**WORKING PAPER** · NO. 2023-76

# Profitability Context and the Cross-Section of Stock Returns

*Alex G. Kim and Valeri V. Nikolaev*

JUNE 2023

Becker Friedman Institute
FOR ECONOMICS AT UCHICAGO

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

# Profitability Context and the Cross-Section of Stock Returns

Alex G. Kim* Valeri V. Nikolaev†

May 25, 2023

**Abstract**

Asset pricing models implicitly assume that firm characteristics are context-free. At the same time, companies provide a substantial narrative context that helps investors to put numeric information in perspective. Management may discuss non-quantitative factors that influence performance, such as changes in competitive strategies, future plans, etc. We study the importance of contextual information for asset pricing by focusing on the narrative context surrounding profitability numbers. We use machine learning to incorporate contextual information into the measurement of profitability. Context-adjusted profitability has a superior ability to explain expected returns, both statistically and economically, compared to conventional operating profitability. Further, the context-adjusted profitability factor performs better in portfolio tests and helps to resolve the biggest challenge facing the five-factor model (Fama and French [2015]). Overall, we find that accounting for context adds significant value for investors and can improve the asset pricing models.

**Keywords:** Contextual information, context-based profitability, asset pricing, machine learning, natural language processing, operating profitability, factor models

**JEL Codes:** C13, C45, C55, C58, G11, G12, M41

---

*The University of Chicago, Booth School of Business, alex.kim@chicagobooth.edu

†The University of Chicago, Booth School of Business, valeri.nikolaev@chicagobooth.edu

We appreciate insightful comments from Jongsub Lee, Stefan Nagel and seminar participants at the University of Chicago Workshop. Both authors gratefully acknowledge financial support from the University of Chicago Research Support Center.

## 1 Introduction

Characteristics in empirical asset pricing models are taken out of context. When two companies report identical earnings figures, the interpretation of these figures varies depending on multiple qualitative factors not captured by numeric information. To address this issue, the Security and Exchange Commission (SEC) requires that companies provide narrative context to the reported numbers. The narrative disclosures cover significant economic events such as changes in strategic goals, changes in competition, delays in production, disruptions in the supply chains, trends in demand for products, future plans, and many more. In this paper, we examine the importance of context for asset pricing models. We show that incorporating context enhances the value of numeric characteristics. Notably, accounting for profitability context goes a long way in addressing the five-factor model's biggest challenge – the difficulty in pricing small stocks with extreme growth (Fama and French [2015]).

The goal of the SEC's reporting requirements (provided in MD&A section of the annual and quarterly reports) is to help investors form more accurate expectations of future earnings and cash flows. Characteristics in asset pricing models commonly proxy for these future quantities. More specifically, in Fama and French [2015] five-factor model, profitability and investment characteristics aim to capture expected future profitability and future investment. As Fama and French [2015] note, measuring these expectations poses an important challenge (see also Fama and French [2006]). To address it, the recent literature makes adjustments to reported net income that tend to remove transitory components. In particular, Novy-Marx [2013] advocates the use of gross profit as a proxy for future profitability. Ball et al. [2015], among others, who advocate for the use of operating profitability. More recently, Rouen et al. [2021] use proprietary data to construct "core earnings."

While making quantitative adjustments to accounting earnings based on disclosed line items is a logical step, it does not incorporate the unstructured narrative context within which the reported numbers should be interpreted. For example, Li et al. [2013] show that companies' return on assets is more likely to mean revert when 10-Ks feature extensive discussion of the competition. More recently, Kim and Nikolaev [2023] shows taking contextual information into account helps financial statements' users to make inferences about a firm's future. At the same time, asset pricing models generally disregard this type of information when constructing and interpreting characteristics.

We aim to improve our understanding of the importance of context related to firm characteristics in the cross-section of stock returns by addressing the following questions.

Does contextual information present opportunities to improve the asset pricing models? How important are the interactions between numeric characteristics and their context? Third, can researchers construct meaningful context-adjusted characteristics and context-adjusted factors?

To address these questions, we focus on the context of operating profitability captured by the discussion of operating performance in the MD&A section of annual reports. We study operating profitability (not net income) because it already largely excludes the transitory components and can be viewed as a proxy for core earnings.

Incorporating context into a measure of operating profitability (OP) is a challenge that we address in two steps. First, we quantify the inherently unstructured and highly multi-dimensional nature of narrative context related to operating performance using BERT. BERT is a language model that encodes textual information in the form of dense vectors and is designed to capture the context-based meaning of words and sentences (Devlin et al. [2018]).[1]

Second, we construct a context-based proxy for future operating profitability, i.e., a proxy that incorporates the context surrounding reported profits. To do so, we use the methodology developed in Farrell et al. [2021b] that models regression parameter heterogeneity based on neural networks. Each year, we run a linear cross-sectional regression that uses the operating profit to predict its own one-step ahead value. However, the regression's parameters as modeled as flexible functions of textual input. In particular, the intercept and slope parameters are deep neural nets that use textual vectors supplied by BERT.

The appeal of this approach compared to standard (densely connected) neural network models is that it "forces" the model to focus on learning the *interactions* between narrative information and numeric input, which are the primary focus of our study. Accordingly, the model learns to make an adjustment for the irrelevant (transitory) portion of operating profits based on the narrative information.

In addition to a context-based proxy for profitability, which is designed to exploit the interactions between numeric and textual information, we also construct a "text-only proxy" for future profitability. It follows the same regression-based approach but models the intercept only (as a function of text) while dropping *OP* and its interactions with contextual information. The intercept captures any direct information revealed by the textual information itself and can be interpreted as profitability sentiment.

[1]BERT is pre-trained on a large corpus of Wikipedia texts. Importantly, corporate communications or business press is not a part of BERT's training set, which ensures that the model did not see our textual data during training.

We train the neural network models annually based on a five-year moving window and using the root mean squared error criterion. We then use the parameters from each training and the (textual and numeric) information up to time $t$ to make out-of-sample forecasts $t+1$ operating profitability. These forecasts rely on current profits as the primary input and effectively "correct" it for the context within which current profits are reported.

As a starting point, we verify the value of contextual information by horseracing the three proxies for future profitability (scaled by total assets): (1) a conventional (unadjusted) operating profitability ($OP$), which uses its current value to proxy for future values; (2) a "text-only" proxy ($OP^C$), which only relies on (con)textual information only; and (3) the context-adjusted proxy ($OP^{CN}$), which combines $OP$ with its context in a way that incorporates non-linear interactions between the two. We find that $OP^{CN}$ comes out as a clear winner. Not only does it significantly outperforms both measures, but it also subsumes both of them when the three variables are used to predict a step ahead profits out-of-sample. This is an indication that contextual information is likely valuable from an asset pricing perspective not because it captures the sentiment in the narrative but because it interacts with and thus changes the interpretation of reported profitability.

Equipped with a context-based measure of profitability, we perform a number of asset pricing tests. Our main analysis begins with Fama and MacBeth [1973] regressions that evaluate our profitability proxies in the cross-section of stock returns. When the proxies are used individually, context-adjusted profitability outperforms both the conventional $OP$ and the text-only proxy in predicting stock returns. More importantly, context-adjusted profitability also subsumes the other two proxies when the three variables are used simultaneously. This result is noteworthy because we train the model to predict profitability, not returns, and yet it captures all return-relevant information. We show that these results hold in the entire sample and after excluding Microcaps.

To judge the economic significance of these results, we follow Fama and French [2006] and use predicted values from the Fama-MacBeth regressions to allocate stocks into above- (high) and below-median (low) portfolios. The spread in average (predicted and actual) high-minus-low portfolio returns increases significantly when we use context-based profitability instead of unadjusted OP. In fact, the increment in the spread due to the incorporation of context exceeds the increment from adding OP into the model (that includes other characteristics).

Because Fama and MacBeth [1973] regressions impose the linearity assumption and are equally weighted, we also analyze the performance of value-weighted portfolios. We sort the stocks annually at the end of June into ten portfolios based on (1) context-adjusted profitability, and (2) unadjusted $OP$. When using CAPM, three-factor, four-factor, and five-factor models, we find that alphas for $OP^{CN}$-based portfolios consistently dominate the corresponding alphas for $OP$-based portfolios. We also perform a double sort into quintiles based on context-adjusted profitability and, independently, on operating profitability, to form 5×5 portfolios. We estimate the three-factor model for each portfolio and find that, for each quintile of $OP$, high-minus-low alpha for context-adjusted profitability-based strategy is significantly positive. At the same time, within each $OP^{C}N$ quintile, the $OP$-based high-minus-low strategy generates insignificant alphas. Overall, the results are in line with the regression-based evidence and indicate that context-adjusted profitability carries substantial information beyond operating profitability but not vice versa in the cross-section of stock returns.

The next logical question is whether context-adjusted profitability enables one to improve the five-factor model. Fama and French [2015] state: "The five-factor model's main problem is its failure to capture the low average returns on small stocks whose returns behave like those of firms that invest a lot despite low profitability." As we discuss further, it is plausible that taking the context associated with low profitability when constructing factor portfolios can help explain these anomalous findings. To this end, we construct a factor $RMW^{CN}$ that is based on context-adjusted profitability and compare its performance with the standard profitability-based factor ($RMW$).[2] We start by running the spanning regressions and show that $RMW$ is subsumed by the context-adjusted factor and not the other way around. This confirms that $RMW^{CN}$ contains at least as much information in line with our prior findings.

We then turn to the portfolio tests to evaluate the model. We first construct 5×5 portfolios by sorting stocks into $OP$ and $OP^{CN}$ quintiles and estimate their five-factors model alphas. We show that the model's ability to price these 25 portfolios, as indicated by the GRS statistic and the average absolute value of alpha, is substantially higher when we use the factor based on context-adjusted profitability.

Finally, we revisit the most problematic portfolios in Fama and French [2015] by constructing $5 \times 5$ portfolios based on *Size* and $B/M$ characteristics. Closely in line with their findings, we find that the three-factor model has the highest alpha of -0.51 for small stocks that with the highest growth and that the inclusion of profitability and growth factors reduces this alpha to -0.29. More importantly, when we use the context-adjusted factor, the alpha goes down to -0.13 and loses its statistical significance.

[2] We follow the methodology in Fama and French [2015] and sort stocks into 2×3 portfolios. We sort stocks into two groups based on size and independently sort them into three groups on context-adjusted profitability. The factor returns are the difference in returns between the two high portfolios and two low $OP^{CN}$-based portfolios.

In sum, the context-based profitability factor improves the ability of the five-factor model to price securities and also helps to resolve a long-standing puzzle in empirical asset pricing.

This paper contributes to the literature in three ways. First, we establish an interplay between numeric information and its corresponding context. The asset pricing literature implicitly assumes that firm characteristics are context-free, i.e., have a similar interpretation across firms. We show that the interpretation of numeric characteristics varies with the non-quantitative context within which the characteristics are reported. Taking the context into account adds considerable value and improves models' performance. Second, we propose and validate a way that researchers can use to measure characteristics in a way that takes their context into account. The proposed measure of context-based operating profitability consistently outperforms the unadjusted profitability measure in explaining the cross-section of stock returns. It also subsumes both the profitability and the textual sentiment contained in management's discussions of operations. Third, our study expands the rapidly growing asset pricing literature on machine learning (e.g., Ke et al. [2019], Gu et al. [2020, 2021]). Gu, Kelly, and Xiu [2020] show that machine learning models outperform the traditional regression-based models by allowing for non-linear interactions among the variables. While a number of prior studies used textual information to measure sentiment and predict stock price movements (see Gentzkow et al. [2019] for review), our study highlights the importance of interactions between textual and numeric characteristics and the value of recent machine learning techniques in detecting these interactions (Farrell et al. [2021a]).

