What Can Analysts Learn from Artificial Intelligence about Fundamental Analysis?

Oliver Binz
INSEAD
oliver.binz@insead.edu

Katherine Schipper
Duke University
katherine.schipper@duke.edu

Kevin Standridge
Duke University
kevin.standridge@duke.edu

This version: October 2022

We thank Dan Bens, Qi Chen, Wilbur Chen (discussant), Stephanie Cheng, Scott Dyreng, Mory Elsaify, John Graham, Luzi Hail, Robert Hills, Xu Jiang, Peter Joos, Thomas Keusch, Matt Kubic, Wayne Landsman, Charles Lee, Bill Mayew, Suresh Nallareddy, Gans Narayanamoorthy, Stephen Penman, James Pinnington, Peter Pope, Ayung Tseng, Teri Yohn (the editor), Haifeng You (discussant), two anonymous reviewers, and workshop participants at Duke University, INSEAD, Tulane University, the Institute for Quantitative Investment Research (INQUIRE), the 2018 AAA/Deloitte/J. Michael Cook Doctoral Consortium, the 2019 EAA Annual Meeting, the 2021 Hawaii Accounting Research Conference, the Accounting Design Project and the 2021 University of Miami Winter Research Conference on Machine Learning and Business for helpful comments and suggestions.

What Can Analysts Learn from Artificial Intelligence about Fundamental Analysis?

Abstract

We use machine learning to estimate Nissim and Penman's (2001) (NP) structural framework that decomposes profitability into four levels of increasingly disaggregated profitability drivers. Our analysis has two distinct features: first we apply machine learning to accommodate the nonlinearities that precluded NP from estimating their framework, and second we analyze the financial statement design choices discussed but not analyzed in NP to provide insights for the teaching and practice of fundamental analysis. We find that out-of-sample profitability forecasts obtained by applying machine learning to NP's framework are more accurate than those derived from a random walk and linear estimation, and that investing strategies based on intrinsic values generated from our profitability forecasts yield risk-adjusted returns. With respect to insights for fundamental analysis, we find that focusing on operating activities, core items and five-year-horizon forecasts improves performance while using a long time series of past information impairs performance. We find mixed evidence of benefits from increasingly granular disaggregation of profitability. The benefits of greater model complexity and nonlinear estimation are pronounced for firms with extreme profitability levels and during the beginning and the end of firms' lifecycles.

JEL Classification: C53, G10, M41
Keywords: Financial Statement Analysis, Machine Learning, Earnings Forecasting

1. Introduction

We combine the capabilities of machine learning, a subfield of artificial intelligence, with Nissim and Penman's (hereafter NP, 2001) hierarchical approach to financial statement analysis to estimate and analyze their nonlinear structural model of accounting profitability.[1] The hierarchical-modeling approach to profitability analysis is foundational in, for example, Penman (2012), Palepu and Healy (2012), Wahlen, Baginski, and Bradshaw (2018), Yohn (2020), and Sommers, Easton, and Drake (2021). We first use machine learning to estimate NP's framework and forecast profitability, then use the forecasts to estimate intrinsic values and, finally, compare those values to market prices. To provide insights for the teaching and practice of fundamental analysis, we analyze the effects of financial statement analysis design choices discussed but not analyzed in NP and provide descriptive analyses of firm characteristics associated with better accuracy of profitability forecasts from nonlinear, machine learning estimation of NP's model as compared to random walk forecasts and forecasts from linear (OLS) estimation.

In using machine learning to estimate complex nonlinear relations between the predictors in NP's structural model and future earnings, we start with a theoretically grounded, and therefore intentionally restricted, information set and apply machine learning to find the functional form that uses this information most effectively. While restricting the information set might sacrifice an unknown amount of predictive ability, such a structural approach 1) allows us to provide insights into the choices analysts and equity investors must make in applying NP's framework and 2) can outperform in out-of-sample tests data mining approaches applied without a framework to very large sets of predictors (Bertomeu, Cheynel, Liao, and Milone 2021; Liu 2021). Regardless of

[1] Using machine learning algorithms requires a number of design choices. We describe our design choices in an online appendix.

whether theoretically grounded models such as NP's do, or do not, perform out-of-sample as well as or better than data-mining models, the advantages of theoretical models that incorporate the firm's underlying profitability structure are numerous, including interpretability (Nissim 2021), a reduced risk of overlooking value-relevant information in distorted accounting numbers (Sloan 2019), and a degree of protection from fluctuations in accounting numbers induced by the reporting process (Penman 2010, Chapters 4 & 5). Furthermore, the theoretical structure facilitates systematic analysis of the effects of variation in design choices, for example, discarding versus including transitory/non-core items and the granularity of profitability driver disaggregation.

We first confirm many of NP's findings for their sample period (1963–1999) for our longer sample period (1963–2019). We show that current-period ratios are individually and interactively associated with current and future profitability in nonlinear ways, including S-shaped, U-shaped, and concave patterns. These relations are visible at least 10 years into the future and attenuate over time. Having confirmed the nonlinearities discussed by NP for our sample, we use neural networks to estimate the nonlinear relations of NP's framework and forecast future profitability.

Building on Gerakos and Gramacy (2013) and Li and Mohanram (2014), we benchmark 1- to 10-year-ahead out-of-sample profitability forecasts against forecasts derived from a random walk and linear (OLS) estimation, and generally find the neural-network-based predictions are more accurate. Median absolute ROCE forecast errors of neural network models are approximately 1.2% (10%) lower than those of random walk (OLS) forecasts, a substantial difference as compared to improvements documented in prior research (Fairfield, Sweeney, and Yohn 1996; Fairfield and Yohn 2001; Esplin, Hewitt, Plumlee, and Yohn 2014). This finding is significant given Gerakos and Gramacy's (2013) and Li and Mohanram's (2014) results that a simple random walk outperforms linear models including multiple predictors and Monahan's (2018) related analysis.[2] We document that forecast accuracy improves with greater disaggregation for RNOA forecasts and with a focus on core items for ROCE forecasts. In additional analysis we find that, relative to linear estimation, accommodating nonlinearities improves accuracy the most for shorter-horizon forecasts based on higher levels of disaggregation. Cross-sectional analyses of forecast accuracy support two inferences. First, the benefits of nonlinear estimation are pervasive, not concentrated in subsamples of firms for which forecasting is particularly difficult. Second, the benefits of higher model complexity (i.e., greater disaggregation, focusing on core items, and using more historical information) and nonlinear estimation are pronounced for firms with extreme profitability levels and during the beginning and the end of firms' lifecycles.

Next, taking the perspective of an equity investor who makes investment decisions using financial statement analysis and strives to maximize risk-adjusted returns, we analyze whether hypothetical trading based on intrinsic value estimates derived from our neural network forecasts generates risk-adjusted returns. Black (1986) argues that noise trading causes short-run divergence while information-based trading causes long-run convergence between prices and value. The more price diverges from value, the larger the reward for information-based trading; the implication is that profitable trades are based on models that produce a less noisy estimate of value and help detect divergence of value from price. We use alpha to measure risk-adjusted returns so that we can compare models of different scales, an important consideration in evaluating the relative performance of RNOA-based and ROCE-based models. We estimate alpha for each of 192

[2] Monahan (2018) describes the finding that a simple random walk outperforms linear models including multiple predictors as "a provocative result because it leads to the seemingly absurd conclusion that, within the context of forecasting earnings, there is no value to peer analysis, trend analysis and using conditioning information" (p. 146). Further, "the random-walk model is inconsistent with standard economic assumptions, accounting practice and the manner in which financial statement analysis is practiced and taught … if the random-walk model is the best academics can do, the relevance of the entire literature on forecasting and financial statement analysis is called into question" (p. 205).

valuation models derived from combinations of five financial statement analysis design choices in NP's framework. Hedge portfolios formed based on the value-to-price ratio obtained from neural network estimation of NP's framework achieve an alpha of up to 9.89% annually, which compares favorably to excess return estimates reported in previous research (e.g., Piotroski 2000). We find that alphas increase with a focus on operating activities and are highest for a forecast horizon of five years.

To address the possibility that positions taken by the best-performing models concentrate in small, illiquid, costly-to-trade stocks, we use only NYSE/AMEX firms and value-weighted portfolio returns. Further, a plot of cumulative abnormal buy-and-hold returns for the short and long portfolios shows that hedge portfolio performance derives from a widening spread, not from excessively negative performance of the short portfolio. This evidence indicates the returns we document are not linked to features of the trading environment or to short positions.

Our research contributes to two literatures. The first is the structural earnings forecasting literature that tests whether forecasting models grounded in accounting-based valuation models such as NP's yield more accurate out-of-sample forecasts as compared to simpler, possibly atheoretical models. Fairfield et al. (1996) find that the income statement disaggregation scheme prescribed by the accounting profession (i.e., disaggregating net income into operating income, non-operating income, income taxes, special items, and income from discontinued operations) improves forecast accuracy over simpler models. Similarly, examining individual ratio disaggregation levels in NP's framework, Fairfield and Yohn (2001), Soliman (2008), and Esplin et al. (2014) find that disaggregating RNOA into turnovers and margins and disaggregating ROE into RNOA and the financial leverage effect improves earnings prediction. These researchers did not estimate NP's model (or another nonlinear structural model). Instead, they analyzed portions of the model or applied linear approximations or both; the existence of this research speaks to the importance both of NP's structural model as a framework for analyzing profitability and of estimating and analyzing the model holistically, including its nonlinear structure.

We contribute to the structural earnings forecasting literature in three related and distinct ways. First, as argued and demonstrated by NP, evaluating their model necessitates accommodating its essential nonlinear structure. Given the technology available at the time, previous research evaluating NP's model approximates these nonlinearities via linear estimation. By using machine learning to estimate NP's model, we extend prior research, document substantial improvements over linear estimators, and highlight both the parts of NP's model and the types of firms for which consideration of nonlinearities matters most. Second, while prior research uses linear estimation to evaluate components of NP's model, we evaluate the model as a whole. This holistic analysis is essential for deriving implications for financial statement analysis practice by linking performance improvements to variation in specific financial statement analysis design choices, holding other design choices constant. Third, to the best of our knowledge, we are the first to examine how focusing on operating activities and core items, extending the forecast horizon, and using more past information affect forecasting performance within NP's framework.

The second literature we contribute to tests whether machine learning techniques can be used to increase earnings forecast accuracy, and if so, how best to exploit these techniques in varying contexts. While Callen, Kwan, Yip, and Yuan (1996) fail to find evidence that machine learning improves firm-level time-series earnings forecasting models, more recent research, for example, Gerakos and Gramacy (2013), Hunt, Myers, and Myers (2019), Anand, Brunner, Ikegwu, and Sougiannis (2020), van Binsbergen, Han, and Lopez-Lira (2020), Cao and You (2020), and Chen, Cho, Dou, and Lev (2022) documents improvements for panel data models. The predictor selection in these papers is purely empirical and atheoretical, a data-mining-style approach that NP describe as "trawling through the data without structure" (p. 125). These papers exploit the strengths of machine learning for processing high-dimensionality data sets and arbitrary nonlinearities among very large sets of predictors while avoiding (or at least mitigating) overfitting by applying regularization. They do not, by design, estimate a structural model such as NP's.

From the perspective of providing insights for teaching and practical applications, a purely empirical, atheoretical approach has two interrelated weaknesses that our approach, grounded in NP's structural model, addresses. First, in accounting settings, where double-entry bookkeeping creates collinear variables, unstructured data-mining approaches can produce good predictive performance at the cost of counterintuitive and even uninterpretable results (e.g., Armstrong (2001) and Bertomeu (2020)).[3] While predictive power is in and of itself highly desirable in practice, teaching and future research must rest on an understanding of the variables that drive the predictive power. A purely empirical, atheoretical approach to understanding these causal effects evaluates how removing one variable at a time affects prediction accuracy (e.g., Chen et al. 2022). However, the results of removing one variable at a time from a set of collinear predictor variables are often uninterpretable (for example, predictions will hardly change if common equity is dropped from a predictor set that also includes assets and liabilities). While it is easy to see the problem in a simple example, combining an atheoretical approach with a nonlinear algorithm like a neural network can make it impossible for a researcher to determine the incremental value of any one variable because the actual degree of collinearity is unknown. While we too iteratively remove variables, our approach does not suffer from the same interpretability problem because NP's

[3] For example, applying an atheoretical approach, Chen et al. (2022) find that the two most important predictors for the sign of future earnings changes are last year's accumulated deficit and the percent change in current liabilities.

hierarchical theoretical structure creates well-defined reference groups based on the level of disaggregation. That is, the variables in the level 2 disaggregation are incremental to those in level 1, the variables in level 3 are incremental to those in level 2, and so on. This clear reference group structure supports sharp insights for practice and teaching as to how best to use the NP framework.

A second and related weakness, specifically applicable in our setting, is that using machine learning methods in an atheoretical and unstructured data-mining approach does not readily lend itself to the systematic analysis of financial statement analysis design choices, including those discussed but not analyzed by NP. This analysis is a key feature of our investigations.

2. Methodology

2.1 Structural Accounting–Based Valuation Models and Financial Statement Analysis Design Choices

Structural accounting-based valuation models express a firm's equity value as a function of expected accounting outcomes. Manipulating the definition of an expected return ($\rho_w = (\mathrm{V}_t^E + \mathrm{D}_t)/\mathrm{V}_{t-1}^E$) under the assumptions of constant expected returns and terminal convergence ($lim_{T\to\infty}\, \mathrm{V}_{t+T}^E/\rho_w^T = 0$) yields the discounted dividend model $\mathrm{V}_0^E = \sum_{t=1}^T \mathrm{D}_t\,/\rho_w^t$, where ρ_w denotes the (constant) expected or required return to common equity, V^E denotes value to equity holders, D denotes dividends, and all variables in periods after t denote expected outcomes. In practice, most analysts predict accounting earnings, not dividends which are discretionary. Using the clean surplus relation $\mathrm{CSE}_t = \mathrm{CSE}_{t-1} + \mathrm{CNI}_t - \mathrm{D}_t$ and an additional terminal convergence condition ($lim_{T\to\infty}\, \mathrm{CSE}_{t+T}/\rho^T = 0$), we reformulate the discounted dividend model into the residual income model $\mathrm{V}_0^E = \mathrm{CSE}_0 + \sum_{t=1}^T \mathrm{RE}_t\,/\rho_w^t$, where CSE denotes the book value of shareholders' equity, CNI comprehensive income, and $\mathrm{RE}_t = \mathrm{CNI}_t - \mathrm{CSE}_{t-1} \times (\rho_w - 1)$

residual income (Preinreich 1938; Edwards and Bell 1961; Ohlson 1995).[4]

The residual income model links value to expected accounting outcomes without providing guidance for forecasting those outcomes.[5] NP's framework provides a solution by decomposing residual income into profitability drivers and relating past to future drivers. In Subsections 2.1.1 through 2.1.5, we discuss how NP's structural framework supports residual income valuation through profitability analysis and explain our approach to investigating five financial statement analysis design choices on which theory is silent. In applying NP's framework, an analyst must make these choices on empirical, not theoretical grounds; we provide evidence to inform those choices.

2.1.1 Level of Disaggregation

NP's structural framework is agnostic about the functional form of the residual income process; instead, it disaggregates residual income into its drivers and relates past to future drivers empirically. Specifically, $\mathrm{RE} = (\mathrm{ROCE} - \rho_w + 1) \times \mathrm{CSE}_{t-1}$, where ROCE (= CNI/CSE) denotes return on common equity, yielding a ratio-based formulation of the residual income valuation model:

$$\mathrm{V}_0^E = \mathrm{CSE}_0 + \sum_{t=1}^{\infty}(\mathrm{ROCE} - \rho_w + 1) \times \mathrm{CSE}_{t-1} \times \rho_w^{-t}. \quad (1)$$

Ratios and income statement accounts without subscripts denote period t amounts. Figure 1 decomposes ROCE into four levels of increasing granularity/disaggregation:

Level 1: ROCE = ROTCE × MSR: ROTCE denotes return on total common equity (= (CNI + MII)/(CSE + MI)), MSR minority sharing ratio (= $\frac{\mathrm{CNI/(CNI + MII)}}{\mathrm{CSE/(CSE + MI)}}$), MII minority (noncontrolling)

[4] Both US GAAP and IFRS require firms to display comprehensive income, which satisfies the clean surplus relation.
[5] While Ohlson (1995) assumes residual income is linked linearly through time to derive further theoretical implications, Dechow, Hutton, and Sloan (1999) do not find evidence supporting this assumption in annual US data.

interest income, and MI minority (noncontrolling) interest.

Level 2: ROTCE = RNOA + FLEV × SPREAD: RNOA denotes return on net operating assets (= OI/NOA), FLEV financial leverage (= NFO/CSE), SPREAD the spread between RNOA and net borrowing cost (= RNOA − NBC), OI operating income, NOA net operating assets (= OA − OL), NFO net financial obligations (= FO − FA), OA operating assets, OL operating liabilities, FO financial obligations, FA financial assets, NBC net borrowing cost (= NFE/NFO), and NFE net financial expense.

Level 3: RNOA = Sales PM × ATO + Other items/NOA: Sales PM denotes sales profit margin (= OI from Sales/Sales) and ATO asset turnover (= Sales/NOA).

Level 4: Sales PM × ATO = Sales PM* × ATO* + OLLEV × OLSPREAD: Sales PM* denotes modified profit margin after considering implicit charges on supplier credit (= (Core OI from Sales + io)/Sales), ATO* modified asset turnover (= Sales/OA), OLLEV operating liability leverage (= OL/NOA), OLSPREAD the spread between return on operating assets and the implicit interest on operating liabilities (= (OI + io)/OA − io/OL), and io the implicit interest charge on operating liabilities.

The analysis reveals eight drivers of ROCE, as shown in Equation (2):

$$\begin{aligned} \text{ROCE} \quad = \quad & \text{MSR} \times [\text{Sales PM}^{*} \times \text{ATO}^{*} + \frac{\text{Other Items}}{\text{OA}} + \text{OLLEV} \times \text{OLSPREAD} \\ & + \text{FLEV} \times (\text{RNOA} - \text{NBC})]. \end{aligned} \tag{2}$$

The choice of disaggregation level is made empirically as a tradeoff between the risk of information loss from less disaggregation and the risk of introducing noise from more disaggregation.[6] If idiosyncratic variation of ratio components cannot be used to increase forecast accuracy there is no gain from disaggregation. Indeed, Gerakos and Gramacy (2013) and Li and Mohanram (2014) document that a simple random walk outperforms linear models with a larger (more disaggregated) set of predictors in out-of-sample earnings forecasting.[7] Our first hypothesis, stated in null form, is as follows:

Hypothesis 1. *Higher-level ratio disaggregation does not change model performance.*

2.1.2 Core versus Transitory Items

While some of the eight ROCE drivers in NP's framework, such as ATO, are persistent, others, such as RNOA deriving from unusual operating income, are mean reverting (transitory). Excluding transitory components could enhance forecasting performance by decreasing prediction-irrelevant noise or impair performance because of information loss; that is, the treatment of non-core/transitory items is a distinct financial statement analysis design choice involving a tradeoff between information loss and noise. Acknowledging this, NP adjust their decomposition of ROCE as shown in Equation (3):

$$\begin{aligned}\text{ROCE} \quad = \quad & \text{MSR} \times [\text{Core Sales PM}^{*} \times \text{ATO}^{*} + \frac{\text{Core Other Items}}{\text{OA}} + \frac{\text{UOI}}{\text{OA}} + \text{OLLEV} \\ & \times \text{OLSPREAD} + \text{FLEV} \times (\text{Core RNOA} - \text{Core NBC} + \frac{\text{UOI}}{\text{NOA}} - \frac{\text{UFE}}{\text{NFO}})],\end{aligned} \tag{3}$$

where Core Sales PM* denotes modified profit margin from core sales (= (Core OI from Sales + io)/Sales), UOI unusual operating income, Core RNOA core return on net operating assets (= Core OI from Sales/NOA + Core Other Items/NOA), Core NBC core net borrowing cost (= Core

[6] If a ratio's components are not perfectly correlated, each component might exhibit idiosyncratic value-relevant variation that is lost in aggregation.

