idiosyncratic volatility from the CAPM (21 days)	ivol_capm_21d	Low Risk	Level	0.36	0.01	0.24
				(0.25)	(0.15)	(0.15)
Idiosyncratic volatility from the CAPM (252 days)	ivol_capm_252d	Low Risk	Level	0.25	−0.06	−0.16
				(0.24)	(0.15)	(0.17)
Idiosyncratic volatility from the Fama-French 3-factor model	ivol_ff3_21d	Low Risk	Level	0.22	−0.02	–
				(0.24)	(0.14)	
Idiosyncratic volatility form the q-factor model	ivol_hxz4_21d	Low Risk	Level	0.25	0.09	0.26*
				(0.24)	(0.15)	(0.15)
Cash flow volatility	ocfq_saleq_std	Low Risk	Level	0.07	−0.12	0.08
				(0.23)	(0.13)	(0.19)
Maximum daily return	rmax1_21d	Low Risk	Level	0.06	−0.27**	0.11
				(0.23)	(0.13)	(0.16)
Highest 5 days of return	rmax5_21d	Low Risk	Level	−0.27	−0.32**	−0.40**
				(0.24)	(0.14)	(0.16)
Return volatility	rvol_21d	Low Risk	Level	0.29	0.10	0.20
				(0.24)	(0.16)	(0.16)
Years 6-10 lagged returns, nonannual	seas_6_10na	Low Risk	Change	−0.10	0.25*	0.08
				(0.29)	(0.14)	(0.19)
Share turnover	turnover_126d	Low Risk	Level	−0.21	−0.03	−0.03
				(0.23)	(0.13)	(0.17)
Number of zero trades with turnover as tiebreaker (1 month)	zero_trades_21d	Low Risk	Level	−0.41*	−0.23	−0.37**
				(0.23)	(0.14)	(0.18)
Number of zero trades with turnover as tiebreaker (6 months)	zero_trades_126d	Low Risk	Level	−0.23	−0.03	−0.05
				(0.23)	(0.13)	(0.17)
Number of zero trades with turnover as tiebreaker (12 months)	zero_trades_252d	Low Risk	Level	−0.41*	0.12	−0.18
				(0.22)	(0.13)	(0.16)
Current to high price over last year	prc_highprc_252d	Momentum	Change	2.53***	2.06***	2.21***
				(0.27)	(0.18)	(0.20)
Residual momentum t-6 to t-1	resff3_6_1	Momentum	Change	1.49***	1.22***	1.33***
				(0.26)	(0.14)	(0.20)
Residual momentum t-12 to t-1	resff3_12_1	Momentum	Change	2.05***	1.62***	1.74***
				(0.24)	(0.14)	(0.18)

Continued on next page

Table A.1 – *continued from previous page*

Anomaly	JKP Variable	JKP Theme	Type	Complexity	Size → Complexity	IVOL → Complexity
Price momentum t-3 to t-1	ret_3_1	Momentum	Change	1.73***	1.28***	1.35***
				(0.24)	(0.13)	(0.18)
Price momentum t-6 to t-1	ret_6_1	Momentum	Change	2.45***	1.89***	1.99***
				(0.29)	(0.16)	(0.19)
Price momentum t-9 to t-1	ret_9_1	Momentum	Change	2.80***	2.19***	2.18***
				(0.33)	(0.18)	(0.21)
Price momentum t-12 to t-1	ret_12_1	Momentum	Change	2.72***	2.27***	2.26***
				(0.32)	(0.18)	(0.20)
Year 1-lagged return, nonannual	seas_1_1na	Momentum	Change	3.15***	2.51***	2.59***
				(0.32)	(0.19)	(0.23)
Change sales minus change inventory	dsale_dinv	Profit Growth	Change	0.57***	0.53***	0.48***
				(0.21)	(0.11)	(0.17)
Change sales minus change receivables	dsale_drec	Profit Growth	Change	0.05	−0.01	−0.05
				(0.22)	(0.12)	(0.17)
Change sales minus change SGA	dsale_dsga	Profit Growth	Change	1.11***	0.57***	0.69***
				(0.29)	(0.16)	(0.20)
Change in quarterly return on assets	niq_at_chg1	Profit Growth	Change	1.38***	1.36***	1.11***
				(0.23)	(0.14)	(0.17)
Change in quarterly return on equity	niq_be_chg1	Profit Growth	Change	1.38***	1.31***	1.14***
				(0.24)	(0.14)	(0.18)
Standardized earnings surprise	niq_su	Profit Growth	Change	1.29***	1.26***	1.15***
				(0.24)	(0.14)	(0.17)
Change in operating cash flow to assets	ocf_at_chg1	Profit Growth	Change	0.75***	0.65***	0.71***
				(0.25)	(0.13)	(0.18)
Price momentum t-12 to t-7	ret_12_7	Profit Growth	Change	1.76***	1.21***	1.43***
				(0.26)	(0.16)	(0.20)
Labor force efficiency	sale_emp_gr1	Profit Growth	Change	0.57**	0.29**	0.29*
				(0.22)	(0.12)	(0.18)
Standardized revenue surprise	saleq_su	Profit Growth	Change	0.83***	0.90***	0.92***
				(0.23)	(0.14)	(0.18)
Year 1-lagged return, annual	seas_1_1an	Profit Growth	Change	0.78***	0.47***	0.54***
				(0.22)	(0.13)	(0.17)
Tax expense surprise	tax_gr1a	Profit Growth	Change	1.09***	0.83***	0.79***
				(0.24)	(0.14)	(0.17)

