Estimating the Private Value of Financial Statement Statistics

by

Russell Lundholm
University of British Columbia
Russell.lundholm@sauder.ubc.ca

and

Xin Zheng
University of British Columbia
xin.zheng@sauder.ubc.ca

November 14, 2023

Abstract

We develop a method for estimating the private value of knowing the future realization of some financial statistic and then apply the measure to the familiar ratios arising from the Dupont decomposition of return on equity. The estimation is grounded in the standard rational expectations model, adapted to accommodate relative risk aversion, and produces an investor's willingness to pay for the signal. The method can accommodate different levels of investable wealth, multiple assets, and any information system that produces signals about those assets. To illustrate the use of this measure, we show that knowing next year's return on equity, given that the investor already knows the current value, is worth six times more than knowing the value of next year's sales growth. And, as predicted by the Dupont model, we find the value of knowing next year's operating asset turnover depends crucially on the level of the operating profit margin. Finally, we show that knowing next year's leverage is practically worthless. Given that investors face trade-offs when deciding where to expend effort in financial statement analysis, these estimates can help them to know where to allocate their time.

Estimating the Private Value of Financial Statement Statistics

1. Introduction

Consider the problem of an effort-constrained investor who is trying to forecast a firm's future fundamental value using the standard tools of financial statement analysis. The investor could read marketing reports to assess the firm's expected future profit margin, she could talk with management to learn about future investments, or she could study industry reports to learn about excess capacity, as just a few examples. But investors face time and resource constraints. When deciding how much effort to expend forecasting one part of a firm's performance, the investor necessarily trades off learning about another part of the firm's performance. How much effort an investor expends on learning about some aspect of the firm depends on what the investor already knows, and on how valuable learning the new thing is expected to be. This paper gives theoretical and empirical evidence to help with this problem by measuring the private value of knowing different financial statement statistics. We answer questions of the form, if you are trying to forecast F, and you already know Z, how much would you pay to know Y?

Many different literatures in accounting indirectly address our question, but none of them directly ask how information changes an investor's portfolio, and hence the value of that information. The market inefficiency literature documents the returns that could have been earned by trading on publicly available information, but doesn't estimate the private value of learning something new that other investors don't know. The forecasting literature shows how to use existing financial data to forecast future financial data, but doesn't produce an estimate of the value of doing so. Short-window return studies can be seen as estimates of the value of knowing the earnings announcement prior to its public release. However, this method only works for financial statistics that are reported in isolation, unconfounded by other value-relevant information. We discuss specific papers in the next section.

We assume the investor starts with complete knowledge of all the current financial statement statistics and then gathers information to inform herself about the next year's realization of those same statistics. For example, we assume that the investor knows the current operating profit margin, but expends effort to learn about next year's operating profit margin. Because we cannot possibly measure how precisely an investor can estimate next year's financial statistics, we go to the extreme and assume that the effort expended results in perfect foresight of next year's value of that statistic. We then measure the value of knowing next year's result, given that the investor knows this year's result. While perfect foresight is unlikely to be attained, it gives a crisp upper bound on the value of information, and it abstracts away from the heterogeneity of investors. Practically speaking, it is likely that an analyst has some assessment of how much effort would be required to reach a given level of precision on some forecasting statistic; our estimates give an upper bound on how valuable expending that effort would be.

Our method of estimating the value of information can be applied to any number of tradeable assets and any collection of signals about those assets' payoffs. Because financial statement analysis is primarily a firm-specific exercise, and for tractability, we begin by estimating the value of learning about a single stock at a time. While valuing information about only one asset may appear overly restrictive, if the asset payoffs are independent then the value of learning about multiple assets is simply the sum of the value of learning about each asset individually. In addition, in section 5.5 we consider how adding the ability to trade an industry ETF as a second tradeable asset might facilitate hedging, and how this changes the value of firm-specific information.

The financial statistics that we consider are the components of the standard Dupont decomposition of return on equity (e.g., sales growth, operating margin, asset turnover, leverage). Limiting the investor's ability to only these financial statistics is not overly restrictive, as the firm's fundamental value can be expressed as a function of these values (see Nissim and Penman 2001). Further, these financial statistics are the mainstay of financial statement analysis and are ubiquitous in analyst reports. We use this model to structure our empirical findings, starting with the two drivers of firm value – growth and profitability – and then systematically estimating the value of knowing the different sub-components of these statistics.

Our estimate of the value of information is taken from the models in Alles and Lundholm (1993) and Admati and Pfleiderer (1987). These papers derive a 'willingness to pay' statistic that is denominated in units of the riskless asset (e.g., money) and is estimable using standard regression techniques. This is the amount the investor would be willing to pay, ex ante, to learn the information. The statistic is remarkably simple; it compares the posterior variance of next year's stock return, given the new information, with the posterior variance without the information. For instance, the model expresses the value of knowing next year's sales growth, given that you already know this year's value, as the ratio of the posterior variance of next year's stock return conditioned on the current and future sales growth to the posterior variance conditioned on only the current period sales growth.[1]

As financial statement analysis textbooks emphasize, the meaning of many financial statistics depend on the levels of other financial statistics. For instance, growth in equity is only

[1] The model estimates the 'private value' of the statistic – the value of being the only one in the market who knows it, given that everyone knows the 'public' information. What the model doesn't estimate is the value of information given that some fraction of other agents in the model also know the information. See Hellwig and Veldkamp (2009) for a discussion of this more complicated problem.

valuable if the firm's return on equity exceeds its cost of capital. A higher asset turnover ratio is a good thing if the profit margin is positive, but is a bad thing if the profit margin is negative. We refer to these interactive effects as providing "context" to the analysis of any one financial statistic, and we make predictions about when the value of perfect foresight on one statistic will depend on the level of another statistic.

We find a number of interesting contrasts. It is six times more valuable to know next year's ROE than it is to know next year's sales growth. In fact, knowing nothing more than next year's ROE is worth 72 percent of the value of knowing the entire battery of statistics that come out of the Dupont model. Most of the value of ROE comes from the operating contribution (the return on net operating assets), rather than the financing contribution. And most of the value of knowing the return on net operating assets comes from knowing the operating profit margin. Further, knowing the future value of the operating asset turnover increases as the operating profit margin increases, consistent with the multiplicative nature of these two ratios. Finally, knowing next year's leverage ratio is practically worthless.

We replicate these results by estimating the value of different statistics separately for each industry. We find that the central tendencies across the distribution of industry results are very close to the pooled results. We also find that, even at the 75^{th} percentile, foreknowledge of sales growth is worth less than the 25^{th} percentile of the value of knowing future ROE. Finally, we find that the value of foresight into the future ROE is highest in the Consumer-Nondurables industry (e.g., food, clothing) and lowest in the Consumer-Durables industry (e.g., cars, dishwashers).

Lastly, we ask how the value of information changes when the investment opportunity set changes. In particular, we allow our hypothetical investor to trade in the firm and in an ETF that holds the firm. The ETF might offer hedging opportunities that could make the value of the firm-specific information more valuable. Alternatively, the existence of the ETF might make prices more informative, lowering the value of being privately informed. We find mixed results. Focusing again on the value of knowing future ROE, we find that the value increases significantly in the Utilities, Telecom and Finance industries, but decreases significantly in the Manufacturing, Healthcare, and Consumer Durables industries.

In the next section we summarize the related literature, and in section three we formally develop our measure of the value of information. We describe the sample in section four and present our results in section five. Section six concludes.

