This article was downloaded by: [101.5.72.193] On: 02 January 2022, At: 21:50
Publisher: Institute for Operations Research and the Management Sciences (INFORMS)
INFORMS is located in Maryland, USA

Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Tail-Heaviness, Asymmetry, and Profitability Forecasting by Quantile Regression

Hui Tian , Andrew Yim , David P. Newton

To cite this article:
Hui Tian , Andrew Yim , David P. Newton (2021) Tail-Heaviness, Asymmetry, and Profitability Forecasting by Quantile Regression. Management Science 67(8):5209-5233. https://doi.org/10.1287/mnsc.2020.3694

Full terms and conditions of use: https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2020, INFORMS

Please scroll down for article—it is on subsequent pages

With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE

ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Tail-Heaviness, As mmetr , and Profitabilit Forecasting b Quantile Regression

Hui Tian,[a,b] Andrew Yim,[c] David P. Newton[d]

[a] School of Management and Economics, Beijing Institute of Technology, Beijing 100081, People's Republic of China; [b] PBC School of Finance, Tsinghua University, Beijing 100083, People's Republic of China; [c] Faculty of Finance, Cass Business School, City, University of London, London EC1Y 8TZ, United Kingdom; [d] School of Management, University of Bath, Bath BA2 7AY, United Kingdom

Contact: tianh@pbcsf.tsinghua.edu.cn (HT); a.yim@city.ac.uk, https://orcid.org/0000-0002-8063-6572 (AY); d.p.newton@bath.ac.uk, https://orcid.org/0000-0003-0470-8113 (DPN)

Received: July 29, 2017
Revised: November 16, 2018; August 18, 2019; April 8, 2020
Accepted: April 27, 2020
Published Online in Articles in Advance: December 7, 2020

https://doi.org/10.1287/mnsc.2020.3694

Copyright: © 2020 INFORMS

Abstract. We show that quantile regression is better than ordinary-least-squares (OLS) regression in forecasting profitability for a range of profitability measures following the conventional setup of the accounting literature, including the mean absolute forecast error (MAFE) evaluation criterion. Moreover, we perform both a simulated-data and an archival-data analysis to examine how the forecasting performance of quantile regression against OLS changes with the shape of the profitability distribution. Considering the MAFE and mean squared forecast error (MSFE) criteria together, we see that the quantile regression is more accurate relative to OLS when the profitability to be forecast has a heavier-tailed distribution. In addition, the asymmetry of the profitability distribution has either a U-shape or an inverted-U-shape effect on the forecasting accuracy of quantile regression. An application of the distributional shape analysis framework to cash flow forecasting demonstrates the usefulness of the framework beyond profitability forecasting, providing additional empirical evidence on the positive effect of tail-heaviness and supporting the notion of an inverted-U-shape effect of asymmetry.

History: Accepted by Shiva Rajgopal, accounting.
Supplemental Material: The online appendix is available at https://doi.org/10.1287/mnsc.2020.3694.

Ke words: heavy tails • distributional shape • profitability forecast • quantile regression

1. Introduction

It is in the interest of different parties, including investors, analysts, and companies themselves, to obtain more accurate profitability forecasts. Companies have experienced extreme profits and losses more often in recent decades.[1] This is likely to impact the distributional shape of profitability, increasing the difficulty in forecasting profitability accurately.

To formulate forecasts as accurately as possible, sophisticated market participants are likely to resort to statistical methods. Ordinary-least-squares (OLS) regression is a very popular choice, if not the prevalent choice. The least squares method has a very long history dating back to 1795 (Courgeau 2012). In contrast, quantile regression (QR), an alternative approach based on the least absolute deviation (LAD) method, was developed only four decades ago by Koenker and Bassett (1978). Unlike the least squares method, the LAD method is not sensitive to outliers (Chen et al. 2008). Despite this advantage, QR applications in finance and accounting remain not popular.[2] However, QR has long been considered an attractive method in areas such as medicine, survival analysis, and economics (Yu et al. 2003).

In this study, we conduct a series of analyses to examine whether the QR approach to profitability forecasting can be more accurate than the OLS approach, and if so, under what distributional shape of profi

Prior research on profitability forecasting examines these traditional measures of profitability, because they are the inputs to accounting-based valuation

accuracy of QR to tail-heaviness and asymmetry using archival data. The data used comes from the sample we use for the out-of-sample testing in the forecasting analysis. Unlike the simulated experiments, where it is straightforward to compute distributional shape measures based on many draws of simulated profitability, archival data does not allow this luxury. Even when some firms have sufficiently long time series to give reliable estimates, the data requirement would induce a severe survivorship bias. Therefore, we estimate the tail-heaviness and asymmetry measures based on the profitability distribution across different firms for each industry-year. This is consistent with the cross-sectional approach to forecasting, which assumes that there is enough similarity across different firms to warrant pooling them together for forecasting,

The above is not the only difference between the simulated and archival data. There are several. For example, in the archival data, the individual firms' absolute and squared forecast errors used for computing the forecasting accuracy measures are based on a full model consistent with Fairfield et al. (2009), instead of the simple first-order autoregressive model assumed in the simulated experiments. Moreover, the archival data comes from an in-sample estimation step using a rolling window of data available from the previous 10 years, whereas the corresponding step in the simulated experiments uses only one prior period of simulated data.

Given such differences, it is not obvious that the insights from the simulated experiments would be robust enough to hold also in the archival data. We, however, find a varying degree of support for the insights. In both the unweighted and the weighted regressions pooling all profitability measures together, the effect of tail-heaviness on the accuracy of QR forecasts relative to OLS is significantly positive across all the tail-heaviness measures controlling for any one of the asymmetry measures. There is also clear support for a positive effect of tail-heaviness from the individual-profitability regressions for CbOP (cash flow approach) and ROE and moderate support from those for OP, CbOP (balance-sheet approach), and RNOA.

Considering the differences between the simulated and archival data, we view the above finding from the archival data as generally corroborating the simulation results of the tail-heaviness effect. Similarly, in the archival-data analysis, the pooled regression and the ROE results show strong support for an inverted-U-shape effect of asymmetry, whereas three of the six profitability measures provide strong to moderate support for a U-shape effect, with the remaining two having no significant effect whatsoever. These results echo the not-so-consistent effect of asymmetry found in the simulated-data analysis.

To demonstrate the usefulness of the distributional shape analysis framework beyond profitability forecasting, we apply the framework to examine the out-of-sample forecasting of cash flows from 1990 to 2015 studied by Nallareddy et al. (2020). We show that the tail-heaviness, measured by the kurtosis, of the yearly cash flow distribution across all firms has a positive effect on the incremental forecasting accuracy of QR, whereas the asymmetry, measured by the skewness coefficient, has an inverted-U-shape effect. We also analyze various subsamples that exclude firms likely to have contributed to the tail-heaviness and asymmetry of the cash flow distribution. By confining our analysis to these subsamples, we expect to see a somewhat weaker relation between the incremental forecasting accuracy and the distributional properties. The subsample findings are largely consistent with our expectation. All in all, the results of the cash flow distributional shape analysis for the full sample and the various subsamples are in line with the earlier findings for profitability forecasting.

To our knowledge, we are the first to provide large-sample evidence of the effects of the profitability distributional shape on the accuracy of QR forecasts relative to OLS using both simulated and archival data. Related prior simulation studies were done 20–30 years ago. They primarily focus on the LAD estimators, rather than out-of-sample forecasts, or otherwise on the small-sample forecasting performance or use a simulation setup that has a maximum of 1,000 draws repeated for only 20 times (Dielman 1986, Mitra 1987, Dielman and Rose 1994). In contrast, our setup has 2,500 draws repeated for 500 times for each set of the distribution type and parameter combinations. Most importantly, none of the prior studies has considered asymmetry jointly with tail-heaviness. We examine both aspects of the distributional shape using two four-parameter distribution families that allow controlling not only the location and scale but also the tail and skewness properties separately. These families are the stable and the inverse hyperbolic sine (IHS) distributions (McDonald and Turley 2011; Nolan 2013, 2019).

In making the contribution above, we develop a framework of conducting simulated-data and archival-data analysis of the profitability distributional shape and its relation to forecasting accuracy under both the MAFE and MSFE criteria. This includes the use of various new measures, such as the incremental and relative forecasting accuracy measures55377lRracyioi

We are also the first to document the higher accuracy of QR forecasts, compared with OLS forecasts, across four new profitability measures and two traditional measures (including ROE). In contrast, a recent paper by Evans et al. (2017) focuses on comparing model-based forecasts of ROE (both LAD and OLS) to analysts' explicit forecasts of ROE. They do not at the same time examine the profitability distributional shape's effects on the accuracy of QR forecasts nor consider both the MAFE and MSFE criteria.

This study also adds to the debate on the reasons behind analyst forecast bias (Gu and Wu 2003, Basu and Markov 2004) by clarifying the roles of earnings skewness and the assumption of an absolute loss function (or MAFE minimization objective) for analysts. An MAFE minimization objective is very plausible, and the distributions of profitability (as a kind of scaled earnings) indeed are often skewed.[3] However, neither of these is necessary to explain why analysts are likely to have formulated their forecasts based on a median rather than a mean forecast (estimated with quantile and OLS regressions, respectively). Even when analysts have a quadratic loss function and the objective is to minimize MSFE, they can still find it gainful to use a QR forecasting approach under the circumstances of heavy-tailed distributions, even without skewness.