Continued on next page

Table A.1 – *continued from previous page*

Anomaly	JKP Variable	JKP Theme	Type	Complexity	Size → Complexity	IVOL → Complexity
Coefficient of variation for dollar trading volume	dolvol_var_126d	Profitability	Level	0.23	0.21*	−0.01
				(0.22)	(0.13)	(0.16)
Return on net operating assets	ebit_bev	Profitability	Level	0.69***	0.34**	0.32*
				(0.26)	(0.15)	(0.18)
Profit margin	ebit_sale	Profitability	Level	0.47**	0.10	0.06
				(0.20)	(0.13)	(0.15)
Pitroski F-score	f_score	Profitability	Other	1.15***	0.62***	0.84***
				(0.21)	(0.11)	(0.16)
Return on equity	ni_be	Profitability	Level	0.61**	0.14	0.18
				(0.26)	(0.16)	(0.18)
Quarterly return on equity	niq_be	Profitability	Level	1.05***	0.51***	0.55***
				(0.23)	(0.15)	(0.21)
Ohlson O-score	o_score	Profitability	Other	0.37	0.21	0.06
				(0.24)	(0.14)	(0.20)
Operating cash flow to assets	ocf_at	Profitability	Level	0.82***	0.45***	0.57***
				(0.27)	(0.16)	(0.20)
Operating profits-to-book equity	ope_be	Profitability	Level	0.47*	−0.07	0.26
				(0.24)	(0.15)	(0.20)
Operating profits-to-lagged book equity	ope_bel1	Profitability	Level	0.49*	−0.02	0.23
				(0.27)	(0.15)	(0.20)
Coefficient of variation for share turnover	turnover_var_126d	Profitability	Level	0.69***	0.30**	0.46***
				(0.19)	(0.13)	(0.16)
Capital turnover	at_turnover	Quality	Level	0.25	−0.06	0.12
				(0.21)	(0.16)	(0.16)
Cash-based operating profits-to-book-assets	cop_at	Quality	Level	0.72***	0.33**	0.67***
				(0.24)	(0.16)	(0.17)
Cash-based operating profits-to-lagged book assets	cop_atl1	Quality	Level	0.76***	0.35**	0.68***
				(0.25)	(0.16)	(0.19)
Change gross margin minus change sales	dgp_dsale	Quality	Change	0.54**	0.50***	0.67***
				(0.21)	(0.13)	(0.17)
Gross profits-to-assets	gp_at	Quality	Level	0.13	−0.12	0.10
				(0.24)	(0.18)	(0.18)