## 2 Profitability context and expected stock returns

Fama and French [2006] propose a unifying framework that helps to interpret the link between expected returns and firm characteristics. It relies on the following valuation equation:

$$M_t = \sum_{\tau=1}^{\infty} \frac{\mathbb{E}_t(Y_{t+\tau} - \Delta B_{t+\tau})}{(1+r)^{\tau}} \tag{1}$$

where $M_t$ is the market value of equity at the end of year $t$, $Y_t$ is the year $t$ earnings, $\Delta B_t$ is the change in book value of equity over the year $t$, which measures reinvestment of earnings.

The equation predicts that expected stock returns $r$ will vary predictably with the $B_t/M_t$ ratio, future expected profits, and future expected investments. In particular, controlling for $B_t/M_t$ and expected investment, higher expected profitability implies higher expected returns. This framework also implies that variables that are helpful in predicting expected profitability or expected investment should also be useful in predicting stock returns.

While Fama and French [2006] find support for these predictions in cross-sectional regressions, the economic magnitudes of the combined effect of profitability and investment is relatively small economically. Furthermore, they find that *simple proxies*, such as the current level of profitability, dominate more complex proxies, which rely on predicted future profitability values based on a broader set of firm characteristics in the first stage regressions.[3]

Novy-Marx [2013] made an important step forward by noting that the above expression is stated in terms of firms' "true profitability," whereas the accounting earnings (net income) reported at the bottom of the income statement is often a misleading proxy for true earnings. Among other things, accounting earnings contain transitory components uninformative about future profits; earnings also treat R&D as an expense, whereas economically, it is an investment that increases future profits.

To address this issue, he proposes to focus on the top profitability number in the income statement and shows that Gross Profit is a much more powerful predictor of returns, as compared to net income. A number of studies have shown that profitability proxy can be further improved by making more refined adjustments (e.g., Ball et al. [2015, 2016], Rouen et al. [2021]).

In contrast to these studies, the focus on adjustments for 'softer' (non-quantitative) factors that nevertheless alter the interpretation of quantitative characteristics, e.g., profitability, has been missing from the literature. At the same time, a narrative discussion that allows investors to evaluate how rapidly the profits are likely to mean revert should be helpful in assessing the implications of current profits for future profits. In particular, a drop in current profits due to a temporary demand shock should be interpreted very differently from a similar drop accompanied by significant changes to the competitive landscape. Such information is inherently hard to communicate quantitatively due to its non-systematic and unstructured nature. At the same time, it provides the context that affects the interpretation of numeric profitability values and what they imply for future stock returns.

[3]Explanations for this finding boil down to considerable noise in the fitted values relative to true expectations. At the same time, current earnings is the most powerful predictor of future earnings and hence likely the most useful proxy.

### 2.1 Constructing a context-based measure of profitability

The challenge we are facing is to encode and compress highly multidimensional narrative information in a useful way, one that allows us to refine the interpretation of numeric data. We turn to machine learning and natural language processing to tackle this challenge.

We break the problem into two steps. Step one employs a pre-trained BERT (or Bidirectional Encoder Representations from Transformers) language model to encode the portion of management's discussion related to operating performance (located in the MD&A part of 10-K) in the form of 768-dimensional vectors. Unlike uni-dimensional proxies such as tone or sentiment, text vectors encode the meaning of the text.

One way to proceed from here could be to feed the profitability numbers and textual embeddings as inputs into a standard densely connected feed-forward neural network and train it to predict future profitability while allowing for rich interactions between textual and numeric data. This approach, however, is likely to face noise-related issues akin to those associated with more complex proxies in Fama and French [2006].[4] Intuitively, the network would have to be very deep and would require vast amounts of data to train to discover the interactions between textual and numeric inputs. We thus follow an approach that imposes further parametric restrictions, as we discuss next.

Specifically, in step two, we combine the simplicity of a univariate regression model with the power of a neural network by following the approach in Farrell et al. [2021a,b]. Their methodology treats parameters in a regression model as deep neural nets allowing them to be non-parametric functions of multidimensional input. Accordingly, we take the simplest possible predictive regression that uses $OP$ to predict itself and hence measure expected profits:

$$OP_{it+1} = \theta_0 + \theta_1 OP_{it} + \epsilon_{it}. \tag{2}$$

We use operating profitability because it already incorporates key quantitative adjustments to net income that aim to improve its accuracy as a proxy for future profits.

However, we further allow the model to "interpret" the reported $OP_t$ values differently, depending on their narrative context $c_{it}$ (text vector):

$$OP_{it+1} = \theta_0(c_{it}) + \theta_1(c_{it}) OP_{it} + \epsilon_{it} \tag{3}$$

[4] Recall that proxies based on multiple predictors to construct profitability expectations are dominated by the levels of (past) profitability.

where $\boldsymbol{\theta}(\boldsymbol{c_{it}})$ is a flexible nonparametric function that represents a deep ANN.

The primary advantage of this structure is that, during the training, the model is effectively "forced" to learn the interactions between text and numbers, i.e., whether and to what extent the level of $OP_{it}$ needs to be qualified (relative to its average value) when it comes to predicting $OP_{it+1}$. For example, when the context suggests the transitory nature of reported numbers, the model will learn to place less weight on the current level of $OP$ and vice versa. Additionally, the model also adjusts the intercept, i.e., the level we expect unconditionally.[5]

To estimate $\boldsymbol{\theta} = \{\theta_0, \theta_1\}$, we minimize the the solve following and obtain $\widehat{\boldsymbol{\theta}}$:

$$\widehat{\boldsymbol{\theta}} = \underset{\tilde{\theta} \in \mathcal{F}_{DNN}}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \left(OP_{it+1} - \theta_0(c_{it}) - \theta_1(c_{it})OP_{it}\right)^2 \tag{4}$$

where $\mathcal{F}_{DNN}$ is the parameter space of $\theta$ generated by a deep neural network. Note that the solution to this problem enables us to estimate $\mathbb{E}_t[OP_{it+1}|\Omega_t]$ where $\Omega_t$ is the set of information available to the public at time $t$.

Based on these steps, we generate *context-based operating profitability*, $OP^{CN}$, where the superscript indicates that it combines contextual and numeric information:

$$OP_{it}^{CN} \equiv \mathbf{E}(OP_{it+1}|OP_{it}, c_{it}; \widehat{\theta}(.)) = \widehat{\theta}_0(c_{it}) + \widehat{\theta}_1(c_{it})OP_{it} \tag{5}$$

Note that $OP_{it}^{CN}$ can be decomposed into two parts, both of which depend on $c_{it}$: one that varies with $OP_{it}$ and one that does not. The invariant part can be interpreted as the sentiment that is extracted by the model from the $c_t$ about future profits, whereas the variable part captures the *interactions* between the $OP$ characteristic and its context. Because the latter is of ultimate interest for this study, we also need to control for sentiment (in some specifications). To do so, we directly construct a sentiment proxy by dropping the $OP_t$ term and re-training the model with the intercept only, such that:

$$OP_{it}^{C} \equiv \mathbf{E}(OP_{it+1}|c_{it}; \widehat{\theta}'(.)) = \widehat{\theta}_0'(c_{it}) \tag{6}$$

where the superscript "C" highlights that the expectation is formed based on contextual information only.

[5] Alternatively, we can interpret intercepts as adjusting the level of $OP$ and slopes as adjusting the sensitivity of $OP$ to the deviations from the intercept.

# 3 Implementation and training

## 3.1 Extracting Context Information with BERT

We extract narrative information about profitability from MD&A sections of all available annual reports. The objective of MD&A is to provide the relevant context that helps investors to interpret operating performance. We use a keyword list by Kothari et al. [2009] and identify sentences that explicitly discuss profitability and retrieve these sentences. We also retrieve one sentence that precedes and one that follows the sentences explicitly mentioning profitability to expand the context.

For each firm-year, we pool and process the identified sentences with BERT – a large language model developed by Google (Devlin et al. [2018]). The model is particularly well-suited here as, in addition to encoding the meaning of words themselves, it is also designed to capture the context within which words are used.

We use the BERT-base uncased model pre-trained on a large corpus of Wikipedia text.[6] Although BERT's pre-trained parameters can be fine-tuned for a variety of applications, doing so is unnecessary in our case (and is effectively achieved) due to several new layers of parameters added on top when we train the model. The last stage hidden state processed by BERT contains a 768-dimensional numeric vector that encodes textual information. This vector is associated with a special classification (CLS) token, which is always placed at the beginning of the input text; this vector summarizes the information in the given input texts.

We drop special characters, tabular information, and graphics. Whenever a sentence contains numbers, we replace it with placeholder tokens (instead of dropping) to enable the model to learn about the presence of numeric information and its position. BERT can process 512 tokens ("words") simultaneously. The profitability-related sentences, combined, are often longer than 512 tokens. We thus divide each document into multiple 512-token passages and run them through BERT independently. Following conventional practice, we then take the average of the textual embeddings associated with the CLS tokens to aggregate information. We refer to these embeddings vectors as $\boldsymbol{c_{it}}$.

## 3.2 Design of Feedforward Deep Neural Network

Textual vectors $\boldsymbol{c_{it}}$ serve as input into a deep neural net that we train to uncover the adjustments for profitability context. The model's design is illustrated in Figure 1. Following the input layer with 768 neurons, the model includes three hidden layers

[6] While there are domain-specific BERT models such as FinBERT (Huang et al. [2022]), we refrain from their use because their training set may overlap with corporate filings.

with 256, 64, and 16 neurons, respectively, and one output layer with two neurons.[7] The output layer produces the two structural parameters in the predictive regression (4). The model is trained based on root mean squared error (RMSE) given by the difference between $\widehat{\theta}_{0it}(c_{it}) + \widehat{\theta}_{1it}(c_{it})OP_{it})$ and $OP_{it+1}$. All neurons except those in the output layer are activated with a ReLU activation function.

We train the model to predict one year ahead profitability based on a rolling window spanning four years ($t-4, t-3, t-2$, and $t-1$). We next use time $t$ data to construct an out-of-sample forecast of $OP_{i,t+1}$. This ensures that our model does not see any test data during the training. Specifically, as our sample period is 1995-2020, we first train the model based on 1995-1998 data; subsequently, we use the profitability and its context information from 1999 to predict $OP$ in 2000. [8] We randomly assign 30% of each training sample as a validation set. To reduce overfitting, we use a 20% dropout rate. We stop the training when the loss function tested on the validation set does not improve by at least 0.001 for ten consecutive training epochs.[9] We set the learning rate to 0.001 and use Adam optimizer. [10]

This procedure generates the estimates $\theta_{0,it}$, $\theta_{1,it}$, and $\theta'_{0,it}$ used to construct time $t$ forecasts, $OP^{CN}_{it}$ and $OP^{C}_{it}$, of $E(OP_{it+1})$ as discussed previously.

## 4 Data

Our sample consists of US-listed securities for the period 1995-2020; the SEC mandated the electronic submission of corporate filings in 1994. We include all firms with a valid MD&A extracted from the 10-K filings. Firm characteristics are taken from Compustat, and stock returns are from CRSP. We require that the stocks are traded in NYSE, Amex, and Nasdaq, and only include ordinary common shares. We exclude financial sector firms (firms with one-digit Standard Industry Classification code six). When delisting returns are missing on CRSP, we set them to -30% for NYSE and Amex stocks and -55% for Nasdaq stocks (Shumway [1997], Shumway and Warther [1999]). The sample consists of firms with non-missing book value of equity, market value of equity, previous one-month return, previous one-year buy-and-hold return, total assets growth, and op-

[7]The number of hidden layers and the number of neurons in each layer can be changed. Our results are qualitatively the same with several other specifications.