2. Related Literature

2.1. Firm Profitabilit Forecasts

Profitability is a key indicator of company performance and is widely used as an input for valuation. Traditional measures of profitability include ROE and RNOA. Freeman et al. (1982) show that there is regression toward the mean in ROE and establish that extreme ROEs are more transitory than moderate ones. Fama and French (2000) provide evidence that mean reversion in firm profitability is a robust phenomenon and suggest that changes in profitability and earnings are to some degree predictable. In a simple partially adjusted model using U.S. data, they find an estimated rate of mean reversion around 38% p.a. Similar results are documented by Allen and Salim (2005), who report a mean reversion rate of 25% p.a. in the UK market. We follow Fairfield et al. (2009) in using a forecasting model that captures the mean-reversion pattern of profitability conditional on the deviation of a firm's profitability from the median profitability benchmark (Freeman et al. 1982, Fama and French 2000).

Besides ROE and RNOA, we consider several alternative measures of profitability: GP, OP, and CbOP. They are the gross profit, operating profit, and cash-based operating profit, deflated by the total assets lagged by one year. *GP* is the sales minus the cost of goods sold. *OP* is defined as the gross profit minus the selling, general, and administrative expenses reported (i.e., the Compustat-adjusted selling, general, and administrative expenses with the expenditures on research and development subtracted in order to undo this adjustment by Compustat). Two versions of *CbOP* are obtained by purging accruals from the operating profit, with the accruals constructed using the cash flow approach or the Sloan (1996) balance-sheet approach (see Ball et al. 2016, p. 44). Panel A of Table 1 summarizes the definitions of the profitability measures examined in this study, which are consistent with prior literature (Fairfield et al. 2009; Novy-Marx 2013; Ball et al. 2015, 2016).

The GP, OP, and CbOP have received considerable attention because of their predictive power in explaining the cross section of stock returns (Novy-Marx 2013; Ball et al. 2015, 2016; Fama and French 2015, 2016, 2017; Akbas et al. 2017). Novy-Marx (2013) find that GP can explain most earnings-related cross-sectional anomalies in stock returns. Ball et al. (2015), however, show that OP has a much stronger link with stock returns than GP. The usefulness of OP in explaining the cross section of stock returns has led to its inclusion as a new factor in the latest five-factor asset pricing model (Fama and French 2015, 2016, 2017). Adding to the success of OP, Ball et al. (2016) show that CbOP outperforms OP in predicting the cross section of stock returns, explaining two anomalies related to accruals and profitability measures that include accruals.

The literature above relates the current profitability to the stock return of the following year. Our interest in the profitability measures comes from their potential for valuation. Because valuation is forward-looking in nature, this study focuses on the forecasts of the measures, rather than their realized current levels.

2.2. QR vs. OLS

We propose constructing point forecasts of profitability using QR, as opposed to the common practice of using OLS regression.[4] Specifically, we focus on the QR for = 0.5 (i.e., the 50th percentile), which is also referred to as the median regression. This special case of QR uses the absolute error loss criterion, as opposed to the squared-error loss criterion upon which OLS regression is based. Median regression has the advantage of being more robust to outliers than OLS regression (Cameron and Trivedi 2005).

Similarly, QR is a more robust alternative for accommodating dependent variables with skewed distributions (Ov2(than)-[(6262125r7b.962)062610.850.160.1sc0.0235Tc-

heavier than the Gaussian distribution (Westfall 2014). All of the profitability measures are leptokurtic. The skewness coefficient column reports the adjusted Fisher-Pearson standardized moment coefficient of skewness. All the earnings and size measures are positively skewed (i.e., skewed to the right—with a longer right tail than the left). With the deflation by some size measures, all the profitability measures are negatively skewed.

3.3. Results of the Forecasting Anal sis

Table 3 presents the forecasting analysis results comparing the alternative approach by economy-wide QR to the benchmark approach by economy-wide OLS regression. We obtain strong evidence showing significantly positive forecast improvements for all the profitability measures. This holds not only for the mean forecast improvements but also for the median. The levels of significance are consistently high (all at the 1% level).

Table 3. Profitability Forecast Improvements of Economy-Wide QR over Economy-Wide OLS Regression

	Value		*p*-value
GP			
Mean	0.138%	***	0.000
Median	0.121%	***	0.000
OP			
Mean	0.063%	***	0.000
Median	0.096%	***	0.000
CbOP_BS			
Mean	0.068%	***	0.000
Median	0.061%	***	0.000
CbOP_CF			
Mean	0.071%	***	0.000
Median	0.070%	***	0.000
RNOA			
Mean	0.120%	***	0.000
Median	0.193%	***	0.000
ROE			
Mean	0.415%	***	0.000
Median	1.492%	***	0.000

Notes. This table reports the profitability forecast improvements of economy-wide QR (the alternative approach) over economy-wide OLS regression (the benchmark approach). The forecast improvement (FI) is measured through a matched-pair comparison of the absolute forecast errors (AFE) from the two competing approaches. A positive FI means the AFE from the benchmark approach is larger than that from the alternative approach. Both forecasting approaches use the same set of predictor variables like those in Fairfield et al. (2009). Regardless of the forecasting approaches, the underlying predictor variable *PREDGSL* (i.e., the predicted growth in sales) is constructed in the same way by industry-specific OLS regression. Industries are defined using the first-digit SIC. Firm-specific forecasts are obtained in two steps. First, the coefficients of a forecasting model are estimated for each year from 1989 to 2018 on a rolling basis using data available in the previous 10 years. Next, the estimated coefficients are applied on a firm's data of the previous year to obtain a firm-specific forecast for the current year. The mean and median FIs of all firm-years in the sample are reported for different profitability measures (see Table 1 for the definitions of the measures). The test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed-rank test.

***Statistical significance at the 1% level; **5% level; *10% level.

We perform a number of additional analyses to ensure that our results are not sensitive to various methodological and sample choices and can extend beyond profitability forecasting. The analyses are summarized in Appendix A in the online appendix.

4. Distributional Shape Anal sis: Research Design

The purpose of the distributional shape analysis is to examine whether, as prior research suggests, the accuracy of QR forecasts relative to OLS forecasts is related to the distributional shape of profitability characterized by its tail-heaviness and asymmetry. We consider both the MAFE and the MSFE criteria in this examination.

Below, we report the research design of the analysis based on data collected from simulated experiments. We introduce two measures of forecasting accuracy for the simulated data and define different measures of tail-heaviness and asymmetry. The simulation procedure is described in Appendix B in the online appendix. The results of this simulated-data analysis and the verification of the key findings using archival data are reported in the next section.

4.1. Forecasting Accurac Measures

Because profitability with a more asymmetric distribution or heavier tails is likely to be harder to forecast, we do not expect QR forecasts to become more accurate in those situations in an absolute sense. Instead, we focus on assessing whether QR forecasts are relatively more accurate than OLS forecasts as the tail-heaviness and asymmetry change, considering both the MAFE and MSFE criteria. To do so, we consider measures that benchmark the forecasting performance of QR under the MAFE criterion against that of OLS under MSFE.

To see why we consider such measures, first note that, when confined to predictions within the training sample, the OLS's mean forecast is by design optimal under the MSFE criterion; similarly, the QR's median forecast is by design optimal under the MAFE criterion. For a hold-out test sample, QR forecasts can be more accurate than OLS forecasts even under the MSFE criterion, and the other way around under the MAFE criterion. Nonetheless, due to the ways these forecasts are designed, we expect them to tend to prevail in out-of-sample testing under the criteria they are optimal for in-sample prediction. Suppose that a forecasting approach performs very competitively even under a criterion unfavorable to it and also

has expectedly superior performance under a criterion favorable to it. An alternative forecasting approach cannot analogously achieve similarly strong performance under the two criteria. Then it is reasonable to consider the former forecasting approach to be relatively more accurate.

More precisely, we look at the statistical test result on the FIs in each simulated experiment. Then, out of the 500 experiments for each distribution type and parameter combination, we count the percentage of the times a forecasting approach prevails under the criterion favorable to it. To determine whether QR prevails in an experiment under the MAFE criterion, we compute the FIs for the 2,500 draws of next-period profitability in the experiment like what we do in the forecasting analysis reported in Table 3. Then we perform a statistical test to see if the mean FI is positive at the 0.01 significance level using the t-test.[9] Similarly, we do this to see whether OLS prevails in an experiment under the MSFE criterion with the forecast improvements (FIs) redefined as the SFE of the QR forecast minus that of the OLS. Counting the results over the 500 experiments, we obtain the following measures for each distribution type and parameter combination:

pct.QR.Prevail	Percentage of the times where quantile regression prevails under the MAFE criterion;
pct.OLS.Prevail	Percentage of the times where OLSprevails under the MSFE criterion.

We also consider the counterparts of these measures by replacing the t-test with the Wilcoxon signed-rank test. This is a test on the median FI. Thus, the counterpart measures are better described as under the median AFE (MdAFE) and median SFE (MdSFE) criteria, respectively. Figure 2 in the online appendix illustrates the empirical cumulative distributions of the p-value of the Wilcoxon (signed-rank) test and the t-test from the 500 experiments for a moderately heavy-tailed, highly skewed stable distribution.