2. Literature Review

The residual income valuation model (Ohlson 1995) expresses the value of an equity in terms of forecasts of future accounting statistics. In particular, the value of a firm can be written in terms of the forecasted future ROE and forecasted growth in common equity (CSE). This observation motivates a large accounting literature aimed at forecasting these two drivers of value. A common approach is to use the Dupont decomposition to write ROE or CSE in terms of other more fundamental financial statistics, and see how well these underlying financial statistics predict the future. For instance, a number of studies find that the asset turnover ratio is more persistent than the profit margin (Fairfield and Yohn (2001), Nissim and Penman (2001), Penman and Zhang (2003) and Soliman (2008)). Amir, Kama, and Livnat (2011) extend this result by showing that, while the asset turnover ratio is more persistent, changes in the profit margin are more predictive of changes in future RNOA (something they label as "conditional persistence"). If we summarize this literature as 'the search for good forecasting variables,' then the earliest systematic work is Ou and Penman (1989) who used a battery of financial statement statistics taken from textbooks to predict the sign of next year's earnings change, and then used this statistic to predict future returns. More recently Chen, Cho, Dou and Lev (2022) address this same question using machine learning techniques.

Other studies have allowed estimates of a variable's persistence to vary with some other attribute. Vorst and Yohn (2018) use a life cycle model to predict ROE and growth in CE, among other things. They find that the persistence of these variables changes dramatically depending on the life cycle stage a firm is in. Jackson, Plumlee, and Roundtree (2018) decompose a firm's ROE into market, industry and firm-specific parts and find that letting each part have its own persistence parameter improves the forecast. Esplin, Hewitt, Plumlee and Yohn (2014) focus on the benefit of decomposing ROE into operating and financing components. They find that this distinction only matters when the forecaster takes a two-step approach, first estimating the persistence of each component separately and then putting the two forecasts together using the Dupont model (as opposed to fitting a single model of ROE as a function of the operating and financing components).

These papers tell the user how to map existing financial data into future values of the same statistic, or of a more value-relevant statistic, like future ROE. But if the user could learn more about the future value of some financial statistic, what would be the most beneficial statistic to learn about? A variable could be very persistent, and predict future ROE well, but still not be useful in predicting future stock returns. In that case it would not make sense for the investor to spend time trying to forecast the future value of this variable. Further, in our model we assume the investor knows the current value of any financial statistic in question. In this case learning about the future value of a highly persistent variable would be worth little because knowing its current value already provides most of the available information for free.

Our work is also related, but different than, the market inefficiency literature in accounting. One measure of the value of a signal is the abnormal return an investor could earn by forming investment portfolios based on the signal. Importantly, this literature only conditions on current publicly available information in order to show that these returns are realizable. This literature doesn't answer the question of what it would be worth to develop private information about the future value of some financial statistic. Finally, our results regarding the impact of allowing trade in the firm and in its industry ETF are related to the findings in Bhojraj, Mohanram, and Zhang (2020). They show that the introduction of a sector ETF improves the efficient pricing of the firms held by the ETF, while the introduction of a broad-based ETF harms the efficient pricing of its constituent firms. These findings help explain our mixed results when we add the ETF to the investment set.

Our work is related to the vast accounting literature, beginning with Beaver (1968), that measures the short-window returns around some announcement, typically an earnings announcement. This approach to measuring the importance of a signal effectively estimates a beta from the regression of the short-window return on the signal in question. If an investor knew the value of the signal before the announcement window, then clearly a bigger beta would be more valuable to an investor. But that is where the similarity in approaches ends. First, short-window returns are only good measures when the signal in question is released in isolation. This is rarely the case, even with earnings announcements, but is certainly not the case with the more focused financial statistics that we study. For the same reason, a short-window test cannot measure the value of a set of signals unless they happen to be released all at the same time. Our method allows us to estimate the value of any collection of signals in isolation or in sets. Second, a short-window test tacitly controls for other factors by keeping the return window small, but this approach never specifies what exactly is being controlled for. Our test explicitly states the variables being controlled for, and other nuance variables are naturally controlled for by their presence in the error terms of both the full and reduced models. Finally, our approach explicitly pins down the set of tradeable assets and the wealth of the investor; a short-window approach is silent on these issues.

Our work is similar to Nieuwerburgh and Veldkamp (2010) who study a setting where learning is constrained and so the investor has to trade off learning a lot about a few assets or a little about many assets. How does the investor choose how many assets to learn about? They model learning as a reduction from some fixed capacity and find that, depending on how learning consumes capacity, they can support various equilibrium allocations of effort. In some cases investors specialize in a few firms and ignore learning about the others, and in other equilibria investors spread their learning out across the maximum number of firms. In contrast, we constrain the number of risky assets the investor can learn about, so this is not a choice, and then we measure which firm-specific signals are most valuable to learn about. In effect, the Nieuwerburgh and Veldkamp (2010) paper assumes that the investor already knows what to study about a firm, she just doesn't know which firms to study.

Farboodi, Singal, Veldkamp, and Venkateswaran (2022) also measure the value of information, deriving a measure similar to the one presented here. They consider an economy where investors can only trade in a few classic portfolios: growth firms, value firms, small firms, large firms, or the entire market portfolio. In this setting they estimate the value of IBES growth forecasts, aggregating together the forecasts for each portfolio. They find that these aggregated growth forecasts are most valuable for investors specializing in large growth stocks. They also estimate the value of perfect foresight of next quarter's GDP growth rate, and find that it is most valuable for investors who only trade in the small stock portfolio. Our estimation method is similar to theirs, but we measure the value of very different information sets, and we estimate the value of information at the firm level rather than at the economy level.

3. Model of the Value of Information

We begin with a generic expression for the value of information derived from the standard rational expectations framework. Assume there is a single riskless asset serving as the numeraire (e.g., money). The investor trades in N risky assets with payoff vector denoted by *F*. There is a public signal vector *Z* that is commonly observed and is informative about some or all of the payoffs in *F*. In addition, by expending costly effort on financial statement analysis, the investor can observe the private signal vector *Y* which is also informative about some or all of the payoffs in *F*. All random variables are jointly normally distributed. The investor's posterior uncertainty about *F*, after having observed *Y* and *Z*, is given by the NxN covariance matrix $Var(F|Y,Z)$. If the investor does not expend effort to acquire the private signal *Y* then her NxN posterior covariance matrix is $Var(F|Z)$.

The value of information in this setting is the cost Φ that equates the utility of wealth achieved by trading on both signal vectors Y and Z, but having to pay Φ for Y, with the utility of wealth achieved by trading on only the public sector Z, but foregoing the private information cost. As is standard in the rational expectations literature, we assume investors have a negative exponential utility function with risk tolerance parameter ρ. Later we let ρ vary with the wealth level of the investor, effectively turning an estimate based on constant absolute risk aversion into an estimate based on constant relative risk aversion (as discussed in detail later).

In this setting the value of information can be shown to be

$$\Phi = \left(\frac{\rho}{2}\right) log \left\{\frac{\det[Var(F|Z)]}{\det[Var(F|Y,Z)]}\right\}.$$

See Admati and Pfliederer (1987) or Alles and Lundholm (1993) for the derivation. The value of information equation is remarkably general. It allows for any number of risky assets and any number of signals, with arbitrary relations between them. Φ is denominated in units of the riskless asset and so, in theory, we can compute an actual dollar value for the Y signal. As discussed later, we transform the utility function to have constant relative risk aversion by backing into the level of risk tolerance that is implied by a given level of wealth. While we do not take the *level* of this estimate too seriously, it provides a useful way to rank-order the value of different statistics, or to make relative statements about them. Finally, Φ is always positive because, in a normal random variable setting, posterior variances can only get smaller when conditioning on more information.