[7] This idea is consistent with the principle of Ockham's razor. William of Ockham, a 14th-century logician, argued that greater model complexity increases the possibility for error. In our context, this principle suggests that simpler (less disaggregated) models might outperform more complex (more disaggregated) ones.

NFE/NFO), and UFE unusual financial expense. Equation (3) identifies eight relatively more persistent drivers of ROCE: MSR, FLEV, Core NBC, ATO*, Core Sales PM*, Core Other Items/OA, OLLEV, and OLSPREAD.

There are at least three considerations as to whether including versus excluding items labeled transitory/non-core (i.e., UOI/OA, UFE/NFO) improves forecasting. First, accounting requirements may produce transitory income items with predictive ability. Penman and Zhang (2002) argue that conservative accounting rules can generate future-period (accounting) benefits while decreasing current-period income; for example, recording a current-period impairment loss implies an increase in future accounting performance. The impairment loss, classified as transitory/non-core, would be relevant for predicting future earnings. Second, while models that include persistent operating items and exclude transitory non-operating items should (theoretically) produce better forecasts, both theory (Dye 2002) and empirical research (Barnea, Ronen, and Sadan 1976; Kinney and Trezevant 1997; Givoly, Hayn, and D'Souza 2000; McVay 2006) suggest managers sometimes manipulate income statement presentation to blur the core/non-core distinction. Third, the distinction between transitory/non-core and persistent/core income items arises at least partly from the firm's business model. The empirical measures used by NP and in this paper are based on Compustat data definitions applied to all entities, which may result in an imperfect separation of core from non-core items for at least some firms. Thus, whether a financial statement analysis design choice to focus on core items improves forecasts is an empirical question. Our second hypothesis, stated in null form, is as follows:

Hypothesis 2. *Excluding transitory items does not change model performance.*

2.1.3 Amount of Past Information/Number of Lags to Use in Predictions

An analyst must decide how much past information to consider. Using more lags of historical data increases both the amount of information supporting predictions and the likelihood that a firm's activities have changed sufficiently to reduce the signal-to-noise ratio. Our third hypothesis, stated in null form, is as follows:

Hypothesis 3. *Using more predictor lags does not change model performance.*

2.1.4 Including versus Excluding Financing Activities

NP propose a model simplification based on provisions in US GAAP and IFRS that require recognition or disclosure of market (fair) values of net financial obligations (NFO). If NFO fair value equals fundamental value, Equation (1) can be simplified as follows:

$$\mathrm{V}_0^E = \mathrm{NOA}_0 - \mathrm{NFO}_0 + \sum_{t=1}^{\infty} (\,\mathrm{RNOA} - \rho_W + 1) \times \mathrm{NOA}_{t-1} \times \rho_W^{-t}, \tag{4}$$

where NOA denotes value of net operating assets, NFO value of net financial obligations, ρ_W weighted average cost of capital (WACC) ($\rho_W = \rho_w \times \frac{\mathrm{V}_0^E}{(\mathrm{V}_0^E + \mathrm{NFO}_0)} + \rho_D \times \frac{(1-\tau)\times\mathrm{NFO}_0}{(\mathrm{V}_0^E + \mathrm{NFO}_0)}$), ρ_D cost of debt, and τ marginal tax rate. Equation (4) reduces the forecasting inputs to the five drivers of RNOA attributable to common shareholders: ATO*, Sales PM*, Other Items/OA, OLLEV, and OLSPREAD. An analyst chooses between the simplified model and the full model.

As in the case of the core versus non-core distinction, it is unclear *ex ante* whether the simplified Equation (4) model improves model performance; that is, it is an empirical question whether past financing activities are informative about future operating activities. While some view the valuation implications of operating and financing activities as mutually independent (Penman 2012; Li, Richardson, and Tuna 2014), the agency cost literature suggests a possible association between operating and financing activities. For example, managers of firms close to bankruptcy might take excessive operating risks, because of the call option-like payoff structure of equity (Jensen and Meckling 1976). Our fourth hypothesis, stated in null form, is as follows:

Hypothesis 4. *Focusing on operating activities does not change model performance.*

2.1.5 Forecast Horizon and Terminal Value

Equations (1) and (4) specify indefinite-horizon forecasts. For practicality, analysts assume firms reach steady-state residual income growth to estimate terminal value and choose a forecast horizon T accordingly (Penman 2012; Nissim 2019). Applying a terminal growth rate g and forecast horizon T to Equations (1) and (4) yields

$$\mathrm{V}_0^E = \mathrm{CSE}_0 + \sum_{t=1}^{T} \frac{(\mathrm{ROCE}_t - \rho_w + 1) \times \mathrm{CSE}_{t-1}}{\rho_w^t} + \frac{(\mathrm{ROCE}_{T+1} - \rho_w + 1) \times \mathrm{CSE}_T}{\rho_w^T \times (\rho_w - g)} \tag{5}$$

and

$$\begin{aligned} \mathrm{V}_0^E = \quad & \mathrm{NOA}_0 - \mathrm{NFO}_0 + \sum_{t=1}^{T} \frac{(\mathrm{RNOA}_t - \rho_W + 1) \times \mathrm{NOA}_{t-1}}{\rho_W^t} \\ & + \frac{(\mathrm{RNOA}_{T+1} - \rho_W + 1) \times \mathrm{NOA}_T}{\rho_W^T \times (\rho_W - g)}. \end{aligned} \tag{6}$$

The analyst must choose T on practical empirical grounds. Applying a terminal growth rate g is effectively a special case of making discrete forecasts for individual periods after T: it is equivalent to making a forecast for each period that is exactly g percent larger than in the previous period. While analysts can incorporate more information in making individual forecasts for each period, doing so will improve performance only if using the additional information improves upon a naïve terminal growth rate approach. The more distant the forecasted period in question, the lower the likelihood this criterion is met. While 1-, 5-, and 10-year horizons are common in practice, Koller, Goedhart, and Wessels (2020, p. 270) refer to 5- and 7-year horizons and Nissim (2019) finds that a 10-year horizon yields the terminal value estimate closest to observed price in the terminal year. Our fifth hypothesis, stated in null form, is as follows:

Hypothesis 5. *Longer prediction horizons do not change model performance.*

We acknowledge the practical necessity of other financial statement analysis design choices, including the reinvestment rate, discount rates ρ_w and ρ_W, and growth rate g. Because the number of models to be estimated grows multiplicatively with the number of choices, we focus on the five design choices described in NP that we believe to be most important and most directly connected to accounting information. We invoke Miller and Modigliani (1961) principles and set the reinvestment rate to one. In line with historical GDP figures, we assume a 2% growth rate (e.g., Penman and Sougiannis 1998). To ensure consistency with our asset pricing tests, we estimate discount rates using the Fama and French (2015) five-factor model.

Equations (5) and (6) link value to ratios, and financial statement analysis links past realizations to future realizations of ratios. Because both links or relations are nonlinear, it is necessary to use methods that can accommodate complex nonlinear associations to apply financial statement analysis in forecasting and value estimation. The next section describes such a tool: neural networks, a machine learning algorithm.

2.2 Neural Networks

A neural network generalizes estimators such as OLS to model complex nonlinear relations among independent and (possibly multiple) dependent variables through a layered system of equations. The basic building block of a neural network is a neuron, a function that takes in variables as inputs, combines them through a linear equation, and transforms the output of that equation through a (typically) nonlinear function known as an activation function. OLS and logistic regression are examples of single neurons that create a linear relation among independent variables and transform the relation by multiplying by one and $(1 + e^x)^{-1}$, respectively.

A neural network organizes relations among inputs into layers. Each layer, including the input layer (independent variables) and output layer (dependent variables), consists of a series of neurons. Layers between the input and output layers are hidden layers. A neural network with one (multiple) hidden layer(s) is referred to as shallow (deep). While the connections between layers can be set arbitrarily, the most common implementations are sequential models (feedforward neural networks) that fully connect each neuron in the preceding layer to each neuron in the next layer. In other words, the neurons of one layer become the independent variables that are the inputs for the neurons in the next layer. We use a fully connected sequential model with multiple variables in the input and output layers, constant activation functions, and a fixed number of neurons in each hidden layer. By combining the activation functions with a series of layers, the neural network can model complex nonlinear relations without the need to specify a functional form.

Panels A to C of Figure 2 compare three special cases of neural networks: a single-layered network with one predictor and one outcome variable; a single-layered network with eight predictors and one outcome variable; and a multi-layered model with two hidden layers, 10 neurons per layer, eight predictors, and five outcome variables. Circles symbolize neurons, and lines indicate connections among neurons. The Panel A model resembles the random walk model recommended by Watts and Leftwich (1977), who find that it outperforms more complicated earnings prediction schemes in time-series regressions. The Panel B model resembles the model of Hou, Van Dijk, and Zhang (2012), who argue that earnings prediction can be enhanced by adding predictors such as accruals and the book-to-market ratio in a pooled cross-sectional model. The Panel C model resembles a more general neural network with the potential to model the nonlinearities and interactive effects of NP's framework by adding hidden layers.

As previously noted, neural networks can estimate arbitrarily complex functions among independent and dependent variables. While in principle the same outcome is achievable by including higher-order polynomials and interaction terms in linear regression models, the dimensionality of this problem quickly makes estimation infeasible. A simple linear regression with 10 independent variables would estimate 1 + 10 = 11 parameters. Including the squares and cubes of each independent variable increases the number of parameters to be estimated to 31 [= 1 + 10 + 10 + 10]. Including interactions among these 30 independent variables increases the number of parameters to $1 + 10 + 10 + 10 + 29! = 8.84 \times e^{30}$. Once the number of parameters exceeds the number of observations in the dataset, either the model is not estimable via OLS, or the estimates will behave poorly (Huber 1973). In contrast, neural networks allow researchers to capture higher-order and interactive relations without explicitly specifying them. The combination of hidden layers connected through nonlinear activation functions approximates such relations automatically (Hornik, Stinchcombe, and White 1989; Cybenko 1989; Tsang, Cheng, and Liu 2017), reducing computing and implementation time, limiting subjective research design choices, and making neural networks prime candidates for modeling the complex nonlinear relations between the past and future value-determining fundamentals discussed in the previous section. As Figure 3 shows, the popularity (measured by web queries analyzed via Google Trends) of neural networks has increased substantially relative to that of other prominent machine learning algorithms previously used in the accounting literature, such as lasso regressions and random forests. We implement the neural network using Google's TensorFlow API. The online appendix details our machine learning design choices.

2.3 Model Performance Evaluation

We assess model performance two ways to provide two distinct kinds of evidence on NP's structural model and on the five financial statement analysis design choices previously discussed. We first compare the forecast accuracy of neural network estimation of NP's framework to that of a random walk and OLS estimation.[8] This assessment is independent of the firm's trading environment and is used to test Hypotheses 1, 2, and 3. However, assessing whether RNOA or ROCE can be forecasted more accurately does not provide evidence of informativeness to an equity investor. Therefore, our second, complementary assessment is based on future portfolio returns, adopting the premise that NP's framework is designed to aid in equity valuation and the assumption that an equity investor wishes to maximize risk-adjusted return.

The portfolio-returns approach allows us to compare ROCE-based models that combine effects of operating and financing activities with RNOA-based models that abstract from financing activities (Hypothesis 4) and to analyze the effects of forecast horizon choices (Hypothesis 5). Given their differing distributional properties as shown in Tables 1 and 3, it is unclear how to analyze the relative performance of ROCE versus RNOA models in other ways, such as comparing absolute forecast error or bias (Hou et al. 2012; Evans, Njoroge, and Yong 2017). Also, if macroeconomic shocks affect earnings outcomes (Ball, Sadka, and Sadka 2009; Bonsall, Bozanic, and Fischer 2013), earnings expectations—not earnings realizations—determine stock prices.

[8] We use scaling to enhance comparability of observations from differently sized firms and to make the observations invariant to inflation. We use CSE or NOA at the beginning of the period in which the forecast is made as scalars for our ROCE and RNOA predictions. This approach addresses a technical problem: analysts forecasting earnings divided by a scalar that will realize after the year in which the forecast is made are forecasting both the earnings number and the scalar. It is therefore unclear whether their forecast accuracy derives from accurate earnings forecasts, accurate scalar forecasts, or accurate forecasts of the covariance between earnings and the scalar. By using a scalar that is realized before the forecast is made, we can attribute forecast accuracy to the earnings forecast. Given that we aim to assess machine learning's usefulness in earnings prediction, this approach is appropriate for our analyses. We thank Stephen Penman for pointing out this issue to us.

We measure portfolio performance as alpha from a Jensen, Black, and Scholes (1972) time-series asset pricing model. We calculate each model's alpha as follows:

1. Calculate a firm's value-to-price ratio. Value is computed using Equations (5) and (6). Valuation inputs are obtained using neural networks under the financial statement analysis design choices described in Section 2.1. To ensure that all information is available at the time of portfolio formation in year t, forecasting models are estimated using data up to year t − Lead, where Lead is the number of ROCE or RNOA leads that the model is designed to forecast.[9] Figure 4 shows the timeline. To ensure that differences in model performance are driven by fundamentals and not by differences in sample composition, we require that the estimation sample includes firm-year observations for which 10 leads and five lags for each predictor variable used in a level 4 disaggregation are available.

2. Following Fama and French (1993), at the end of June in year t + 1, form decile portfolios using the value-to-price ratios based on accounting data for fiscal year t. The first portfolios are formed in 1988 (1963 plus a maximum lead of 10 years plus a maximum lag of five years plus a minimum sample estimation period of 10 years).

3. Calculate future annual value-weighted portfolio returns from July in year t + 1 to June in year t + 2.[10] The portfolio holding period is shown in Figure 4.[11]

4. Calculate excess returns by subtracting returns on the risk-free security and the lowest decile portfolio from the return on the highest decile portfolio.

5. Regress excess return on the risk factors in Fama and French (2015). Alpha is the regression intercept estimate.

In the absence of an agreed-upon asset pricing model, we do not aim to provide evidence on a novel market anomaly (i.e., mispricing). We use the results of portfolio returns tests to provide evidence on the decision usefulness to equity investors of different financial statement analysis design choices, consistent with Jackwerth and Slavutskaya's (2018) observation that asset pricing

[9] If the parameters to be estimated vary predictably by industry, estimating the model by industry would improve forecasting. However, Damodaran (2007) argues that such variation is not predicted by valuation theory, and Fairfield, Ramnath, and Yohn (2009) do not find that estimating earnings forecasting models by industry yields more accurate predictions.
[10] As noted by Loughran and Ritter (2000), using value-weighted instead of equal-weighted returns helps ensure that results are not driven by small, illiquid stocks that are costly to trade. Tests based on value-weighted returns, therefore, offer less power to detect mispricing.
[11] Our results are robust to using monthly instead of annual returns.

models can be used to assess *relative* performance of models incorporating different fundamentals. We are agnostic about individual models' alpha magnitudes, and we focus on comparing *relative* alpha magnitudes across models to test the hypotheses developed in Section 2.1. That said, portfolio returns tests require the choice of a factor model. Results reported in the tables are based on the five factors in Fama and French (2015). Untabulated analyses based on an unadjusted excess return model, the CAPM, the Fama and French (1993) three-factor model, and the Carhart (1997) four-factor model support inferences generally similar to those discussed in Section 4.2.

3. Data and Descriptive Evidence

3.1 Data

Following NP, we use annual Compustat data from 1963 to 2019 for NYSE and AMEX firms. We retain observations with five lags for all required variables and non-negative values for CSE, NOA, OA, and OL at fiscal-year beginning and end to ensure both that our results are not driven by different sample compositions across models and meaningful values for computed ratios. Stock return and asset pricing factor data are from CRSP and Ken French's website.

3.2 Summary Statistics

Table 1 Panel A presents descriptive statistics for price and valuation anchors, i.e., shareholders' equity (CSE), net operating assets (NOA), and net financial obligations (NFO). Following Fama and French (1993), we measure price as market value of equity on the last day of June in the year succeeding the fiscal year of portfolio formation. The mean and median values of price exceed the mean and median values of CSE and NOA, indicating that the market, on average, expects firms to earn positive future residual income. For most firms NFO is positive, which means their financial obligations exceed their financial assets. However, financial assets exceed financial obligations for firms in the first percentile of the NFO distribution. Table 1 Panel B reports descriptive statistics for the profitability drivers in NP's framework. The mean and median values are similar to those presented in NP's Table 1; the standard deviations are higher, because NP winsorize all variables while we do not. In contrast to NP, we aim to predict profitability out-of-sample, not to report descriptive evidence. Therefore, to mimic a realistic forecasting environment as closely as possible, our algorithms need to be capable of handling outliers.

Table 2 presents correlations for selected ratios, with Pearson (Spearman) correlations above (below) the diagonal. Several correlations of profitability drivers with ROCE and RNOA are economically and statistically significant, suggesting the predictive usefulness of NP's framework. ROCE and ROTCE are close to perfectly correlated, while RNOA exhibits idiosyncratic variation, with Pearson (Spearman) correlation with ROCE and ROTCE equal to 0.02 (0.89). The divergence between Pearson correlations (which are heavily affected by outliers) and Spearman correlations (which are not) points to the practical need for algorithms that are capable of handling outliers. Sales PM is more strongly correlated with contemporaneous non-core profitability measures than Core Sales PM, while the opposite holds for Core RNOA. While Fairfield and Yohn (2001) find turnovers are better profitability predictors than margins, in Table 2 margins are more strongly correlated with contemporaneous profitability than turnovers. OLLEV correlates more strongly with RNOA than with ROCE, a potential by-product of cross-sectional variation in financial leverage confounding the univariate relation between OLLEV and ROCE.

3.3 Nonlinearities

Figure 5 Panels A to H present visual evidence on the relation between future profitability and contemporaneous ratios, by plotting median portfolio ROCE in periods t to t+10 by ROCE, FLEV, SPREAD, ATO, Sales PM, OLLEV, OLSPREAD, and RNOA decile in period t. Except for OLLEV, the plots suggest nonlinear relations. Plotted relative to future ROCE, current ROCE, OLSPREAD, and RNOA have an S-shaped association, FLEV and ATO have a U-shaped association, and SPREAD and Sales PM have a concave association. These functional forms are visible over 10 years and attenuate over time. Figure 6 Panels A to D present examples of interactive relations across ratios in predicting future ROCE. The surfaces obtained from plotting ATO and Sales PM, FLEV and OLSPREAD, FLEV and Sales PM, and ATO and OLSPREAD on the X and Y axes and lead 1 ROCE on the Z axis exhibit curvatures that are visually different from the straight plane observed under linear, non-interactive relations. In sum, the visual evidence in Figures 5 and 6 suggests nonlinearities in the dynamic relations across several fundamental ratios and subsequent profitability.