We consider two forecasting accuracy measures that benchmark the performance of QR under the MAFE criterion against that of OLS under MSFE. The *incremental forecasting accuracy* of QR is

$$IncrAccur \quad pct.QR.Prevail \quad pct.OLS.Prevail.$$

Because *pct.OLS.Prevail* represents the prevalence of OLS over QR under the MSFE criterion, the lower this measure, the more competitive the forecasting performance of QR under this criterion unfavorable to it. If QR and OLS do similarly well under the criteria favorable to them, respectively, then *IncrAccur* should be close to zero. If *IncrAccur* increases above zero50.1(z79(t)380

The literature does not have a universally accepted definition of the tails of a distribution. We use Taleb's definition for its simplicity (Taleb 2017, 2018

Table 4. Descriptive Statistics of the Simulated and Archival Data for the Distributional Shape Analysis

Panel A: Simulated data

Variable	Mean	Std. dev.	Min.	Median	Max.
Forecasting accuracy (Wilcoxon-test based):					
pct.QR.Prevail	0.319	0.313	0.012	0.185	0.972
pct.OLS.Prevail	0.033	0.026	0.000	0.026	0.102
IncrAccur	0.286	0.334	0.074	0.154	0.972
RelAccur	2.148	2.333	1.792	2.001	6.879
Forecasting accuracy (t-test based):					
pct.QR.Prevail	0.338	0.323	0.014	0.198	0.982
pct.OLS.Prevail	0.175	0.139	0.000	0.124	0.612
IncrAccur	0.164	0.251	0.106	0.057	0.924
RelAccur	0.518	1.566	1.792	0.342	6.820
Distributional properties (mean-aggregated):					
Mean%Extremes	0.136	0.102	0.000	0.137	0.319
Kurtosis†	3.884	1.897	1.104	4.294	7.247
Tail asymmetry‡	0.002	0.193	0.279	0.046	0.278
Mean-less-median‡	0.006	0.464	0.684	0.114	0.683
Skewness coefficient‡	0.097	1.253	2.451	0.315	2.275
sd(*Pro t.*)†	0.798	1.325	0.224	0.264	8.127
Distributional properties (median-aggregated):					
Mean%Extremes	0.135	0.104	0.000	0.140	0.320
Kurtosis†	3.092	1.635	1.102	3.142	7.204
Tail asymmetry‡	0.003	0.194	0.279	0.000	0.278
Mean-less-median‡	0.009	0.467	0.685	0.143	0.686
Skewness coefficient‡	0.110	1.107	2.320	0.313	2.162
sd(*Pro t.*)†	0.530	0.719	0.198	0.242	4.702

Notes. This panel gives an overview of the 6,751 observations of profitability-industry-years used in the archival-data distributional shape analysis. The sample is constructed from the firm-year observations used in the out-of-sample tests reported in Table 3. A minimum of 20 firms in each industry-year is required to avoid unreliable estimates of the profitability distributional properties. The industry classification is based on two-digit SIC. The forecasting accuracy measures are computed for each profitability measure using the forecasting performance of the QR and OLS approaches for each firm aggregated across all firms in an industry-year based on the MAFE and RMSFE criteria, respectively. The measures of the distributional properties are computed for each profitability measure based on all firms in an industry-year. See panel B of Table 1 for details of the variable definitions.

†Measures in log value; ‡measures in cube-root value.

Panel B: Archival data

Variable	Mean	Std. dev.	Min.	Median	Max.
Size of industry-year	82.2	82.7	20	50	631
Forecasting accuracy:					
r.QR.Prevail	1.027	0.067	0.817	1.016	2.130
r.OLS.Prevail	0.996	0.043	0.571	1.000	1.246
IncrAccur	0.031	0.104	0.422	0.018	1.559
RelAccur	0.298	1.003	4.137	0.181	13.167
Distributional properties:					
Mean%Extremes	0.076	0.236	0.000	0.000	2.174
Kurtosis†	1.552	0.475	0.274	1.521	3.583
Tail asymmetry‡	0.011	0.302	0.523	0.000	0.497
Mean-less-median‡	0.064	0.671	1.142	0.378	1.129
Skewness coefficient‡	0.034	0.896	1.741	0.126	1.568
sd(*Pro t.*)†	1.978	0.386	3.238	1.964	0.824

Notes. This panel gives an overview of the 1,024 observations used in the simulated-data distributional shape analysis based on data from 512,000 simulated experiments (500 experiments for each set of the distribution type and parameter combination over 4 distribution types and 256 parameter combinations). The forecasting accuracy measures are computed based on the Wilcoxon-test- or t-test-based forecasting performance of the QR and OLS approaches in each group of 500 simulated experiments of the 4×256 sets of the distribution type and parameter combination. The measures of the distributional properties are computed based on the 2,500 draws of the simulated next-period firm profitability to be forecast in each simulated experiment. Presented in this panel are these measures mean- or median-aggregated to the distribution type–parameter combination level. See panel B of Table 1 for the definitions of the forecasting accuracy and distributional property measures.

†Variables in log value; ‡variables in cube-root value.

- Distribution type fixed effects = Effects of whether the distribution is positively or negatively skewed stable or IHS;
- = Error term.

Driven by goodness-of-fit considerations, the log values of kurtosis and sd(*Pro t.*) and the cube-root values of the *Asymmetric* measures are used in the regression. The cube-root transformation works much like the log transformation but accepts and maintains negative values (Cox 2011). We control for sd(*Pro t.*) because not all the measures involve the deflation by the sample standard deviation and, even when some do, deflation alone is not likely to remove the influence completely.

Table 5 show the results of the simulated-data regression analysis at the mean-aggregated level for the pooled regressions. Without an exception, the effect of tail-heaviness on the incremental forecasting accuracy of QR is significantly positive across all the combinations of *Heavy* and *Asymmetric* measures and for both the Wilcoxon-test- and t-test-based definitions of *IncrAccur*.

For asymmetry, we focus on the shape of its effect on the incremental forecasting accuracy of QR. The effect has a U-shape with the minimum around $_2/2\ _3$ if the coefficient $_3$ of the *Asymmetric*^{2} term is significantly positive (an inverted-U-shape if significantly negative). The results in the table show that the shape of the asymmetry effect is consistently a U-shape throughout.

The shape of the asymmetry effect is not as consistent throughout Table 6, where the results for the relative forecasting accuracy of QR for the pooled regressions are presented. However, it is still highly consistent when confining to only the Wilcoxon-test- or only the t-test-based results. The asymmetry effect has a U-shape in the former but an inverted-U-shape in the latter. This mixed result is in sharp contrast to the highly consistent significantly positive effect of tail-heaviness in Table 6.

The individual-distribution regression results are presented in Tables A1 and A2 in the online appendix. Regardless of the distributions (stable or IHS) and measures (*IncrAccur* or *RelAccur*), the results are highly consistent with the corresponding pooled-regression results. In an untabulated analysis, we have examined also the median-aggregated versions of the pooled and individual-distribution regressions, and the results are very similar.

The findings above continue to hold in the regression analysis at the experimental level, where the *IncrAccur* or *RelAccur* is regressed on the experimental-level profitability distributional properties with robust standard errors adjusted for clustering by distribution type–parameter combination. The effect of tail-heaviness continues to be significantly positive without an exception. The shape of the asymmetry effect again is typically opposite for the Wilcoxon-test- vs. the t-test-defined *RelAccur*. In the interest of space, we do not tabulate these highly similar results.

In Table A3 (in the online appendix), we report the regression results of the building blocks, *pct.QR.Prevail* and *pct.OLS.Prevail*, of the incremental and relative forecasting accuracy measures defined based on the Wilcoxon (signed-rank) test. Panel A of the table shows the findings for the pooled sample of the stable and the IHS distributions. The breakdown of *IncrAccur* or *RelAccur* into its building blocks reveals that *pct.OLS.Prevail* (i.e., the percentage of the times where OLS prevails under the MSFE criterion) always decreases with the tail-heaviness measures. By contrast, *pct.QR.Prevail* (i.e., the percentage of the times where QR prevails under the MAFE criterion) always increases with the tail-heaviness measures. This supports the notion that heavy-tailed profitability distributions are driving the superior forecasting performance of QR under the MAFE criterion reported in Table 3.

Table A4 (in the online appendix) presents the regression results of *pct.QR.Prevail* and *pct.OLS.Prevail* defined based on the t-test. Panel A of the table again shows that *pct.QR.Prevail* increases with the tail-heaviness measures, whereas *pct.OLS.Prevail* decreases with the measures (except for the insignificant findings when *Asymmetric* is tail asymmetry). Therefore, the effect of tail-heaviness on the building blocks of the incremental and relative forecasting accuracy measures is highly consistent, regardless of the statistical test used to define the measures.

The finding of a U-shape effect of asymmetry on *pct.QR.Prevail* is also highly consistent among the regression results of the Wilcoxon-test- or the t-test-based measure. However, the shape of the asymmetry effect on *pct.OLS.Prevail* is opposite between the regression results reported in panels A of Tables A3 and A4 (inverted-U-shape in the former and U-shape in the latter). The difference again explains why we need both tests to see the not-so-robust effect of asymmetry and the highly robust effect of tail-heaviness.