To develop some intuition for this expression, suppose there is only one risky asset with payoff F. In this case $Var(F|Z)$ and $Var(F|Y,Z)$ are scalars. As the signal Y becomes more informative, given Z, it reduces the posterior variance $Var(F|Y,Z)$ and Φ increases. Alternatively, if the signal Y is redundant given Z, then $Var(F|Y,Z)=V(F|Z)$, and Φ is zero. The log function shows that the marginal value of additional information is positive but decreasing. Finally, information is more valuable to a more risk tolerant investor because she is willing to take larger positions for a given level of posterior variance.

Now suppose there are N risky assets available to trade. The numerator and denominator in Φ are now the determinants of the posterior variance matrices (so that the ratio is once again a scalar). Suppose that the N assets' payoffs are independent, and that Y consists of N signals, one for each asset, that are independent of each other. In this case the determinant is the product of the diagonal in the posterior variance matrix, where each term is the posterior variance of the asset payoff conditioned on its associated signal. Since each diagonal element is a posterior variance, not surprisingly, more precise information on any asset increases the value of information.[2] However, if the asset payoffs or the signals are correlated, then the impact of changing any one part of the information system can have a quite complicated impact on the value of information. Such complications play out in the computation of the determinants of the two variance matrices.

To operationalize this framework, we need to specify *F*, the asset payoff that investors collect information about, and we need to propose some *Y* and *Z* signals that have a plausible relation with *F*. We assume that investors have a one-year horizon, so that *F* is the gross realized return $RET_1 = (P_1 + D_1 - P_0)$, where P_0 and P_1 are prices at time 0 or time 1, respectively, and D_1 is the dividend paid during the period.[3] For *Y* signals, we use the future realized date 1 values of some common accounting statistics, and for *Z* signals we use the date 0 values of these same statistics plus a few control variables. Thus, our characterization of financial statement analysis is a process that moves the investor from knowing the public date 0 value of some important statistic to privately knowing the date 1 future realization of that statistic. For instance, to estimate the value of perfect foresight about ROE_1 we compare the posterior variance of RET_1 given ROE_1 and ROE_0, with the posterior variance of RET_1 given only ROE_0.

We pick our accounting statistics based on how the residual income model relates changes in financial statement data to changes in value. Specifically, we show in the Appendix

---

[2] Taking the log of this product produces the result that in this case the value of Y is the sum of the value of each of its elements

[3] In setting $F = P_1 + D_1 – P_0$, note that the relevant uncertain variables are P_1 and D_1; P_0 is known at the beginning of the period.

that RET_1 is increasing in beliefs about the future ROE_1 and, conditional on ROE_1 being greater than the discount rate, is increasing in beliefs about the growth in common shareholders' equity (CSE_1). Finally, these two forces are multiplicative, so that expected growth in CSE_1 is particularly valuable when expected future ROE_1 is large (and visa versa). Thus, knowing the future realization of ROE_1 or CSE_1 has direct implications for RET_1.

Along with ROE_1 and CSE_1, we measure the value of different components of these statistics. As is standard in financial statement analysis, we express *ROE* using the advanced Dupont model (e.g. Lundholm and Sloan 2023, pp. 107) as

$ROE = RNOA + LEV*(RNOA – NBC),$ where

RNOA = return on net operating assets = net operating income/net operating assets,
LEV = financial leverage = net financial obligations/*CSE*, and
NBC = net borrowing costs = net financial expense/net financial obligations.

In addition, label *(RNOA – NBC)* as the spread, or *SPD*, and label the combined value *LEV*(RNOA – NBC)* as financing, or *FIN*, so that *ROE = RNOA + FIN*. Finally, *RNOA* can be expressed as the product of the operating profit margin (*OPM*) and the operating asset turnover (*OAT*). The exact construction for these values using COMPUSTAT data are given in the next section.

For the growth in *CSE*, note that this is equivalent to the growth in net operating assets if *LEV* remains constant, and is equivalent to growth in sales (Sg) if *LEV* and the *OPM* remain constant. To the extent that the growth rates in common equity, total assets, or sales differ, it is because of the changes in these other financial statement items. Finally, note that one reason CSE grows is that the firm earns positive net income. Thus, growth in CSE does not cleanly identify growth as a separate element from profitability. Further, other reasons for changes in CSE are capital market transactions, such as share issuances, repurchases, or dividends. These actions have immediate and know impacts on returns. For these reasons, Sales growth is the more common measure of 'growth' as a separate concept from profitability.

For each of these accounting statistics, we compute the value of knowing the date 1 value at date 0, given that the investor already knows the date 0 value. By assuming perfect one-year-ahead foresight, our estimates are the upper bounds of the value of information. Clearly the reality of financial statement analysis lies somewhere between knowing only the publicly available value at date 0 and having perfect foresight of the next year's date 1 value.

The size of the payoff vector F determines the investor's available opportunity set. We consider two different specifications of what the investor can trade: a single stock, or a single stock and its industry ETF. [4] Note that this specification choice can affect the value of information even if it doesn't change the information set available to the investor. For example, the investor's value of a firm-specific piece of information may depend on whether the investor can hedge out the industry component from the future return. Indeed, when on industry ETF is launched, Lundholm and Zheng (2021) show that analysts change the mix of firms they follow and their recommendation style in a way that aids their clients use of the ETF as a hedge.

3.1 Estimation

To operationalize the Φ equation we need estimates of the elements in $Var(F|Y,Z)$ and $Var(F|Z)$, the two posterior variance matrices in the value of information function. In a linear regression framework, the regression residuals' mean squared error with n degrees of freedom is

---

[4]A one or two firm investment opportunity set sounds restrictive. However, if the payoffs are independent, the value of information in an N-firm investment set is simply the sum of the individual values. To see this, note that, with independent payoffs in the F vector, the determinants of V(F|Y,Z) and V(F|Z) are just the product of the diagonal elements (i.e., the trace of each matrix). The ratio of these products can be rearranged to be the product of the ratio of each firm's relative values, and taking the log gives the sum of values of information over the N assets.

the maximum likelihood estimator of the posterior variance (MSE). Define a firm's date 1 annual return (RET_1) as the 12-month stock return ending 4 months after the date 1 fiscal year end. For the single asset setting we run two regressions, one that regresses RET_1 on the date 0 and date 1 values of the accounting statistic of interest – call this the *full model* - and one that regresses RET_1 on only the date 0 value of the same statistic – call this the *reduced model*. We then take the log of the ratio of the estimated MSEs and plug this into Φ to get the value of knowing the date 1 realization of the accounting statistic, given that you already know the date 0 value.

In the case of multiple assets, we also need estimates of the off-diagonal elements in the posterior matrices. We estimate these covariances directly from the vectors of residuals. For example, with two risky assets $Var(F|Y,Z)$ and $Var(F|Z)$ are both 2x2 symmetric matrices. The first diagonal element of $Var(F|Y,Z)$ is estimated by regressing the first asset's return on Y and Z, the second diagonal element is estimated by regressing the second asset's return on Y and Z, and the off-diagonal element is estimated by computing the covariance between the residuals from the two regressions. This results in a 2x2 variance-covariance matrix for the full model based on Y and Z. $Var(F|Z)$ is estimated the same way, but only using Z as the regressor in the reduced model. With the two 2x2 posterior variance matrices in hand, taking the determinant of each yields the necessary ratio of scalars to compute Φ.