Continued on next page

Table A.1 – *continued from previous page*

Anomaly	JKP Variable	JKP Theme	Type	Complexity	Size → Complexity	IVOL → Complexity
Gross profits-to-lagged assets	gp_atl1	Quality	Level	0.24	−0.08	0.14
				(0.25)	(0.16)	(0.18)
Mispricing factor: performance	mispricing_perf	Quality	Other	1.16***	0.81***	0.93***
				(0.28)	(0.17)	(0.21)
Number of consecutive quarters with earnings increases	ni_inc8q	Quality	Change	0.59***	0.42***	0.59***
				(0.19)	(0.11)	(0.15)
Quarterly return on assets	niq_at	Quality	Level	1.01***	0.55***	0.44**
				(0.26)	(0.15)	(0.19)
Operating profits-to-book assets	op_at	Quality	Level	0.31	0.07	0.32*
				(0.25)	(0.16)	(0.18)
Operating profits-to-lagged book assets	op_atl1	Quality	Level	0.32	0.05	0.28
				(0.24)	(0.16)	(0.19)
Operating leverage	opex_at	Quality	Level	−0.08	0.03	0.19
				(0.18)	(0.15)	(0.16)
Quality minus junk: composite	qmj	Quality	Other	0.81***	0.38***	0.38**
				(0.24)	(0.15)	(0.18)
Quality minus junk: growth	qmj_growth	Quality	Change	0.69***	0.41***	0.54***
				(0.27)	(0.15)	(0.18)
Quality minus junk: profitability	qmj_prof	Quality	Level	0.73***	0.18	0.45***
				(0.24)	(0.14)	(0.17)
Quality minus junk: safety	qmj_safety	Quality	Level	0.69***	0.48***	0.28
				(0.24)	(0.14)	(0.19)
Assets turnover	sale_bev	Quality	Level	0.37*	0.28*	0.35**
				(0.20)	(0.15)	(0.18)
Market correlation	corr_1260d	Seasonality	Level	−0.05	0.56***	0.09
				(0.21)	(0.15)	(0.18)
Coskewness	coskew_21d	Seasonality	Level	0.38*	0.27**	0.24
				(0.21)	(0.13)	(0.17)
Net debt issuance	dbnetis_at	Seasonality	Change	0.05	0.21*	0.22
				(0.20)	(0.12)	(0.15)
Kaplan-Zingales index	kz_index	Seasonality	Other	−0.14	−0.48***	−0.12
				(0.22)	(0.18)	(0.19)
Change in long-term investments	lti_gr1a	Seasonality	Change	−0.61**	–	−0.44
				(0.29)		(1.35)
Taxable income-to-book income	pi_nix	Seasonality	Level	0.27	0.03	−0.03
				(0.19)	(0.12)	(0.15)

Continued on next page

Table A.1 – *continued from previous page*

Anomaly	JKP Variable	JKP Theme	Type	Complexity	Size → Complexity	IVOL → Complexity
Years 2-5 lagged returns, annual	seas_2_5an	Seasonality	Change	0.63***	0.21	0.39**
				(0.23)	(0.14)	(0.17)
Years 6-10 lagged returns, annual	seas_6_10an	Seasonality	Change	0.35	0.02	−0.05
				(0.24)	(0.14)	(0.18)
Years 11-15 lagged returns, annual	seas_11_15an	Seasonality	Change	−0.15	−0.17	−0.02
				(0.23)	(0.14)	(0.19)
Years 11-15 lagged returns, nonannual	seas_11_15na	Seasonality	Change	0.46**	0.21	0.01
				(0.22)	(0.14)	(0.18)
Years 16-20 lagged returns, annual	seas_16_20an	Seasonality	Change	−0.32	−0.00	−0.33
				(0.28)	(0.16)	(0.21)
Change in short-term investments	sti_gr1a	Seasonality	Change	−0.19	−0.20	−0.52**
				(0.22)	(0.21)	(0.23)
Amihud measure	ami_126d	Size	Level	−0.43**	0.23**	−0.38***
				(0.18)	(0.11)	(0.14)
Dollar trading volume	dolvol_126d	Size	Level	−0.43***	0.12	−0.43***
				(0.16)	(0.12)	(0.14)
Market equity	market_equity	Size	Level	−0.85***	–	−0.71***
				(0.19)		(0.15)
Price per share	prc	Size	Level	−1.05***	−0.89***	−0.85***
				(0.27)	(0.16)	(0.18)
RD to market	rd_me	Size	Level	−0.60*	−0.19	−0.62**
				(0.34)	(0.18)	(0.26)
Idiosyncratic skewness from the CAPM	iskew_capm_21d	Short-Term Reversal	Level	−0.13	−0.13	0.08
				(0.16)	(0.10)	(0.14)
Idiosyncratic skewness from the Fama French 3 factor model	iskew_ff3_21d	Short-Term Reversal	Level	−0.21	−0.16	−0.02
				(0.15)	(0.10)	(0.15)
Idiosyncratic skewness from the q-factor model	iskew_hxz4_21d	Short-Term Reversal	Level	−0.20	−0.14	−0.02
				(0.17)	(0.09)	(0.15)
Short term reversal	ret_1_0	Short-Term Reversal	Change	−1.46***	−1.08***	−0.92***
				(0.27)	(0.15)	(0.19)
Highest 5 days of return scaled by volatility	rmax5_rvol_21d	Short-Term Reversal	Level	−0.34*	−0.31***	−0.15
				(0.20)	(0.11)	(0.17)