Panels B and C of Tables A3 and A4 show the findings for the stable and the IHS distribution separately, which are very similar to those for the pooled sample discussed above.

5.2. Regression Anal sis of Archival Data

In the archival data used for the distributional shape analysis, distributional properties are estimated for each profitability measure using all firms in each industry-year. The industry classification is based on the two-digit SIC. The firm-year observations used to construct the industry-year observations come from the sample for out-of-sample testing reported in Table 3.

Table 5. Incremental Forecasting Accuracy and Profitability Distributional Shape: Simulated-Data Analysis at the Mean-Aggregated Level (Pooled Sample of Both Stable and IHS Distributions)

IncrAccur	Wilcoxon-test based						t-test based					
Asymmetric	Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡		Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡	
Heavy	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†
Heavy	1.814***	0.170***	1.233***	0.114***	1.554***	0.143***	1.676***	0.159***	1.521***	0.142***	1.566***	0.148***
	(0.038)	(0.004)	(0.035)	(0.003)	(0.045)	(0.005)	(0.024)	(0.002)	(0.026)	(0.003)	(0.025)	(0.003)
	t = 47.330	t = 44.716	t = 34.963	t = 33.000	t = 34.545	t = 31.211	t = 70.898	t = 66.436	t = 57.674	t = 54.490	t = 61.825	t = 57.012
Asymmetric	0.071	0.049	0.013	0.006	0.003	0.001	0.053	0.033	0.003	0.006	0.004	0.009**
	(0.058)	(0.060)	(0.016)	(0.017)	(0.007)	(0.007)	(0.036)	(0.038)	(0.012)	(0.013)	(0.004)	(0.004)
	t = 1.226	t = 0.814	t = 0.773	t = 0.335	t = 0.461	t = 0.133	t = 1.482	t = 0.872	t = 0.215	t = 0.491	t = 1.139	t = 2.176
Asymmetric^2	7.348***	7.682***	1.449***	1.497***	0.145***	0.151***	1.225***	1.507***	0.309***	0.359***	0.034***	0.037***
	(0.206)	(0.212)	(0.028)	(0.028)	(0.004)	(0.005)	(0.127)	(0.132)	(0.021)	(0.021)	(0.003)	(0.003)
	t = 35.639	t = 36.294	t = 52.260	t = 53.521	t = 32.447	t = 32.095	t = 9.638	t = 11.373	t = 14.925	t = 16.965	t = 13.680	t = 13.914
sd(*Pro t.*)†	0.143***	0.089***	0.137***	0.101***	0.072***	0.025***	0.128***	0.077***	0.128***	0.082***	0.114***	0.066***
	(0.003)	(0.004)	(0.002)	(0.003)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	t = 47.912	t = 25.068	t = 59.618	t = 36.058	t = 19.113	t = 6.366	t = 69.188	t = 34.614	t = 74.381	t = 38.917	t = 53.717	t = 29.158
Intercept	0.453***	1.047***	0.416***	0.811***	0.437***	0.929***	0.303***	0.856***	0.298***	0.793***	0.301***	0.812***
	(0.015)	(0.021)	(0.010)	(0.016)	(0.014)	(0.022)	(0.009)	(0.013)	(0.008)	(0.012)	(0.008)	(0.012)
Profitability FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,024	1,024	1,024	1,024	1,024	1,024	1,024	1,024	1,024	1,024	1,024	1,024
Adjusted R^2	0.893	0.885	0.935	0.931	0.883	0.870	0.928	0.920	0.936	0.930	0.934	0.925

Notes. This table is based on observations from 512,000 simulated experiments mean-aggregated to the distribution type–parameter combination level. In each experiment, 2,500 draws of next-period firm profitability are simulated by applying the intercept and slope parameters and 2,500 independent draws of the error term of a first-order autoregressive model on 2,500 draws of current-period firm profitability, which were simulated from iterative applications of the model on a simulated firm profitability seed. The draws of the error term at different stages and the draw of the seed are independent draws from a stable or IHS distribution with their tail and skewness parameters set to different values. 500 experiments are run for each of the 256 parameter combinations and 4 distribution types (i.e., a positively or negatively skewed stable or IHS distribution). Sample distributional properties of profitability are measured using the 2,500 draws of the simulated next-period firm profitability to be forecast in each experiment. See Appendix B in the online appendix for further details of the simulated experiments. The regression model in this table is $IncrAccur = {}_0 + {}_1 Heavy + {}_2 Asymmetric + {}_3 Asymmetric^2 + {}_4 \text{sd}(Pro\ t.) + \text{Distribution type fixed effects} + \ $, where the following experiment-level explanatory variables are mean-aggregated to the distribution type–parameter combination level: *Heavy* = *Mean%Extremes* or kurtosis†; *Asymmetric* = tail asymmetry,‡ mean-less-median,‡ or skewness coefficient‡; sd(Profit.)† = standard deviation of the sample distribution of profitability. See panel B of Table 1 for further details on the variable definitions. For brevity, the coefficient of the distribution type fixed effect is omitted from all the panels and the coefficients of the intercept and sd(*Pro t.*) are omitted from the panels for individual distributions. Standard errors are reported in parentheses.

†Variables in log value; ‡in cube-root value; ***statistical significance at the 1% level; **5% level; *10% level.

Table 6. Relative Forecasting Accuracy and Profitability Distributional Shape: Simulated-Data Analysis at the Mean-Aggregated Level (Pooled Sample of Both Stable and IHS Distributions)

RelAccur	Wilcoxon-test based						t-test based					
Asymmetric	Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡		Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡	
Heavy	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†
Heavy	13.615***	1.296***	10.253***	0.967***	12.035***	1.128***	9.193***	0.873***	11.151***	1.042***	10.649***	1.024***
	(0.166)	(0.016)	(0.158)	(0.015)	(0.227)	(0.023)	(0.205)	(0.020)	(0.213)	(0.021)	(0.206)	(0.020)
	t = 82.170	t = 78.990	t = 65.073	t = 63.457	t = 53.102	t = 48.115	t = 44.932	t = 43.895	t = 52.400	t = 49.795	t = 51.815	t = 51.322
Asymmetric	0.988***	0.826***	0.357***	0.296***	0.117***	0.083**	0.092	0.019	0.068	0.003	0.002	0.030
	(0.251)	(0.260)	(0.074)	(0.075)	(0.034)	(0.037)	(0.310)	(0.315)	(0.099)	(0.103)	(0.031)	(0.032)
	t = 3.938	t = 3.181	t = 4.852	t = 3.931	t = 3.388	t = 2.237	t = 0.295	t = 0.059	t = 0.688	t = 0.026	t = 0.070	t = 0.958
Asymmetric^2	51.131***	53.219***	9.269***	9.560***	0.974***	1.002***	17.614***	16.148***	4.149***	3.781***	0.482***	0.473***
	(0.891)	(0.911)	(0.124)	(0.124)	(0.023)	(0.024)	(1.100)	(1.104)	(0.167)	(0.170)	(0.020)	(0.020)
	t = 57.376	t = 58.418	t = 74.850	t = 77.324	t = 43.222	t = 41.598	t = 16.008	t = 14.627	t = 24.805	t = 22.274	t = 23.575	t = 23.088
sd(*Pro t.*)†	1.002***	0.588***	0.941***	0.631***	0.493***	0.127***	0.859***	0.581***	0.861***	0.529***	1.064***	0.735***
	(0.013)	(0.015)	(0.010)	(0.012)	(0.019)	(0.021)	(0.016)	(0.019)	(0.014)	(0.017)	(0.017)	(0.017)
	t = 77.527	t = 38.402	t = 91.682	t = 51.048	t = 25.903	t = 6.186	t = 53.822	t = 31.285	t = 62.139	t = 31.146	t = 61.604	t = 42.077
Intercept	2.913***	7.439***	2.582***	5.949***	2.779***	6.671***	1.203***	4.250***	1.209***	4.836***	1.030***	4.561***
	(0.064)	(0.092)	(0.047)	(0.073)	(0.072)	(0.111)	(0.079)	(0.112)	(0.063)	(0.100)	(0.065)	(0.095)
Profitability FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,024	1,024	1,024	1,024	1,024	1,024	1,024	1,024	1,024	1,024	1,024	1,024
Adjusted R^2	0.959	0.956	0.973	0.972	0.939	0.930	0.862	0.857	0.892	0.884	0.889	0.887

Notes. This table is based on observations from 512,000 simulated experiments mean-aggregated to the distribution type–parameter combination level. In each experiment, 2,500 draws of next-period firm profitability are simulated by applying the intercept and slope parameters and 2,500 independent draws of the error term of a first-order autoregressive model on 2,500 draws of current-period firm profitability, which were simulated from iterative applications of the model on a simulated firm profitability seed. The draws of the error term at different stages and the draw of the seed are independent draws from a stable or IHS distribution with their tail and skewness parameters set to different values. 500 experiments are run for each of the 256 parameter combinations and 4 distribution types (i.e., a positively or negatively skewed stable or IHS distribution). Sample distributional properties of profitability are measured using the 2,500 draws of the simulated next-period firm profitability to be forecast in each experiment. See Appendix B in the online appendix for further details of the simulated experiments. The regression model in this table is $RelAccur = {}_0 + {}_1 Heavy + {}_2 Asymmetric + {}_3 Asymmetric^2 + {}_4 \text{sd}(Pro\ t.) + \text{Distribution type fixed effects} + \ $, where the following experiment-level explanatory variables are mean-aggregated to the distribution type–parameter combination level: *Heavy* = *Mean%Extremes* or kurtosis†; *Asymmetric* = tail asymmetry‡, mean-less-median‡, or skewness coefficient‡; sd(*Pro t.*)† = standard deviation of the sample distribution of profitability. See panel B of Table 1 for further details on the variable definitions. For brevity, the coefficient of the distribution type fixed effect is omitted from all the panels and the coefficients of the intercept and sd(*Pro t.*) are omitted from the panels for individual distributions. Standard errors are reported in parentheses.