The final unknown in equation Φ is the risk tolerance parameter ρ. This parameter follows from the constant *absolute* risk aversion assumption that is standard in these models. However, it is generally believed that constant *relative* risk aversion is more behaviourally descriptive; the general function for utility with a constant relative risk aversion is given as

$U(c) = c^{1-\sigma}/1 - \sigma$, where σ is the parameter of relative risk aversion.[5] To back into a reasonable value of absolute risk tolerance ρ we follow Farboodi, Singal, Veldkamp, and Venkateswara (2022) and assume a specific level of wealth and relative risk aversion, and then back into the level of absolute risk aversion that equates with this. In particular, we assume the investor has a *relative* risk aversion parameter of 2 and either $1M or $100M of wealth to invest. We label the $1M investor as a *retail investor* and the $100M as an *institutional investor*. For the $1M investor, this implies an *absolute* risk aversion parameter of 1/72359, or risk tolerance parameter of ρ equal to 72,359; for the $100M investor this implies ρ equals 5,428,881.[6] Higher levels of wealth imply greater risk tolerance, consistent with relative risk aversion.[7] While this is clearly arbitrary – other risk tolerance parameters would produce different dollar value estimates of the value of a signal - all of our results are based on the value of one type of information *relative* to another type, so changing the estimated value of ρ does not impact these conclusions.

We estimate the regressions needed to produce the variance estimates at two different levels of aggregation: 1) the pool of all firms, as in Chattopadhyay, Lyle, and Wang (2022), Lewellen (2015), and Lyle, Callen, and Elliott (2013); and 2) the pool of all firms in an industry for each of the Fama/French 38 industries, as in Lyle and Wang (2015). Smaller, more

---

[5] Harel, Francis and Harpaz (2018, section 12.1) compare all the major utility functions used in economics, finance, and accounting, concluding that "The attractiveness of the power utility function is increased by the fact it has constant relative risk aversion (CRRA) over all outcomes." And going back to some of the original work in this area, Friend and Blume (1975) conduct a large survey of house wealth and risk-taking behaviour and conclude "The empirical results in Section III indicate that the assumption of constant proportional risk aversion for households is a fairly accurate description of the market place."

[6] With a relative risk aversion parameter of two, a utility function with constant relative risk aversion is given by the function u(c)=-1/c. Equate this with the absolute risk aversion utility function -exp(-α*c) and solve for the absolute risk aversion parameter α=ln(c)/c. Insert c=$1,000,000 and invert to get risk tolerance of 72,359. The assumption that relative risk aversion is two is supported by Gandelman and Hernández-Murillo (2015), who assert that the literature generally finds a relative risk aversion parameter between one and three.

[7] For intuition, the $1M investor with a risk tolerance of ρ = 72,359, if offered a gamble between winning $0 or winning $1M, has a certainty equivalent is $50,651. For the $100M investor with risk tolerance of ρ = 5,428,881, the same 0/$1M gamble has a certainty equivalent of $475,000.

homogeneous pools of data may yield better specified regression estimates, but at the expense of having fewer observations to estimate the parameters. In addition, research has found that industry-level models can lead to worse predictions than fully pooled models (Fairfield, Ramnath, and Yohn 2009). Note that the level of aggregation for estimation purposes does not correspond to the number of assets the investor can trade in or learn about. Using the full cross-section of data to estimate the value of information does not imply that the investor learns about or can trade all these assets; rather, each firm-year is assumed to be an iid draw from the distribution of errors used to estimate the posterior variances.

For example, we can use the fully pooled time-series cross section of data to estimate the value of information about a single firm, or we could restrict the sample to only firms in the same industry. Similarly, we could estimate the value of information when then investor can trade the firm and an industry ETF using the fully pooled sample, or we could restrict the sample to only firms in the same industry.

To be abundantly clear, consider the case where there are two tradable assets – the firm and the firm's ETF – and we want to use the full panel of data to estimate the value of the firm-specific accounting statistics in the Y vector given the Z vector of control variables. The script would be 1) using the full panel of data, regress the firm's return on vectors Y and Z, and record the MSE, 2) regress the ETF's return on the same set of firm-specific variables in Y and Z, and record the MSE, 3) from the residual vectors in steps 1 and 2, compute the covariance, 4) construct the estimated variance-covariance matrix consisting of the two diagonal variances and the off-diagonal covariance and take the determinant, label this as DetYZ, 5) repeat steps 1-4 but using only the Z vector as regressors, labeling the determinant as DetZ, and finally 6) compute Φ using $\rho/2$ times log(DetZ/DetYZ).

4. The Data

The sample for the pooled and industry level results is constructed from all firm-years between 1973 and 2021 in the COMPUSTAT database with the financial information given below. These are then matched with the CRSP database to get each firm-year's annual returns. The intersection of these two databases yields our final pooled sample consisting of 44,146 firm-year observations.

The sample that includes the ETF returns is considerably smaller because ETFs did not come into existence until the end of 1998. We intersect the pooled sample with the CRSP mutual fund database and Thomson Reuters Mutual Fund Holdings (S12) database to identify whether a firm is held by any equity ETF that year. We then manually identify which ETFs have an industry focus and eliminate all others (see Lundholm and Zheng 2021 for details). If a firm is included in multiple ETFs, we select the ETF that has the most concentrated holdings in the firm. The final sample has 8714 firm-years with the necessary financial statement data, firm stock returns, and ETF returns.

4.1 Dependent variable

Our estimate of value compares the mean squared errors from regressions with and without the variable (or variables) being valued. The dependent variable in these regressions is the firm's future annual stock return, ending four months after the next year's fiscal end (RET_1). When we are estimating the value of a variable in the economy with the ETF available, we also estimate regressions using the ETF return over the same period as the dependent variable. Table 1 gives the descriptive statistics for the firm returns and the ETF returns (ETFRET_1). The mean firm annual return is 15.3 percent. This might appear high, but the data cleaning process described below removes many unusual and poorly preforming firms. The annual return is skewed upward, with a median value of 7.91 percent. The ETF sample is for a later period of time, and so isn't directly comparable, but note that the ETF return has considerably less variation than the firm returns.

4.2 Variables of Interest

We estimate the value of 10 different accounting statistics derived from the Dupont model, as presented in standard financial statement analysis textbooks. These variables make up the *Y* vector in the model.

We require the financial statement data necessary to deconstruct a firm's return on equity into its constituent parts as defined by the Dupont model. To map from raw financial statement data to the Dupont model elements we follow Esplin et al (2014).[8] We also replicate their rules for screening out unusual data item values. While these screens might appear extreme, it has been shown repeatedly in the financial statement analysis literature that without these screens, the Dupont model produces results that offer little predictive value.

For each item the COMPUSTAT codes are given in lower case. The first step is to delete observations with negative common equity (ceq), negative revenue (revt), negative total assets (ta), missing income before extraordinary items income (ib), or missing pre-tax income (pi).

Next, we set to zero some unusual items if they are missing. In particular, we change to zero any missing values for cash and short-term investments (che), debt in current liablities (dlc), long term debt (dltt), discontinued operations (do), preferred dividends (dvp), preferred

[8] We diverge from Esplin et al (2014) in one respect. They compute ROE using net income before taxes as the numerator, while we use an after-tax number. We believe that one of the benefits of the advanced Dupont model that we use is its allocation of tax expense between operating and financing activities.

dividends in arrears (dvpa), interest and related income (idit), other income and advances (ivao), preferred stock (pstk), special items (spi), preferred treasury stock (tstkp), income tax (txt), advertising expense (xad), extraordinary items (xi), interest and related expenses (xint), and research and development expenses (xrd).