[8]Similarly, our last training session is based on the 2015-2018 period. We then take the data from 2019 to predict the profitability in 2020.

[9]We set the maximum training to 250 epochs. However, our model does not reach the maximum epoch allowed for any of the training samples.

[10]These parameters can also be changed at the discretion of researchers. Slightly changing these parameters yield similar results.

erating profit. Lastly, we require that each MD&A contains at least one sentence that discusses the results of operations, i.e., how profitable the company is[11]. We use firm characteristics from Compustat with a six-month lag relative to stock returns to ensure that accounting information is public and is incorporated into stock prices.

We use the conventional definition of operating profitability ($OP$): revenues less cost of goods sold, selling, general, and administrative expenses, and, following Fama and French [2015], also subtract interest expense. [12] The book value of equity is the total shareholders' equity, plus deferred taxes, investment tax credits, post-retirement benefit liabilities, and less preferred stock. When shareholders' equity is missing, we impute common share value or total assets minus total liabilities. When other balance sheet items (deferred taxes and investment credits) are missing on Compustat, we set their values to zero. Preferred stock is measured, depending on data availability, based on its redemption value, liquidation value, or carrying value, using this order.

For Fama and MacBeth [1973] regressions, we update momentum ($r_{0,1}$ and $r_{2,12}$) and market capitalization ($\log(ME)$) each month, and book value of equity ($\log(BE/ME)$) and investment ($Investment$) each quarter. Profitability measures are updated annually because the corresponding MD&A texts are published annually. We perform the regression analysis for the samples of all and all-but-microcap stocks. Microcap stocks refer to stocks that are below the 20th percentile of the NYSE market value, obtained from Ken French's data repository. To restrict the influence of outliers, we truncate one percent of independent variables in the analyses (Ball et al. [2015, 2016]). In portfolio analysis, we rebalance portfolios on June 30th of each year.

Table 1 presents descriptive statistics for the variables used in the analyses. Operating profitability ($OP$) has a mean of 0.15 and a standard deviation of 0.62. On the other hand, $OP^{CN}$, which is our context-adjusted profitability measure, has a mean of 0.12 and a considerably lower standard deviation of 0.29. This fact is consistent with context-based adjustments removing transitory variation from the reported profits. The three profitability measures are positively correlated with each other (untabulated). $OP$ and $OP^{CN}$ has a Pearson correlation of 0.63 (significant under 1% significance level). $OP$ and $OP^{C}$ has a correlation of 0.03 (significant under 1% significance level).

[11]This restriction aims to ensure the validity of the extracted information. It is highly unlikely to be binding for valid observations since the objective of MD&A is to discuss the results of operations

[12]We also employ the definition of Ball et al. [2015], which adds back R&D expense and does not adjust for interest and obtain quantitatively similar results.

## 5 Validating Context-Adjusted Profitability

In this section, we validate the predictive ability of the three profitability measures. To do so, we run simple ordinary least squares regressions of future operating profitability on our proxies and a constant term. Note that the analysis is out of sample as the model trained without seeing the predicted values. We cluster standard errors at the firm and year levels and report *t*-statistics.

We report the results in Table 2. The first column shows that the standard (unadjusted) $OP$ exhibits a coefficient of 0.41. Unsurprisingly, it is highly significant but is relatively far from unity, which is in line with the presence of significant transitory components in the unadjusted measure. The adjusted $R^2$ for this model is 22.49%. The second column features the text-only proxy of profitability $OP^C$ which captures the sentiment in contextual information. The slope coefficient in this case is somewhat lower, 0.32, and it has an order of magnitude lower t-value.[13] Accordingly,the adjusted $R^2$ in the second column drops to 3.03% only. The third column tests our context-adjusted profitability measure, which delivers the strongest predictive power. While the forecasts are performed out of sample, the coefficient on $OP^{CN}$ is close to unity and the adjusted $R^2$ is with 28.33%.

Furthermore, we include all three profitability proxies in one linear regression, the context-based proxy, $OP^{CN}$, subsumes the two other measures and remains highly significant. At the same time, $OP$ and $OP^C$ exhibit a large drop in their slope coefficients, which become statistically insignificant. The adjusted $R^2$ improves only slightly compared to column (3) (from 28.33% to 29.66%) in line with $OP^{CN}$ capturing almost all the relevant information about future profits contained in the three proxies.[14]

In sum, we show that the context-adjusted profitability, $OP^{CN}$, is a considerably more powerful predictor of future profitability, subsuming the unadjusted profitability measure in the prediction task. This result in itself is not very surprising as the model was trained to excel in this task. Still, it validates the value of contextual information when interacting with the numeric data and also sets the stage for subsequent asset pricing tests.

[13]Note that $OP^C$ does not consider any numeric information in its construction and that $OP^C$ and $OP$ are not mechanically correlated. Even so, we observe a significant positive association between current $OP^C$ and future $OP$, which implies that text information only can reasonably predict future profitability as well.

[14]In untabulated tests we use standardized variables and obtain quantitatively identical results.

# 6 The cross-section of stock returns

## 6.1 Fama-MacBeth regressions

Our main analysis starts by examining the three measures of profitability by running Fama and MacBeth [1973] regressions. Following prior studies (e.g., Novy-Marx [2013], Ball et al. [2015]), we regress monthly stock returns on our proxies for profitability and the following control variables: stock return for the prior month, stock return for the prior year (skipping the most recent month), log of the market value of equity, log of book-to-market ratio, and investment. We multiply the regression coefficients by 100 to ease interpretation.

Table 3, Panel A, presents the results for the entire sample. In line with prior literature, Column 1 shows that operating profitability ($OP$) is a statistically significant predictor of stock returns with a slope of 0.91 ($t$=3.01). We note that the t-values throughout the table are somewhat lower compared to prior studies because our sample starts in 2000 (electronic filings are available since 1995, and we require five years of data to pre-train the first model), whereas many studies use data starting in 1963.

Column 2 replaces $OP$ with text-based proxy $OP^C$. Here, we find a statistically insignificant coefficient on $OP^C$ and a noticeable drop in the $R^2$ from 3.94 % to to 3.28%. This implies that text information surrounding the disclosed earnings alone does not have predictive power with respect to stock returns. In contrast, column 3 is based on context-adjusted profitability, $OP^{CN}$, which has a coefficient of 1.95 with $t$-value=4.23. Consistent with the notion that $OP^{CN}$ is a less noisy measure of expected profitability, its coefficient is considerably higher than in column 1, and so is the t-value. R-squared also exhibits a noticeable increase (from 3.94% to 4.45%). The signs and significance of the coefficients on other characteristics remain quantitatively similar and line up with those in prior studies.

Column 4 performs a horserace between $OP$ and $OP^{CN}$ by concluding them simultaneously. The coefficient on $OP$ drops sharply in absolute value and loses its statistical significance ($t$=-0.67), whereas the coefficient on $OP^{CN}$ increases in value and remains statistically significant ($t$=2.29). Furthermore, when we include all three proxies of profitability in the fifth column, only the coefficient on $OP^{CN}$ remains statistically significant ($t$=2.34). In other words, we find that context-based profitability subsumes information in the other two measures.

Table 3, Panel B, repeats the same analysis after we exclude microcaps from our sample. A stock is considered a microcap if its market value is below the 20th percentile of the NYSE market value distribution. We find a set of very similar results. As in Panel

A, $OP^{CN}$ dominates the other profitability measures based on statistical significance and also the explanatory power. It also subsumes information in the other two measures.

Overall, our findings imply that taking the context of profitability into account yields a superior proxy for expected profitability. The result that the context-adjusted proxy subsumes information in the other proxies is noteworthy because the model was not trained to predict returns. To the extent, each predictor is subject to its own (independent) measurement error, one would expect to see some incremental value added by each of the profitability proxies. In contrast, we observe that $OP^{CN}$ is a dominant proxy.

*Sub-periods* Figure 2 plots $t$-values associated with operating profitability and context-adjusted profitability based on a 10-year rolling window averages of Fama and MacBeth [1973] regressions. The horizontal axis shows the end of this window. Since our context-adjusted profitability is available starting from 2000, our first estimation period ends in January 2010. As of that time, both operating profitability and context-adjusted profitability show relatively high levels of statistical significance with $t$-values around 4. However, as time goes by and more towards 2020, the $t$-values associated with operating profitability decline in their magnitude, reaching 1.30 in 2020. At the same time, the $t$-values associated with context-adjusted profitability persist at the same level throughout the period.

Overall, our evidence suggests that the results documented in Table 3 hold for different subperiods in our sample. Furthermore, the evidence suggests that while operating profitability is losing its predictive power over time, this is not the case for the proxy of profitability that takes context information into account.

*Economic magnitudes*. To place the results in perspective, it is useful to assess the economic significance of the superior information content of $OP^{CN}$. To do so, we follow the approach in Fama and French [2006]. We first partition our samples based on the fitted values from the Fama-MacBeth regressions. Specifically, we use the full sample average slopes reported in columns 1-3 of Table 3, Panel A, to obtain the predicted monthly returns for the next year across three models featuring different profitability proxies. The predicted values of monthly returns are then used to assign stocks into high (above-median) vs. low (below-median) expected return portfolios. We then calculate the difference between the average predicted returns on high vs. low portfolios. We calculate the same difference (spread) for the average actual returns.

Table 4 presents the high minus low spreads on equally-weighted (EW) and value-weighted (VW) portfolio returns. The first two columns are based on the spreads in predicted returns, followed by two columns of spreads in the actual returns. The last two columns report t-values associated with the actual spreads. This discussion will focus on value-weighted portfolios, although the results are consistent with equally based portfolios. Raw "0" in the table reports a benchmark model that is based on Fama-MacBeth regression (not tabulated) without profitability but with all other controls. For this model, the value-weighted predicted spread between high and low portfolios is 0.40, whereas the actual value-weighted spread is 0.35. Row 1 adds operating profitability ($OP$) into the model and re-computes the value-weighted spreads. The spread on high minus low predicted returns increases from 0.40 to 0.45, whereas the spread in actual returns increases from 0.35 to 0.42. The increase in spreads is economically significant and is in line with $OP$ being an important determinant of expected returns.

Raw 2 replaces $OP$ with the text-only profitability proxy, $OP^C$. Not surprisingly, we observe a decline in spread across both predicted and actual portfolios. This happens because the text-only measure is a weak proxy for expected profitability. However, the results differ starkly in row 3, where we use $OP^{CN}$ as a proxy for profitability. We observe that the value-weighted spread in high minus low portfolio returns predicted by the cross-sectional model increases from 0.45 (the case of $OP$) to 0.54 (recall that benchmark is 0.40). In the case of actual returns, the spread in the high minus low portfolio increases from 0.42 (the case of $OP$) to 0.51. In both cases, the spread increments realized due to the context-based adjustment (i.e., moving from $OP$ to $OP^{CN}$) exceed the corresponding increments from adding $OP$ into the model. Overall, the evidence indicates that the value of the adjustments to operating profitability to incorporate its narrative context is not only a statistical gain but also a non-trivial economic gain that can be seen in large and well-diversified portfolios.

## 6.2 Portfolio Sorts

Portfolio tests provide an alternative and potentially more robust method to evaluate the predictive ability of profitability. Specifically, portfolio sorts do not impose linearity nor make parametric assumptions as compared to Fama-MacBeth regressions. Additionally, linear regressions over-emphasize economically small stocks due to equal weighting.

*Single sorts.* Table 5 reports the analysis based on single sorts and compares context-adjusted profitability to operating profitability. Note that Equation (1) does not imply that profitability necessarily predicts returns without controlling for other factors such as book-to-market. Therefore, a sort based on only context-adjusted profitability should not necessarily produce a spread. We tabulate excess returns over the risk-free rate, CAPM alphas, three-factor model alphas (FF3), four-factor model alphas (FF4), and five-factor alphas (FF5), based on decile portfolios sorted on operating profitability. Portfolios are value-weighted to mitigate the undue influence from small stocks.[15] We also present long-short portfolio returns and their corresponding *t*-values.