†Variables in log value; ‡in cube-root value; ***statistical significance at the 1% level; **5% level; *10% level.

A minimum of 20 firms in each industry-year is required to avoid unreliable estimates of the distributional properties.

The regression model is

$$
\begin{aligned}
DepVar \quad & {}_0 + {}_1 Heavy + {}_2 Asymmetric \\
& + {}_3 Asymmetric^2 + {}_4 \mathrm{sd}(Profit.) \\
& + \text{Profitability fixed effects (only} \\
& \quad \text{for the pooled all-profitability regression)} \\
& + \text{First-digit SIC industry fixed effects} \\
& + \text{Year fixed effects} + \varepsilon, \qquad (12)
\end{aligned}
$$

where the *Heavy* and *Asymmetric* measures are the same set as in the simulated-data analysis. The two forecasting measures for *DepVar* are still referred to as *IncrAccur* and *RelAccur*. However, they are redefined as follows for the archival-data analysis:

$$
\begin{aligned}
IncrAccur \quad & fir.QR.Prevail \quad fir.OLS.Prevail, \\
RelAccur \quad & \log fir.QR.Prevail \quad \log fir.OLS.Prevail ,
\end{aligned}
$$

where $r.QR.Prevail = \mathrm{mean}(AFE_{\mathrm{OLS}})/\mathrm{mean}(AFE_{\mathrm{QR}})$ is the forecast improvement ratio (FIR) of QR under the MAFE criterion, and $r.OLS.Prevail = [\mathrm{mean}(SFE_{\mathrm{QR}})/\mathrm{mean}(SFE_{\mathrm{OLS}})]^{1/2}$ is the forecast improvement ratio of OLS under the root mean squared forecast error (RMSFE) criterion. The mean( ) operation in the forecast improvement ratios is taken over all firms in an industry-year. We use the RMSFE criteria, which has the same ranking as MSFE, to define $r.OLS.Prevail$ so that its scale is comparable to $r.QR.Prevail$, and hence, the meaning of *IncrAccur* as their difference is more intuitive.

Industry and year fixed effects are included in the regression. Robust standard errors adjusted for clustering by profitability-industry-year are reported in parentheses in the result tables. Because the observations for each profitability measure are at the industry-year level with the industry classification based on the two-digit SIC, we use the broader first-digit SIC to define the industry for the industry fixed effects and robust standard errors.

Panel B of Table 4 provides the descriptive statistics of the archival data used for the distributional shape analysis. The mean and median sizes of each industry-year are 82.7 and 50 firms, respectively. This variable provides the weights for the size-weighted regressions reported in Table 7, in addition to the unweighted regressions.

RMSFE should be an evaluation criterion more favorable to OLS. However, the mean $r.OLS.Prevail$ is below one (0.996), whereas the mean $r.QR.Prevail$ is above one (1.027). This necessarily results in a positive mean *IncrAccur*, suggesting that on average the forecasting accuracy of QR is higher relative to OLS, just like in the simulated data.

The mean and median of the asymmetry measures are nonzero, also like in the simulated data. Note that the kurtosis reported in the panel and used in the regressions is in log scale. Therefore, its median at 1.521 is equivalent to a value of 4.577 in the original scale. This suggests that over half of the industry-years have profitability distributions with tails heavier than the Gaussian. However, with a minimum at 0.274 for kurtosis in log scale, there should be cases with tails lighter than the Gaussian, which do not exist at all in the simulated data. This could be a reason for expecting results somewhat different from the simulated-data analysis.

Table 7 shows the results of the archival-data analysis at the industry-year level for *IncrAccur* as the dependent variable. In panel A where the results for the pooled all-profitability regressions are reported, the effect of tail-heaviness on the incremental forecasting accuracy of QR is significantly positive across all the combinations of *Heavy* and *Asymmetric* measures, as well as for both the unweighted and size-weighted regressions. This highly consistent result also appears in panel E for the individual-profitability regressions for CbOP_CF (except for the *Mean%Ex-*

mated coefficients for the tail-heaviness effect a
significant, they have a positive sign (except for the
kurtosis-skewness oef ficient co
weighted and size-weighted regr51.41(e)12.8(s)12(s)12(i)13.5(o)17

Table 7. Incremental Forecasting Accuracy and Profitability Distributional Shape: Archival-Data Analysis at the Mean-Aggregated Industry-Year Level with Industry-Year Distributional Properties

Panel A: Pooled sample of all profitability measures

	Unweighted						Size-weighted					
Asymmetric	Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡		Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡	
Heavy	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†	*Mean% Extremes*	Kurtosis†
Heavy	0.025***	0.021***	0.026***	0.021***	0.022***	0.028***	0.026***	0.021***	0.028***	0.021***	0.027***	0.028***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
	t = 4.953	t = 7.698	t = 5.175	t = 7.571	t = 3.695	t = 7.152	t = 5.954	t = 8.533	t = 6.409	t = 8.544	t = 5.394	t = 8.870
Asymmetric	0.019***	0.019***	0.008***	0.008***	0.003**	0.004***	0.017***	0.017***	0.008***	0.008***	0.002*	0.003***
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	t = 4.161	t = 4.148	t = 3.862	t = 3.720	t = 2.229	t = 2.622	t = 4.424	t = 4.586	t = 4.762	t = 4.746	t = 1.918	t = 2.708
Asymmetric^2	0.063***	0.062***	0.019***	0.016***	0.005*	0.008**	0.094***	0.087***	0.023***	0.020***	0.001	0.008***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.00)	(0.00)	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)
	t = 2.867	t = 2.827	t = 3.605	t = 3.007	t = 1.673	t = 2.024	t = 4.704	t = 4.397	t = 5.365	t = 4.639	t = 0.360	t = 2.831
sd(*Pro t.*)†	0.013***	0.019***	0.014***	0.020***	0.015***	0.020***	0.006*	0.002	0.005	0.001	0.005	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	t = 3.210	t = 4.528	t = 3.533	t = 4.819	t = 3.575	t = 4.853	t = 1.819	t = 0.593	t = 1.469	t = 0.230	t = 1.484	t = 0.020
Intercept	0.069***	0.112***	0.069***	0.112***	0.082***	0.125***	0.024**	0.065***	0.024**	0.065***	0.034***	0.082***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Profitability FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,751	6,751	6,751	6,751	6,751	6,751	6,751	6,751	6,751	6,751	6,751	6,751
Adjusted R^2	0.136	0.141	0.136	0.140	0.134	0.139	0.169	0.173	0.170	0.174	0.164	0.170

Notes. The industry-year observations used in this table are constructed from the firm-year observations used in the out-of-sample tests reported in Table 3. A minimum of 20 firms in each industry-year is required to avoid unreliable estimates of the profitability distributional properties. The industry classification is based on the two-digit SIC. The regression model is *IncrAccur* = $_0$ + $_1$ *Heavy* + $_2$ *Asymmetric* + $_3$ *Asymmetric*^{2} + $_4$ sd(*Pro t.*) + Profitability fixed effects (only for the pooled all-profitability regression) + First-digit SIC industry fixed effects + Year fixed effects + , where *IncrAccur* is redefined as the forecast improvement ratio of QR under the MAFE criterion minus the forecast improvement ratio of OLS under the RMSFE; *Heavy* = *Mean% Extremes* or kurtosis†; *Asymmetric* = tail a07.71.22512631838055.1(digit)-243.2.8(effects)-296.4((only)-292.6(for-8907002Tc/F51.1808TL7.9702-7.97020390.1038165.372842m[307d)-297.9m]T-TJ4(less-mediTJ13(,/F111Tf

Table 7. (Continued)