It would be difficult or even infeasible to specify these (visually) nonlinear functional forms in a linear model based on accounting or financial statement analysis intuition, which makes flexible machine learning algorithms the appropriate estimation tool. While other algorithms such as Random Forests or Gradient Boosted Trees can handle nonlinearities, we use NN because it is readily able to approximate any functional form and, as evidenced by Figure 3, it is widely used (Hornik et al. 1989; Schmidhuber 2015; Huang, Jin, Gao, Thung, and Shen 2016). We acknowledge that other algorithms (or some combination thereof) might produce more accurate profitability forecasts. However, our goal is to estimate nonlinear relations within NP's framework to derive implications for financial statement analysis, not to analyze multiple machine learning algorithms to determine which one yields the most accurate predictions.[12]

4. Evaluating Hypotheses 1, 2, and 3 Using Forecast Errors

We first analyze whether estimating NP's framework via neural networks yields more

[12] Examples of papers that take an algorithm-comparison approach include Hunt et al. (2019) and Anand et al. (2020).

accurate out-of-sample profitability forecasts than a random walk. Previous research (e.g., Gerakos and Gramacy (2013) and Li and Mohanram (2014), finds that a random walk tends to yield more accurate predictions than other more complex earnings forecasting models, justifying the random walk as a benchmark in our setting. Table 3 Panels A and B show 1- to 10-year-ahead median absolute out-of-sample forecast errors for ROCE and RNOA for each of the NP models estimated via neural networks and median absolute random walk out-of-sample forecast errors (line 1 of the table).[13,14] Standard errors are computed following Mann and Whitney (1947). The table shows that most of the NP models generally outperform the random walk; accuracy differences are significant at the 1% level. The improvement in accuracy is substantial, as compared to the improvements obtained by analyzing components of NP's framework reported by Fairfield et al. (1996), Fairfield and Yohn (2001), and Esplin et al. (2014). Several 1-year-ahead ROCE models outperform the random walk by more than half a percentage point of ROCE. For both ROCE and RNOA, the model improvements relative to the random walk increase in the forecast horizon. For example, at the 9-year horizon, several models outperform the random walk by more than a percentage point for ROCE and by more than half a percentage point for RNOA.

---

[13] We analyze median instead of mean absolute forecast errors because they are less affected by outliers; our results for mean forecast errors are even stronger. Further, following Call, Hewitt, Shevlin, and Yohn (2016) and Jackson, Plumlee, and Rountree (2018), in Online Appendix Tables O1 and O2, we generally find that the proportion of cases for which neural network estimation of NP's framework yields more accurate ROCE and RNOA forecast than linear and random walk estimation exceeds 50%. The improvements are more pronounced relative to OLS than to random walk estimation, which is not surprising since OLS's mean squared error objective function gives relatively more weight to avoiding individual large forecast errors while the random walk's prediction that things stay as they were tends to be approximately right in most cases but terribly wrong in some cases.

[14] As discussed in the online appendix, our setting requires that we estimate 640 forecasting models (32 sets of independent variables × 2 dependent variables × 10 leads) for 31 years, for a total of 19,840 neural networks. Results of those 640 models are reported in Table 3. Using the 640 forecasting models, we construct intrinsic value estimates from forecasts over 1-, 5-, and 10-year horizons. This analysis yields 192 valuation models (32 sets of independent variables × 2 dependent variables × 3 forecast horizons).

4.1 Financial Statement Analysis Design Choices

Table 4 analyzes the relation between three financial statement analysis design choices and forecasting performance for ROCE (Panel A) and RNOA (Panel B) by regressing 1- to 10-year-ahead median absolute forecast errors on indicators for these design choices. *Level 2* (*3*, *4*) is an indicator that the profitability forecasting model uses the level 2 (3, 4) disaggregation illustrated in Figure 1. *Core* is an indicator that the model excludes transitory items. *Lag 1* (*3*, *5*) is an indicator that the model uses ratios from the current and preceding 1 (3, 5) years.

4.1.1 Hypothesis 1: Disaggregation Level

We find strong evidence that increasing the ratio disaggregation level improves model performance for RNOA forecasts. Relative to a level 1 disaggregation, a level 2 disaggregation, which incorporates financial leverage, reduces year-ahead median absolute forecast errors by 1.4%. The magnitude of improvement declines over time; its statistical significance persists up to 10 years ahead. Level 3 and level 4 disaggregations provide somewhat smaller improvements over level 1 disaggregation, and the statistical significance declines more quickly. In contrast, we find median absolute ROCE forecast accuracy appears to deteriorate with disaggregation.

4.1.2 Hypothesis 2: Core versus Transitory Items

We find evidence that focusing on core items (excluding transitory items) improves model performance for ROCE forecasts one and five years ahead but not for other horizons and not for RNOA forecasts. This finding suggests, consistent with the discussion in Section 2.1.2, that, as a practical matter, items labeled “non-core” can have predictive ability, perhaps because financial reporting rules and/or management’s reporting decisions blur the core/non-core distinction.

4.1.3 Hypothesis 3: Historical Information

We find no evidence that using more historical information improves forecast accuracy. If anything, the results indicate that using more historical information *decreases* ROCE and RNOA forecast accuracy, especially for forecasts using financial statement information from three to five years back.

Viewed as a whole, we believe the results of testing Hypotheses 1, 2, and 3 suggest meaningful forecasting benefits from increasingly granular disaggregation of recent reporting outcomes, but not much benefit (and perhaps forecast accuracy losses) from using more lags of highly aggregated past realizations. Results with respect to including only core items are more mixed: our results suggest a focus on core items does not improve RNOA forecasts and may, for some horizons, improve ROCE forecasts.

4.2 Importance of Nonlinearities for Profitability Forecasting

Figure 1 and Equations (2) and (3) reveal multiplicative relations between several ratios and profitability, with the number of nonlinear relations increasing in the level of disaggregation. Figures 5 and 6 provide visual evidence of nonlinear associations between ratios and profitability that attenuate over a 10-year horizon. This evidence suggests nonlinearities should affect the choices of forecast horizon and the number of historical periods to include in the short run, and possibly not in the long run.

To provide evidence on the importance of considering nonlinearities, we re-estimate the 640 models analyzed in Table 3 using OLS and present their median absolute ROCE and RNOA forecast errors in Table 5. All OLS models have reliably larger median absolute forecast errors, as compared to neural network models. The differences often exceed several percentage points for both ROCE and RNOA forecasts and appear most pronounced for long-horizon forecasts. Table 6 analyzes these differences by pooling the forecast errors in Tables 3 and 5 and extending the Table 4 analysis by interacting the indicators for financial statement analysis design choices with *NN*, an indicator that the model is estimated via a neural network rather than OLS.

In Table 6, *NN*'s main effect is significantly negative for most horizons for ROCE and RNOA forecast errors, indicating that considering nonlinearities improves forecast accuracy on average. We find evidence that considering nonlinearities incrementally improves forecast accuracy for models built on higher levels of disaggregation, especially for level 3 disaggregation for short-horizon ROCE and RNOA forecasts, and for long-horizon RNOA forecasts of models focusing on core items. Consistent with our previous findings concerning the benefits of using more historical information, Table 6 shows the benefit of using neural network estimation over OLS models is partially offset for models using more past financial statement data.

4.3 Cross-Sectional Analysis

To shed light on the types of firms for which financial statement analysis design choices and nonlinear estimation are especially important, we perform two cross-sectional analyses using firm characteristics that make forecasting more difficult measured three ways: extreme profitability outcomes, operating in a competitive industry, and lifecycle stage.[15] First, in Table 7 Panel A (Panel B), we report results of a median regression of 1-year-ahead absolute ROCE (RNOA) forecast errors of all neural network models (summarized in Table 3) on indicators for financial statement analysis design choices interacted with firm-characteristics indicators.[16] To ensure our inferences derive from variation in the interactive effect between financial statement analysis design choices and firm characteristics rather than variation in firm characteristics per se,

[15] With respect to lifecycle stage, evidence in Vorst and Yohn (2018) and Lyle, Vorst and Yohn (2021) suggests a life-cycle component of earnings.
[16] Median regressions are also known as a least absolute deviation (LAD) regression or a quantile regression at the 50^{th} percentile.

we include firm-year fixed effects. We suppress the slope coefficients of different financial statement analysis design choices, because these coefficients mirror the findings in Table 4.

Second, in Table 8 Panel A (Panel B), we pool the 1-year-ahead absolute ROCE (RNOA) forecast errors of all neural network models with those of all OLS models (summarized in Table 5) for each firm-year and regress the errors on an indicator that the model is estimated via the neural network (*NN*) interacted with the firm-characteristics indicators described before. To ensure that our inferences derive from variation in the interactive effect between the estimation method and firm characteristics rather than variation in firm characteristics or financial statement analysis design choices per se, we include firm-year-model fixed effects.[17] Consistent with Table 6, we find that the coefficient on *NN* is significantly negative across all columns, highlighting that the improvements of nonlinear over linear estimations are pervasive and not concentrated in small pockets of firms.

4.3.1 Extreme Profitability Outcomes

As illustrated in Figure 5, mean reversion in extreme ROCE deciles is greater than in middle deciles. Results in Table 7 Panels A and B column (1) interact the financial statement analysis design choice indicators with *Outlier*, an indicator that the firm-year observation's ROCE is in an extreme ROCE decile. As evidenced by negative coefficients on interactions of *Outlier* with *Core* and indicators for higher disaggregation levels, focusing on core items and more granular ratio disaggregation help predict extreme profitability outcomes. In contrast, consistent with the notion that extreme profitability outcomes are unusual and therefore less predictable using more historical information, the coefficients on *Outlier*'s interaction with *Lags 1* to *5* are

[17] By model we mean an indicator variable for the set of financial statement analysis design choices underlying the neural network, i.e., a unique identifier for each permutation of level of disaggregation, focus on core items, and the number of historical lags.

increasingly positive. The negative coefficients on interactions of *Outlier* with *NN* in Table 8 Panels A and B column (1) suggest a more pronounced forecasting advantage from nonlinear estimation (as compared to linear estimation) when accounting performance is extreme.

4.3.2 Competition

Competition can erode profit margins and thereby drive mean reversion in profitability (Bertrand 1883). Table 7 Panels A and B column (2) shows results from interacting the financial statement analysis design choice indicators with *Competition*, an indicator that the firm operates in a competitive industry. We find some evidence that higher levels of ratio disaggregation (using the past five years of financial statement data) are especially important for predicting ROCE (RNOA) of firms in competitive industries. Table 8 Panel A column (2) suggests less forecasting advantage from nonlinear estimation (as compared to linear estimation) for firms in competitive industries, possibly because profitability prediction is easier when the profitability of all firms within an industry exhibits mean reversion.

4.3.3 Corporate Lifecycle

We examine the relative importance of considering nonlinearities at five stages of a firm's lifecycle, measured following Dickinson (2011): *Introduction*, *Growth*, *Maturity*, *Shakeout*, and *Decline*. Table 7 Panels A and B columns (4) to (8) show results when we interact the financial statement analysis design choice indicators with lifecycle-stage indicators. We generally find that higher model complexity does not improve performance, and for some financial statement design choices even degrades performance during the introduction, growth, and maturity stages, while greater disaggregation and a focus on core items provides advantages in the shakeout and decline stages. We find mixed evidence with respect to using more historical information. Using one to three years of additional financial statement data helps during the decline stage, while using five years of additional financial statement data always hurts prediction. Lastly, Table 8 Panels A and B columns (4) to (8) show that nonlinear estimation's advantage over linear estimation is more pronounced during the introduction, shakeout, and decline stages, and less pronounced during the growth and maturity stages, consistent with the intuition that profitability prediction is more difficult during the earlier and later stages of the corporate lifecycle than during the mid-stages.

5. Evaluating Hypotheses 3 and 4 Using Alphas

Table 9 reports Fama and French (2015) alphas computed as described in Section 2.3. Panel A (Panel B) presents results for ROCE-based (RNOA-based) models built on Equations (5) and (6) and compares their performance to that of a random walk–based model. Following Newey and West (1987), we compute heteroscedasticity- and autocorrelation-robust standard errors with a lag order of $4 \times (29/100)^{2/9} \approx 3$. ROCE-based models generally perform poorly, with alphas often lower than those of the ROCE-based random walk model and sometimes significantly negative. For RNOA-based models, neural network models generally outperform the random walk model.

5.1 Financial Statement Analysis Design Choices

5.1.1 Hypothesis 4: Operating Activities

As previously noted, the results in Table 4 cannot test Hypotheses 4 and 5 because of the nature of the dependent variable (median absolute forecast error). To test Hypothesis 4 we compare alphas in Table 9 Panels A and B by counting the number of times a RNOA-based model outperforms a ROCE-based model using the same financial statement design choices. RNOA-based models outperform their ROCE-based counterparts for 87 of 96 models, providing strong evidence that a focus on operating activities improves model performance. This evidence is consistent with the notion put forward by Penman (2012) and others that value primarily derives from firms' operating activities rather than from the way firms arrange to finance the resources needed to perform these activities, and is not consistent with an agency-cost framework that links the valuation implications of operating activities with financing decisions.

5.1.2 Hypothesis 5: Forecast Horizon

We test Hypothesis 5 by counting the number of times a ROCE-based (RNOA-based) model with forecast horizons of 1, 5, and 10 years ahead outperforms models using the same financial statement design choices applied to alternative forecast horizons. For ROCE-based (RNOA-based) models the 5-year horizon outperforms for 18 (23) of the 32 models, indicating a 5-year horizon generally does best. This evidence suggests a tradeoff: while short-horizon forecasts tend to be more accurate, switching from individual forecasts to a terminal growth rate too early increases the risk of failing to incorporate value-relevant information. While a long forecast horizon mitigates that risk, those forecasts tend to be increasingly imprecise and eventually constitute more noise than signal.

5.2 Trading Cost

We designed hypothetical trading strategies to reduce or eliminate certain obstacles to implementation: including NYSE/AMEX stocks and excluding NASDAQ stocks, conditioning on firms with at least five lags of fundamental data, using value-weighted returns, and taking positions in June of the following year to ensure all accounting information is available. Nevertheless, the alphas presented might be unrealistic if returns are mostly from short positions. Short-selling is expensive, and shares to short might be unavailable in the slow-moving over-the-counter short market (Lee and So 2015).

We evaluate the importance of short positions using the model with the highest alpha, i.e., a 5-year horizon RNOA-based level 3 disaggregation model that does not focus on core items and incorporates five years of past data. Figure 7 Panel A, which plots the neural network model's cumulative risk-adjusted long and short portfolio returns over 1988–2019, provides visual evidence that the returns are driven by the long side. Indeed, while the cumulative return difference between the two portfolios increases over time, there is no evidence of extreme negative returns for the short portfolio, and the cumulative return on the long portfolio exceeds that on the short portfolio in every year. Figure 7 Panel B repeats the analysis when we estimate the model using OLS. The returns spread between the short portfolio and the long portfolio is positive and smaller, resulting in a cumulative return less than half that of the neural network model. This result provides additional evidence that an estimation approach that accommodates the nonlinearities inherent in NP’s framework is a main driver of model performance.

6. Conclusion

We estimate the structural profitability framework that Nissim and Penman (2001) proposed but did not estimate because estimation methods available at the time were insufficient for the framework’s nonlinear structure. We resolve the nonlinear estimation issue by using neural networks, a widely used machine learning algorithm that can capture arbitrarily complex relations across variables, to forecast firm-specific profitability using NP’s framework. We use the forecasts to estimate intrinsic values and apply those intrinsic value estimates in value-to-price-based hypothetical trading strategies. To provide insights for the teaching and practice of fundamental analysis, we explore the effects of variation in five financial statement analysis design choices that analysts must make on empirical rather than theoretical grounds. We aim to analyze NP’s structural framework, not to search empirically for the best profitability predictors, as might be done, for example, by applying machine learning to a large set of predictors atheoretically, without a framework that specifies the information to be considered. Such an approach focuses on prediction, not explanation, as discussed by Bertomeu (2020), for example.

We find that profitability forecasts derived by estimating NP's framework via neural networks are in general substantially more accurate than those derived from either a random walk or linear estimation and that hypothetical trading on intrinsic value estimates based on forecasts from estimating NPs structural model yields substantial risk-adjusted returns. With respect to the effects of design choices, we find that a focus on operating activities and core items and a forecast horizon of five years generally improves model performance. We find mixed evidence that using higher levels of disaggregation improves forecast accuracy and evidence that using more historical information appears not to improve forecast accuracy and may even harm it. Cross-sectional analyses suggest the benefits of nonlinear estimation are pervasive, not concentrated in small pockets of firms for which forecasting is particularly difficult. That said, higher model complexity and nonlinear estimation provide forecasting advantages for firms with extreme profitability levels and during the earlier and later stages of firms' lifecycles, but not the mid-stages. The models' hedge portfolio returns appear to be driven by long positions in firms trading on major exchanges, providing support for the practicability of the trading strategy.

Variable Definitions Appendix (Δ denotes change over the fiscal year)

Variable	Definition
ATO	Asset turnover: Sales/NOA
ATO*	Modified asset turnover: Sales/OA
CNI	Comprehensive net income: Compustat: NI – DVP + ΔMSA + ΔRECTA
Core NBC	Core net borrowing cost: Core NFE/NFO
Core NFE	Core net financial expense: Compustat: (XINT − IDIT) × (1 − MTR) + DVP
Core OI from Sales	Core operating income from sales: OI from Sales − UOI
Core Other Items	Core other items: Other Items − UOI
Core RNOA	Core return on net operating assets: Core OI from Sales/NOA
Core Sales PM	Core sales profit margin: Core OI from Sales/Sales
Core Sales PM*	Modified core sales profit margin: (Core OI from Sales + io)/Sales
Core SPREAD	Core financial leverage spread: Core RNOA − Core NBC
CSE	Common equity: Compustat: CEQ + TSTKP − DVPA
FLEV	Financial leverage: NFO/CSE
io	Implicit interest charge on operating liabilities: R_f × (OL − TXDITC)
MIB	Minority (noncontrolling) interest book value: Compustat: MIB
MII	Minority (noncontrolling) interest income: Compustat: MII
MSA	Marketable security adjustment: Compustat: MSA
MSR	Minority sharing ratio: $(CNI/(CNI + MII)) \times (CSE/(CSE + MIB))^{-1}$ (if CNI, MII, CSE, MIB ≥ 0, else 1)
MTR	Marginal tax rate: Top statutory federal tax rate plus 2% percent average state tax rate. The top federal statutory corporate tax (in percent): 52 (1963), 50 (1964), 48 (1965-1967), 52.8 (1968-1969), 49.2 (1970), 48 (1971-1978), 46 (1979-1986), 40 (1987), 34 (1988-1992), 35 (1993-2017), and 21 (2018).
NBC	Net borrowing cost: NFE/NFO
NFE	Core net financial expense: Core NFE − ΔMSA
NFO	Net financial obligations: Compustat: (DLC + DLTT + PSTK − TSTKP + DVPA) − (CHE + IVAO)
NOA	Net operating assets: NFO + CSE + MIB
OA	Operating assets: Compustat: AT − CHE − IVAO
OI	Operating income: NFE + CNI + MII
OI from Sales	OI − Other Items
OL	Operating liabilities: OA − NOA
OLLEV	Operating liability leverage: OL/NOA
OLSPREAD	Operating leverage spread: (OI + io)/OA − io/OL
Other Items	Compustat: ESUB
P	Price at the end of the June in year following the fiscal year: Compustat: PRCCF × CSHO
R_f	Risk-free rate: One-year Treasury bill yield during the year
RNOA	Return on net operating assets: OI/NOA
ROCE	Return on common equity: CNI/CSE
ROTCE	Return on total common equity: (CNI + MII)/(CSE + MIB)
Sales	Compustat: SALE
Sales PM	Sales profit margin: OI from Sales/Sales
Sales PM*	Modified sales profit margin: (OI from Sales + io)/Sales
SPREAD	Financial leverage spread: RNOA − NBC
TXDITC	Deferred taxes and investment tax credit: Compustat: TXDITC
UFE	Unusual financial expense: Compustat: ΔMSA
UOI	Unusual operating income: Compustat: (NOPI + SPI) × (1 − MTR) − ESUB + XIDO + ΔRECTA

References

Anand, V., R. Brunner, K. Ikegwu, and T. Sougiannis. 2020. Predicting Profitability Using Machine Learning. Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3466478.