From this partially-cleaned data we construct the aggregated financial statement items necessary for the Dupont decomposition. These are given as

Common shareholders' equity cse=ceq+tstkp-dvpa;
Financial assets fa=che+ivao;
Financial expense fe=xint+dvp;
Financial Obligations fo=dlc+dltt+pstk-tstkp+dvpa;
Financial revenue fr=idit;
Net financial obligations nfo=fo-fa;
Net operating assets noa=nfo+cse;
Effective tax rate eft=txt/pi;
Net operating income noi=ib-dvp +nfe;
Net financial expense nfe=(xint-idit)*(1-eft)+dvp; and
Net income available to common naic = ib - dvp ;[9]

We compute ratios based on average balance sheet items. Denoting the prior year's value with an 'l' proceeding the name, we get the Dupont model objects:

Return on equity ROE= naic/((cse+lcse)/2);
Return on net operating assets RNOA=noi/((noa+lnoa)/2);
Leverage LEV=(nfo+lnfo)/(cse+lcse);
Net borrowing costs NBC=nfe/((nfo+lnfo)/2);
Spread SPD=rnoa-nbc;
Financing FIN=roe-rnoa;
Operating profit margin OPM=noi/revt; and
Operating asset turnover OAT=revt/((noa+lnoa)/2).

[9] We delete the five observations where net operating assets equals zero. In addition, there are 15 observations with pretax income of zero; for these we set their effective tax rate (eft) to zero. Finally, there are 24 observations with net financial obligation of zero; for these we set net financial expense to zero.

To be precise, we compute the decomposition of Return on Average Common Equity before Extraordinary Items and Preferred Dividends.

At the level of the Dupont elements, Esplin et al (2104) then apply two last screens on the data. They delete any observation where noa<10, cse<1, |nfo|<5, |rnoa|>1, |roe|>1, nbc<0, or nbc>1.

Recall that SPD = RNOA – NBC. We can now express the decomposition of ROE as

$$
\begin{aligned}
ROE &= RNOA + FIN \\
&= OPM * OAT + LEV*SPD
\end{aligned}
$$

These are the seven performance statistics that we measure the value of. In addition, we assess three growth variables:

Sales growth Sg=revt/lrevt – 1,
Total asset growth TAg=at/lat – 1, and
Common Shareholders' Equity growth CSEg = cse/lcse – 1.

Finally, we winsorize the current and lagged values of the 10 accounting statistics at 1 percent and 99 percent.

Table 1 gives descriptive statistics. The median ROE is 10.92 percent, composed of RNOA of 9.15 percent and FIN = 0.7 percent. The median OPM is 5.23 percent and the median OAT is 1.79. Note that the 25th percentile of SPD is negative; in these cases FIN has a negative impact on ROE. Median Sales growth is 8.38 percent which might appear large, but recall that the data cleaning process above removed firms with unusually poor performance.

4.3 Control Variables

These are the Z variables from the model. In all cases we include the lagged value(s) of the variable(s) whose value is being estimated. We are assuming the investor knows the current value of the accounting statistic when assessing the value of knowing its future value. In addition, we include three control variables in all the regressions, measured at the end of the current fiscal year (date 0): the value of common shareholders' equity (CSE_0), the book-to-market ratio (BM_0) and the annual stock return (RET_0). These variables are intended to capture basic firm characteristics – size, value, and momentum - that affect the firm's expected annual return, but are not driven by information that arrives during the upcoming year. These variables are commonly used in models of expected returns.

Table 1 shows that, at the beginning of the return window, the median firm has $164M in common shareholders' equity (CSE) and a book-to-market (BM) of 0.5833.

4.4 Correlations

Table 2 gives the correlations between all the variables of interest. Looking down the first column we see that RET_1 is positively correlated with the ROE_1 that is reported during the period. It is also positively correlated with every profitability and growth statistic. RET_1 is negatively related to the prior period's return RET_0, showing a modest reversal in momentum. RET_1 is negatively related to CSE_0 and positively related to BM_0, consistent with the size and value factors.

Not surprisingly, the second column of Table 2 shows that ROE_1 is positively correlated with the concurrent operating statistics $RNOA_1$, OPM_1, and OAT_1, and with FIN_1 and SPD_1. It is negatively related to LEV_1. The Dupont model shows that leverage contributes to ROE only when the spread is positive and we saw in the descriptive statistics that the spread is negative more than 25 percent of the time. ROE_1 is also positively correlated with the three growth measures. Note in particular the 0.4158 correlation between ROE_1 and $CSEg_1$. As discussed earlier, one of the main reasons for a change in CSE is net income, causing the strong positive correlation between these variables. For this reason, even though growth in CSE is a driver of the valuation, sales growth is the more commonly used measure of growth.

Column 4 of Table 2 shows the trade-off between OPM and OAT, as is often discussed in financial statement analysis textbooks. Similarly, column 7 shows the negative relationship between leverage and the spread. Finally, columns 9-11 show high correlations between the three growth statistics.

5. Results

5.1 Pooled Results

We begin by discussing the value estimates for a single accounting statistic in the context of a one firm investment opportunity set. We estimate these values by pooling the cross-section and time-series data. Later we report results estimated by industry. We label the realized and publicly known values of the accounting statistics as the date 0 values, and we measure of value of having perfect foresight of the date 1 upcoming realization of the same statistic.

Table 3 gives the results from the pooled data. Panel A presents three important results that provide context for the more detailed analysis that follows. The first row of Panel A estimates the value of perfect foresight of ROE_1, given that the investor already knows the realized value of ROE_0 and the three control variables listed above. The ratio of the full model R-squared to the reduced model R-squared is 10.126 (=0.0740/0.0073). We report this ratio because R-squared is a common measure of fit, but the statistic that actually goes into the value of information estimate is the ratio of MSEs, given in the next column. The result is that our hypothetical retail investor (with \$1M of wealth) values next year's ROE_1 at \$2,515 and the hypothetical institutional investor (with \$100M of wealth) values it at \$188,680. Given that forecasting next year's ROE_1 (or net income) is a ubiquitous task in equity valuation and analysis, these value estimates give us a baseline to compare other more focused statistics. We will return to this base model repeatedly. Finally, we do not rely heavily on the level of these value estimates; rather, we discuss the relative value of one statistic over another, effectively eliminating the dependence on the risk tolerance parameter for a given level of wealth.

The second row of Panel A in Table 3 estimates the value of foreknowledge of sales growth Sg_1 given that the investor knows last year's sales growth Sg_0 and the three control variables. The value of this foresight is only $447 for the retail investor and $33,540 for the institutional investor. It is roughly six times more valuable to know next year's ROE_1 than it is to know next year's sales growth Sg_1. This is a surprising result given the attention that forecasting sales growth gets in financial statement analysis textbooks (e.g. Lundholm and Sloan 2023).

The final row in Panel A of Table 3 considers the value of knowing the whole collection of accounting statistics identified from the Dupont decomposition and three different measures of growth. The estimate of perfect foresight of all these variables together, given their date 0 values and the three control variables, represents the most value we could possibly get out of the Dupont framework. The value estimate for the whole collection of date 1 statistics is $3,478 for the retail investor and $260,975 for the institutional investor. Comparing these estimates with the value of knowing ROE_1 illustrates how useful ROE is as a summary measure of value creation. Simply knowing the future ROE_1 is worth 72 percent of the value of knowing the future values of all the Dupont statistics and growth variables.

Panel B of Table 3 investigates which measure of growth is the most valuable. Comparing the first three rows shows that knowing the future sales growth Sg_1 or total asset growth TAg_1 have roughly the same value, but knowing the growth in Common Shareholders' Equity $CSEg_1$ is much higher. However, as previously discussed, one common reason CSE grows is because the firm earns positive net income. Thus, foreknowledge of the growth in CSE_1 is partly driven by future ROE_1, and we have already shown ROE_1 to be quite valuable. Further, the other main reasons for changes in CSE are capital transactions by the firm, and these have known and immediate impact on the firm's return (for instance, large special dividends reduce the market value of equity by almost exactly the same amount). For this reason, we focus most of our attention on sales growth, as it is a measure of growth that is less confounded with profitability.