	Unweighted						Size-weighted					
Asymmetric	Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡		Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡	
Heavy	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†
Panel B: *GP*												
	(0.03)	(0.01)	(0.03)	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
Asymmetric	0.065***	0.065***	0.017**	0.019**	0.032***	0.037***	0.089***	0.087***	0.027***	0.027***	0.036***	0.042***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Asymmetric^2	0.136*	0.112	0.037**	0.037**	0.049***	0.087***	0.142**	0.117*	0.027*	0.027*	0.050***	0.079***
	(0.07)	(0.08)	(0.02)	(0.02)	(0.01)	(0.02)	(0.07)	(0.07)	(0.01)	(0.01)	(0.01)	(0.01)
	No	No	No	No	No	No	No	No	No	No	No	No
Adjusted R^2	0.201	0.201	0.196	0.195	0.214	0.220	0.242	0.242	0.223	0.224	0.257	0.264
Panel C: *OP*												
Heavy	0.018*	0.012**	0.014	0.012**	0.004	0.002	0.014	0.010**	0.012	0.010**	0.003	0.000
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Asymmetric	0.006	0.007	0.013***	0.013***	0.006**	0.006**	0.003	0.003	0.008**	0.008**	0.004	0.004
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Asymmetric^2	0.045	0.044	0.019*	0.020*	0.018***	0.018**	0.001	0.006	0.014	0.015	0.013**	0.012**
	(0.05)	(0.05)	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)
FE?	No	No	No	No	No	No	No	No	No	No	No	No
Adjusted R^2	0.148	0.150	0.156	0.159	0.156	0.156	0.249	0.250	0.253	0.255	0.254	0.254
Panel D: *CbOP_BS*												
Heavy	0.016**	0.003	0.015**	0.003	0.021**	0.006	0.014**	0.006	0.014**	0.005	0.021***	0.010**
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
Asymmetric	0.003	0.003	0.004	0.004	0.001	0.001	0.005	0.005	0.001	0.001	0.002	0.002
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Asymmetric^2	0.001	0.001	0.003	0.004	0.005	0.003	0.022	0.020	0.001	0.000	0.005	0.005
	(0.03)	(0.03)	(0.01)	(0.01)	(0.00)	(0.01)	(0.03)	(0.03)	(0.01)	(0.01)	(0.00)	(0.00)
FE?	No	No	No	No	No	No	No	No	No	No	No	No
Adjusted R^2	0.137	0.134	0.139	0.136	0.139	0.134	0.263	0.261	0.262	0.260	0.264	0.262

Note. See the note under panel A for variable descriptions and other details.

	Unweighted						Size-weighted					
Asymmetric	Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡		Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡	
Heavy	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†
Panel E: *CbOP_CF*												
Heavy	0.017**	0.014***	0.016*	0.015***	0.009	0.013**	0.019**	0.013***	0.019**	0.013***	0.020**	0.016***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)
Asymmetric	0.003	0.003	0.001	0.001	0.001	0.001	0.006	0.006	0.001	0.001	0.000	0.001
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)

Table 7. (Continued)

	Unweighted						Size-weighted					
Asymmetric	Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡		Taleb's tail asymmetry‡		Mean-less-median‡		Skewness coefficient‡	
Heavy	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†	*Mean%Extremes*	Kurtosis†
Asymmetric^2	0.004	0.004	0.009	0.012	0.007	0.001	0.026	0.021	0.002	0.004	0.001	0.005
	(0.04)	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.01)	(0.01)	(0.00)	(0.01)
Profitability FE?	No	No	No	No	No	No	No	No	No	No	No	No
Adjusted R^2	0.156	0.161	0.157	0.162	0.158	0.161	0.265	0.268	0.264	0.267	0.264	0.268
Panel F: *RNOA*												
Heavy	0.005	0.016***	0.001	0.014**	0.018	0.003	0.015*	0.017***	0.012	0.015***	0.004	0.004
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Asymmetric	0.017*	0.018**	0.002	0.002	0.005	0.005*	0.020***	0.022***	0.007**	0.008**	0.004	0.004*
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Asymmetric^2	0.149***	0.170***	0.031***	0.035***	0.020***	0.013*	0.120***	0.140***	0.017*	0.021**	0.018***	0.014**
	(0.05)	(0.05)	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)
Profitability FE?	No	No	No	No	No	No	No	No	No	No	No	No
Adjusted R^2	0.200	0.206	0.195	0.199	0.198	0.196	0.270	0.275	0.265	0.269	0.267	0.267
Panel G: *ROE*												
Heavy	0.015	0.021**	0.018	0.020**	0.024*	0.047***	0.016	0.029***	0.021*	0.028***	0.026**	0.059***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Asymmetric	0.118***	0.120***	0.047***	0.048***	0.041***	0.039***	0.112***	0.118***	0.048***	0.049***	0.038***	0.037***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Asymmetric^2	0.041	0.014	0.048***	0.044**	0.010	0.014	0.111	0.053	0.053***	0.043***	0.011	0.016
	(0.08)	(0.08)	(0.02)	(0.02)	(0.01)	(0.01)	(0.08)	(0.08)	(0.02)	(0.02)	(0.01)	(0.01)
Profitability FE?	No	No	No	No	No	No	No	No	No	No	No	No
Adjusted R^2	0.375	0.377	0.382	0.383	0.370	0.376	0.412	0.416	0.421	0.424	0.399	0.411

Note. See the note under panel A for variable descriptions and other details.

The pooled regressions in panel A of Table 7 support the notion of an inverted-U-shape asymmetry effect (with 10 of the 12 estimated coefficients of the $Asymmetric^2$ term being significantly negative). However, this finding appears to be driven by the result for ROE in panel G. Across all the regressions for the other profitability measures, either there is no significant asymmetry effect (for CbOP_BS and CbOP_CF in panels D and E) or any significant finding is consistent with a U-shape asymmetry effect (for GP, OP, and RNOA in panels B, C, and F).

The untabulated results for *RelAccur* as the dependent variable are very similar. Nearly all of the regressions for GP and RNOA and half of those for OP have a significantly positive coefficient of the $Asymmetric^2$ term, whereas the regressions for CbOP_BS and CbOP_CF show no significant effect of asymmetry. As in Table 7, the inverted-U-shape asymmetry effect found in the pooled regressions appears to be driven by the regression results for ROE. Additionally, the pooled and individual regressions for the profitability measures show consistent support for a positive tail-heaviness effect. Overall, the evidence from the archival-data analysis confirms the key insight about the tail-heaviness effect from the simulated-data analysis and highlights again the mostly significant but not entirely consistent effect of asymmetry (i.e., can be U-shape or inverted-U-shape).

6. Application to Cash Flow Forecasting

To demonstrate the usefulness of our analysis framework beyond profitability forecasting, we apply the framework to examine the out-of-sample forecasting of cash flows studied by Nallareddy et al. (2020). They find that, under the MSFE criterion and using the OLS approach, the first-order autoregressive model (i.e., using lagged cash flows to forecast cash flows) is more accurate than the forecasting-by-lagged-earnings model (i.e., using lagged earnings to forecast cash flows).

Following Nallareddy et al. (2020), we examine the out-of-sample forecasts of cash flows for the period from 1990 to 2015. We are interested to relate together the annual time series of the cash flow distributional properties and the incremental forecasting accuracy of the QR approach against OLS. Prior research mentions that the cash flow distribution has changed significantly over time (Gassen 2018). In an untabulated analysis, we find a moderate upward trend in the yearly variation in the tail-heaviness of the cash flow distribution across all firms: an OLS regression of the tail-heaviness, measured as kurtosis in log scale, on the year gives a slope coefficient of 0.027 (with a p-value of 0.053).

We compare the QR approach to estimating the first-order autoregressive cash flow forecasting model against the OLS approach. Note that the forecasting-by-lagged-earnings model does not fit into the simple/extended first-order autoregressive structure upon which our analysis framework was developed. Therefore, we do not expect that the QR approach would prevail for this second model or that the (perhaps nonpositive) incremental forecasting accuracy would be associated with the distributional properties of cash flows. Nonetheless, we are interested to know whether to some extent the key insights of our framework might hold after controlling for the cross-sectional variability of the lagged earnings in the distributional shape analysis. Controlling for the variability of this only predictor variable of the second model is important because the variability is likely to adversely impact the forecasting accuracy of both the QR and the OLS approach perhaps unevenly.

We obtain the data of U.S. firms from the Compustat North America annual fundamentals file on WRDS. Consistent with Nallareddy et al. (2020), cash flows (*CF*) are measured as cash flows from operations adjusted for extraordinary items and discontinued operations (derived from cash flow statements). Earnings (*EARN*) are defined as income before extraordinary items and discontinued operations. Both variables are deflated by average total assets. Following them, we exclude observations meeting any of the following criteria: (i) sales of less than \$10 million; (ii) share price of less than \$1; (iii) SIC code in the range of 6000–6999 (i.e., in the financial services sector).[13] This would yield a sample of 110,597 firm-year observations if we also followed them to winsorize all continuous independent variables of the full sample at the 1% and 99% levels. Instead, we mitigate the effects of outliers only at the in-sample estimation stage to finalize the sample used for the regression with a given rolling window of data (e.g., the most recent two years of available data as in Nallareddy et al. (2020)). This alternative approach avoids a look-ahead bias. We truncate the top and the bottom one percent of all continuous variables used in the in-sample regression, rather than winsorize them, to be consistent with the literature our profitability forecasting analysis builds upon. This prevents the clustering of observations around the 1% and 99% levels. To avoid a look-ahead bias, there is no truncation on the sample of the prior-year data for constructing the out-of-sample forecasts and on the sample of the forecasts constructed.

Figure 4 in the online appendix depicts the annual time series of the incremental forecasting accuracy, its forecast improvement ratio components, and the distributional properties of cash flows. The temporal variation of the incremental forecasting accuracy (*IncrAccur*) of the QR approach (against OLS) for the first-order autoregressive cash flow model is shown

in the first chart of the figure. The components of *IncrAccur*, namely, the forecast improvement ratio of QR under the MAFE criterion (*r.QR.Prevail*) and the forecast improvement ratio of OLS under the RMSFE criterion (*r.OLS.Prevail*), are depicted in the second and the third charts of the figure, respectively. Note that *r.QR.Prevail* is above one for nearly all the years, whereas *r.OLS.Prevail* is more evenly spread above and below one. In other words, the QR approach clearly prevails under the MAFE criterion but the OLS approach on average cannot prevail even under the RMSFE criterion more favorable to it. Consequently, *IncrAccur* is positive for most of the years.