Armstrong, J. S. 2001. *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Vol. 30: Springer Science & Business Media.

Ball, R., G. Sadka, and R. Sadka. 2009. Aggregate Earnings and Asset Prices. *Journal of Accounting Research* 47 (5):1097-1133.

Barnea, A., J. Ronen, and S. Sadan. 1976. Classificatory Smoothing of Income with Extraordinary Items. *The Accounting Review* 51 (1):110-122.

Bertomeu, J. 2020. Machine Learning Improves Accounting: Discussion, Implementation and Research Opportunities. *Review of Accounting Studies* 25 (3):1135-1155.

Bertomeu, J., E. Cheynel, Y. Liao, and M. Milone. 2021. Using Machine Learning to Measure Conservatism. Working paper.

Bertrand, J. 1883. Book Review of "Theorie Mathematique De La Richesse Sociale and of Recherches Sur Les Principles Mathematiques De La Theorie Des Richesses". *Journal de Savants* 67:499–508.

Black, F. 1986. Noise. *The Journal of Finance* 41 (3):528-543.

Bonsall, S. B., Z. Bozanic, and P. E. Fischer. 2013. What Do Management Earnings Forecasts Convey About the Macroeconomy? *Journal of Accounting Research* 51 (2):225-266.

Call, A. C., M. Hewitt, T. Shevlin, and T. L. Yohn. 2016. Firm-Specific Estimates of Differential Persistence and Their Incremental Usefulness for Forecasting and Valuation. *The Accounting Review* 91 (3):811-833.

Callen, J. L., C. C. Kwan, P. C. Yip, and Y. Yuan. 1996. Neural Network Forecasting of Quarterly Accounting Earnings. *International Journal of Forecasting* 12 (4):475-482.

Cao, K., and H. You. 2020. Fundamental Analysis Via Machine Learning. Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3706532.

Carhart, M. M. 1997. On Persistence in Mutual Fund Performance. *The Journal of Finance* 52 (1):57-82.

Chen, X., Y. H. Cho, Y. Dou, and B. Lev. 2022. Fundamental Analysis of Xbrl Data: A Machine Learning Approach. *Journal of Accounting Research* 60 (2):467-515.

Cybenko, G. 1989. Approximation by Superpositions of a Sigmoidal Function. *Mathematics of Control, Signals and Systems* 2 (4):303-314.

Damodaran, A. 2007. Valuation Approaches and Metrics: A Survey of the Theory and Evidence. *Foundations and Trends in Finance* 1 (8):693-784.

Dechow, P. M., A. P. Hutton, and R. G. Sloan. 1999. An Empirical Assessment of the Residual Income Valuation Model. *Journal of Accounting and Economics* 26 (1):1-34.

Dickinson, V. 2011. Cash Flow Patterns as a Proxy for Firm Life Cycle. *The Accounting Review* 86 (6):1969-1994.

Dye, R. A. 2002. Classifications Manipulation and Nash Accounting Standards. *Journal of Accounting Research* 40 (4):1125-1162.

Edwards, E., and P. Bell. 1961. The Theory and Management of Business Income: University of California Press, Berkeley.

Esplin, A., M. Hewitt, M. Plumlee, and T. L. Yohn. 2014. Disaggregating Operating and Financial Activities: Implications for Forecasts of Profitability. *Review of Accounting Studies* 19 (1):328-362.

Evans, M. E., K. Njoroge, and K. O. Yong. 2017. An Examination of the Statistical Significance and Economic Relevance of Profitability and Earnings Forecasts from Models and Analysts. *Contemporary Accounting Research*.

Fairfield, P. M., S. Ramnath, and T. L. Yohn. 2009. Do Industry‐Level Analyses Improve Forecasts of Financial Performance? *Journal of Accounting Research* 47 (1):147-178.

Fairfield, P. M., R. J. Sweeney, and T. L. Yohn. 1996. Accounting Classification and the Predictive Content of Earnings. *The Accounting Review* 71 (3):337-355.

Fairfield, P. M., and T. L. Yohn. 2001. Using Asset Turnover and Profit Margin to Forecast Changes in Profitability. *Review of Accounting Studies* 6 (4):371-385.

Fama, E. F., and K. R. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33 (1):3-56.

Fama, E. F., and K. R. French. 2015. A Five-Factor Asset Pricing Model. *Journal of Financial Economics* 116 (1):1-22.

Gerakos, J., and R. Gramacy. 2013. Regression-Based Earnings Forecasts. Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2112137.

Givoly, D., C. Hayn, and J. D'Souza. 2000. Measurement Errors and Information Content of Segment Reporting. *Review of Accounting Studies* 4 (1):15-43.

Hornik, K., M. Stinchcombe, and H. White. 1989. Multilayer Feedforward Networks Are Universal Approximators. *Neural Networks* 2 (5):359-366.

Hou, K., M. A. Van Dijk, and Y. Zhang. 2012. The Implied Cost of Capital: A New Approach. *Journal of Accounting and Economics* 53 (3):504-526.

Huang, L., Y. Jin, Y. Gao, K.-H. Thung, and D. Shen. 2016. Longitudinal Clinical Score Prediction in Alzheimer's Disease with Soft-Split Sparse Regression Based Random Forest. *Neurobiology of Aging* 46:180-191.

Huber, P. J. 1973. Robust Regression: Asymptotics, Conjectures and Monte Carlo. *The Annals of Statistics* 1 (5):799-821.

Hunt, J. O., J. N. Myers, and L. A. Myers. 2019. Improving Earnings Predictions with Machine Learning. Working paper. Available at https://zicklin.baruch.cuny.edu/wp-content/uploads/sites/10/2019/12/Improving-Earnings-Predictions-with-Machine-Learning-Hunt-Myers-Myers.pdf.

Jackson, A. B., M. A. Plumlee, and B. R. Rountree. 2018. Decomposing the Market, Industry, and Firm Components of Profitability: Implications for Forecasts of Profitability. *Review of Accounting Studies* 23 (3):1071-1095.

Jackwerth, J. C., and A. Slavutskaya. 2018. Relative Alpha. Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2439145.

Jensen, M. C., F. Black, and M. S. Scholes. 1972. *The Capital Asset Pricing Model: Some Empirical Tests*: Praeger Publishers, Inc.

Jensen, M. C., and W. H. Meckling. 1976. Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics* 3 (4):305-360.

Kinney, M., and R. Trezevant. 1997. The Use of Special Items to Manage Earnings and Perceptions. *Journal of Financial Statement Analysis* 3:45-53.

Koller, T., M. Goedhart, and D. Wessels. 2020. *Valuation: Measuring and Managing the Value of Companie*. 7 ed: John Wiley & Sons.

Lee, C. M. C., and E. C. So. 2015. Alphanomics: The Informational Underpinnings of Market Efficiency. *Foundations and Trends in Accounting* 9 (2--3):59-258.

Li, K. K., and P. Mohanram. 2014. Evaluating Cross-Sectional Forecasting Models for Implied Cost of Capital. *Review of Accounting Studies* 19 (3):1152-1185.

Li, N., S. Richardson, and İ. Tuna. 2014. Macro to Micro: Country Exposures, Firm Fundamentals and Stock Returns. *Journal of Accounting and Economics* 58 (1):1-20.

Liu, M. 2021. Assessing Human Information Processing in Lending Decisions: A Machine Learning Approach. Working paper. Available at https://www.proquest.com/docview/2427239457?pq-origsite=gscholar&fromopenview=true.

Loughran, T., and J. R. Ritter. 2000. Uniformly Least Powerful Tests of Market Efficiency. *Journal of Financial Economics* 55 (3):361-389.

Lyle, M., P. Vorst and T. L. Yohn. 2021. Firm Life Cycle Classification: Co-movement and Information Transfers. Working paper. available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3436737

Mann, H. B., and D. R. Whitney. 1947. On a Test of Whether One of Two Random Variables Is Stochastically Larger Than the Other. *The Annals of Mathematical Statistics* 18 (1):50-60.

McVay, S. E. 2006. Earnings Management Using Classification Shifting: An Examination of Core Earnings and Special Items. *The Accounting Review* 81 (3):501-531.

Miller, M. H., and F. Modigliani. 1961. Dividend Policy, Growth, and the Valuation of Shares. *The Journal of Business* 34 (4):411-433.

Newey, W. K., and K. D. West. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica: Journal of the Econometric Society*:703-708.

Nissim, D. 2019. Terminal Value. Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3095564.

———. 2021. Earnings Quality. Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3794378.

Nissim, D., and S. H. Penman. 2001. Ratio Analysis and Equity Valuation: From Research to Practice. *Review of Accounting Studies* 6 (1):109-154.

Ohlson, J. A. 1995. Earnings, Book Values, and Dividends in Equity Valuation. *Contemporary Accounting Research* 11 (2):661-687.

Palepu, K. G., and P. M. Healy. 2012. *Business Analysis and Valuation*. 5 ed: Cengage Learning.

Penman, S. 2010. *Accounting for Value*: Columbia University Press.

Penman, S. H. 2012. *Financial Statement Analysis and Security Valuation*: McGraw-Hill New York.

Penman, S. H., and T. Sougiannis. 1998. A Comparison of Dividend, Cash Flow, and Earnings Approaches to Equity Valuation. *Contemporary Accounting Research* 15 (3):343-383.

Penman, S. H., and X.-J. Zhang. 2002. Accounting Conservatism, the Quality of Earnings, and Stock Returns. *The Accounting Review* 77 (2):237-264.

Piotroski, J. D. 2000. Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. *Journal of Accounting Research* 38 (Supplement):1-41.

Preinreich, G. A. D. 1938. Annual Survey of Economic Theory: The Theory of Depreciation. *Econometrica: Journal of the Econometric Society*:219-241.

Schmidhuber, J. 2015. Deep Learning in Neural Networks: An Overview. *Neural Networks* 61:85-117.

Sloan, R. G. 2019. Fundamental Analysis Redux. *The Accounting Review* 94 (2):363-377.

Soliman, M. T. 2008. The Use of Dupont Analysis by Market Participants. *The Accounting Review* 83 (3):823-853.

Sommers, G., P. Easton, and P. Drake. 2021. *Valuation Using Financial Statements*. 2 ed: Cambridge Business Publishers, LLC.

Tsang, M., D. Cheng, and Y. Liu. 2017. Detecting Statistical Interactions from Neural Network Weights. *arXiv preprint arXiv:1705.04977*.

van Binsbergen, J. H., X. Han, and A. Lopez-Lira. 2020. Man Versus Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases. NBER working paper 27843.

Vorst, P. and T. L. Yohn. 2018. Life Cycle Models and Forecasting Growth and Profitability. *The Accounting Review* 93(6): 357-381.

Wahlen, J., S. Baginski, and M. Bradshaw. 2018. *Financial Reporting, Financial Statement Analysis and Valuation: A Strategic Approach*. edited by C. Learning. 9 ed.

Watts, R. L., and R. W. Leftwich. 1977. The Time Series of Annual Accounting Earnings. *Journal of Accounting Research* 15 (2):253-271.

Yohn, T. L. 2020. Research on the Use of Financial Statement Information for Forecasting Profitabiilty. *Accounting and Finance* 60: 3163-3181.

Figure 1. Nissim and Penman (2001) Analysis of ROCE
This figure depicts the analysis of profitability developed in Nissim and Penman (2001). All variables are defined in the appendix.

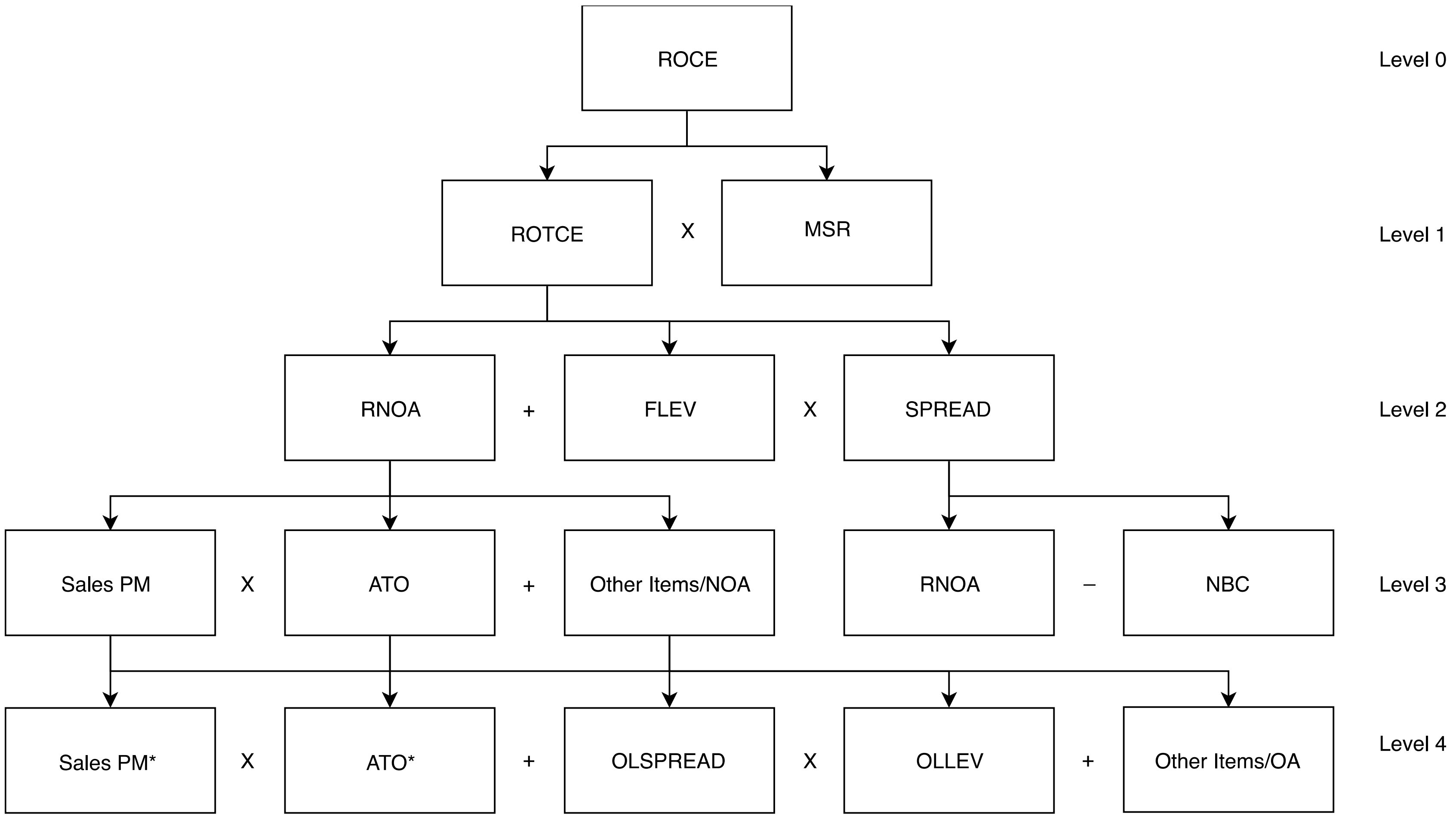
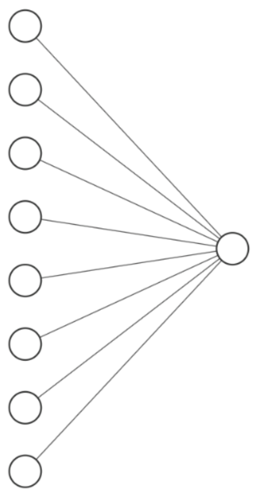


Figure 2. Estimator Comparison

Panels A, B, and C graphically present the architecture of a single-layered neural network with a single predictor, a single-layered neural network with 8 predictors, and a fully connected sequential multi-layered neural network with 1 input layer with 8 neurons, 2 hidden layers with 10 neurons each, and one output layer with 5 neurons.

Panel A. Single-layered Neural Network with a Single Predictor

Panel B. Single-layered Neural Network with Multiple Predictors



Panel C. Multi-layered Neural Network

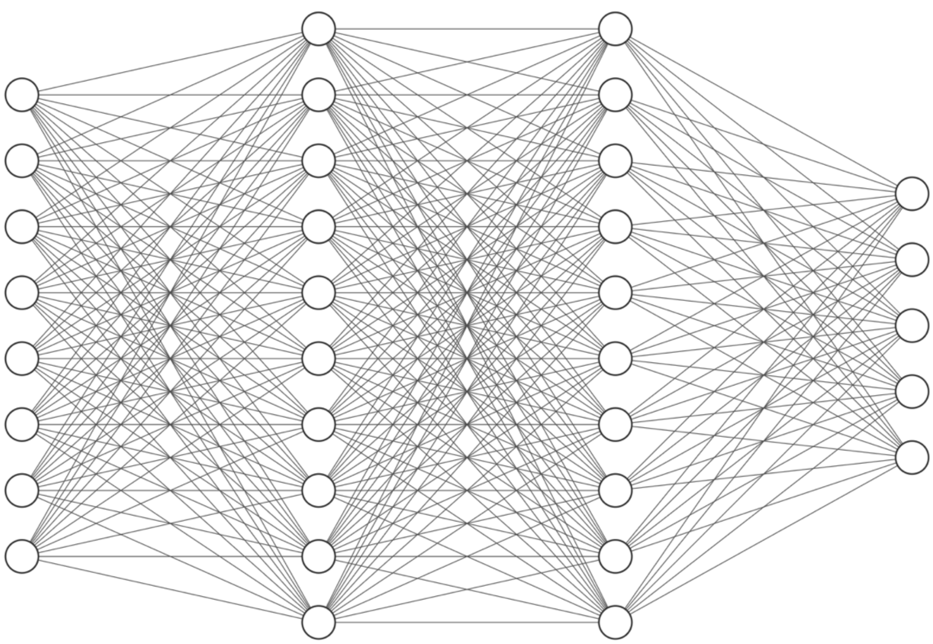


Figure 3. Prominence of Three Machine Learning Algorithms, 2004-2020
This figure depicts the prominence of different machine learning algorithms as measured by web queries analyzed via the web tool Google Trends over the 2004 to 2020 period.