Note also that the number of observations available for the growth estimations in Panels A and B are lower than for the ROE estimations in the first row. This is because the lagged growth variables require data from two prior years. To be sure that our conclusion that foresight about ROE is much more valuable than foresight about growth is correct, and not due to a change in the sample, in Panel B we estimate the value of the three growth variables without requiring the lagged value. In all three cases the value of the growth variable is even lower when its date 0 value is not included in the reduced model.

As derived in the previous section, *ROE* has two parts – the value of operations, as measured by *RNOA*, and the value of financing activities, as measures by *FIN*. Table 3 Panel C shows that the value of knowing $RNOA_1$ is roughly three times higher than the value of knowing FIN_1. This conclusion holds whether or not we control for the date 0 values of the other variables being measured. Note, however, that FIN_1 still has value. For instance, the value of knowing $RNOA_1$ is 76 percent of the value of knowing ROE_1 (1909/2515). While $RNOA_1$ is more valuable than FIN_1, FIN_1 still contributes 24 percent of the value in ROE_1.

Panel D further decomposes *RNOA* into the product of operating profit margin (*OPM*) and operating asset turnover (*OAT*). The results show that knowing the future OPM_1 is roughly four times more valuable than knowing the future OAT_1. This conclusion holds whether the reduced model controls for the variable's own date 0 value or if it controls for the date 0 values of both variables. Finally, note that in the Dupont decomposition, *OPM* and *OAT* have an interactive effect. We explore this in the next section.

Table 3 Panel E completes the analysis of the individual Dupont statistics. *FIN* is the product of leverage (*LEV*) and spread (*SPD*). The results show that foreknowledge of the spread is roughly 10 times more valuable than foreknowledge of leverage. This is unsurprising given that $RNOA_1$ is part of SPD_1. The most surprising result in this panel is the observation that knowing the future leverage LEV_1, with or without controlling for knowledge of the date 0 value, is almost worthless.

In sum, where should an investor with limited time and resources invest their forecasting energy? In general, knowing the future value of operating statistics is more valuable than knowing the future value of financing statistics. In particular, the links from operating profit margin to the return on net operating assets, and then to the return on equity, captures an increasing amount of value. Compared to these profitability statistics, the value of knowing sales growth is much smaller, although not insignificant, while the value of knowing leverage is almost zero.

5.2 Interaction Effects in the Value of Information

The valuation model in the appendix shows that ROE and growth are multiplicative – the higher the value of ROE, the greater the impact of growth (assuming ROE is greater than the cost of capital), and visa-versa. In addition, the Dupont decomposition on ROE includes two interactive terms – the product of operating profit margin and operating asset turnover equals

RNOA and the product of leverage and spread equals FIN. In this section we investigate whether the value of these items exhibit the interactions predicted by theory.

Recall from Table 3 that the unconditional value of knowing Sg_1 is \$447 for the retail investor and \$33,540 for the institutional investor. Table 4 Panel A estimates the value of knowing future sales growth Sg_1 (given Sg_0 and the controls) for different quintiles of ROE_0. For the retail investor, the value of knowing next year's sales growth is \$248 in the first quintile of ROE, increases steadily across the quintiles, and reaches \$903 in the fifth quintile. The value to institutional investors has a similar pattern. This is more than a three-fold increase. Clearly there is an interactive relationship between growth and profitability that impacts the value of knowing future sales growth.

Table 4 Panel B explores the interactive nature of operating profit margin and operating asset turnover. Recall that RNOA = OPM*OAT, so that the higher OPM is expected to be, the greater the impact of OAT on RNOA, and visa versa. Stated differently, if OPM is expected to be around zero, then there is little value in forecasting what OAT might be. The table shows that the value of knowing future operating asset turnover increases over the first four quintiles of current operating profit margin, but drops in the last quintile. Over the first four quintiles of OPM the value of knowing OAT increases almost threefold. However, some other force is at work when the current operating margin is very high (i.e. the fifth quintile). One possibility is that the extreme quintile of operating profit margin exhibits the fastest mean reversion, so this level is not expected to continue into the future, and therefore it's true multiplicative value is smaller.

Table 4 Panel C explores the interactive nature of leverage and spread. Recall that FIN = LEV*SPD, so that higher levels of leverage are more valuable when the spread is higher.

Leverage contributes positively to the overall ROE only if the spread is positive. Recall from Table 3 that the value of knowing future LEV was very low - estimated at $144 for the retail investor and $10,831 for the institutional investor. Table 4 shows that the value of knowing future leverage is weakly increasing over the quintiles of spread. For the retail investor the value goes from $190 in the first quintile to $295 in the fifth quintile; for the institutional investor the value goes from $14,218 to $22,163. While this is weak evidence of a multiplicative relationship, it also shows that the value of LEV_1 is generally low regardless of the value of SPD_0.

The summary of Table 4 is that the predicted interactive effects are generally present. The practical advice is that the user shouldn't bother forecasting sales growth if ROE is expected to be low, and the user shouldn't bother forecasting OAT if OPM is expected to be low. And in all cases there is little value in forecasting LEV_1.

5.3 Estimations by Industry

The pooled results provide straightforward answers to questions about the value of different accounting statistics, but they also assume a homogeneity across many firms that are very different. By estimating the value within industry, we improve the specification of our estimates, but at the cost of producing a lengthy list of results. Rather than replicate all the pooled results, we select only the highlights, and report the distribution of estimates across industries. Then, for the ROE results, we present the estimates separately for the Fama/French 12 industries.

Table 5 gives the distribution of value estimates over 36 industries of the Fama/French 38 (two industries have insufficient data) for the most important accounting statistics. Panel A estimates, by industry, the value of perfect foresight of ROE_1 given ROE_0 and the three control variables. For comparison, the last column in the table shows the estimate from the Table 3 pooled regressions. As seen in the table, the median value is $2,968 for retail investors and $222,659 for institution investors. In addition, the maximum value for retail investors, which occurs in the XXX industry, is $6,764, more than twice as large as the median. Generally, the mean estimate is very close to the median estimate, so the distribution of value estimates is fairly symmetric.

One of the surprising results from the pooled estimates is the relatively low value of knowing future sales growth Sg_1, especially when compared to the value of knowing ROE_1. In Table 5 Panel B we see that this conclusion holds up in the distribution of industry level results. The median value of Sg_1 given Sg_0 and the controls, is only $379 for the retail investors and $28,431 for the institutional investors. Even at the 75th percentile the values are only $811 and $60,837, respectively. Even the 25th percentile of the value of knowing ROE_1 is more than 2.5 times larger than the 75th percentile of the value of knowing Sg_1.

The central tendency of the estimated value of knowing $RNOA_1$, when estimated by industry, is similar to the pooled results. As seen in Table 5 Panel C, it makes up about 75 percent of the value of knowing ROE_1. However, in contrast to the pooled results, Panel D shows that the median value of knowing FIN_1 is $1,495 for retail investors and $112,134 for institutional investors; these value estimates are more than twice the estimated values from the pooled estimates. While the pooled estimates led to the conclusion that knowing the future value of FIN_1 was not very valuable, the industry level estimates put financing on the same footing as operations. Indeed, the median value of knowing FIN_1 is roughly half the median value of knowing ROE_1.

The pooled results showed that the driver of RNOA's value was the value in knowing the future operating profit margin OPM_1. Table 5 Panel E shows that the median value of knowing

OPM_1 is 69 percent of the median value of knowing $RNOA_1$. By comparison, the pooled estimate of the value of OPM_1 was only 59 percent of the value of knowing $RNOA_1$. The industry results confirm the conclusion that the value of knowing $RNOA_1$ is largely due to the value of knowing OPM_1.