The fourth chart in the figure depicts the kurtosis of the cash flow distribution, which shows a moderate upward trend. The skewness coeffia 9 8 . 5 (s 4) 3 0 1 . (t o) - 1 . 1 5 (f t 9 8 . p i c t) 9 u r e 4 2 9 9 2 9 i E

Table 8. Incremental Forecasting Accuracy and Cash Flow Distributional Shape: QR Forecasting vs. OLS

Panel A: Lagged cash flows or lagged earnings as the predictor variable, with short to long in-sample estimation windows

Window (years)	2			4			7		10	
Predictor	CF_{t-1}	CF_{t-1}	$EARN_{t-1}$	CF_{t-1}	CF_{t-1}	$EARN_{t-1}$	CF_{t-1}	CF_{t-1}	CF_{t-1}	CF_{t-1}
Intercept	0.009**	0.088**	0.164***	0.011*	0.100*	0.161***	0.014**	0.066	0.016***	0.053
	(0.004)	(0.038)	(0.041)	(0.006)	(0.049)	(0.048)	(0.006)	(0.048)	(0.005)	(0.041)
Heavy		0.016**	0.007**		0.022***	0.011**		0.011***		0.006*
		(0.006)	(0.003)		(0.005)	(0.004)		(0.003)		(0.003)
Asymmetric		0.001	0.008***		0.006	0.002		0.005		0.004
		(0.006)	(0.002)		(0.006)	(0.004)		(0.005)		(0.005)
Asymmetric^2		0.027**	0.013**		0.026***	0.013**		0.011**		0.003
		(0.011)	(0.006)		(0.006)	(0.005)		(0.005)		(0.006)
sd(*CF*)		0.677**	1.715***		0.62	1.618***		0.465		0.444
		(0.297)	(0.418)		(0.407)	(0.479)		(0.422)		(0.357)
sd($EARN_{t-1}$)			0.315***			0.343***				
			(0.028)			(0.030)				
Observations	26	26	26	26	26	26	26	26	26	26
Adjusted R^2	0	0.232	0.743	0	0.23	0.821	0	0.063	0	0.076

Notes. This panel presents the results of the cash flow distributional shape analysis for the full sample of U.S. firms based on out-of-sample forecasts from 1990 to 2015 (with in-sample estimation data as far back as in 1987). The yearly observations used in this table are constructed from the firm-year observations used for forecasting cash flows out-of-sample with a rolling window of in-sample estimation. Each yearly observation is based on the distributional properties of cash flows, or lagged earnings, for the cross section of firms in a given year and the incremental forecasting accuracy of the QR approach (vs. OLS) to forecasting cash flows for this cross section. The forecasting model used for comparing the QR approach to OLS has the lagged cash flows or the lagged earnings as the only predictor variable (see Nallareddy et al. 2020). The dependent variable of the distributional shape analysis in this table is the incremental forecasting accuracy *IncrAccur* computed yearly for a given forecasting model. *IncrAccur* is defined as the forecast improvement ratio of QR under the MAFE criterion minus the forecast improvement ratio of OLS under the RMSFE criterion. The independent variables in this table are: *Heavy* = kurtosis† of the cash flow distribution of a year; *Asymmetric* = skewness coefficient‡ of the cash flow distribution of a year; sd(*CF*) = standard deviation of the cash flow distribution of a year; sd($EARN_{t-1}$) = standard deviation of the lagged earnings distribution of a year. Cash flows (*CF*) are defined as net cash flow from operating activities less cash flow from extraordinary items and discontinued operations (Compustat: OANCF − XIDOC). Earnings (*EARN*) are defined as income before extraordinary items and discontinued operations (Compustat: IB). Both variables are deflated by total assets (Compustat: AT) averaged over the current and the prior years. See Table 1 for the details of the definitions of *IncrAccur*, kurtosis, and skewness coefficient.

†Variables in log value; ‡in cube-root value; ***statistical significance at the 1% level; **5% level; *10% level.

Panel B: Subsamples for various exclusion criteria (lagged cash flows as the predictor variable and two-year in-sample estimation window)

Subsample excluding	Intangible-intensive firms		Loss firms		Smaller firms		Size-tails firms	
Intercept	0.005**	0.055	0.001	0.196***	0.008**	0.098***	0.009***	0.047*
	(0.002)	(0.040)	(0.002)	(0.055)	(0.003)	(0.031)	(0.003)	(0.023)
Heavy		0.018		0.004		0.010**		0.011***
		(0.012)		(0.014)		(0.004)		(0.003)
Asymmetric		0.003		0.012**		0.008*		0.0002
		(0.007)		(0.005)		(0.004)		(0.005)
Asymmetric^2		0.0005		0.002		0.020*		0.017***
		(0.010)		(0.020)		(0.011)		(0.005)
sd(*CF*)		0.162		1.906***		0.908***		0.345*
		(0.395)		(0.604)		(0.267)		(0.191)
Observations	26	26	26	26	26	26	26	26
Adjusted R^2	0	0.036	0	0.114	0	0.429	0	0.117

Notes. This panel presents the results of the cash flow distributional shape analysis for the subsamples excluding the following firms one at a time: *intangible-intensive firms* defined as the firms in the health, business equipment, telecommunication, and chemical sectors of the Fama-French 12-industry classification, *loss firms* defined as those with negative earnings (*EARN* < 0), *smaller firms* defined as those below the first quartile of the firm size distribution (where firm size is measured by total assets), and *size-tails firms* defined as those outside the 12.5th and the 87.5th percentile of the firm size distribution. The dependent variable in this table is the incremental forecasting accuracy *IncrAccur* computed yearly for the forecasting model with the lagged cash flows as the only predictor variable. See the note below panel A for other details.

Panel B of Table 8 shows the subsample findings based on the first-order autoregressive cash flow model with a two-year in-sample estimation window. Compared with the full-sample result (columns 1 and 2 of panel A), the magnitude or statistical significance of the mean *IncrAccur* and of the estimated coefficients of the *Heavy* and the *Asymmetric*^{2} terms is generally lower. The few exceptions are the statistically more significant mean *IncrAccur* of the same magnitude in column 7 and the statistically more significant coefficients of the *Heavy* and the *Asymmetric*^{2} terms in column 8 (but both coefficients are lower in magnitude).

All in all, we conclude that the results of the distributional shape analysis for the cash flow forecasting models and for the various subsamples are in line with our earlier findings for profitability forecasting.

7. Conclusion

We document that QR performs better than OLS in forecasting profitability for a range of profitability measures under the MAFE criterion. Considering the MAFE and the MSFE (RMSFE) criteria together, we also examine how QR's forecasting performance, benchmarked against OLS's, changes with the shape of the profitability distribution. Specifically, we perform a distributional shape analysis to relate the forecasting accuracy of QR against OLS to the tail-heaviness and asymmetry of profitability distribution. In the simulated-data analysis of this analysis, we find a robust positive effect of tail-heaviness on the accuracy of QR relative to OLS. The finding is strongly to moderately supported by the archival-data results of the pooled and individual profitability (unweighted and size-weighted) regressions.

In the simulated-data analysis, we also find that asymmetry has either a U-shape or inverted-U-shape effect on the accuracy of QR forecasts. Which of these holds depends on (i) whether Wilcoxon-test- or t-test-based evidence is relied upon to determine the prevalence of a forecasting approach under a given evaluation criterion (MAFE or MSFE) and (ii) whether the accuracy measure is the incremental or the relative forecasting accuracy. The archival-data analysis also shows mixed evidence: the effect of asymmetry is mostly significant but not entirely consistent (i.e., can be U-shaped or inverted-U-shaped).

Applying the distributional shape analysis framework to cash flow forecasting, we demonstrate the usefulness of the framework beyond profitability forecasting. The empirical results support the notion of an inverted-U-shape effect of asymmetry and provide additional evidence on the positive effect of tail-heaviness.

In this study, we have only scratched the surface of QR's usefulness by focusing on the median regression as its special case. QR in general can produce optimal estimates/forecasts for asymmetric loss functions (when $\neq 0.5$). Prior research has argued that financial analysts have an asymmetric loss function (Clatworthy et al. 2011). If they do, would they find formulating their forecasts based on QR with $\neq 0.5$ more aligned with their forecasting objective? What is the implied that can be inferred from analyst earnings forecasts? Are the implied 's similar across different types of analyst forecasts (cash flow forecasts, revenue forecasts, etc.)? These are interesting questions left for future research to answer.

Acknowledgments

The authors are indebted to the anonymous associate editor and referees for their valuable comments and suggestions. They thank Yachang Zeng and participants of the 2017 Australasian Finance and Banking Conference (AFBC) at Sydney, Australia; and 2017 European Accounting Association (EAA) Annual Congress at Valencia, Spain for their helpful comments and suggestions. Hui Tian thanks Nottingham University Business School, United Kingdom; and the University of Bath, United Kingdom, where parts of this work were carried out during 2014–2016 and 2016–2018. All remaining errors are the authors' own.