Panel A shows portfolios based on context-adjusted profitability. The monthly excess return on high-minus-low portfolio is positive yet statistically insignificant (22 basis points with a *t*-value of 0.57), however, the avrage excess returns exhibit non-monotonic pattern[16] In contrast, CAPM alphas, three-factor alphas, four-factor alphas, and five-factor alphas all display monotonically increasing patterns. High-minus-low portfolio monthly alphas range from 48 basis points (five-factor model) to 62 basis points (three-factor model) and they are all statistically significant (*t*-values ranging from 2.65 to 3.03). It is noteworthy that we observe a positive and economically significant alpha even after controlling for the profitability factor, which is based on operating profitability (Fama-French five factors model). This result indicates that context-adjusted profitability contains information not reflected by the unadjusted operating profitability.

Panel B tabulates the same analysis except that portfolios are formed based on the conventional *OP* proxy. In this panel, excess returns increase in magnitude (0.63) and reach a level of statistical significance, however, there is a large wedge between the 10th and 9th portfolios. We also observe statistically and economically significant alphas in the case of CAPM, three-factor, and four-factor models. However, in line with our prior findings, the alphas across these models are lower as compared to Panel A. In particular, CAPM alpha is 46 basis points (versus 55 basis points in Panel A), the three-factor alpha is 47 basis points (versus 62 basis points in Panel A), and the four-factor alpha is 41 basis points (versus 58 basis points in Panel A). Finally, alpha in the five-factor model is statistically insignificant, which is not surprising since the profitability factor is now included in the model.

*Double sorts*. To shed more light on the predictive ability of context-adjusted profitability, we also provide independent two-way sorts on context-adjusted profitability and operating profitability, tabulated in Table 6. We focus on three-factor model alphas for two-way sorted 5×5 portfolios. For all quintiles on operating profitability, we observe that the trading strategy based on context-adjusted profitability generates positive high-minus-low returns. These range from 60 basis points (*t*=3.00) for the highest operating profitability quintile to 71 basis points (*t*=3.54) for the lowest operating profitability quintile. Interestingly, we see a negative trend of high-minus-low returns as we move

[15]We find similar results using equal-weighted portfolios (untabulated)

[16]The lack of significant high-minus low spread in excess returns seems to be driven by the low portfolio. In particular, the second and decide exhibits exhibit considerably lower alphas as compared to the first decile.

from the lowest operating profitability quintile to the highest quintile.[17] This result suggests that context information plays a more important role when accounting profitability is low.

In contrast, for all quintiles of context-adjusted profitability, a trading strategy based on operating profitability does not generate significant high-minus-low alphas. Specifically, high-minus-low alphas range from 2 to 33 basis points and all of them are statistically insignificant ($t$-values ranging from 0.02 to 1.52).

In Appendix A, we also perform conditional sorts. In Table A1, Panel A, we first sort the portfolios on operating profitability into quintiles and, subsequently, within each profitability quintile, we sort based on context-adjusted profitability. The three-factor model alphas based on context-adjusted profitability high-minus-low portfolios are positive and statistically significant for each $OP$ quantile. In contrast, in Table A1, Panel B, we first sort portfolios on context-adjusted profitability and then, within each context-adjusted profitability quintile, we sort observations based on operating profitability. The three-factor model alphas for high-minus-low portfolios based on operating profitability are not significantly different from zero within any $OP^{CT}$ quintile. Thus, conditional sorts support our conclusions based on independent sorts.

Overall, portfolio analyses corroborate our Fama and MacBeth [1973] regression results. Context-adjusted profitability carries additional useful information about the cross-section of stock returns as compared to conventional operating profitability but not the other way around.

## 7 Context-based Profitability Factor

In this section, we construct a profitability factor that is based on context-adjusted profitability and explore its ability to price assets.

We follow the steps in Fama and French [2015] and perform $2 \times 3$ independent sorts each year. We sort stocks into two groups based on the median NYSE market capitalization breakpoints provided in Ken French's data library.[18] We independently sort stocks into three groups based on context-adjusted profitability. The high group corresponds to stocks with context-adjusted profitability higher than 70th percentile breakpoints and the low group corresponds to those with context-adjusted profitability lower than 30th percentile breakpoint. The factor return is the difference in the value-weighted average return for two high portfolios and that for two low portfolios.

[17] In fact, the difference of the high-minus-low returns between the highest quintile and the lowest quintile is 11 basis points ($t$=1.96).

[18] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Panel A of Table 7 presents the annualized returns, annualized standard deviations, and the corresponding $t$-values for the factors used in the analysis. We include the five factors suggested by Fama and French [2015]: return less risk-free rate ($MKT$), size ($SMB$), value ($HML$), investment ($CMA$), and profitability ($RMW$). We also include context-adjusted profitability factor, which is denoted as $RMW^{CN}$. Compared to $RMW$, $RMW^{CN}$ has a higher annualized return (12.72% versus 4.80%) and a higher standard deviation (21.17% versus 10.18%). All factor returns are statistically significant at conventional levels. Interestingly, $RMW^{CN}$ has the highest t-statistics across the six factors ($t$=2.75), which is equivalent to having the highest Sharpe ratio.

Panel B of Table 7 presents the correlations among the factor returns. As one would expect, $RMW$ and $RMW^{CN}$ are highly correlated with a correlation coefficient of 0.77. However, it is also clear that contextual information makes the correlation between $RMW^{CN}$ and $RMW$ far from one. $RMW^{CN}$ and $RMW$ share similar relations with other factor returns. Both $RMW^{CN}$ and $RMW$ are positively correlated with $CMA$ and $HML$ and they both are negatively correlated with $MKT$ and $SMB$.

### 7.1 Spanning tests

To examine whether $RMW^{CN}$ and $RMW$ carry sufficiently different sets of information, we perform spanning tests reported in Panel C of Table 7. Models in columns 1 and 2 use monthly $RMW$ as the dependent variable and regress it on the other factor returns (with and without context-based factor). When $RMW$ is regressed on factor returns excluding the $RMW^{CN}$ factor, we observe statistically significant factor loadings on $MKT$, $SMB$, and $HML$ factor returns. We also observe a statistically significant alpha of 0.56 (with $t$=4.54). This result is intuitive and implies that $RMW$ cannot be priced by the other four factors. It changes, however, in column 2, where we include context-based factor $RMW^{CN}$ in addition to the other four factors. The alpha in this regression loses statistical and economic significance (-0.05 with $t$=-0.53) suggesting that the model now fully prices the risks in $RMW$. Additionally, while the coefficient on $RMW^{CN}$ is highly significant, the loadings on other factors drop both in value and statistical significance.

In contrast, columns 3 and 4 use $RMW^{CN}$ as the dependent variable. When we regress $RMW^{CN}$ on the four factors excluding the profitability factor, we also observe significant a significant alpha (1.70 with $t$=7.09). In fact, both economic and statistical significance is substantially higher compared to the case of $RMW$. Furthermore, when we include $RMW$ as an additional independent variable in the fourth column, we now observe both statistically and economically a significant alpha (0.94 with $t$=5.23). This contrasts sharply with column 2 and indicates that the five traditional factors, including

*RMW* do not price the context-adjusted profitability factor.

We conclude that $RMW^{CN}$ in combination with other factors span the *RMW* factor but not the other way around. In contrast, $RMW^{CN}$ contains useful information about asset prices not reflected by the standard five factors including the profitability factor.

### 7.2 Pricing of profitability portfolios

We next use the constructed factor to examine its ability to price portfolios at a more granular level. We start with profitability-based portfolios. Table 8 reports performance for 5 $\times$ 5 portfolios based on operating profitability and context-adjusted profitability. Portfolio sorts are performed independently, in the same manner as in Table 6. For each portfolio, Panel A presents alphas based on (1) the five-factor model (upper part) and (2) based on the five-factor model that relies on the context-adjusted profitability factor $RMW^{CN}$ instead of *RMW*.

When we consider the five-factor model, high-minus-low portfolios based on operating profitability generally do not generate returns that are significantly different from zero. One exception is the second quintile of context-adjusted profitability, where we observe a high-minus-low return of 42 basis points ($t$=2.63). In contrast, high-minus-low portfolio returns based on context-adjusted profitability are significantly positive across all profitability quintiles. As in Table 6, we see a declining pattern in context-adjusted profitability high-minus-low portfolio returns as we move across OP quintiles (56 basis points for the lowest profitability quintile versus 38 basis points for the highest profitability quintile).

The lower part of Panel A presents the five-factor model alphas when we replace *RMW* with $RMW^{CN}$. Now, in contrast to the above, none of the five high-minus-low portfolio returns based on context-adjusted profitability sorts are statistically significant. At the same time, profitability-based strategies also do not generate positive alphas in all context-adjusted profitability quintiles. Overall, this suggests that, in line with spanning tests, context-based profitability factor prices portfolios more accurately. We next test this conclusion more formally.

Panel B of Table 8 present Gibbons, Ross, and Shanken [1989] statistics and the averaged absolute values of alphas corresponding to each set of the 25 portfolios in Panel A. Gibbons, Ross, and Shanken [1989] (hereafter, GRS) statistics tests the probability that alphas generated from the factor models are jointly different from zero. For the five-factor model, the average absolute alpha ($A(|\alpha|)$) across the 25 portfolios is 23 basis points and the GRS statistic is 3.25, which is significantly different from zero.

In contrast, for the model that relies on $RMW^{CN}$ and the other four factors, the average absolute alpha decreases by about 60% to 0.09. The corresponding GRS statistic also decreases to 1.36, which is statistically indistinguishable from zero. Overall, the evidence implies that the five-factor model based on context-adjusted $OP^{CN}$ factor is more effective at pricing the profitability portfolios.

### 7.3 Pricing of the Problematic Portfolios

We next put our model to the test based on the most challenging to price 25 portfolios. Fama and French [2015] show that the smallest stocks with the highest growth present a major problem for both three- and five-factor models. Specifically, they find for stocks in the smallest size and lowest profitability quintiles, three- and five-factor model alphas are -0.49 and -0.29, respectively. Based on their findings, it is plausible that measuring profitability factors with error is responsible for the failure of the five-factor model to price this portfolio. In particular, the extreme portfolio has negative exposure to the profitability factor but to the extent this exposure is underestimated (in absolute terms), a significant negative alpha is expected. Furthermore, the accounting profitability measurement for small firms with extreme growth is likely to be problematic and context-dependent. For example, these firms are likely to invest aggressively in R&D projects, whereas accountants conservatively assume that such investments do not create future profits and thus are expensed.[19]

To probe our model, we construct $5 \times 5$ *size* and *B/M* portfolios based on these characteristics using the NYSE breakpoints. Specifically, as of June 30 each year, we sort stocks into five groups using NYSE market capitalization breakpoints. We independently sort stocks into five groups based on NYSE book-to-market breakpoints. For each of the 25 resulting portfolios, we estimate alphas using Fama and French's three-factor model, five-factor model, and the five-factor model that relies on $RMW^{CN}$ factor.

For each of the three models, Table 9, Panel A. presents the estimated alphas and their corresponding t-statistics. Despite our sample period being considerably shorter, we are able to replicate the problematic portfolio results in Fama and French [2015] very closely. The three-factor model shows a large unexplained negative alpha in the small and low book-to-market portfolio (-0.51 with a *t*-value -4.35). This alpha decreases by about 40% (to -0.29 with a *t*-value -3.09) when we add investment and profitability factors into the model. However, as we replace the $RMW$ factor with our $RMW^{CN}$, the alpha in the problematic portfolio decreases further to -0.13 and loses statistical significance (*t*-value -1.65). Thus, taking profitability context into account goes a pretty long way toward

[19]Further, R&D projects are heterogeneous in terms of realization of future cash flows and their effect on long-term profitability is context-specific. Standard profitability measures cannot capture this context.

improving the performance of the existing asset pricing models.