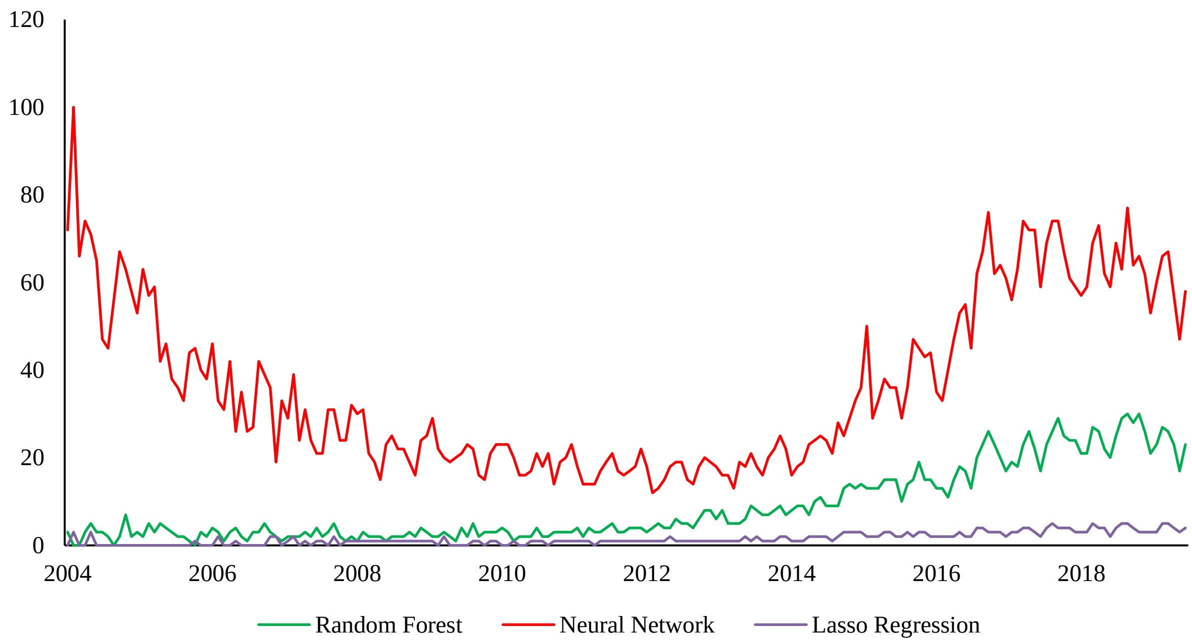


Figure 4. Model Estimation and Forecasting Timeline
This figure depicts the timeline for model estimation, forecasting, and portfolio formation and resolution. The sample period covers 1963 to 2019. The sample includes NYSE and AMEX firms with available data, as described in Section 3.1.

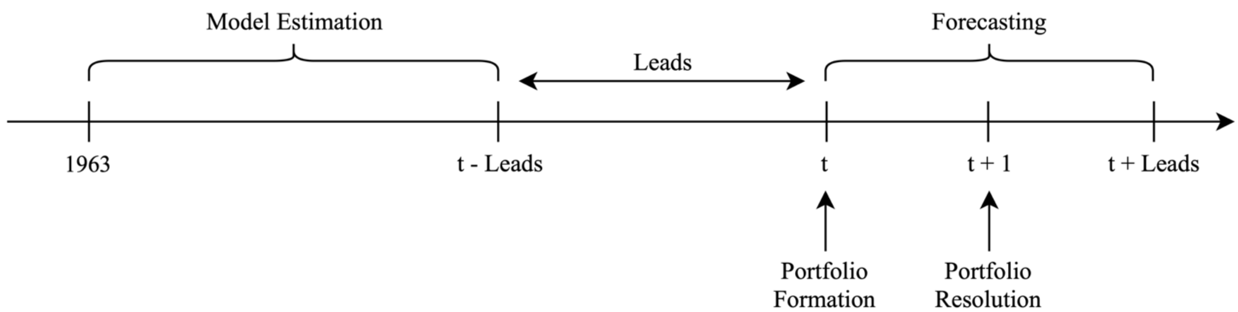
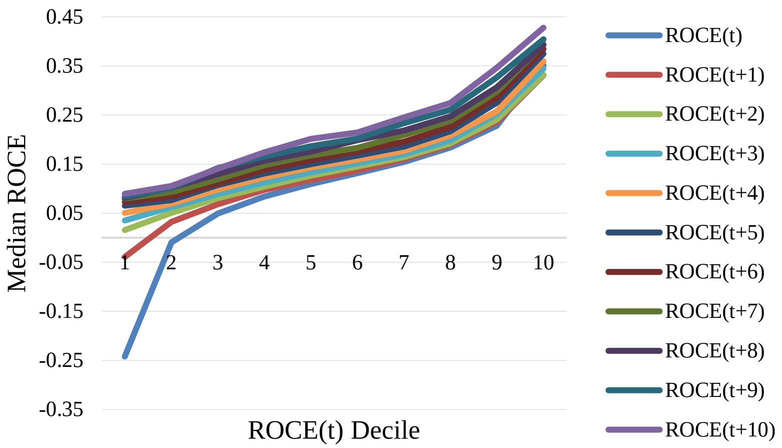


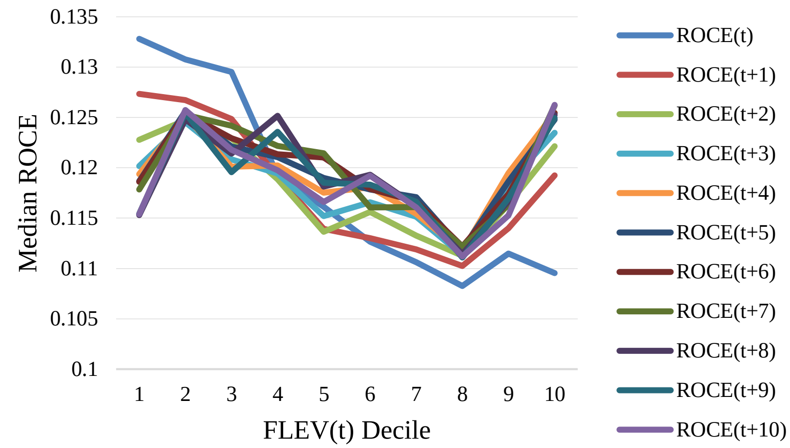
Figure 5. Univariate Time-Series Plots

Panels A to H present median portfolio ROCE in periods t to t+10 by ROCE, FLEV, SPREAD, ATO, Sales PM, OLLEV, OLSPREAD, and RNOA decile in period t. All variables are defined in the appendix.

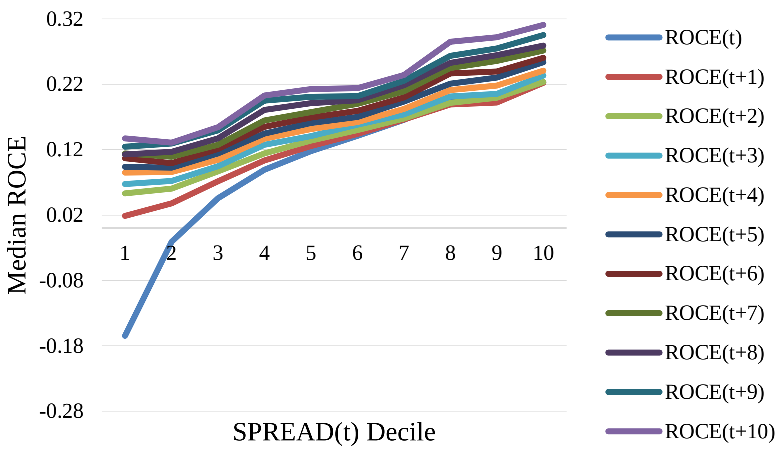
Panel A. Future ROCE by Current ROCE Decile



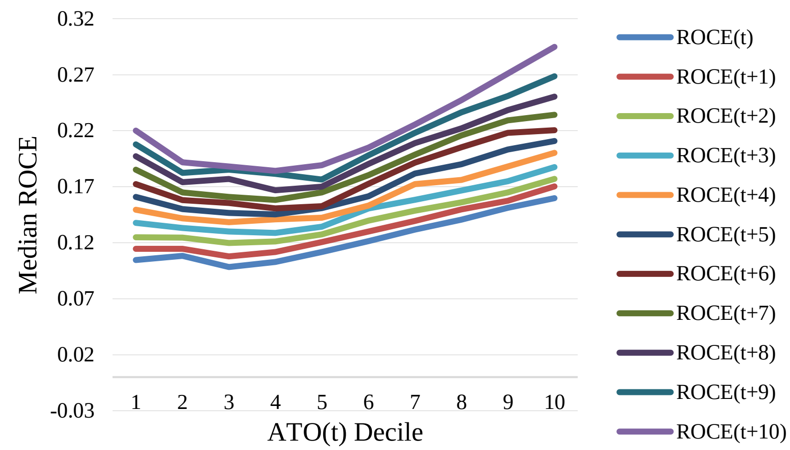
Panel B. Future ROCE by Current FLEV Decile



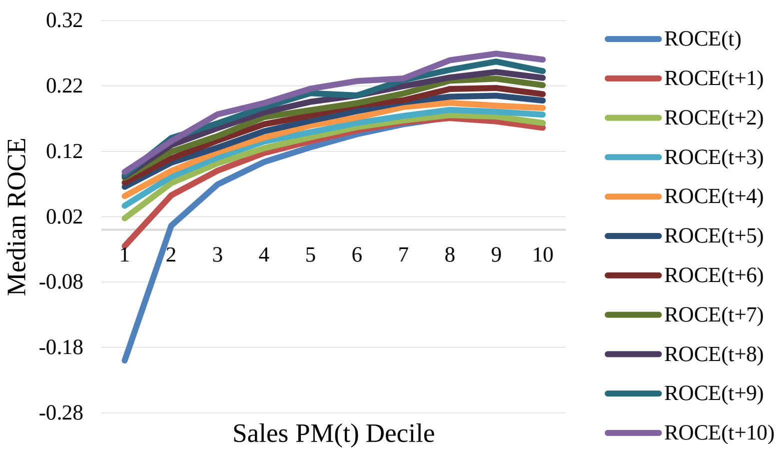
Panel C. Future ROCE by Current SPREAD Decile



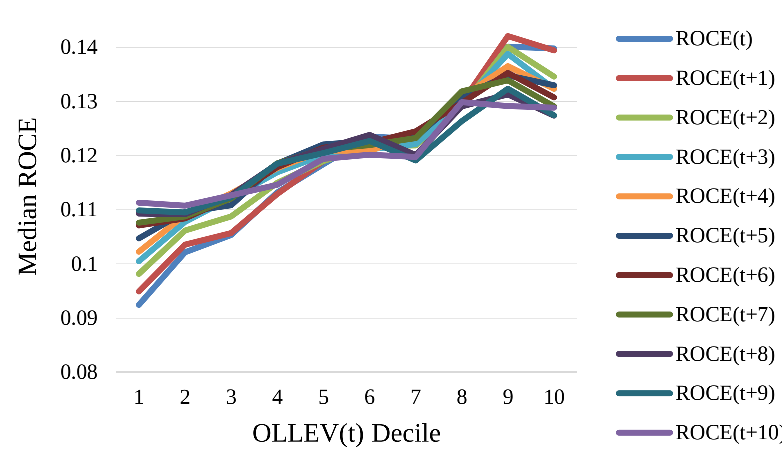
Panel D. Future ROCE by Current ATO Decile



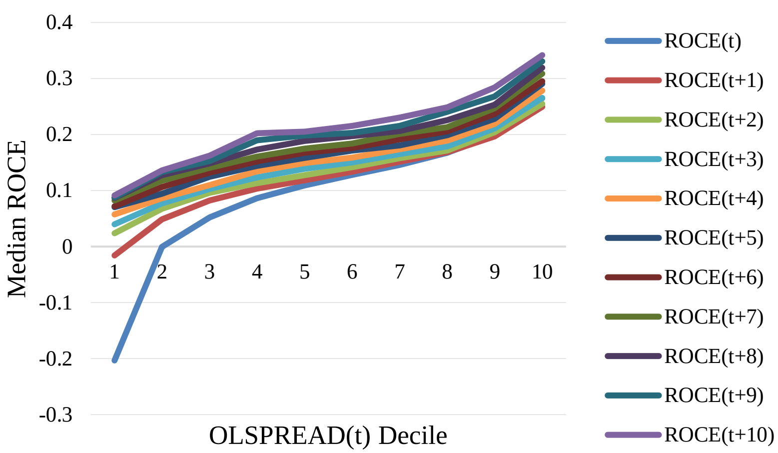
Panel E. Future ROCE by Current Sales PM Decile



Panel F. Future ROCE by Current OLLEV Decile



Panel G. Future ROCE by Current OLSPREAD Decile



Panel H. Future ROCE by Current RNOA Decile

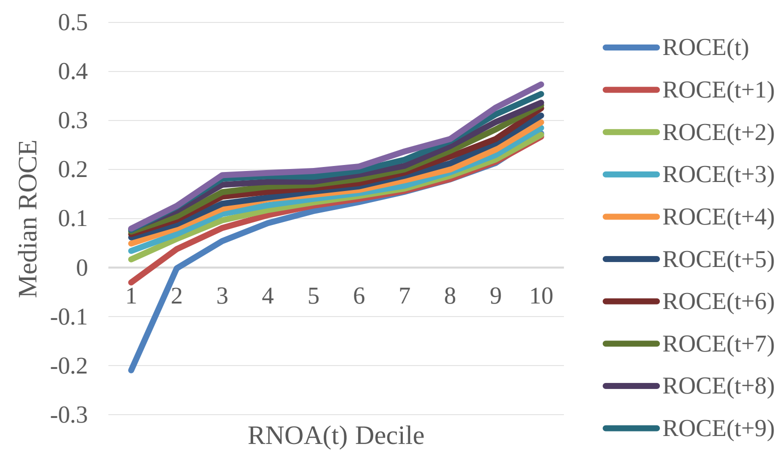
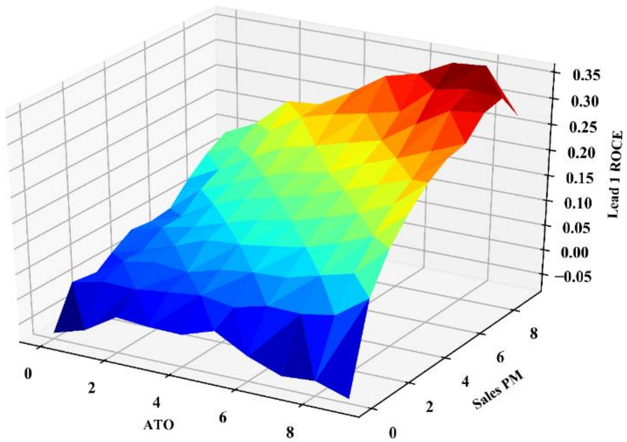


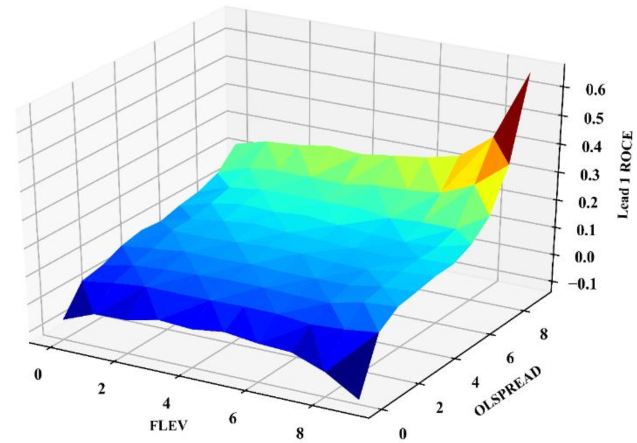
Figure 6. Illustrating Interactive Relationships across Variables in ROCE Prediction

Panels A to D illustrate examples of nonlinear, interactive relationships across various ratios and future ROCE. All variables are defined in the appendix.

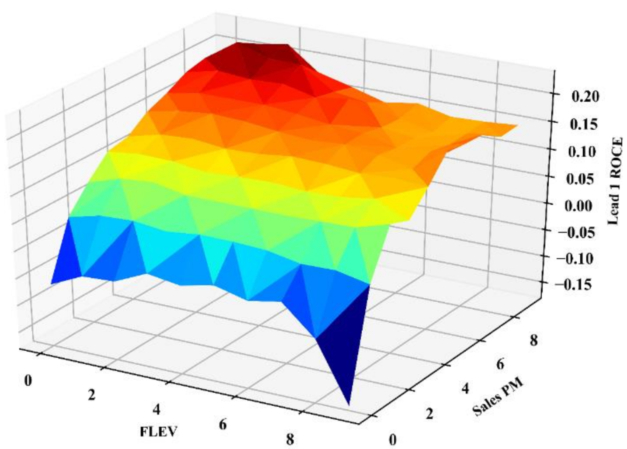
Panel A. ATO, Sales PM, and Lead 1 ROCE



Panel B. FLEV, OLSPREAD, and Lead 1 ROCE



Panel C. FLEV, Sales PM, and Lead 1 ROCE



Panel D. ATO, OLSPREAD, and Lead 1 ROCE

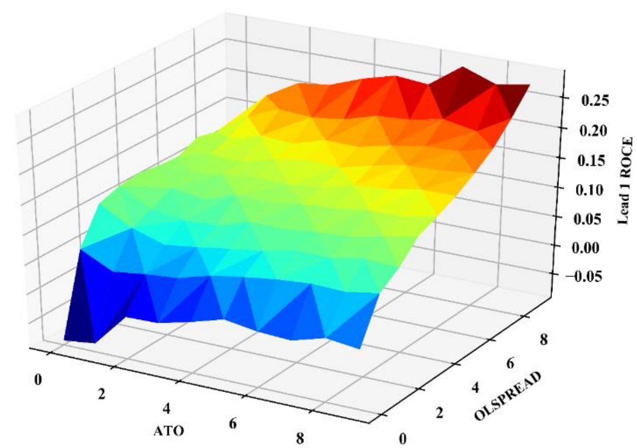
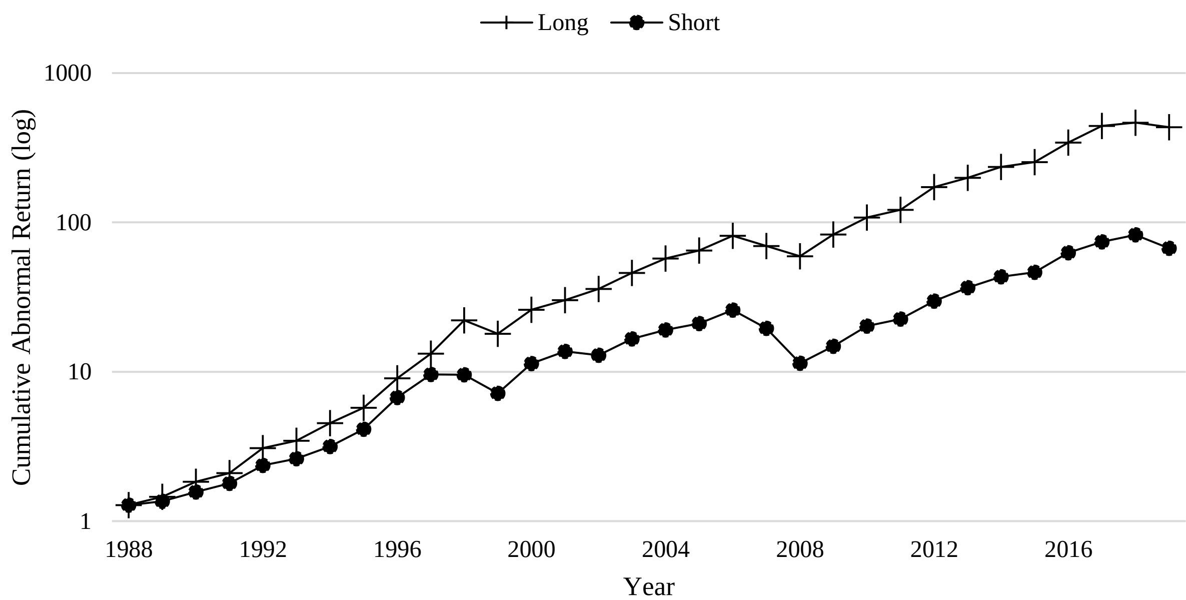


Figure 7. Cumulative Portfolio Long and Short Position Abnormal Returns

Panel A (Panel B) plots value-weighted cumulative abnormal long and short returns derived from portfolios formed based on forecasts of 5 RNOA leads obtained from Neural Network (OLS) models that rely on level 1 ratio disaggregation, do not focus on core items, and use the information in the financial statements of the previous five years.