Endnotes

[1] List of largest corporate profits and losses, 2019. Wikipedia, https://en.wikipedia.org/wiki/List_of_largest_corporate_ profits_and_losses (accessed August 10, 2019).

[2] The applications in finance that we are aware of include return forecasting, portfolio analysis, and risk measurement (Lauridsen 2000, Bassett and Chen 2001, Pohlman and Ma 2010). Recent applications in accounting include forecasting risk in earnings (Konstantinidi and Pope 2016).

[3] Forecasting earnings in practice is often equivalent to forecasting profitability (e.g., Li 2011, Chang et al. 2020). Data samples used to forecast earnings typically include firms of different sizes. Deflation is a technique to control for the size differences.

[7] A firm-year observation's SIC code is identifiable if its value is not missing or otherwise may be imputed based on the nonmissing SIC code of the firm in the nearest future year.

[8] The U.S. Postal Service category comprises all establishments of the U.S. Postal Service as an agency of the executive branch of the U.S. federal government responsible for providing postal service in the United States. The public administration category contains the executive, legislative, judicial, administrative, and regulatory activities of federal, state, local, and international governments.

[9] We have considered also the 0.05 significance level, and the findings are highly similar.

[10] To be precise, in defining *RelAccur*, we set *pct.QR.Prevail* and *pct.OLS.Prevail* to 0.5/500 = 0.001 whenever they have a zero value. Note that, for any given setup of M experiments ($M = 500$ in our case), the lowest nonzero value of *pct.QR.Prevail* and *pct.OLS.Prevail* is $1/M$. So the adjustment above avoids any undefined/infinite value due to the log transformation while maintaining the intended ranking of the *RelAccur* measure.

[11] Nassim N. Taleb, Distinguished Professor of Risk Engineering at the New York University Tandon School of Engineering and the author of the best seller *The Black Swan: The Impact of the Highly Improbable*, defines the fat tails of a perturbed Gaussian distribution to start from the mean minus and plus approximately 2.136 times the standard deviation.

[12] We also have considered two additional asymmetry measures and one additional tail-heaviness measure explained in Appendix C in the online appendix. The inclusion of these measures does not change the highly consistent findings of the tail-heaviness effect. In the interest of space, we omit these measures from the reported tables.

[13] If the SIC code of a firm-year observation is missing, we impute the value based on the nonmissing SIC code of the firm in the nearest future year.

[14] This means that for the 10-year window case, only three years of available cash flows data (i.e., from 1987 to 1989) are used in the in-sample estimation for constructing the 1990 forecasts; only four years of available data (i.e., from 1987 to 1990) are used for constructing the 1991 forecasts; and so forth. See also the explanation in footnote 6.

[15] We downloaded the definition of the Fama-French 12-industry classification on March 31, 2020 from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_12_ind_port.html.

References

Akbas F, Jiang C, Koch PD (2017) The trend in firm profitability and the cross section of stock returns. *Accounting Rev.* 92(5): 1–32.

Allen DE, Salim HM (2005) Forecasting profitability and earnings: A study of the UK market (1982–2000). *Appl. Econom.* 37(17): 2009–2018.

Arnold BC, Groeneveld RA (1992) Skewness and kurtosis orderings: An introduction. Shaked M, Tong YL, eds. *Stochastic Inequalities: Papers from the AMS-IMS-SIAM Joint Summer Research Conference Held in Seattle, Washington, July 1991* (Institute of Mathematical Statistics, Hayward, CA), 17–24.

Ball R, Gerakos J, Linnainmaa JT, Nikolaev VV (2015) Deflating profitability. *J. Financial Econnom.* 117(2):225–248.

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

Bassett GW Jr, Chen H-L (2001) Portfolio style: Return-based attribution using quantile regression. *Empirical Econom.* 26:293–305.

Basu S (1997) The conservatism principle and the asymmetric timeliness of earnings. *J. Accounting Econom.* 24(1):3–37.

Basu S, Markov S (2004) Loss function assumptions in rational expectations tests on financial analysts' earnings forecasts. *J. Accounting Econom.* 38:171–203.

Cameron AC, Trivedi PK (2005) *Microeconometrics: Methods and Applications* (Cambridge University Press, Cambridge, UK).

Carhart MM (1997) On persistence in mutual fund performance. *J. Finance* 52(1):57–82.

Chang WJ, Monahan S, Ouazad A, Vasvari F (2020) The higher moments of future earnings. *Accounting Rev.* Forthcoming.

Chen K, Ying Z, Zhang H, Zhao L (2008) Analysis of least absolute deviation. *Biometrika* 95(1):107–122.

Clatworthy MA, Peel DA, Pope PF (2011) Are analysts' loss functions asymmetric? *J. Forecasting* 31:736–756.

Courgeau D (2012) *Probability and Social Science: Methodological Relationships Between the Two Approaches* (Springer Science & Business Media, Berlin, Heidelberg, Germany).

Cox N (2010) Speaking Stata: The limits of sample skewness and kurtosis. *Stata J.* 10:482–495.

Cox N (2011) Stata tip 96: Cube roots. *Stata J.* 11:149–154.

Diebold F (2015) Forecasting. *Economics, Business, Finance and Beyond* Accessed July 29, 2020, https://www.sas.upenn.edu/~fdiebold/Teaching221/Forecasting.pdf.

Diebold FX, Mariano RS (1995) Comparing predictive accuracy. *J. Bus. Econom. Statist.* 13:253–263.

Dielman TE (1986) A comparison of forecasts from least absolute value and least squares regression. *J. Forecasting* 5(3):189–195.

Dielman TE, Rose EL (1994) Forecasting in least absolute value regression with autocorrelated errors: a small-sample study. *Internat. J. Forecasting* 10(4):539–547.

Doane DP, Seward LE (2011) Measuring skewness: A forgotten statistic? *J. Statist. Education* 19(2):1–8

Evans ME, Njoroge K, Yong KO (2017) An examination of the statistical significance and economic relevance of profitability and earnings forecasts from models and analysts. *Contemporary Accounting Res.* 34(3):1453

Konstantinidi T, Pope PF (2016) Forecasting risk in earnings. *Contemporary Accounting Res.* 33(2):487–525.

Lauridsen S (2000) Estimation of value at risk by extreme value methods. *Extremes* 3:107–144.

Li KK (2011) How well do investors understand loss persistence? *Rev. Accounting Stud.* 16:630–667.

Li KK, Mohanram P (2014) Evaluating cross-sectional forecasting models for implied cost of capital. *Rev. Accounting Stud.* 19: 1152–1185.

Li N, Richardson S, Tuna İ (2014) Macro to micro: Country exposures, firm fundamentals and stock returns. *J. Accounting Econom.* 58(1):1–20.

McDonald JB, Turley P (2011) Distributional characteristics: Just a few more moments. *Amer. Statist.* 65(2):96–103.

Mitra A (1987) Minimum absolute deviation estimation in regression analysis. *Comput. Indust. Engrg.* 12(1):67–76.

Myers JL, Well AD, Lorch RF Jr (2010) *Research Design and Statistical Analysis*. 3rd ed. (Routledge, Abingdon-on-Thames, UK).

Nallareddy S, Sethuraman M, Venkatachalam M (2020) Changes in accrual properties and operating environment: Implications for cash flow predictability. *J. Accounting Econom.* 69(2–3):101313.

Nolan JP (1998) Parameterizations and modes of stable distributions. *Statist. Probab. Lett.* 38:187–195.

Nolan JP (2013) Financial modeling with heavy-tailed stable distributions. *Wiley Interdisciplinary Rev. Comput. Statist.* 6(1):45–55.

Nolan JP (2019) *Stable Distributions: Models for Heavy-Tailed Data* (Birkhauser, Basel, Switzerland).

Olsen CS, Clark AE, Thomas AM, Cook LJ (2012) Comparing least-squares and quantile regression approaches to analyzing median hospital charges. *Acad. Emergency Medicine* 19(7):866–875.

Parzen E (1979) Nonparametric statistical data modeling. *J. Amer. Statist. Assoc.* 74:105–121.

Pohlman L, Ma L (2010) Return forecasting by quantile regression. *J. Investing* 19(4):116–121.

Schröder D, Yim A (2018) Industry effects in firm and segment profitability forecasting. *Contemporary Accounting Res.* 35(4): 2106–2130

Sloan RG (1996) Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Rev.* 71(3): 289–315.

Taleb NN (2017) Probability, Risk, and Extremes (Original Lecture Title: Extreme Events and How to Live with Them), Needham D, Weitzdorfer J., eds. *Extremes (2017 Darwin College Lectures)* (Cambridge University Press, Cambridge, UK), 46–66.

Taleb NN (2018) Overview of fat tails, level 1, finite moments. *The Statistical Consequences of Fat Tails: Real World Preasymptotics, Epistemology and Applications* (STEM Academic Press), 65–87.

Westfall PH (2014) Kurtosis as Peakedness, 1905 – 2014. R.I.P. *Amer. Statist.* 68(3):191–195.

Wilcox RR, Rousselet GA (2018) A guide to robust statistical methods in neuroscience. *Current Protocols Neuroscience* 82(1): 8.42.1–8.42.30.

Yu K, Lu Z, Stander J (2003) Quantile regression: Applications and current research areas. *Statistician* 52(3):331–350.