The final panel in Table 5 confirms the prior result that foreknowledge of leverage is of little value. Even at the 75th percentile the value of LEV_1 is only \$443 for retail investors and \$33,235 for institutional investors.

Collectively the industry-level estimates of value confirm many of the conclusions from the pooled analysis. Foreknowledge of ROE_1 is far more valuable than foreknowledge of sales growth. Knowing next year's leverage is almost worthless. However, at the median, roughly half of the value of knowing ROE_1 comes from operations – $RNOA_1$ – and half from financing – FIN_1, whereas the pooled results skew more in favour of $RNOA_1$.

5.4 Detailed Estimated by Fama/French 12 Industries

In this section, we present the value estimates of ROE_1, given ROE_0 and the three control variables, by industry. To keep the list from being too detailed, we switch to the Fama/French 12 industry classifications. This is largely a descriptive exercise, as we have no ex ante theory for which industries will value foreknowledge of ROE_1 more than others. Table 6 gives the results sorted from highest value estimates to the lowest. Interestingly, the highest values are in Consumer Non-durables (e.g. food and clothing) while the lowest values are in Consumer Durables (e.g. cars and dishwashers). The value of knowing of ROE_1 is about twice as high in Consumer Non-durables than in Consumer durables. The central difference between these two industries is the level of discretion the consumer has in timing their purchases. The table shows that this causes information about ROE_1 to be considerably more valuable when consumers have little discretion. This is possibly because, given the lack of discretion in Consumer Non-durables, growth is well modeled, but profitability is more uncertain, while the opposite is true for Consumer Durables.

5.5 Increasing the investment opportunity set

In this section we consider how enlarging the set of available securities impacts the value of information about the firm. We hold the information constant – firm-specific information about future ROE_1 - but now assume that the investor can also take a position in an industry ETF that holds the firm. As discussed earlier, the introduction of the ETF has potentially conflicting impacts on the value of information. On the one hand, it allows the investor to hedge out a common industry component. All else equal, this makes firm-specific information more valuable. On the other hand, the ability to hedge causes investors to take more extreme positions and this makes price more informative and, consequently, private information less valuable.

The estimation of value is more complicated with multiple assets, as described in section 2. Instead of estimating posterior variances directly from the residuals, we need to estimate the generalized posterior variances, which are the determinants of the posterior variance-covariance matrices. Table 7 gives the results, by Fama/French 12 industry, of estimating the value of ROE_1 given ROE_0 and the controls, when the investor can also take positions in the industry ETF. For comparison we also re-estimate the value from the one-asset case on the same ETF sample. Table 7 is sorted by the difference between the two estimates – the one with the ETF present and the other without it, as shown in the last column. Consider the first row – Telecom. The estimated value of knowing the future ROE_1, in a setting that allows trading in the ETF, is $6,158 for retail investors and $461,980 for institutional investors. These estimates are in contrast to the estimates produced by valuing ROE_1 in the one-asset setting, where they are

$3,124 for the retail investor and $234,370 for the institutional investors (as seen in the second-to last and next-to-last columns). In other words, the ability to hedge using in the ETF increases the value of knowing ROE_1 by $3,034 for retail investors and $227,610 for the institutional investors. Overall, 10 of the 12 industries show increases in the value of knowing ROE_1 when the estimate takes into account that investors can also trade in the ETF. The theoretical reason for information being more valuable in these 10 industries is that the ability to hedge out an industry component adds more value than is lost by having more informative prices.

To help understand why the value of information changes so much in Table 7, contrast Telecom, which has the biggest improvement when the ETF is added to the investment opportunity set, with Utilities, which has the smallest improvement. The second-to-last column in Table 7 shows that the one-asset estimates of the value to a retail investor of knowing future ROE_1 for these two industries are approximately the same ($3,124 for Telecom and $3,027 for Utilities). So what changes for Telecom, but not for Utilities, when we add the ETF to the investment opportunity set? The difference is that the estimated covariance between the firm payoff and the ETF payoff is three times higher in the Telecom industry than in the Utilities industry (untabulated). This makes the ETF a much more effective hedging instrument in Telecom than in Utilities.[10] Consequently, when we add the ETF as a tradeable asset, the value of firm-specific information in the Telecom industry increases much more than in the Utilities industry.

---

[10] To analytically how correlation impacts the value of information, assume that the information vectors Y and Z have no impact on the covariance C between the firm payoff F_i and the ETF payoff F_{ETF}, or on the posterior variance of the ETF payoff. With this, the value of information is proportional to the ratio of determinants, given as $\frac{Var(F_i|Z)*Var(F_{ETF})-C^2}{Var(F_i|Y,Z)*Var(F_{ETF})-C^2}$ which is increasing in C because $Var(F_i|Z)$ is always greater than $Var(F_i|Y,Z)$.

6. Conclusion

Financial statement analysis offers a long list of items to consider when forecasting a firm's future fundaments, but little advice on where to expend the most energy. While our estimates do not take into account the cost of making the forecast, and go to the limit by assuming perfect foresight, they offer a starting place for a more data-driven valuation process. Our empirical findings also show stark contrasts between the value of some common statistics over others.

There are a number of avenues available for future research. Financial statement analysis offers another level of more detailed statistics – inventory turnovers, SGA-to-Sales, the quick ratio – just to name a few popular ratios. Our measure could help establish which of these more nuanced statistics is the most valuable. In addition, our results regarding the impact of allowing trade in an ETF just scratch the surface of investigating the impact of an investment opportunity set with correlated payoffs. In this case it could be that new financial statistics designed to differentiate between common and idiosyncratic payoff components could be particularly valuable.

Appendix

Assume the realized value at time 1 is

$$P_1 = CE_1 + \sum_{t=1}^{\infty} E[RI_{t+1}/(1+r)^t | I_1] \ where,$$

residual income is $RI_{t+1} = NI_{t+1} - r^*CE_t$, r is the discount rate, and I_1 is the information set at time=1. CE_t is common equity at time t, and NI_t is comprehensive income at time t. Writing the same model at time zero, conditioned on information set I_0, gives

$$P_0 = CE_0 + \sum_{t=1}^{\infty} E[RI_t/(1+r)^t | I_0].$$

Taking the difference gives

$$P_1 - P_0 = CE_1 - CE_0 + \sum_{t=1}^{\infty} \frac{E[RI_{t+1}|I_1] - E[RI_t|I_0]}{(1+r)^t}.$$

Denote the net dividend in period 1 as D_1, so that clean surplus gives $CE_1 - CE_0 = NI_1 - D_1$, which gives

$$P_1 + D_1 - P_0 = NI_1 + \sum_{t=1}^{\infty} \frac{E[RI_{t+1}|I_1] - E[RI_t|I_0]}{(1+r)^t}.$$

Further, note that $NI_1 = RI_1 + r^*CE_0$ and $RI_1 = E(RI_1|I_1)$ so we can write the return as

$$P_1 + D_1 - P_0 = rCE_0 + \sum_{t=0}^{\infty} \left[\frac{E[RI_{t+1}|I_1]}{(1+r)^t} - \frac{E[RI_{t+1}|I_0]}{(1+r)^{t+1}} \right].$$

The value r^*CE_0 is the expected return. The realized return is then more or less than this amount because of changes in expectations between I_0 and I_1 about the stream of future residual incomes (adjusted for the one-year difference in the cumulative discount rate).