Panel B evaluates the model's performance across a broader set of portfolios, by computing GRS statistic (Gibbons et al. [1989]) and the average absolute value of alphas across each set of 25 portfolios. For the three-factor alphas, the averaged absolute alphas is 0.125 and the corresponding GRS test statistic is 3.88, which rejects the null that alphas are jointly zero. However, the averaged absolute alphas decrease to 0.094 when we use the five-factor model and GRS statistic also declines to 2.75. Finally, the model with a context-based profitability factor still delivers a considerable improvement, with $A(|\alpha|)$ of 0.068 and GRS statistic of 1.76, which is at the borderline of statistical significance. In fact, both the relative and absolute improvement in these performance metrics from incorporating context is at least as large as that from going from a three- to five-factor model.

### 7.4 Ex Post Maximum Sharpe Ratio

In this subsection, we evaluate how much an investor could gain in mean-variance efficiency from incorporating the context into profitability measurement. We examine ex post mean-variance efficient portfolios by comparing Sharpe ratios associated with different models. Sharpe ratios characterize a feasible set of investment opportunities that investors are facing. We construct ex post portfolios that maximize the mean-to-variance ratio using the results of Table 7. We begin with $MKT$ and gradually expose our portfolio to include the five factors and $RMW^{CN}$. Specifically, we include $RMW^{CN}$ on top of the five factors to show an increase in investment opportunities provided by our new factor.

Table 10 presents optimal factor loadings along with the ex-post Sharpe ratios. The annualized Sharpe ratio for market portfolio during our sample period is 0.46. When we add $SMB$, $HML$, and $CMA$ into the model, the maximum Sharpe ratio increases to 0.77. Now, when we include $RMW$ into the model the Sharpe ratio improves further to 1.04. This means that an investor who trades based on the five factors can achieve an annualized mean-to-variance efficiency ratio of 1.04. When we replace $RMW$ with $RMW^{CN}$, the annualized Sharpe ratio improves to 1.24, which constitutes a 60% improvement relative to the four-factors model and a 19% improvement on top of the conventional five-factor model. Incorporating both $RMW$ and $RMW^{CN}$ into the investment strategy increments the Sharpe by another 0.07 points to 1.31, or a 26% increase relative to the traditional five-factor model (from 1.04 to 1.31).[20]

[20] The difference of the bias-adjusted squared Sharpe ratios between the traditional five factor model (model 5)and the model with context-adjusted profitability (model 6) 0.11 with a p-value of less than 0.01

Overall, this analysis suggests that investors can considerably benefit from incorporating profitability context into their optimal investment portfolios.

## 8 Extending the Prediction Horizon

Our final analysis extends the prediction horizon beyond one year. Specifically, we repeat the Fama-MacBeth regression analysis presented in Table 3, columns 1 and 3, while replacing the corresponding profitability proxies with their lags using six-month increments up to five years into the past.[21] We start the sample in January 2006 to ensure that all observations have at least five years of lagged data. We present the results graphically in Figure 3. Panels A and B plot regression coefficients and their 95% confidence intervals. Panel C plots the corresponding *t*-values.

Panel A shows the plot based on context-adjusted profitability, which corresponds to the specification in column 3 of Table 3. The zero-lag coefficient on $OP^{CN}$ is 1.95 with a *t*-value of 4.23. As profitability lag increases, i.e., information becomes stale over time, the coefficients decay in magnitude. Nevertheless, even after five years, we still observe a positive coefficient on $OP^{CN}$, implying that the profitability context (from five years ago) still appears to carry useful information not captured by the *current* characteristics. Panel C shows *t*-values, which also decay over time, remain statistically significant even at the four-and-a-half-year mark.

Panel B plots the coefficients for operating profitability, corresponding to the specification in column 1 of Table 3. Similar to Panel A, the coefficients decay over time as the information becomes more stale. However, the coefficients are lower in magnitude and reach the borderline of statistical significance after about two years. After four years, they are virtually equal to zero. Panel C also shows that the corresponding *t*-values decay faster than those of context-adjusted profitability. [22]

Finally, we observe that the *t*-values corresponding to $OP^{CN}$ are generally higher than those for *OP*. Further, at lags zero and one, the *t*-values for the two measures are reasonably close. However, as the lag increases, the gap widens, suggesting that the value of information in unadjusted profitability decays more quickly. This further highlights the

(Barillas et al. [2020]).

[21]Note that we keep all other regressors updated as in Table 3 while lagging operating profitability or context-adjusted profitability only. Doing so corresponds to a scenario where investors do not know the current value of profitability while having access to all other information. In conjunction with other equity characteristics, we examine whether stale profitability information can still predict future stock returns.

[22]In Ball et al. [2015], the coefficients on operating profitability remain statistically significant for about five years. The difference with the current result stems from the fact that the predictive power of operating profitability become lower in the 2010s. This can also be confirmed by the results in Figure 2. Our sample is also shorter, which reduces the power of the tests.

value added by corroborating narrative context when measuring profitability or numeric characteristics more generally.

## 9 Conclusion

Characteristics and factors used in asset pricing models generally do not account for the context within which they are reported. At the same time, companies generate large amounts of unstructured non-quantitative information that changes the way numeric information should be interpreted (e.g., discussion of competitive pressures). We study whether incorporating context surrounding the reported numbers improves our ability to measure stocks' characteristics and helps to explain the cross-section of stock returns. In particular, we take advantage of the recent advancements in natural language processing and deep learning to develop a measure of context-adjusted operating profitability. This measure avoids the "one-fits-all" approach and is able to accommodate rich contextual heterogeneity embedded in the required narrative disclosures.

The construction of our measure involves two key building blocks. First, we use BERT to encode and extract contextual information from MD&A statements in a systematic and tractable way. Second, we use the approach in Farrell et al. [2021b] which relies on restricted neural networks to learn the importance of context for the interpretation of numeric data. The model incorporates the rich heterogeneity in profitability context across firms and across time.

We perform Fama-MacBeth regression and portfolio sorts to show that context-adjusted profitability dominates and subsumes the traditional operating profitability measure in explaining the cross-section of future stock returns. We also demonstrate the economic importance of taking the context into account and that our measure carries more information as far as five years into the future.

Factor-based on context-adjusted profitability also delivers superior performance as compared to the traditional *RMW* factor. Most notably, the context-adjusted factor goes a long way in resolving the biggest challenge facing the five-factor model.

Overall, our study is the first to document the importance of context information when measuring asset pricing characteristics and constructing factors. Our context-adjusted proxy for future profitability can be easily adopted by other asset pricing researchers or modified to tackle the context of other characteristics.

## References


R. Ball, J. Gerakos, J. T. Linnainmaa, and V. V. Nikolaev. Deflating profitability. *Journal of Financial Economics*, 117(2):225–248, 2015.

R. Ball, J. Gerakos, J. T. Linnainmaa, and V. Nikolaev. Accruals, cash flows, and operating profitability in the cross section of stock returns. *Journal of Financial Economics*, 121(1): 28–45, 2016.

F. Barillas, R. Kan, C. Robotti, and J. Shanken. Model comparison with sharpe ratios. *Journal of Financial and Quantitative Analysis*, 55(6):1840–1874, 2020.

J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

E. F. Fama and K. R. French. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465, 1992.

E. F. Fama and K. R. French. Profitability, investment and average returns. *Journal of Financial Economics*, 82(3):491–518, 2006.

E. F. Fama and K. R. French. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22, 2015.

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

M. H. Farrell, T. Liang, and S. Misra. Deep neural networks for estimation and inference. *Econometrica*, 89(1):181–213, 2021a.

M. H. Farrell, T. Liang, and S. Misra. Deep learning for individual heterogeneity: An automatic inference framework, 2021b.

M. Gentzkow, B. Kelly, and M. Taddy. Text as data. *Journal of Economic Literature*, 57(3): 535–74, 2019.

M. R. Gibbons, S. A. Ross, and J. Shanken. A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, pages 1121–1152, 1989.

S. Gu, B. Kelly, and D. Xiu. Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5):2223–2273, 2020.

S. Gu, B. Kelly, and D. Xiu. Autoencoder asset pricing models. *Journal of Econometrics*, 222(1):429–450, 2021.

A. H. Huang, H. Wang, and Y. Yang. Finbert: A large language model for extracting information from financial text. *Contemporary Accounting Research*, 2022.

Z. T. Ke, B. T. Kelly, and D. Xiu. Predicting returns with text data. Technical report, National Bureau of Economic Research, 2019.

A. G. Kim and V. V. Nikolaev. Context-based interpretation of financial information.

*SSRN Working Paper*, 2023.

S. P. Kothari, X. Li, and J. E. Short. The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis. *The Accounting Review*, 84(5):1639–1670, 2009.

F. Li, R. Lundholm, and M. Minnis. A measure of competition based on 10-k filings. *Journal of Accounting Research*, 51(2):399–436, 2013.

R. Novy-Marx. The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28, 2013.

E. Rouen, E. C. So, and C. C. Wang. Core earnings: New data and evidence. *Journal of Financial Economics*, 142(3):1068–1091, 2021.

T. Shumway. The delisting bias in crsp data. *The Journal of Finance*, 52(1):327–340, 1997.

T. Shumway and V. A. Warther. The delisting bias in crsp's nasdaq data and its implications for the size effect. *The Journal of Finance*, 54(6):2361–2379, 1999.

# Appendix A1. Dependent Double-Sort Portfolios

This table reports three-factor model alphas two-way sorted by context-adjusted profitability and operating profitability. In Panel A, we first sort the stocks on operating profitability into quintiles and, conditioning on the sorts, we sort each quintile into another quintiles on context-adjusted profitability. In Panel B, we first sort the stocks on context-adjusted profitability into quintiles and, conditioning on the sorts, we sort each quintile into another quintiles on operating profitability. We construct value-weighted portfolios on June 30th of each year and hold the portfolio for one year. The sample starts in January 2001 and ends in October 2021.

**Panel A**. Context-Adjusted Profitability Conditioned on Operating Profitability

| Context-Adjusted | Profitability (Condition) | | | | | | |
|---|---|---|---|---|---|---|---|
| Profitability | Low | 2 | 3 | 4 | High | High-Low | *t*-value |
| Low | -0.35 | -0.33 | -0.41 | -0.22 | -0.11 | 0.24 | [1.49] |
| 2 | -0.18 | -0.11 | -0.23 | -0.16 | -0.09 | 0.09 | [0.98] |
| 3 | -0.10 | -0.06 | -0.13 | 0.11 | 0.15 | 0.25 | [1.72] |
| 4 | -0.13 | 0.13 | 0.22 | 0.18 | 0.21 | 0.34 | [1.88] |
| High | 0.23 | 0.20 | 0.12 | 0.20 | 0.22 | -0.01 | [-0.03] |
| High - Low | 0.58 | 0.53 | 0.53 | 0.42 | 0.33 | | |
| *t*-value | [2.98] | [2.85] | [2.88] | [2.56] | [2.12] | | |

**Panel B**. Operating Profitability Conditioned on Context-Adjusted Profitability

| | Context-Adjusted Profitability (Condition) | | | | | | |
|---|---|---|---|---|---|---|---|
| Profitability | Low | 2 | 3 | 4 | High | High-Low | *t*-value |
| Low | -0.19 | -0.21 | -0.20 | -0.18 | 0.16 | 0.35 | [2.39] |
| 2 | -0.15 | -0.19 | -0.13 | 0.18 | 0.23 | 0.38 | [2.51] |
| 3 | -0.12 | -0.01 | 0.05 | 0.12 | 0.29 | 0.41 | [2.55] |
| 4 | -0.11 | 0.18 | 0.11 | 0.16 | 0.19 | 0.30 | [1.80] |
| High | -0.13 | -0.07 | 0.08 | 0.13 | 0.26 | 0.39 | [2.56] |
| High - Low | 0.06 | 0.14 | 0.28 | 0.31 | 0.10 | | |
| *t*-value | [0.14] | [0.45] | [1.73] | [1.86] | [0.35] | | |

## Figure 1. Feedforward Deep Neural Network

This figure illustrates the feedforward deep neural network model that we use to obtain context-adjusted profitability. The input layer comprises context neurons extracted from the last hidden stage of the BERT base-uncased model. There are three hidden layers with 256, 64, and 16 neurons each. The input layer and the three hidden layers are fully connected with each other. The parameter layer comprises two parameters from the ordinary least squares regression $OP_{it+1} = \theta_0 + \theta_1 OP_{it} + \varepsilon_i t$. We get context-adjusted parameters $\widehat{\theta}_{0it}$ and $\widehat{\theta}_{1it}$ and combine them with the real data. We use RMSE as a loss function. The model iterates training following the rules that we introduce in Section 2.