Panel A. Neural Network Estimation



Panel B. OLS Estimation

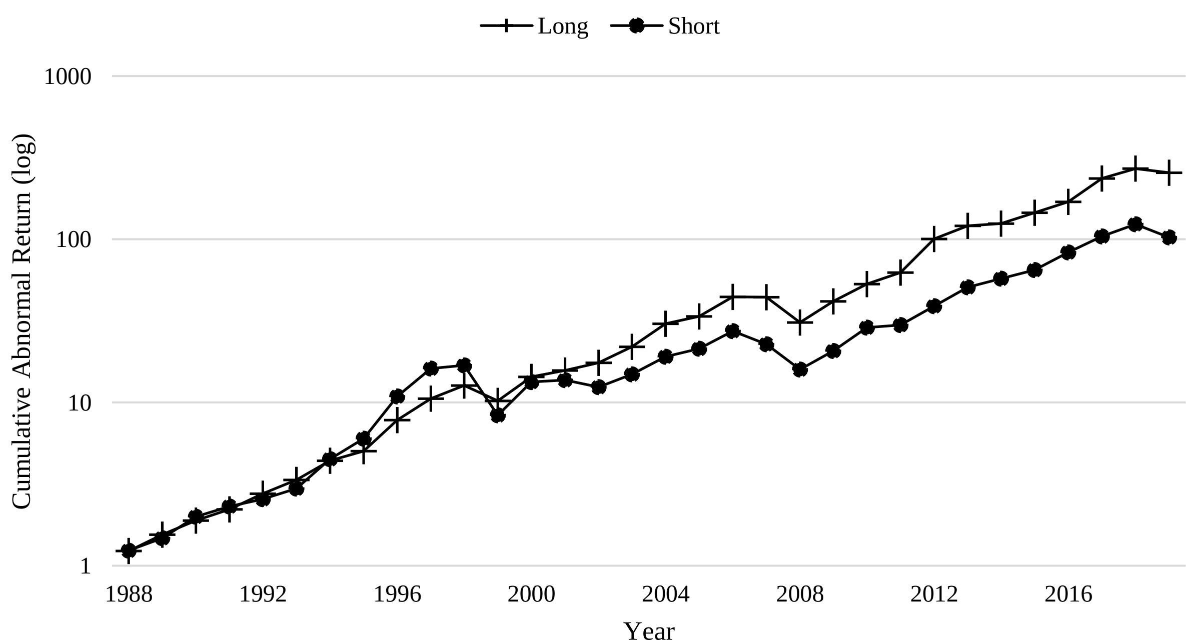


Table 1. Descriptive Statistics

Panels A and B present descriptive statistics for valuation anchors and profitability drivers. All variables are defined in the appendix.

Panel A. Anchors

Variable	N	Mean	StD	P1	P25	Median	P75	P99
P	57,454	4,446	18,506	3	68	423	2,110	75,716
CSE	57,454	1,743	7,881	1	57	270	1,047	24,242
NOA	57,454	3,230	18,290	3	80	400	1,664	43,428
NFO	57,454	1,460	15,174	-1,459	3	71	542	18,875

Panel B. Profitability Drivers from NP's Framework

Variable	N	Mean	StD	P1	P25	Median	P75	P99
ROCE	57,454	0.09	1.39	-0.86	0.05	0.12	0.18	0.74
MSR	57,454	1.00	0.11	0.81	1.00	1.00	1.00	1.10
ROTCE	57,454	0.10	1.37	-0.85	0.05	0.12	0.18	0.72
RNOA	57,454	0.05	16.16	-0.47	0.05	0.10	0.16	0.75
FLEV	57,454	0.97	8.16	-0.73	0.06	0.44	0.98	8.47
SPREAD	57,454	0.01	16.32	-1.02	-0.01	0.04	0.11	1.27
Sales PM	57,454	-0.03	5.46	-0.48	0.02	0.05	0.10	0.36
ATO	57,454	2.76	15.31	0.20	1.11	1.92	2.93	13.80
Other Items/NOA	57,454	0.00	0.19	-0.02	0.00	0.00	0.00	0.06
NBC	57,454	0.04	2.33	-0.66	0.03	0.05	0.07	0.79
Sales PM*	57,454	-0.01	5.26	-0.46	0.03	0.06	0.11	0.46
ATO*	57,454	1.53	2.10	0.10	0.79	1.34	1.93	5.88
OLLEV	57,454	1.03	60.05	0.07	0.27	0.40	0.60	4.82
OLSPREAD	57,454	0.02	0.71	-0.45	-0.01	0.02	0.06	0.28
Other Items/OA	57,454	0.00	0.05	-0.01	0.00	0.00	0.00	0.04
Core Sales PM	57,454	-0.02	4.95	-0.35	0.03	0.05	0.09	0.31
Core Other Items/NOA	57,454	0.00	0.58	-0.22	-0.01	0.00	0.01	0.28
Core RNOA	57,454	0.04	16.21	-0.36	0.05	0.09	0.15	0.67
Core NBC	57,454	0.05	1.93	-0.60	0.03	0.05	0.07	0.76
Core SPREAD	57,454	-0.01	16.33	-0.95	-0.01	0.03	0.10	1.13
Core Sales PM*	57,454	-0.01	4.95	-0.33	0.03	0.06	0.11	0.43
Core Other Items/OA	57,454	0.00	0.15	-0.12	-0.01	0.00	0.01	0.16
UOI/NOA	57,454	0.00	0.52	-0.26	-0.01	0.00	0.01	0.22
UFE/NFO	57,454	-0.01	1.31	-0.05	0.00	0.00	0.00	0.04

Table 2. Correlation Matrix

The table presents correlations among selected variables. Pearson (Spearman) correlations are above (below) the diagonal. * indicates statistical significance at the 5% level. All variables are defined in the appendix.

#	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	ROCE	1.00	0.99*	0.02*	0.24*	0.02*	0.02*	0.00	0.00	0.00	0.08*	0.02*	0.02*	0.00	0.01*	0.10*	0.00
2	ROTCE	1.00*	1.00	0.02*	0.26*	0.02*	0.02*	0.00	0.00	0.00	0.08*	0.02*	0.02*	0.00	0.02*	0.10*	0.00
3	RNOA	0.89*	0.89*	1.00	0.00	0.99*	0.01*	-0.05*	0.00	0.07*	0.04*	0.01*	1.00*	0.00	0.99*	-0.10*	0.00
4	FLEV	-0.06*	-0.07*	-0.34*	1.00	0.00	0.00	-0.01*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	SPREAD	0.74*	0.74*	0.79*	-0.32*	1.00	0.01*	-0.05*	-0.14*	0.07*	0.04*	0.01*	0.99*	-0.12*	1.00*	-0.10*	-0.08*
6	Sales PM	0.66*	0.66*	0.60*	0.06*	0.53*	1.00	0.00	0.00	0.00	0.05*	0.84*	0.01*	0.00	0.01*	0.01	0.00
7	ATO	0.23*	0.23*	0.37*	-0.43*	0.26*	-0.36*	1.00	0.00	0.66*	-0.01	0.00	-0.04*	0.00	-0.04*	-0.20*	0.00
8	NBC	-0.06*	-0.06*	-0.07*	0.26*	-0.52*	-0.10*	0.02*	1.00	0.00	0.00	0.00	0.00	0.83*	-0.10*	0.00	0.56*
9	OLLEV	0.16*	0.16*	0.26*	-0.25*	0.18*	-0.11*	0.48*	-0.03*	1.00	0.00	0.00	0.08*	0.00	0.08*	-0.28*	0.00
10	OLSPREAD	0.74*	0.74*	0.78*	-0.23*	0.70*	0.66*	0.11*	-0.19*	0.20*	1.00	0.04*	0.03*	0.00	0.03*	0.08*	0.00
11	Core Sales PM	0.56*	0.56*	0.50*	0.08*	0.45*	0.89*	-0.40*	-0.11*	-0.09*	0.59*	1.00	0.01*	0.00	0.01*	-0.01	0.00
12	Core RNOA	0.79*	0.80*	0.89*	-0.34*	0.71*	0.52*	0.39*	-0.07*	0.28*	0.71*	0.58*	1.00	0.00	0.99*	-0.13*	0.00
13	Core NBC	-0.05*	-0.05*	-0.07*	0.27*	-0.51*	-0.10*	0.02*	0.97*	-0.04*	-0.19*	-0.11*	-0.08*	1.00	-0.12*	0.00	0.00
14	Core SPREAD	0.64*	0.64*	0.69*	-0.32*	0.91*	0.45*	0.26*	-0.54*	0.20*	0.63*	0.50*	0.77*	-0.55*	1.00	-0.13*	0.00
15	UOI/NOA	0.34*	0.34*	0.37*	-0.10*	0.30*	0.31*	0.08*	-0.01*	-0.02*	0.26*	0.02*	0.08*	-0.01*	0.06*	1.00	0.00
16	UFE/NFO	-0.02*	-0.02*	-0.01	0.00	-0.06*	-0.01*	0.01	0.09*	0.00	-0.01*	-0.01	0.00	-0.01	0.00	-0.01	1.00

Table 3. Median Absolute Forecast Errors from Neural Network Estimation of NP's Framework Compared to a Random Walk

Panel A (Panel B) presents out-of-sample 1-year-ahead to 10-year-ahead median absolute ROCE (RNOA) forecast errors derived from neural network models based on various combinations of financial statement analysis design choices within Nissim and Penman's (2001) profitability framework. Standard errors are computed following Mann and Whitney (1947). *, **, and *** indicate that the median absolute forecast error is significantly smaller at the 10%, 5%, and 1% levels, respectively, when the forecast is based on neural network models as compared to a random walk forecast.

Panel A. ROCE

	Median Absolute ROCE Forecast Error Lead									
	1	2	3	4	5	6	7	8	9	10
Random Walk	0.0638	0.0874	0.1017	0.1150	0.1277	0.1374	0.1500	0.1658	0.1793	0.1931
Disaggregation level 1										
Lag 0, Core 0	0.0626	0.1350	0.0909***	0.1048***	0.1172***	0.1290***	0.1394***	0.1534***	0.1656***	0.1813***
Lag 1, Core 0	0.0596***	0.0774***	0.0916***	0.1052***	0.1164***	0.1296***	0.1403***	0.1534***	0.1676***	0.1817***
Lag 3, Core 0	0.0856	0.0784***	0.0914***	0.1054***	0.1186***	0.1302***	0.1439***	0.1538***	0.1699***	0.1896
Lag 5, Core 0	0.0618***	0.0794***	0.0940***	0.1059***	0.1178***	0.1295***	0.1443*	0.1602**	0.1767	0.1918
Lag 0, Core 1	0.0590***	0.0815***	0.0959***	0.1044***	0.1184***	0.1289***	0.1393***	0.1548***	0.1654***	0.1810***
Lag 1, Core 1	0.0591***	0.1046	0.0912***	0.1049***	0.1153***	0.1282***	0.1389***	0.1549***	0.1657***	0.1850**
Lag 3, Core 1	0.0646	0.0788***	0.0934***	0.1048***	0.1157***	0.1280***	0.1408***	0.1579***	0.1725*	0.1908
Lag 5, Core 1	0.0608***	0.0781***	0.0926***	0.1065***	0.1177***	0.1294***	0.1460	0.1553***	0.1775	0.1883
Disaggregation level 2										
Lag 0, Core 0	0.0593***	0.0786***	0.0914***	0.1041***	0.1172***	0.1294***	0.1404***	0.1534***	0.1685***	0.1796***
Lag 1, Core 0	0.0604***	0.0781***	0.0934***	0.1071***	0.1166***	0.1300***	0.1395***	0.1540***	0.1659***	0.1826***
Lag 3, Core 0	0.0611***	0.0807***	0.0947***	0.1059***	0.1181***	0.1301***	0.1413***	0.1629	0.1748	0.1906
Lag 5, Core 0	0.0602***	0.0812***	0.0937***	0.1105***	0.1197***	0.1343*	0.1484	0.1537***	0.1774	0.1910
Lag 0, Core 1	0.0565***	0.0899	0.0906***	0.1023***	0.1153***	0.1282***	0.1401***	0.1535***	0.1651***	0.1825***
Lag 1, Core 1	0.0580***	0.0756***	0.0912***	0.1039***	0.1158***	0.1304***	0.1401***	0.1534***	0.1667***	0.1837***
Lag 3, Core 1	0.0593***	0.0779***	0.0956***	0.1071***	0.1208***	0.1318***	0.1429***	0.1574***	0.1821	0.1922
Lag 5, Core 1	0.0689	0.0796***	0.0940***	0.1109***	0.1185***	0.1348	0.1498	0.1635	0.1863	0.2001
Disaggregation level 3										
Lag 0, Core 0	0.0594***	0.0792***	0.0914***	0.1046***	0.1175***	0.1288***	0.1426***	0.1554***	0.1736*	0.1849***
Lag 1, Core 0	0.0602***	0.0773***	0.0924***	0.1053***	0.1175***	0.1287***	0.1425***	0.1573***	0.1739	0.1868**
Lag 3, Core 0	0.0624***	0.0828***	0.0941***	0.1089***	0.1219***	0.1275***	0.1480	0.1649	0.1791	0.1830**
Lag 5, Core 0	0.0634**	0.0832***	0.0939***	0.1087***	0.1210***	0.1395	0.1535	0.1573**	0.1820	0.1913
Lag 0, Core 1	0.0565***	0.0750***	0.0903***	0.1035***	0.1164***	0.1298***	0.1395***	0.1535***	0.1698***	0.1860**
Lag 1, Core 1	0.0595***	0.0762***	0.0917***	0.1042***	0.1166***	0.1294***	0.1437***	0.1566***	0.1688***	0.1864*
Lag 3, Core 1	0.0580***	0.0795***	0.0933***	0.1097***	0.1193***	0.1301***	0.1411***	0.1645	0.1791	0.2065
Lag 5, Core 1	0.0603***	0.0793***	0.1027	0.1114**	0.1236***	0.1353*	0.1586	0.1623	0.1874	0.1941
Disaggregation level 4										
Lag 0, Core 0	0.0597***	0.0793***	0.0941***	0.1052***	0.1160***	0.1270***	0.1405***	0.1545***	0.1706***	0.1827***
Lag 1, Core 0	0.0699	0.0792***	0.0948***	0.1049***	0.1158***	0.1301***	0.1407***	0.1587***	0.1762	0.1908
Lag 3, Core 0	0.0615***	0.0814***	0.0957***	0.1085***	0.1212***	0.1362	0.1508	0.1695	0.1845	0.2107
Lag 5, Core 0	0.0681	0.0840***	0.1013	0.1139	0.1223***	0.1339**	0.1578	0.1766	0.1966	0.2243
Lag 0, Core 1	0.0567***	0.0762***	0.0920***	0.1043***	0.1145***	0.1297***	0.1436***	0.1519***	0.1692***	0.1847**
Lag 1, Core 1	0.0575***	0.0785***	0.0921***	0.1045***	0.1175***	0.1300***	0.1419***	0.1629	0.1699***	0.1948
Lag 3, Core 1	0.0585***	0.0797***	0.0963***	0.1037***	0.1167***	0.1343	0.1486	0.1702	0.1882	0.2018
Lag 5, Core 1	0.0615***	0.0815***	0.0989	0.1124	0.1163***	0.1390	0.1475	0.1791	0.1861	0.2006