Continued on next page

Table A.1 – *continued from previous page*

Anomaly	JKP Variable	JKP Theme	Type	Complexity	Size → Complexity	IVOL → Complexity
Total skewness	rskew_21d	Short-Term Reversal	Level	−0.16	−0.03	0.07
				(0.17)	(0.10)	(0.15)
Assets-to-market	at_me	Value	Level	−0.56**	−0.45**	−0.63***
				(0.24)	(0.18)	(0.21)
Book to market equity	be_me	Value	Level	−0.82***	−0.44**	−0.75***
				(0.28)	(0.18)	(0.23)
Book to market enterprise value	bev_mev	Value	Level	−0.66**	−0.57***	−0.74***
				(0.28)	(0.19)	(0.23)
Net stock issues	chcsho_12m	Value	Change	0.25	−0.06	−0.09
				(0.21)	(0.14)	(0.16)
Debt to market	debt_me	Value	Level	−0.59**	−0.55***	−0.74***
				(0.27)	(0.17)	(0.21)
Dividend yield	div12m_me	Value	Level	−0.30	0.02	0.36
				(0.19)	(0.20)	(0.26)
Ebitda to market enterprise value	ebitda_mev	Value	Level	−0.20	−0.70***	−0.60***
				(0.24)	(0.16)	(0.18)
Equity duration	eq_dur	Value	Level	−0.63**	−0.73***	−0.87***
				(0.27)	(0.15)	(0.18)
Net equity issuance	eqnetis_at	Value	Change	0.48**	0.03	0.13
				(0.22)	(0.13)	(0.16)
Equity net payout	eqnpo_12m	Value	Level	0.09	−0.11	−0.19
				(0.19)	(0.14)	(0.17)
Net payout yield	eqnpo_me	Value	Level	0.45*	−0.10	−0.03
				(0.23)	(0.14)	(0.17)
Payout yield	eqpo_me	Value	Level	−0.14	−0.22	−0.43**
				(0.25)	(0.14)	(0.18)
Free cash flow to price	fcf_me	Value	Level	0.27	0.28**	0.17
				(0.24)	(0.14)	(0.22)
Intrinsic value to market	ival_me	Value	Level	−0.46*	−0.55***	−0.59***
				(0.24)	(0.14)	(0.20)
Net total issuance	netis_at	Value	Change	0.40*	0.23*	0.18
				(0.23)	(0.12)	(0.16)
Earnings to price	ni_me	Value	Level	−0.03	−0.69***	−0.49**
				(0.26)	(0.15)	(0.19)
Operating cash flow to market	ocf_me	Value	Level	−0.25	−0.40***	−0.42**
				(0.24)	(0.15)	(0.19)

Continued on next page

Table A.1 – *continued from previous page*

Anomaly	JKP Variable	JKP Theme	Type	Complexity	Size → Complexity	IVOL → Complexity
Sales to market	sale_me	Value	Level	−0.70***	−0.61***	−0.60***
				(0.26)	(0.17)	(0.19)

Notes: Entries are CAPM alphas of the high-complexity anomaly return minus the low-complexity anomaly return, in percent per month. Factor classifications are based on the Change/Level/Other classification described in Section 3.3. The size-controlled specification excludes market equity, and the idiosyncratic-volatility-controlled specification excludes the 21-day Fama–French three-factor idiosyncratic volatility characteristic. Newey–West standard errors are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.