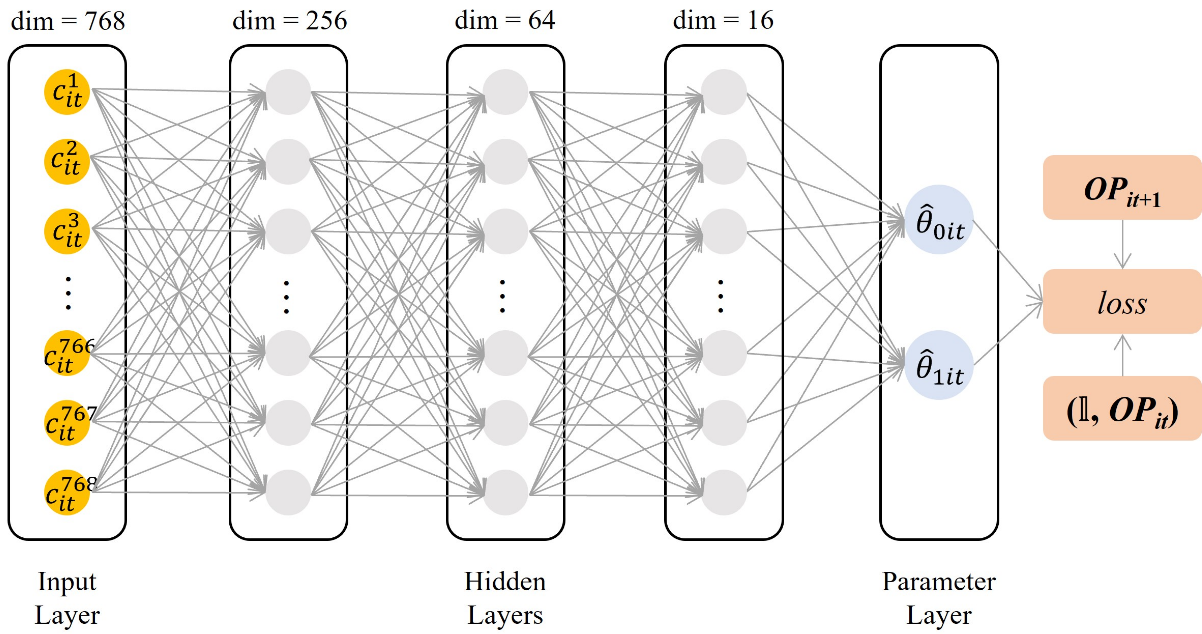



Figure 1. Feedforward Deep Neural Network

## Figure 2. Subsample Analysis of Fama and MacBeth [1973] Regressions

This figure illustrates the subsample analysis of Fama and MacBeth [1973] regressions in Table 3. We use a rolling sample period of ten years starting from 2000 January to 2020 December. The graph plots the *t*-values of regression slopes on operating profitability and context-adjusted profitability, respectively. We regress future stock returns on operating profitability or context-adjusted profitability, while controlling for other equity characteristics used in the main analysis (columns (1) and (2) of Table 3).

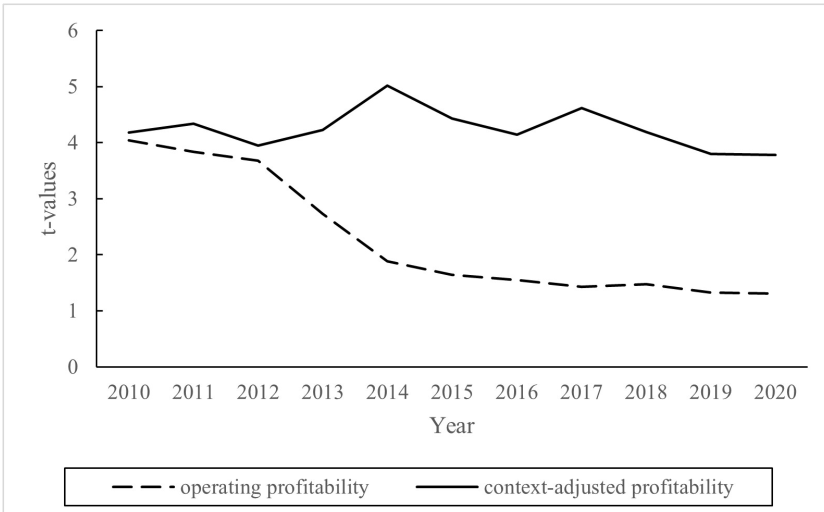
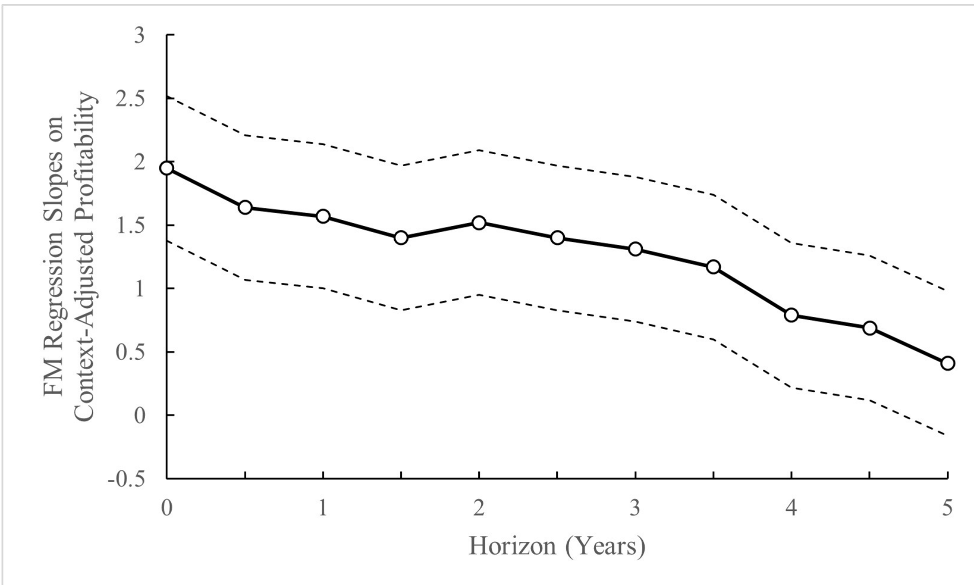


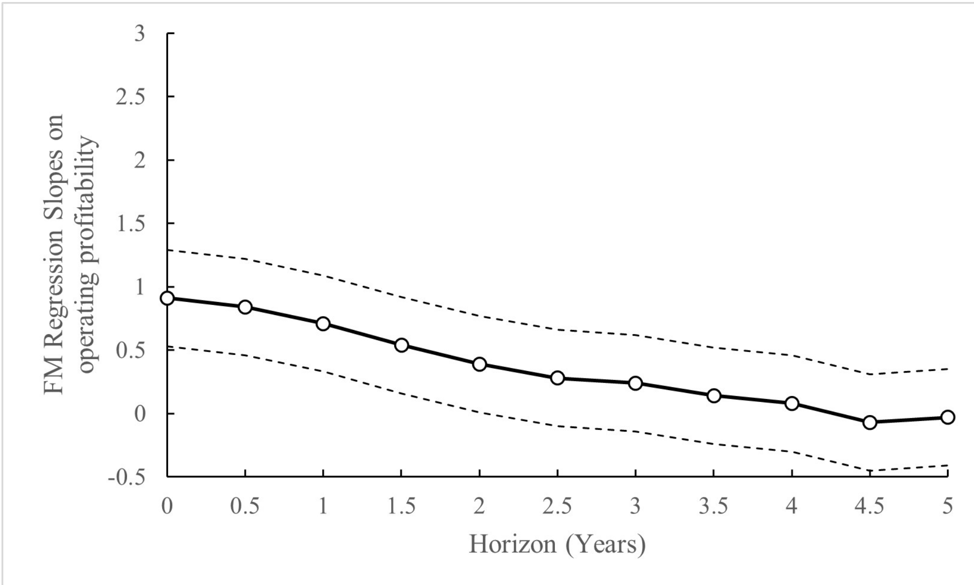

Figure 2. Subsample Analysis of Fama and MacBeth [1973] Regressions

## Figure 3. Extending Predictive Horizons

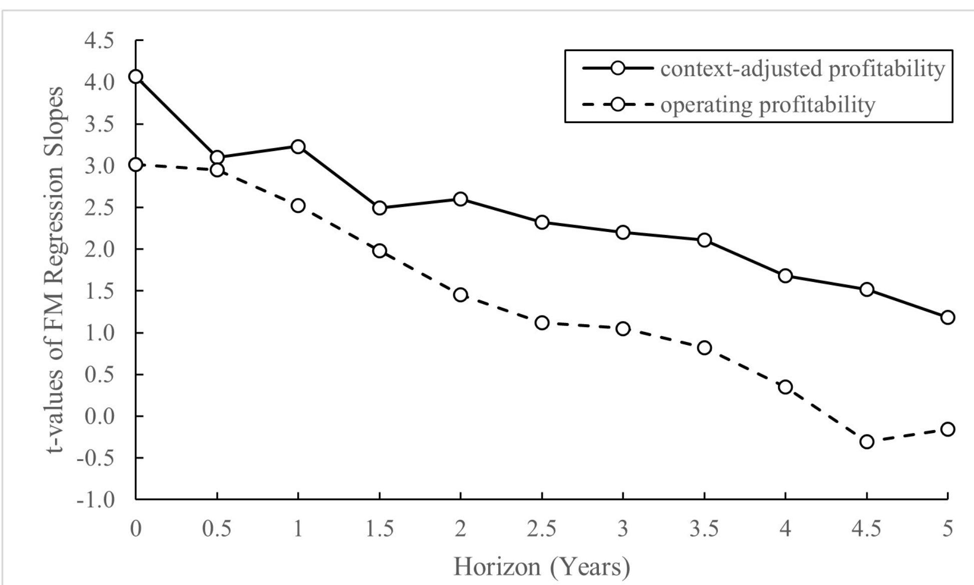
This figure plots average Fama and MacBeth [1973] regression slopes and their 95% confidence intervals. Regressions are estimated using observations from January 2006 to October 2021. We use samples from 2006 so that all observations have lagged terms of up to five years. We lag operating profitability or context-adjusted profitability up to five years using six-month increments. Other regressors are identical to those in Table 3 amd are updated accordingly. In Panel A, we lag context-adjusted profitability. In Panel B, we lag operating profitability. In Panel C, we plot $t$-values associated with regression coefficients.