Panel B. RNOA

	Median Absolute RNOA Forecast Error Lead									
	1	2	3	4	5	6	7	8	9	10
Random Walk	0.0462	0.0632	0.0743	0.0841	0.0937	0.1005	0.1097	0.1217	0.1306	0.1399
Disaggregation level 1										
Lag 0, Core 0	0.0570	0.0670	0.0762	0.0838	0.0924	0.1019	0.1079	0.1171	0.1261	0.1375
Lag 1, Core 0	0.0557	0.0663	0.0748	0.0827	0.0917	0.1015	0.1097	0.1183	0.1280	0.1385
Lag 3, Core 0	0.0571	0.0679	0.0754	0.0846	0.0931	0.1008	0.1097	0.1188	0.1306	0.1465
Lag 5, Core 0	0.0566	0.0692	0.0758	0.0866	0.0951	0.1033	0.1132	0.1242	0.1339	0.1459
Lag 0, Core 1	0.0567	0.0675	0.0758	0.0838	0.0924	0.1023	0.1105	0.1181	0.1276	0.1377
Lag 1, Core 1	0.0611	0.0686	0.0755	0.0834	0.0922	0.0997	0.1086	0.1176	0.1264	0.1378
Lag 3, Core 1	0.0568	0.0674	0.0764	0.0841*	0.0936	0.1025	0.1091	0.1191	0.1298	0.1424
Lag 5, Core 1	0.0575	0.0693	0.0771	0.0858	0.0954	0.1030	0.1127	0.1429	0.1317	0.1461
Disaggregation level 2										
Lag 0, Core 0	0.0427***	0.0565***	0.0680***	0.0776***	0.0876***	0.0959***	0.1055***	0.1158*	0.1230***	0.1348
Lag 1, Core 0	0.0430***	0.0572***	0.0685***	0.0769***	0.0880***	0.0964**	0.1072	0.1157**	0.1241**	0.1377
Lag 3, Core 0	0.0439***	0.0590***	0.0685***	0.0792***	0.0880***	0.0993	0.1103	0.1165	0.1274	0.1408
Lag 5, Core 0	0.0459	0.0595***	0.0710**	0.0835	0.0908	0.1004	0.1127	0.1236	0.1318	0.1484
Lag 0, Core 1	0.0414***	0.0554***	0.0659***	0.0760***	0.0850***	0.0954***	0.1042***	0.1166	0.1242*	0.1353
Lag 1, Core 1	0.0415***	0.0559***	0.0664***	0.0787***	0.0887***	0.0973	0.1061	0.1144**	0.1274	0.1365
Lag 3, Core 1	0.0430***	0.0567***	0.0692***	0.0791***	0.0896*	0.0975	0.1092	0.1220	0.1311	0.1485
Lag 5, Core 1	0.0446***	0.0592***	0.0722	0.0828	0.0905	0.1000	0.1096	0.1202	0.1403	0.1424
Disaggregation level 3										
Lag 0, Core 0	0.0425***	0.0561***	0.0669***	0.0782***	0.0862***	0.0960***	0.1052**	0.1155**	0.1276	0.1359
Lag 1, Core 0	0.0429***	0.0566***	0.0683***	0.0774***	0.0871***	0.0957**	0.1066*	0.1183	0.1284	0.1379
Lag 3, Core 0	0.0438***	0.0580***	0.0698***	0.0809**	0.0898***	0.1042	0.1064	0.1220	0.1326	0.1500
Lag 5, Core 0	0.0466	0.0605***	0.0718	0.0810	0.0924	0.1018	0.1124	0.1212	0.1423	0.1483
Lag 0, Core 1	0.0416***	0.0547***	0.0665***	0.0758***	0.0858***	0.0955***	0.1057**	0.1170	0.1258	0.1370
Lag 1, Core 1	0.0412***	0.0569***	0.0674***	0.0790***	0.0873***	0.0955**	0.1068	0.1189	0.1305	0.1403
Lag 3, Core 1	0.0421***	0.0567***	0.0701***	0.0815*	0.0948	0.1021	0.1132	0.1248	0.1401	0.1552
Lag 5, Core 1	0.0439***	0.0591***	0.0729	0.0838	0.0920	0.1063	0.1079	0.1234	0.1380	0.1545
Disaggregation level 4										
Lag 0, Core 0	0.0429***	0.0573***	0.0687***	0.0781***	0.0865***	0.0940***	0.1041***	0.1151***	0.1274	0.1367
Lag 1, Core 0	0.0430***	0.0567***	0.0699***	0.0785***	0.0879***	0.0980**	0.1062**	0.1174	0.1319	0.1435
Lag 3, Core 0	0.0446**	0.0586***	0.0684***	0.0818	0.0900*	0.1009	0.1122	0.1238	0.1376	0.1485
Lag 5, Core 0	0.0462	0.0605*	0.0737	0.0827	0.0924	0.1068	0.1137	0.1384	0.1468	0.1593
Lag 0, Core 1	0.0419***	0.0548***	0.0674***	0.0773***	0.0871***	0.0944***	0.1052***	0.1150***	0.1254*	0.1404
Lag 1, Core 1	0.0486	0.0571***	0.0681***	0.0775***	0.0857***	0.0966**	0.1090	0.1184	0.1304	0.1461
Lag 3, Core 1	0.0435***	0.0602***	0.0708**	0.0804***	0.0890***	0.1015	0.1135	0.1166	0.1362	0.1534
Lag 5, Core 1	0.0618	0.0595***	0.0756	0.0822	0.0937*	0.1101	0.1218	0.1303	0.1412	0.1562

Table 4. Association between Financial Statement Analysis Design Choices and Median Absolute Profitability Forecast Errors from Neural Network Estimation of NP's Framework

Panel A (Panel B) regresses median absolute ROCE (RNOA) forecast errors on indicators for levels of ratio disaggregation (*Level 2 to 4*), an indicator that the model focuses on core items (*Core*), and indicators for lags of financial statement information (*Lags 1 to 5*). Robust t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A. ROCE

	Median Absolute ROCE Forecast Error Lead									
	1	2	3	4	5	6	7	8	9	10
Level 2	-0.004	-0.009	0.000	0.001	0.001	0.002**	0.001	0.001	0.003*	0.002
	(-1.160)	(-1.216)	(0.428)	(1.508)	(0.849)	(2.275)	(1.286)	(0.539)	(2.003)	(1.135)
Level 3	-0.004	-0.010	0.001	0.002*	0.002**	0.002	0.005***	0.004**	0.007***	0.004
	(-1.462)	(-1.359)	(0.825)	(1.951)	(2.597)	(1.632)	(3.055)	(2.085)	(4.890)	(1.332)
Level 4	-0.002	-0.009	0.003***	0.002*	0.000	0.003**	0.005***	0.010***	0.010***	0.013***
	(-0.780)	(-1.240)	(2.808)	(1.917)	(0.444)	(2.700)	(3.178)	(3.784)	(5.176)	(3.489)
Core	-0.004**	-0.003	0.000	-0.001	-0.001*	0.000	-0.001	0.001	-0.000	0.001
	(-2.135)	(-0.694)	(0.229)	(-1.008)	(-1.770)	(0.279)	(-0.688)	(0.514)	(-0.153)	(0.439)
Lags 1	0.002	-0.006	0.000	0.001	-0.000	0.001	0.000	0.003	0.001	0.004*
	(1.269)	(-0.831)	(0.286)	(1.302)	(-0.185)	(0.941)	(0.297)	(1.381)	(0.570)	(1.924)
Lags 3	0.005*	-0.007	0.002**	0.003***	0.002***	0.002*	0.004***	0.009***	0.010***	0.013***
	(1.901)	(-1.025)	(2.413)	(3.106)	(2.916)	(2.052)	(3.189)	(4.274)	(6.368)	(4.082)
Lags 5	0.004**	-0.006	0.004***	0.006***	0.003***	0.006***	0.010***	0.010***	0.015***	0.015***
	(2.538)	(-0.886)	(2.967)	(6.178)	(3.319)	(4.415)	(6.101)	(3.409)	(7.792)	(4.051)
Observations	32	32	32	32	32	32	32	32	32	32
R-Squared	0.359	0.224	0.536	0.705	0.589	0.621	0.772	0.687	0.852	0.669

Panel B. RNOA

	Median Absolute RNOA Forecast Error Lead									
	1	2	3	4	5	6	7	8	9	10
Level 2	-0.014***	-0.010***	-0.007***	-0.005***	-0.005***	-0.004***	-0.002**	-0.004	-0.001	-0.001
	(-13.532)	(-23.032)	(-10.394)	(-9.361)	(-8.002)	(-3.678)	(-2.274)	(-1.619)	(-0.422)	(-0.651)
Level 3	-0.014***	-0.011***	-0.007***	-0.005***	-0.004***	-0.002*	-0.002*	-0.002	0.004**	0.003**
	(-13.230)	(-22.615)	(-10.088)	(-7.505)	(-4.864)	(-1.778)	(-1.850)	(-0.746)	(2.732)	(2.302)
Level 4	-0.011***	-0.010***	-0.006***	-0.005***	-0.004***	-0.002	0.001	-0.000	0.005***	0.006***
	(-4.965)	(-17.591)	(-6.880)	(-10.091)	(-7.446)	(-1.072)	(0.373)	(-0.054)	(3.628)	(4.302)
Core	0.001	-0.000	0.000	-0.000	0.000	0.000	0.001	0.001	0.000	0.001
	(0.691)	(-1.537)	(0.229)	(-0.390)	(0.458)	(0.212)	(0.693)	(0.531)	(0.394)	(1.140)
Lags 1	0.001	0.001	0.000	0.000	0.001	0.001	0.001	0.001	0.003**	0.003**
	(1.208)	(1.543)	(0.721)	(0.808)	(1.109)	(0.599)	(1.573)	(1.241)	(2.344)	(2.601)
Lags 3	0.001	0.002***	0.002**	0.003***	0.003***	0.004***	0.004***	0.004**	0.007***	0.011***
	(1.164)	(3.703)	(2.742)	(4.746)	(3.698)	(3.584)	(3.838)	(2.640)	(5.479)	(7.005)
Lags 5	0.005**	0.003***	0.004***	0.005***	0.005***	0.007***	0.007***	0.012***	0.012***	0.013***
	(2.208)	(8.766)	(5.469)	(7.611)	(8.257)	(4.593)	(4.582)	(4.058)	(7.145)	(8.140)
Observations	32	32	32	32	32	32	32	32	32	32
R-Squared	0.822	0.972	0.896	0.900	0.826	0.684	0.645	0.599	0.823	0.850

Table 5. Median Absolute Forecast Errors from Neural Network Estimation of NP's Framework Compared to Linear Estimation

Panel A (Panel B) presents out-of-sample 1-year-ahead to 10-year-ahead median absolute ROCE (RNOA) forecast errors of linear (OLS) models based on various combinations of financial statement analysis design choices within Nissim and Penman's (2001) profitability framework. Standard errors are computed following Mann and Whitney (1947). *, **, and *** indicate that the median absolute neural network profitability forecast error reported in Table 3 is significantly smaller at the 10%, 5%, and 1% levels, respectively, than the OLS forecast errors reported in this table.

Panel A. ROCE

	Median Absolute ROCE Forecast Error Lead									
	1	2	3	4	5	6	7	8	9	10
Disaggregation level 1										
Lag 0, Core 0	0.0937***	0.0898	0.1227***	0.1248***	0.1324***	0.2254***	0.1556***	0.2161***	0.2370***	0.2920***
Lag 1, Core 0	0.0932***	0.0899***	0.1229***	0.1250***	0.1329***	0.2246***	0.1566***	0.2170***	0.2359***	0.2885***
Lag 3, Core 0	0.0920***	0.0906***	0.1227***	0.1246***	0.1329***	0.2223***	0.1580***	0.2161***	0.2365***	0.2888***
Lag 5, Core 0	0.0916***	0.0906***	0.1228***	0.1246***	0.1327***	0.2226***	0.1589***	0.2163***	0.2360***	0.2872***
Lag 0, Core 1	0.0937***	0.0898***	0.1227***	0.1248***	0.1324***	0.2254***	0.1556***	0.2161***	0.2370***	0.2920***
Lag 1, Core 1	0.0932***	0.0899	0.1229***	0.1250***	0.1329***	0.2246***	0.1566***	0.2170***	0.2359***	0.2885***
Lag 3, Core 1	0.0920***	0.0906***	0.1227***	0.1246***	0.1329***	0.2223***	0.1580***	0.2161***	0.2365***	0.2888***
Lag 5, Core 1	0.0916***	0.0906***	0.1228***	0.1246***	0.1327***	0.2226***	0.1589***	0.2163***	0.2360***	0.2872***
Disaggregation level 2										
Lag 0, Core 0	0.0909***	0.1422***	0.1569***	0.1522***	0.1374***	0.2399***	0.1719***	0.2273***	0.2301***	0.2924***
Lag 1, Core 0	0.0905***	0.1381***	0.1565***	0.1470***	0.1370***	0.2336***	0.1683***	0.2241***	0.2259***	0.2851***
Lag 3, Core 0	0.0890***	0.1303***	0.1477***	0.1399***	0.1361***	0.2263***	0.1655***	0.2204***	0.2242***	0.2761***
Lag 5, Core 0	0.0889***	0.1236***	0.1388***	0.1353***	0.1367***	0.2149***	0.1645***	0.2156***	0.2191***	0.2675***
Lag 0, Core 1	0.0867***	0.1377***	0.1530***	0.1473***	0.1345***	0.2376***	0.1644***	0.2266***	0.2270***	0.2877***
Lag 1, Core 1	0.0851***	0.1311***	0.1504***	0.1408***	0.1347***	0.2320***	0.1627***	0.2235***	0.2249***	0.2821***
Lag 3, Core 1	0.0839***	0.1245***	0.1424***	0.1348***	0.1338***	0.2250***	0.1631***	0.2211***	0.2248***	0.2744***
Lag 5, Core 1	0.0839***	0.1182***	0.1338***	0.1307***	0.1342***	0.2125***	0.1622***	0.2163***	0.2196***	0.2675***
Disaggregation level 3										
Lag 0, Core 0	0.0924***	0.1426***	0.1579***	0.1532***	0.1379***	0.2393***	0.1717***	0.2266***	0.2312***	0.2922***
Lag 1, Core 0	0.0923***	0.1404***	0.1596***	0.1502***	0.1378***	0.2342***	0.1683***	0.2225***	0.2260***	0.2854***
Lag 3, Core 0	0.0918***	0.1362***	0.1537***	0.1464***	0.1380***	0.2278***	0.1687***	0.2189***	0.2255***	0.2797***
Lag 5, Core 0	0.0924***	0.1287***	0.1442***	0.1416***	0.1381***	0.2170***	0.1668***	0.2149***	0.2195***	0.2683***
Lag 0, Core 1	0.0915***	0.1425***	0.1579***	0.1526***	0.1382***	0.2400***	0.1721***	0.2281***	0.2313***	0.2925***
Lag 1, Core 1	0.0924***	0.1402***	0.1593***	0.1496***	0.1379***	0.2345***	0.1684***	0.2223***	0.2269***	0.2844***
Lag 3, Core 1	0.0930***	0.1363***	0.1538***	0.1460***	0.1385***	0.2306***	0.1684***	0.2219***	0.2267***	0.2772***
Lag 5, Core 1	0.0924***	0.1292***	0.1442***	0.1416***	0.1380***	0.2154***	0.1666***	0.2156***	0.2209***	0.2662***
Disaggregation level 4										
Lag 0, Core 0	0.0867***	0.1348***	0.1500***	0.1449***	0.1361***	0.2415***	0.1624***	0.2241***	0.2253***	0.2875***
Lag 1, Core 0	0.0851***	0.1196***	0.1399***	0.1355***	0.1367***	0.2314***	0.1612***	0.2232***	0.2241***	0.2796***
Lag 3, Core 0	0.0873***	0.1177***	0.1341***	0.1346***	0.1402***	0.2199***	0.1893***	0.2232***	0.2290***	0.2675***
Lag 5, Core 0	0.0875***	0.1145***	0.1307***	0.1337***	0.1407***	0.2113***	0.1887***	0.2227***	0.2258***	0.2647***
Lag 0, Core 1	0.0849***	0.1317***	0.1471***	0.1439***	0.1368***	0.2397***	0.1615***	0.2249***	0.2277***	0.2878***
Lag 1, Core 1	0.0850***	0.1200***	0.1399***	0.1359***	0.1390***	0.2320***	0.1628***	0.2248***	0.2269***	0.2792***
Lag 3, Core 1	0.0886***	0.1168***	0.1361***	0.1351***	0.1443***	0.2286***	0.1915***	0.2290***	0.2328***	0.2714***
Lag 5, Core 1	0.0892***	0.1131***	0.1315***	0.1349***	0.1442***	0.2165***	0.1916***	0.2255***	0.2297***	0.2688***

Panel B. RNOA

	Median Absolute RNOA Forecast Error Lead									
	1	2	3	4	5	6	7	8	9	10
Disaggregation level 1										
Lag 0, Core 0	0.0866***	0.0971***	0.1094***	0.1220***	0.1362***	0.1508***	0.1724***	0.2011***	0.2219***	0.2507***
Lag 1, Core 0	0.0871***	0.0970***	0.1089***	0.1216***	0.1356***	0.1504***	0.1717***	0.2000***	0.2200***	0.2503***
Lag 3, Core 0	0.0866***	0.0963***	0.1082***	0.1210***	0.1355***	0.1500***	0.1710***	0.1997***	0.2194***	0.2498***
Lag 5, Core 0	0.0865***	0.0958***	0.1084***	0.1210***	0.1353***	0.1502***	0.1711***	0.1998***	0.2197***	0.2500***
Lag 0, Core 1	0.0866***	0.0971***	0.1094***	0.1220***	0.1362***	0.1508***	0.1724***	0.2011***	0.2219***	0.2507***
Lag 1, Core 1	0.0871***	0.0970***	0.1089***	0.1216***	0.1356***	0.1504***	0.1717***	0.2000***	0.2200***	0.2503***
Lag 3, Core 1	0.0866***	0.0963***	0.1082***	0.1210***	0.1355***	0.1500***	0.1710***	0.1997***	0.2194***	0.2498***
Lag 5, Core 1	0.0865***	0.0958***	0.1084***	0.1210***	0.1353***	0.1502***	0.1711***	0.1998***	0.2197***	0.2500***
Disaggregation level 2										
Lag 0, Core 0	0.0826***	0.0984***	0.0996***	0.1196***	0.1316***	0.1445***	0.1874***	0.1981***	0.1905***	0.3038***
Lag 1, Core 0	0.0799***	0.0951***	0.0976***	0.1160***	0.1292***	0.1401***	0.1696***	0.2237***	0.2140***	0.3826***
Lag 3, Core 0	0.0767***	0.0912***	0.0955***	0.1143***	0.1283***	0.1376***	0.1650***	0.2221***	0.2146***	0.3724***
Lag 5, Core 0	0.0774***	0.0909***	0.0960***	0.1153***	0.1305***	0.1387***	0.1660***	0.2331***	0.2188***	0.3849***
Lag 0, Core 1	0.0685***	0.0837***	0.0996***	0.1172***	0.1329***	0.1419***	0.1601***	0.2856***	0.2432***	0.4116***
Lag 1, Core 1	0.0671***	0.0819***	0.0984***	0.1145***	0.1292***	0.1407***	0.1786***	0.2888***	0.2824***	0.4275***
Lag 3, Core 1	0.0671***	0.0809***	0.0987***	0.1155***	0.1322***	0.1422***	0.1790***	0.2805***	0.2725***	0.4129***
Lag 5, Core 1	0.0676***	0.0814***	0.1004***	0.1165***	0.1338***	0.1445***	0.1805***	0.2848***	0.2744***	0.4178***
Disaggregation level 3										
Lag 0, Core 0	0.1083***	0.1033***	0.1125***	0.1223***	0.1285***	0.1436***	0.1783***	0.1549***	0.1782***	0.1716***
Lag 1, Core 0	0.0977***	0.0968***	0.1054***	0.1149***	0.1221***	0.1369***	0.1728***	0.1520***	0.1736***	0.1697***
Lag 3, Core 0	0.0914***	0.0946***	0.1014***	0.1122***	0.1219***	0.1386***	0.1865***	0.1691***	0.1940***	0.2000***
Lag 5, Core 0	0.0892***	0.0926***	0.1007***	0.1109***	0.1214***	0.1380***	0.1853***	0.1706***	0.1974***	0.2059***
Lag 0, Core 1	0.1075***	0.1107***	0.1180***	0.1276***	0.1356***	0.1496***	0.1819***	0.1974***	0.2062***	0.2338***
Lag 1, Core 1	0.1012***	0.1052***	0.1126***	0.1227***	0.1306***	0.1446***	0.1757***	0.2173***	0.2067***	0.2681***
Lag 3, Core 1	0.0975***	0.1028***	0.1083***	0.1202***	0.1292***	0.1442***	0.1807***	0.2224***	0.2154***	0.2849***
Lag 5, Core 1	0.0951***	0.1042***	0.1078***	0.1199***	0.1313***	0.1486***	0.1866***	0.2422***	0.2386***	0.3151***
Disaggregation level 4										
Lag 0, Core 0	0.0771***	0.0795***	0.0899***	0.1053***	0.1320***	0.1454***	0.1747***	0.1595***	0.2078***	0.2242***
Lag 1, Core 0	0.0747***	0.0831***	0.0935***	0.1104***	0.1415***	0.1601***	0.2130***	0.1839***	0.2315***	0.2421***
Lag 3, Core 0	0.0822***	0.0930***	0.1089***	0.1284***	0.1631***	0.1852***	0.2350***	0.2372***	0.3076***	0.3151***
Lag 5, Core 0	0.0873***	0.0998***	0.1147***	0.1317***	0.1642***	0.1850***	0.2403***	0.2461***	0.3145***	0.3197***
Lag 0, Core 1	0.0768***	0.0853***	0.0922***	0.1073***	0.1385***	0.1564***	0.1901***	0.4367***	0.3581***	0.6634***
Lag 1, Core 1	0.0839***	0.0881***	0.0963***	0.1151***	0.1476***	0.1650***	0.2092***	0.4265***	0.3527***	0.6538***
Lag 3, Core 1	0.0979***	0.1055***	0.1119***	0.1375***	0.1676***	0.1885***	0.2318***	0.5051***	0.3969***	0.7282***
Lag 5, Core 1	0.1011***	0.1124***	0.1181***	0.1397***	0.1671***	0.1893***	0.2391***	0.5397***	0.4277***	0.7753***