Panel A. Fama and MacBeth [1973] Regression Coefficients on Context-Adjusted Profitability




Panel B. Fama and MacBeth [1973] Regression Coefficients on Operating Profitability




Panel C. *t*-values of Fama and MacBeth [1973] Regression Coefficients

## Table 1. Descriptive Statistics

This table reports the descriptive statistics of variables used in further analyses. Our sample period starts in 1995 and ends in 2020. $r_{0,1}$ is a one-month lagged return; $r_{2,12}$ is a one-year lagged return calculated after skipping one month; log(*ME*) is the natural logarithm of the monthly market capitalization; log(*BE*/*ME*) is the natural logarithm of the book-to-market ratio; *Investment* is asset growth ratio; *OP* is calculated by subtracting cost of goods sold and sales expense less R&D expense from revenue, deflated by total assets; $OP^{CN}$ is the expectation of profitability in t+1 formed at time t using the approach of Farrell et al. [2021b]. We require our observations to be traded on NYSE, Amex, and Nasdaq, and be a common share. We exclude financial firms and observations without valid financial information.

| | | | | Percentile | | |
|---|---|---|---|---|---|---|
| Variables | N | Mean | SD | $25^{th}$ | $50^{th}$ | $75^{th}$ |
| $r_{0,1}$ | 63,621 | 0.01 | 0.18 | -0.07 | 0.00 | 0.07 |
| $r_{2,12}$ | 63,621 | 0.13 | 0.73 | -0.21 | 0.04 | 0.30 |
| log(*ME*) | 63,621 | 5.85 | 2.02 | 4.39 | 5.79 | 7.21 |
| log(*BE*/*ME*) | 63,621 | -0.70 | 0.86 | -1.20 | -0.64 | -0.15 |
| *Investment* | 63,621 | 0.07 | 0.82 | 0.00 | 0.02 | 0.07 |
| *OP* | 63,621 | 0.15 | 0.62 | 0.02 | 0.20 | 0.34 |
| $OP^{C}$ | 63,621 | 0.10 | 0.03 | 0.08 | 0.10 | 0.12 |
| $OP^{CN}$ | 63,621 | 0.12 | 0.29 | 0.06 | 0.17 | 0.25 |

# Table 2. Predictive Ability of Expectations about Profitability

This table reports the predictive ability of three different estimates of future profitability. The dependent variable is the realized value of operating profitability of year *t*+1. We run simple OLS regressions without control variables. All independent variables are of year *t*. *t*-values of are reported within parentheses and standard errors are clustered at firm and year level.

| | Dependent Variable = $OP_{it+1}$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $OP_{it}$ | 0.4086 | | | 0.0932 |
| | [65.10] | | | [0.95] |
| $OP^{C}{}_{it}$ | | 0.3238 | | 0.1895 |
| | | [2.95] | | [0.49] |
| $OP^{CN}{}_{\mathrm{it}}$ | | | 0.9958 | 0.8448 |
| | | | [59.26] | [37.25] |
| Adjusted $R^2$ | 22.49% | 3.03% | 28.33% | 29.66% |

# Table 3. Fama-MacBeth Regressions

This table reports monthly Fama and MacBeth [1973] regressions and Newey-West $t$-values with lag=3. We use returns from January 2001 to October 2021. $r_{0,1}$ is the lagged monthly return, $r_{2,12}$ is the lagged yearly return after skipping a month. Investment is asset growth ratio. $OP$ is (total revenue – cost of goods sold – (sales, administrative expense – R&D expense)) scaled by total asset. log($ME$) is the natural logarithm of the market value, log($BE/ME$) is the natural logarithm of the book-to-market ratio, $OP^{CN}$ is context-adjusted profitability. Microcap stocks are below the 20th percentile of the NYSE market value. All independent variables are trimmed at 1% level.

**Panel A**. All Stocks

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $r_{0,1}$ | -2.26 | -2.02 | -2.43 | -2.46 | -2.49 |
| | [-3.09] | [-2.73] | [-3.42] | [-3.49] | [-3.53] |
| $r_{2,12}$ | -0.10 | -0.05 | -0.13 | -0.15 | -0.14 |
| | [-0.30] | [-0.14] | [-0.38] | [-0.44] | [-0.42] |
| log($ME$) | -0.16 | -0.12 | -0.18 | -0.19 | -0.18 |
| | [-2.98] | [-1.88] | [-3.57] | [-3.74] | [-3.75] |
| log($BE/ME$) | 0.02 | -0.05 | -0.05 | -0.04 | -0.03 |
| | [0.17] | [-0.39] | [-0.39] | [-0.34] | [-0.22] |
| *Investment* | -1.70 | -1.35 | -1.68 | -1.70 | -1.73 |
| | [-4.52] | [-3.55] | [-4.45] | [-4.47] | [-4.60] |
| $OP$ | 0.91 | | | -0.32 | -0.41 |
| | [3.01] | | | [-0.67] | [-0.83] |
| $OP^{C}$ | | 4.17 | | | 5.94 |
| | | [0.34] | | | [0.48] |
| $OP^{CN}$ | | | 1.95 | 3.26 | 3.47 |
| | | | [4.23] | [2.29] | [2.34] |
| Average Adj $R^2$ | 3.94% | 3.28% | 4.45% | 4.68% | 4.73% |

**Panel B**. All but Microcap Stocks

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $r_{0,1}$ | -1.26 | -1.02 | -1.48 | -1.65 | -1.51 |
| | [-1.96] | [-1.55] | [-2.35] | [-2.64] | [-2.40] |
| $r_{2,12}$ | 0.01 | 0.04 | -0.01 | -0.02 | -0.01 |
| | [0.03] | [0.13] | [-0.04] | [-0.05] | [-0.04] |
| log($ME$) | -0.16 | -0.10 | -0.17 | -0.17 | -0.17 |
| | [-2.85] | [-1.55] | [-3.51] | [-3.51] | [-3.68] |
| log($BE/ME$) | -0.02 | -0.01 | -0.10 | -0.04 | -0.06 |
| | [-0.14] | [-0.04] | [-0.85] | [-0.36] | [-0.52] |
| *Investment* | -2.02 | -1.59 | -2.01 | -2.02 | -2.01 |
| | [-5.03] | [-3.98] | [-5.00] | [-5.05] | [-5.02] |
| $OP$ | 0.99 | | | -0.17 | -0.36 |
| | [3.11] | | | [-0.36] | [-0.72] |
| $OP^{C}$ | | -1.16 | | | -1.18 |
| | | [-0.10] | | | [-0.09] |
| $OP^{CN}$ | | | 2.13 | 3.14 | 3.52 |
| | | | [3.27] | [2.31] | [2.47] |
| Average Adj $R^2$ | 4.64% | 3.79% | 5.18% | 5.39% | 5.66% |

## Table 4. Economic Magnitudes

Using the averaged monthly Fama and MacBeth [1973] regression slopes from Table 3, we calculate monthly fitted returns of each stock. Then we partition the sample into two groups based on monthly fitted (predicted) returns. Then we compute the spread between the two groups based on predicted returns and actual returns. We use regression specifications from Panel A, Table 3, columns (1), (2), and (3). The benchmark model refers to where we use one-month momentum, one-year momentum, size, book-to-market, and investment as regressors. We also report t-statistics of actual return spreads based on their time-series standard errors.

| | Predicted Spread | | Actual Spread | | t(Actual Spread) | |
|---|---|---|---|---|---|---|
| Regression | EW | VW | EW | VW | EW | VW |
| 0 Benchmark (no OP) | 0.38 | 0.40 | 0.44 | 0.35 | 2.78 | 2.13 |
| 1 (OP) | 0.43 | 0.45 | 0.50 | 0.42 | 3.52 | 3.25 |
| 2 ($OP^{C}$) | 0.39 | 0.42 | 0.44 | 0.36 | 3.30 | 2.16 |
| 3 ($OP^{CN}$) | 0.51 | 0.54 | 0.59 | 0.51 | 4.48 | 3.59 |

# Table 5. Single Sort Portfolios

This table reports value-weighted excess returns over market returns, CAPM alphas, three-factor model alphas, and five-factor model alphas for portfolios sorted by contextuality factor. We construct value-weighted portfolios on June 30th of each year based on $OP^{CN}$ (Panel A) and $OP$ (Panel B) and hold the portfolio for one year. The sample starts in January 1999 and ends in October 2020. $t$-values are reported in parentheses.

**Panel A**. Sorted on Context-Adjusted Profitability

| | | Alphas | | | |
|---|---|---|---|---|---|
| | Excess Return | CAPM | FF3 | FF4 | FF5 |
| Low | 0.67 | -0.27 | -0.33 | -0.28 | -0.25 |
| 2 | 0.48 | -0.15 | -0.34 | -0.25 | -0.15 |
| 3 | 0.44 | -0.16 | -0.23 | -0.15 | -0.13 |
| 4 | 0.75 | -0.17 | -0.24 | -0.22 | -0.07 |
| 5 | 0.64 | -0.13 | -0.19 | -0.15 | -0.13 |
| 6 | 0.74 | -0.04 | -0.11 | -0.04 | 0.05 |
| 7 | 0.67 | -0.00 | -0.04 | -0.03 | 0.11 |
| 8 | 0.85 | 0.15 | 0.19 | 0.18 | 0.16 |
| 9 | 0.92 | 0.22 | 0.28 | 0.29 | 0.18 |
| High | 0.89 | 0.28 | 0.29 | 0.30 | 0.23 |
| High - Low | 0.22 | 0.55 | 0.62 | 0.58 | 0.48 |
| *t*-value | [0.57] | [2.96] | [3.03] | [2.98] | [2.65] |

**Panel B**. Sorted on Profitability

| | | Alphas | | | |
|---|---|---|---|---|---|
| | Excess Return | CAPM | FF3 | FF4 | FF5 |
| Low | 0.39 | -0.37 | -0.32 | -0.30 | 0.03 |
| 2 | 0.35 | -0.48 | -0.47 | -0.43 | -0.07 |
| 3 | 0.78 | -0.26 | -0.38 | -0.37 | 0.25 |
| 4 | 0.77 | -0.02 | -0.13 | -0.15 | 0.17 |
| 5 | 0.59 | -0.13 | -0.17 | -0.16 | 0.11 |
| 6 | 0.81 | 0.12 | 0.05 | 0.03 | 0.20 |
| 7 | 0.68 | 0.06 | 0.01 | 0.04 | 0.17 |
| 8 | 0.84 | 0.15 | 0.13 | 0.13 | 0.20 |
| 9 | 0.65 | -0.01 | -0.03 | -0.01 | -0.04 |
| High | 1.02 | 0.09 | 0.15 | 0.11 | 0.38 |
| High - Low | 0.63 | 0.46 | 0.47 | 0.41 | 0.35 |
| *t*-value | [1.77] | [2.53] | [2.62] | [2.58] | [1.34] |

# Table 6. Double-Sort Portfolios and Three-Factor Alphas

This table reports three-factor model alphas independently sorted by context-adjusted profitability and operating profitability. We construct value-weighted portfolios on June 30th of each year and hold the portfolio for one year. The sample starts in January 2001 and ends in October 2021.

| Context-Adjusted Profitability | Profitability Low | 2 | 3 | 4 | High | High-Low | *t*-value |
|---|---|---|---|---|---|---|---|
| Low | -0.49 | -0.41 | -0.40 | -0.39 | -0.36 | 0.13 | [0.53] |
| 2 | -0.20 | -0.09 | -0.03 | -0.31 | 0.09 | 0.29 | [1.35] |
| 3 | -0.08 | -0.01 | -0.13 | 0.02 | 0.25 | 0.33 | [1.52] |
| 4 | -0.17 | 0.11 | 0.22 | 0.26 | 0.09 | 0.26 | [1.12] |
| High | 0.22 | 0.26 | 0.23 | 0.25 | 0.24 | 0.02 | [0.02] |
| High - Low | 0.71 | 0.67 | 0.63 | 0.64 | 0.60 | | |
| *t*-value | [3.54] | [3.05] | [2.97] | [2.99] | [3.00] | | |

## Table 7. Spanning Regressions

In Panel A, we present annualized average returns of six factors: market return minus the risk-free rate (*MKT*), size (*SMB*), value (*HML*), investment (*CMA*), profitability (*RMW*), and context-adjusted profitability ($RMW^{CN}$). To calculate $RMW^{CN}$ factor returns, we follow Fama and French [1992] and form six portfolios based on size and context-adjusted profitability factors. Panel B reports Pearson correlations among the factors. Finally, in Panel C, we tabulate spanning regression results. *t*-values are reported in parentheses.