Table 6. Association Between Financial Statement Analysis Design Choices and Nonlinearities in Neural Network Estimation of Nissim and Penman's (2001) Framework

Panel A (Panel B) regresses median absolute ROCE (RNOA) forecast errors reported in Panel A (Panel B) of Table 3 (neural network estimation) and Table 6 (OLS estimation) on indicators for ratio disaggregation (*Level 2 to 3*), an indicator that the model focuses on core items (*Core*), and indicators for lags of financial statement information (*Lags 1 to 5*) interacted with an indicator that the model is estimated via a neural network rather than OLS (*NN*). Main coefficients for design choices are suppressed to enhance readability. Robust t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. ROCE

	Lead									
	1	2	3	4	5	6	7	8	9	10
NN	-0.030***	-0.002	-0.038***	-0.028***	-0.016***	-0.105***	-0.017***	-0.069***	-0.075***	-0.121***
	(-10.236)	(-0.194)	(-12.960)	(-11.053)	(-13.061)	(-30.903)	(-5.170)	(-26.726)	(-36.857)	(-34.432)
Level 2 × NN	0.002	-0.049***	-0.024***	-0.015***	-0.002**	-0.002	-0.007***	-0.004*	0.015***	0.012***
	(0.475)	(-6.368)	(-9.292)	(-6.832)	(-2.163)	(-0.726)	(-2.762)	(-2.002)	(7.838)	(4.265)
Level 3 × NN	-0.004	-0.057***	-0.030***	-0.021***	-0.003***	-0.004	-0.007***	-0.001	0.017***	0.012***
	(-1.319)	(-7.352)	(-11.834)	(-11.759)	(-3.658)	(-1.434)	(-2.742)	(-0.677)	(9.686)	(3.278)
Level 4 × NN	0.003	-0.040***	-0.013***	-0.011***	-0.007***	-0.001	-0.014***	0.002	0.019***	0.026***
	(1.034)	(-5.120)	(-5.080)	(-5.452)	(-4.665)	(-0.169)	(-2.878)	(0.563)	(7.893)	(5.958)
Core × NN	-0.003	-0.001	0.001	0.001	-0.001	-0.000	0.000	-0.000	-0.001	0.001
	(-1.409)	(-0.240)	(0.917)	(0.557)	(-1.318)	(-0.129)	(0.017)	(-0.107)	(-0.666)	(0.534)
Lags 1 × NN	0.002	-0.001	0.002	0.005***	-0.001	0.006***	0.002	0.005*	0.003*	0.010***
	(1.364)	(-0.094)	(1.097)	(2.900)	(-0.497)	(2.872)	(0.391)	(1.991)	(1.726)	(3.839)
Lags 3 × NN	0.006*	0.002	0.009***	0.010***	0.001	0.013***	-0.002	0.012***	0.012***	0.025***
	(1.941)	(0.222)	(4.552)	(5.533)	(0.842)	(5.605)	(-0.471)	(4.710)	(5.582)	(6.927)
Lags 5 × NN	0.005**	0.007	0.017***	0.015***	0.002	0.025***	0.005	0.016***	0.020***	0.033***
	(2.467)	(0.936)	(5.978)	(7.237)	(1.202)	(7.851)	(1.077)	(4.701)	(7.971)	(7.537)
Observations	64	64	64	64	64	64	64	64	64	64
R-Squared	0.952	0.912	0.988	0.983	0.976	0.996	0.870	0.990	0.990	0.990

Panel B. RNOA

	Lead									
	1	2	3	4	5	6	7	8	9	10
NN	-0.032***	-0.029***	-0.032***	-0.036***	-0.040***	-0.046***	-0.056***	-0.017	-0.053***	-0.018
	(-9.883)	(-8.564)	(-11.397)	(-10.451)	(-11.111)	(-9.917)	(-7.531)	(-0.503)	(-2.705)	(-0.362)
Level 2 × NN	-0.001	-0.002	0.003**	0.000	0.000	0.005*	-0.004	-0.056**	-0.019	-0.140***
	(-0.269)	(-0.596)	(2.239)	(0.096)	(0.027)	(1.811)	(-0.719)	(-2.453)	(-1.454)	(-3.851)
Level 3 × NN	-0.026***	-0.015***	-0.006**	-0.002	0.004	0.005*	-0.012***	0.008	0.023*	0.022
	(-9.289)	(-6.172)	(-2.366)	(-0.760)	(1.564)	(1.713)	(-3.272)	(0.331)	(1.725)	(0.685)
Level 4 × NN	-0.009**	-0.007	-0.000	-0.005	-0.021***	-0.023***	-0.045***	-0.142***	-0.099***	-0.234***
	(-2.138)	(-1.637)	(-0.011)	(-1.097)	(-4.615)	(-4.137)	(-5.877)	(-3.330)	(-4.652)	(-3.491)
Core × NN	0.000	-0.002	-0.003	-0.003	-0.004	-0.004	-0.001	-0.098***	-0.052***	-0.143***
	(0.131)	(-0.717)	(-1.188)	(-1.253)	(-1.358)	(-1.158)	(-0.130)	(-4.407)	(-4.606)	(-4.027)
Lags 1 × NN	0.003	0.002	0.002	0.001	0.001	0.000	-0.004	-0.006	-0.007	-0.014
	(0.868)	(0.544)	(0.459)	(0.321)	(0.178)	(0.009)	(-0.579)	(-0.204)	(-0.380)	(-0.283)
Lags 3 × NN	0.002	0.001	0.000	-0.001	-0.002	-0.002	-0.008	-0.021	-0.019	-0.027
	(0.525)	(0.304)	(0.092)	(-0.166)	(-0.480)	(-0.473)	(-1.090)	(-0.671)	(-1.206)	(-0.544)
Lags 5 × NN	0.005	0.001	0.001	0.001	-0.001	-0.001	-0.008	-0.023	-0.023	-0.038
	(1.076)	(0.269)	(0.334)	(0.148)	(-0.230)	(-0.122)	(-1.025)	(-0.704)	(-1.325)	(-0.728)
Observations	64	64	64	64	64	64	64	64	64	64
R-Squared	0.947	0.933	0.947	0.946	0.966	0.962	0.964	0.828	0.928	0.828

Table 7. Cross-Sectional Analysis of the Importance of Different Financial Statement Analysis Design Choices

Panel A (Panel B) presents fully interacted firm-year-model level median regressions of 1-year-ahead absolute ROCE (RNOA) NN forecast errors on firm-year fixed effects and indicators for ratio disaggregation (*Level 2 to 3*), an indicator that the model focuses on core items (*Core*), and indicators for lags of financial statement information (*Lags 1 to 5*) interacted with indicators that the firm-year observations' ROCE is an extreme ROCE decile (*Outlier*), that the firm operates in a competitive industry (*Competition*), and for different stages of a firm's lifecycle (*Introduction*, *Growth*, *Maturity*, *Decline*, *Shakeout*) measured following Dickinson (2011). We suppress the main coefficients of different financial statement analysis design choices to enhance readability. Robust t-statistics clustered at the firm level are reported in parentheses. Standard errors are clustered by firm. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. ROCE

	Cross Sectional Variable						
	Outlier	Competition	Introduction	Growth	Maturity	Decline	Shakeout
Level 2 × Cross Sectional Variable	-0.003***	-0.000	0.000	0.000**	0.000**	-0.006***	-0.002***
	(-7.103)	(-1.315)	(0.769)	(2.041)	(2.296)	(-5.064)	(-4.551)
Level 3 × Cross Sectional Variable	-0.003***	-0.000***	0.001**	0.000**	0.000	-0.004***	-0.002***
	(-5.593)	(-2.655)	(2.426)	(2.217)	(1.423)	(-3.176)	(-6.057)
Level 4 × Cross Sectional Variable	-0.002***	-0.001***	0.002***	0.000	0.000	-0.003***	-0.002***
	(-4.479)	(-4.346)	(4.830)	(1.075)	(0.515)	(-2.612)	(-3.961)
Core × Cross Sectional Variable	-0.005***	0.000	-0.001	0.000**	0.001***	-0.007***	-0.002***
	(-10.279)	(0.894)	(-1.480)	(2.042)	(4.239)	(-6.446)	(-5.596)
Lags 1 × Cross Sectional Variable	-0.000	0.000	0.000	0.000	0.000	-0.001*	-0.000
	(-0.694)	(1.367)	(0.124)	(0.310)	(0.607)	(-1.887)	(-0.591)
Lags 3 × Cross Sectional Variable	0.001**	0.000	-0.001	-0.000	0.000	-0.001	0.000
	(2.292)	(0.152)	(-1.181)	(-0.803)	(1.365)	(-0.807)	(0.390)
Lags 5 × Cross Sectional Variable	0.003***	-0.000	0.001**	-0.000***	-0.000	0.002**	0.001***
	(5.283)	(-1.480)	(2.396)	(-3.166)	(-0.333)	(2.281)	(2.850)
Observations	991,680	991,680	991,680	991,680	991,680	991,680	991,680
R-Squared	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Panel B. RNOA

	Cross Sectional Variable						
	Outlier	Competition	Introduction	Growth	Maturity	Decline	Shakeout
Level 2 × Cross Sectional Variable	-0.014***	-0.000	0.001	0.001***	-0.000	-0.003**	-0.002***
	(-9.645)	(-0.485)	(1.422)	(3.131)	(-1.151)	(-2.208)	(-3.238)
Level 3 × Cross Sectional Variable	-0.014***	-0.000	0.002**	0.001***	-0.000	-0.002*	-0.003***
	(-9.533)	(-0.987)	(2.252)	(3.197)	(-1.215)	(-1.780)	(-3.850)
Level 4 × Cross Sectional Variable	-0.012***	-0.001	0.003***	0.001***	-0.001*	-0.001	-0.002***
	(-7.977)	(-1.258)	(2.970)	(2.872)	(-1.746)	(-0.635)	(-3.042)
Core × Cross Sectional Variable	-0.003***	0.000	-0.001	0.000	0.001***	-0.006***	-0.002***
	(-7.884)	(1.605)	(-1.362)	(0.489)	(5.762)	(-6.592)	(-5.255)
Lags 1 × Cross Sectional Variable	-0.001*	-0.000	-0.000	-0.000	0.000	-0.001*	-0.000
	(-1.669)	(-0.335)	(-0.161)	(-0.072)	(1.051)	(-1.874)	(-0.517)
Lags 3 × Cross Sectional Variable	-0.001***	-0.000	-0.001	-0.000	0.000	-0.002**	0.000
	(-2.607)	(-0.858)	(-1.503)	(-0.940)	(1.418)	(-2.236)	(1.059)
Lags 5 × Cross Sectional Variable	0.002***	-0.001***	0.000	-0.000	0.000	-0.001	0.000
	(4.539)	(-3.189)	(0.562)	(-1.249)	(0.024)	(-0.729)	(1.617)
Observations	991,680	991,680	991,680	991,680	991,680	991,680	991,680
R-Squared	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 8. Cross-Sectional Analysis of the Importance of Nonlinear Estimation

Panel A (Panel B) presents fully interacted firm-year-model level median regressions of 1-year-ahead absolute ROCE (RNOA) forecast errors on firm-year-model fixed effects and an indicator that the model is estimated via a neural network rather than OLS (*NN*) interacted with indicators that the firm-year observations' ROCE is an extreme ROCE decile (*Outlier*), that the firm operates in a competitive industry (*Competition*), and for different stages of a firm's lifecycle (*Introduction*, *Growth*, *Maturity*, *Decline*, *Shakeout*) measured following Dickinson (2011). Robust t-statistics clustered at the firm level are reported in parentheses. Standard errors are clustered by firm. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. ROCE

	Cross Sectional Variable						
	Outlier	Competition	Introduction	Growth	Maturity	Decline	Shakeout
Cross Sectional Variable × NN	-0.083***	0.009***	-0.031***	0.005***	0.005***	-0.040***	-0.008***
	(-35.364)	(7.285)	(-9.313)	(5.017)	(4.474)	(-8.864)	(-4.429)
NN	-0.018***	-0.029***	-0.023***	-0.025***	-0.026***	-0.023***	-0.023***
	(-31.530)	(-29.442)	(-31.897)	(-30.902)	(-28.233)	(-32.502)	(-31.371)
Observations	1,983,360	1,983,360	1,983,360	1,983,360	1,983,360	1,983,360	1,983,360
R-Squared	0.001	0.000	0.000	0.000	0.000	0.000	0.000

Panel B. RNOA

	Cross Sectional Variable						
	Outlier	Competition	Introduction	Growth	Maturity	Decline	Shakeout
Cross Sectional Variable × NN	-0.046***	0.000	-0.011***	0.003***	0.002***	-0.022***	-0.008***
	(-26.016)	(0.293)	(-4.910)	(4.560)	(3.835)	(-6.865)	(-6.495)
NN	-0.028***	-0.032***	-0.031***	-0.032***	-0.033***	-0.031***	-0.031***
	(-63.466)	(-48.781)	(-67.412)	(-62.478)	(-56.110)	(-67.881)	(-65.808)
Observations	1,983,360	1,983,360	1,983,360	1,983,360	1,983,360	1,983,360	1,983,360
R-Squared	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 9. Alphas Based on Neural Network Estimation of NP's Framework

The table presents Fama and French (2015) alphas for portfolios based on profitability forecasts derived from various combinations of financial statement analysis design choices within Nissim and Penman's (2001) profitability framework. Financial statement analysis design choices include ratio disaggregation (*Level 2 to 3*), focus on core items (*Core*), and lags of financial statement information (*Lag 1 to 5*). Standard errors are computed following Newey and West (1987) with a lag order of 3. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A. ROCE

	Core 0, Level 1	Core 1, Level 1	Core 0, Level 2	Core 1, Level 2	Core 0, Level 3	Core 1, Level 3	Core 0, Level 4	Core 1, Level 4	Random Walk
Lags 0, Leads 1	-0.0277	-0.0469**	-0.0528**	-0.0328**	-0.0501***	-0.0295	-0.0465**	-0.0365	
Lags 0, Leads 5	-0.0345	-0.0059	-0.0154	-0.0185	-0.0268	-0.0066	-0.0095	-0.0232	
Lags 0, Leads 10	-0.0044	0.0230	-0.0229	-0.0618*	0.0024	-0.0255	-0.0351	-0.0268	
Lags 1, Leads 1	-0.0366**	-0.0451*	-0.0489**	-0.0406**	-0.0535*	-0.0297	-0.0440	-0.031*	
Lags 1, Leads 5	-0.0041	-0.0056	-0.0376	-0.0082	-0.0073	-0.0563	0.0122	-0.0031	
Lags 1, Leads 10	-0.0626**	-0.1068***	-0.0123	-0.0526	-0.0575***	-0.0006	-0.0195	-0.0345	-0.04233
Lags 3, Leads 1	-0.0593*	-0.0260	-0.0306	-0.0312	-0.0421*	-0.0255	-0.0316	-0.0425*	
Lags 3, Leads 5	-0.0030	-0.0141	-0.0296	-0.0101	-0.0340	-0.0192	-0.0357	-0.0438	
Lags 3, Leads 10	-0.0701**	0.0069	-0.0637**	-0.0432*	-0.0032	-0.0186	-0.0439*	-0.0043	
Lags 5, Leads 1	-0.0423**	-0.0327**	-0.043**	-0.0273	-0.0331	-0.0356	-0.0365	-0.0290	
Lags 5, Leads 5	-0.0178	0.0199	-0.0527**	-0.0230	-0.0343	-0.0424	-0.0168	-0.0361	
Lags 5, Leads 10	-0.0297	-0.061**	-0.0242	-0.0287	-0.0782**	-0.0491	-0.0447	-0.0727*	

Panel B. RNOA

	Core 0, Level 1	Core 1, Level 1	Core 0, Level 2	Core 1, Level 2	Core 0, Level 3	Core 1, Level 3	Core 0, Level 4	Core 1, Level 4	Random Walk
Lags 0, Leads 1	0.0546**	0.0558**	0.0178	0.0091	0.0208	0.0173	0.0139	0.0197	
Lags 0, Leads 5	0.0708**	0.0782***	0.0402*	0.0406**	0.0522**	0.0392*	0.0352**	0.0488***	
Lags 0, Leads 10	0.0332	0.0521	-0.0250	-0.0029	0.0054	0.0116	-0.0042	-0.0442	
Lags 1, Leads 1	0.0524**	0.0309	0.0113	0.0129	0.0232	0.0108	0.0186	0.0073	
Lags 1, Leads 5	0.0367**	0.0158	0.0446*	0.0444*	0.0477***	0.0549***	0.0608***	0.0343	
Lags 1, Leads 10	0.0256*	-0.0080	-0.0073	0.0326**	0.0052	-0.0047	-0.0265	0.0009	0.0212
Lags 3, Leads 1	0.0453*	0.0512**	0.0073	0.0217	0.0158	0.0135	0.0352*	0.0203	
Lags 3, Leads 5	0.0704	0.0707**	0.0108	0.0425**	0.053***	0.0185	0.0219	0.0257	
Lags 3, Leads 10	0.0333	0.0051	0.0161	-0.0038	-0.0264	-0.0621**	-0.0199	-0.0393***	
Lags 5, Leads 1	0.0491*	0.0445*	0.0126	0.0112	0.0297	0.0188	0.0199	0.0053	
Lags 5, Leads 5	0.0989***	0.0586*	0.0041	0.0162	0.0242*	0.0150	0.0135	0.0308	
Lags 5, Leads 10	0.0125	0.0378	0.0036	0.0197	0.0020	-0.0629**	0.0073	0.0280	