**Panel A**. Descriptive Statistics

| | *MKT* | *SMB* | *HML* | *CMA* | *RMW* | $RMW^{CN}$ |
|---|---|---|---|---|---|---|
| Annualized Returns (%) | 7.20 | 1.68 | 3.36 | 3.00 | 4.80 | 12.72 |
| Annual Standard deviation (%) | 15.69 | 11.60 | 10.88 | 6.96 | 10.18 | 21.17 |
| *t*-value | 1.94 | 2.20 | 1.70 | 1.99 | 2.00 | 2.75 |

**Panel B**. Correlation Matrix

| | *MKT* | *SMB* | *HML* | *CMA* | *RMW* | $RMW^{CN}$ |
|---|---|---|---|---|---|---|
| *MKT* | 1 | | | | | |
| *SMB* | 0.27 | 1 | | | | |
| *HML* | -0.00 | 0.05 | 1 | | | |
| *CMA* | -0.24 | 0.07 | 0.57 | 1 | | |
| *RMW* | -0.37 | -0.49 | 0.39 | 0.24 | 1 | |
| $RMW^{CN}$ | -0.41 | -0.42 | 0.41 | 0.31 | 0.77 | 1 |

**Panel C**. Spanning Regressions

| Dep Var= | *RMW* | | $RMW^{CN}$ | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Alpha | 0.56 | -0.05 | 1.70 | 0.94 |
| | [4.54] | [-0.53] | [7.09] | [5.23] |
| *MKT* | -0.21 | -0.04 | -0.45 | -0.17 |
| | [-6.55] | [-1.77] | [-7.36] | [-3.56] |
| *SMB* | -0.23 | -0.05 | -0.51 | -0.19 |
| | [-4.76] | [-1.38] | [-5.31] | [-2.64] |
| *HML* | 0.19 | -0.12 | 0.86 | 0.61 |
| | [3.36] | [-2.70] | [7.90] | [7.58] |
| *CMA* | -0.01 | 0.09 | -0.26 | -0.26 |
| | [-0.05] | [1.49] | [-1.61] | [-2.19] |
| *RMW* | | | | 1.35 |
| | | | | [15.00] |
| $RMW^{CN}$ | | 0.36 | | |
| | | [15.00] | | |
| *N* | 249 | 249 | 249 | 249 |
| Adjusted $R^2$ | 29.39% | 63.84% | 41.27% | 59.92% |

# Table 8. Double Sorts and Five-Factor Alphas

Panel A reports five-factor model alphas independently sorted by context-adjusted profitability and operating profitability. We include four factors and $RMW^{CN}$ factor in the first analysis and exchange the $RMW$ factor with our $RMW^{CN}$ factor in the second analysis. Panel B reports test statistics. A($|\alpha|$) is the averaged regression intercepts; GRS is the Gibbons, Ross, and Shanken [1989] test statistic which measures the probability that all alphas are jointly zero for a given model. We construct value-weighted portfolios on June 30th of each year and hold the portfolio for one year. The sample starts in January 2001 and ends in October 2021.

**Panel A**. Alphas

| | Monthly Alphas | | | | | | t-values | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Profitability | | | | | | Profitability | | | | | |
| $OP^{CN}$ | Low | 2 | 3 | 4 | High | H-L | Low | 2 | 3 | 4 | High | H-L |
| **Five-Factor Model** | | | | | | | | | | | | |
| Low | -0.25 | -0.18 | -0.19 | -0.17 | -0.12 | 0.13 | -1.85 | -1.43 | -0.85 | -0.78 | -0.63 | 1.00 |
| 2 | -0.22 | 0.22 | 0.29 | 0.19 | 0.20 | 0.42 | -1.73 | 1.78 | 1.72 | 1.56 | 1.58 | 2.63 |
| 3 | 0.12 | 0.21 | 0.19 | 0.53 | 0.22 | 0.10 | 0.51 | 1.60 | 1.39 | 2.88 | 1.59 | 0.92 |
| 4 | 0.26 | 0.10 | 0.15 | 0.29 | 0.25 | -0.01 | 1.60 | 1.12 | 1.18 | 2.02 | 1.90 | -0.08 |
| High | 0.31 | 0.27 | 0.27 | 0.22 | 0.26 | -0.05 | 2.03 | 1.82 | 1.90 | 1.76 | 1.86 | -0.16 |
| High-Low | 0.56 | 0.45 | 0.46 | 0.39 | 0.38 | | 3.07 | 2.58 | 2.67 | 2.51 | 2.45 | |
| **Four-Factor Model + Context-Adjusted Profitability Factor** | | | | | | | | | | | | |
| Low | -0.19 | -0.10 | -0.02 | -0.07 | -0.03 | 0.16 | -1.39 | -1.28 | -0.05 | -0.12 | -0.09 | 1.28 |
| 2 | -0.04 | 0.14 | 0.14 | 0.15 | 0.12 | 0.16 | -0.12 | 1.25 | 1.13 | 1.30 | 1.03 | 1.32 |
| 3 | 0.01 | 0.07 | 0.11 | 0.16 | 0.10 | 0.09 | 0.05 | 0.40 | 0.85 | 1.36 | 0.90 | 0.48 |
| 4 | -0.13 | 0.14 | 0.04 | 0.06 | 0.04 | 0.17 | -1.48 | 1.19 | 0.11 | 0.12 | 0.08 | 1.36 |
| High | -0.11 | -0.13 | 0.09 | 0.05 | 0.08 | 0.19 | -1.35 | -1.26 | 0.18 | 0.10 | 0.13 | 1.45 |
| High-Low | 0.08 | -0.03 | 0.11 | 0.12 | 0.11 | | 0.19 | -0.08 | 0.95 | 1.02 | 0.83 | |

**Panel B**. Test Statistics

| Model | GRS | A($\lvert\alpha\rvert$) |
|---|---|---|
| Four Factors + $RMW$ | 3.25 | 0.23 |
| Four Factors + $RMW^{CN}$ | 1.36 | 0.09 |

## Table 9. Size-B/M Portfolios, $OP$, and $OP^{CN}$

This table reports the alphas of the portfolios sorted on size and book-to-market. Stocks are assigned to five size groups using NYSE market cap breakpoints (available at Ken French's website) and independently assigned to five book-to-market groups using NYSE breakpoints (also available at Ken French's website). The portfolios are formed on June 30th of each year. In Panel A, we regress the monthly returns of each portfolios on monthly factors to get alphas. We get alphas for the three-factor model, the five-factor model, and the four-factor model with context-adjusted profitability factor. We also report t-statistics for each alpha. In Panel B, we calculate the GRS statistics and the averaged absolute alphas for the three factor models.

**Panel A**. Portfolio Sorts

| | Monthly Alphas | | | | | t-values | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Book-to-Market | Low | 2 | 3 | 4 | High | Low | 2 | 3 | 4 | High |
| | Three Factor Model | | | | | | | | | |
| Small | -0.51 | 0.01 | 0.03 | 0.15 | 0.16 | -4.35 | 0.04 | 0.32 | 2.45 | 2.67 |
| 2 | -0.20 | -0.05 | -0.11 | 0.10 | -0.12 | -2.98 | -0.60 | -1.36 | 1.19 | -1.20 |
| 3 | -0.08 | 0.08 | 0.01 | 0.03 | 0.13 | -1.02 | 0.99 | 0.02 | 0.35 | 2.37 |
| 4 | 0.15 | -0.11 | 0.06 | 0.13 | -0.15 | 2.05 | -1.54 | 0.98 | 2.31 | -2.15 |
| Big | 0.21 | 0.05 | -0.12 | -0.15 | -0.22 | 2.98 | 0.45 | -1.99 | -2.10 | -2.33 |
| | Five Factor Model | | | | | | | | | |
| Small | -0.29 | 0.04 | 0.05 | 0.13 | 0.11 | -3.09 | 0.33 | 0.54 | 2.28 | 2.09 |
| 2 | -0.13 | -0.03 | -0.12 | 0.04 | -0.05 | -1.52 | -0.29 | -1.48 | 0.41 | -0.62 |
| 3 | -0.15 | -0.11 | 0.01 | 0.08 | 0.06 | -1.94 | -1.35 | 0.03 | 0.83 | 1.19 |
| 4 | 0.11 | 0.12 | 0.02 | 0.11 | -0.08 | 1.63 | 1.46 | 0.08 | 1.40 | -1.05 |
| Big | 0.10 | 0.15 | -0.09 | 0.02 | -0.15 | 1.48 | 2.05 | -0.15 | 0.08 | -1.57 |
| | Four-Factor Model + Context-Adjusted Profitability Factor | | | | | | | | | |
| Small | -0.13 | 0.08 | 0.03 | 0.10 | 0.08 | -1.65 | 0.82 | 0.35 | 1.10 | 0.92 |
| 2 | -0.05 | -0.06 | -0.11 | 0.02 | -0.02 | -0.58 | -0.52 | -1.01 | 0.12 | -0.06 |
| 3 | -0.10 | -0.04 | 0.02 | 0.05 | 0.04 | -1.35 | -0.35 | -0.42 | 0.63 | 0.54 |
| 4 | 0.08 | 0.09 | 0.02 | 0.08 | -0.06 | 1.21 | 1.03 | 1.11 | 0.89 | -0.85 |
| Big | 0.12 | 0.14 | -0.06 | 0.01 | -0.08 | 1.53 | 1.66 | -0.61 | 0.03 | -1.03 |

**Panel B**. GRS Statistics

| | Three-Factor | Five-Factor | Four-Factor + $OP^{CN}$ |
|---|---|---|---|
| GRS | 3.88 | 2.75 | 1.76 |
| $A(\|\alpha\|)$ | 0.125 | 0.094 | 0.068 |

## Table 10. Ex-Post Sharpe Ratios

This table reports the maximum possible Sharpe ratio ex-post and optimal portfolio weights. We start from the market return less risk-free returns (*MKT*) and include the factors sequentially. In the sixth iteration, we also add $RMW^{CN}$ factor. The sample starts in January 2001 and ends in October 2021.

| | Optimal Loadings | | | | | | Sharpe Ratio |
|---|---|---|---|---|---|---|---|
| | *MKT* | *SMB* | *HML* | *CMA* | *RMW* | $RMW^{CN}$ | |
| 1 | 100% | | | | | | 0.46 |
| 2 | 93% | 7% | | | | | 0.47 |
| 3 | 48% | 3% | 49% | | | | 0.55 |
| 4 | 26% | 3% | 17% | 54% | | | 0.77 |
| 5 | 24% | 2% | 8% | 36% | 30% | | 1.04 |
| 6 | 29% | 1% | 9% | 40% | | 21% | 1.24 |
| 7 | 27% | 1% | 5% | 30% | 23% | 14% | 1.31 |