To get closer to common accounting variables, write $RI_t = CE_{t-1}(ROE_t - r)$ so that the equation says that realized returns are driven by changes in expectations about future *ROE* and future CE. In other words, returns differ from expectations because of changes in beliefs about future profitability and future growth. This also highlights that realized returns depend on the *interactive* effect of profitability and growth.

As a simple example, suppose $E(ROE_t|I_1) = ROE_1 \ and \ E(CE_t|I_1) = CE_1$ for $t \geq 1$, and $E(ROE_t|I_0) = r \ and \ E(CE_t|I_0) = CE_0, for \ all \ t$. In other words, at time 0 the investor believes ROE_t and CE_t will remain at the date 0 values forever, and at time 1 she believes they will remain at the date 1 values forever. In this case the return is

$$RET_1 = rCE_0 + (ROE_1 - r)\left[CE_0 + \frac{CE_1}{r(1+r)}\right].$$

Again, realized returns depend on the expected return, changes in expectations about profitability and changes in expectations about growth in *CE*, and profitability and growth are multiplicative.

References

Admati, A., and P. Pfleiderer. Viable Allocations of Information in Financial Markets. *Journal of Economic Theory* **43**, 76–115 (1987).

Alles, M. and R. Lundholm. On the Optimality of Public Signals. *The Accounting Review* 26, 107 (1993).

Amir, E., Kama, I. and Livnat, J. Conditional versus unconditional persistence of RNOA components: implications for valuation. *Review of Accounting Studies* 16, 302–327 (2011).

Beaver, W. 1968. The Information Content of Annual Earnings Announcements. *Journal of Accounting Research*. 6:67-92.

Bhojraj, S., P. Mohanram, P. and S. Zhang. ETFs and Information Transfer Across Firms. *Journal of Accounting and Economics* (2020) doi:10.1016/j.jacceco.2020.101336.

Chattopadhyay, A., M. Lyle, and C. Wang. 2022. Expected Stock Returns Worldwide: A Log-Linear Present-Value Approach. *The Accounting Review* 97, no. 2 (March 2022): 107–133.

Chen, X., Cho, Y.H., Dou, Y. and Lev, B., 2022. Predicting future earnings changes using machine learning and detailed financial data. Journal of Accounting Research, 60(2), pp.467-515.

Esplin, A., M Hewitt, M Plumlee, M., and T. Yohn. Disaggregating operating and financial activities: implications for forecasts of profitability. *Review of Accounting Studies* **19**, 328–362 (2014).

Fairfield, P. M., S. Ramnath, and T. L. Yohn. Do industry-level analyses improve forecasts of financial performance? *Journal of Accounting Research* 47 (1): 147–178. (2009).

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

Farboodi, M., and D. Singal, L. Veldkamp, and V. Venkateswaran. 2022. Valuing Financial Data. NBER working paper. http://www.nber.org/papers/w29894.

Gandelman, N. and R. Hernández-Murillo. Risk aversion at the country level. Federal Reserve Bank of St. Louis working paper. (2014). http://research.stlouisfed.org/wp/2014/2014-005.pdf

Hellwig, C. and L. Veldkamp, 2009. Knowing What Others Know: Coordination Motives in Information Acquisition. *Review of Economic Studies*. 76:223–251.

Jackson, A. B., Plumlee, M. A. & Rountree, B. R. Decomposing the market, industry, and firm components of profitability: implications for forecasts of profitability. *Review of Accounting Studies* 23, 1071–1095 (2018).

Lewellen, J. The cross-section of expected stock returns. *Critical Finance Review* 4 (1): 1–44 (2015).

Lundholm, R., and R. Sloan. 2023. Equity Valuation and Analysis. https://www.lundholmandsloan.com.

Lundholm, R. and X. Zheng. The Changing Nature of Financial Analysis in the Presence of ETFs. University of British Columbia working paper. (2021).

Lyle, M., J. Callen, and R. Elliott. Dynamic risk, accounting-based valuation and firm fundamentals. *Review of Accounting Studies* 18 (4): 899–929. (2013).

Lyle, M., and C. Wang. The cross section of expected holding period returns and their dynamics: A present value approach. *Journal of Financial Economics* 116 (3): 505–525. (2015).

Lyle, M. R. and Yohn, T. L. Fundamental Analysis and Mean-Variance Optimal Portfolios. *The Accounting Review*, 96, 303–327. (2021).

S. Nieuwerburgh, and L. Veldkamp. Information Acquisition and Under-diversification. *The Review of Economic Studies* 35. (2010).

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

Ohlson, J. Earnings, Book Values, and Dividends in Equity Valuation. *Contemporary Accounting Reearch* 11, 661–687 (1995).

Penman, S. and X. Zhang. Accounting conservatism, the quality of earnings, and stock returns. *The Accounting Review* 77, 237–264 (2002).

Soliman, M. The Use of DuPont Analysis by Market Participants. *Accounting Review* 83, 823–853 (2008).

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

Table 1: Descriptive Statistics for Financial Statement and Market Data

Dependent Variable	N	Mean	Minimum	25th percentile	50th percentile	75th percentile	Maximum
Future Annual Return (RET(1))	44,146	0.1530	-0.9965	-0.1667	0.0791	0.3483	15.9114
Variables from the Dupont model							
Return on Equity ROE(1)	44,146	0.0805	-0.6740	0.0318	0.1092	0.1712	0.5182
Return on Net Operating Assets RNOA(1)	44,146	0.0928	-0.4313	0.0464	0.0915	0.1429	0.5622
Operating Profit Margin OPM(1)	44,146	0.0578	-0.4860	0.0221	0.0523	0.0965	0.4388
Operating Asset Turnover OAT(1)	44,146	2.1716	0.1400	1.0546	1.7957	2.7055	10.1366
Financing Contribution FIN(1)	44,146	-0.0115	-0.5350	-0.0333	0.0074	0.0376	0.3480
Leverage LEV(1)	44,146	0.8640	-0.7181	0.1722	0.5524	1.1164	9.0691
Spread SPD(1)	44,146	0.0236	-0.5766	-0.0209	0.0289	0.0816	0.5292
Sales Growth Sg(1)	44,146	0.1236	-0.4387	-0.0013	0.0838	0.1973	1.3069
Total Asset Growth TAg(1)	44,146	0.1241	-0.3134	-0.0080	0.0672	0.1753	1.5187
Common Equity Growth CSEg(1)	44,146	0.1230	-0.5700	-0.0069	0.0817	0.1807	1.8981
Control Variables							
Current year Annual Return RET(0)	44,146	0.1497	-0.9851	-0.1661	0.0729	0.3382	22.8745
Current year Common Equity CSE(0)	44,146	1070.640	5.0730	47.0900	164.2630	682.2720	21409
Current year Book-to-Market BM(0)	44,146	0.8018	0.0061	0.3420	0.5833	0.9632	48.9308
Other Variables							
RET(1) in ETF sample	8,714	0.1315	-0.9438	-0.1152	0.0993	0.3083	15.9114
ETFRET(1) in ETF sample	8,714	0.0883	-0.8171	-0.0442	0.1058	0.2208	1.9872

Table 1 notes: The perfect foresight date 1 value is shown as variable(1); the known date 0 value is shown as variable(0). Variable definitions and constructions are in Section 3.

Table 2: Pearson Correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) RET(1)	1												
(2) ROE(1)	0.1687	1											
	<.0001												
(3) RNOA(1)	0.1401	0.7634	1										
	<.0001	<.0001											
(4) OPM(1)	0.0991	0.6096	0.6341	1									
	<.0001	<.0001	<.0001										
(5) OAT(1)	0.0483	0.1529	0.2578	-0.1643	1								
	<.0001	<.0001	<.0001	<.0001									
(6) FIN(1)	0.0875	0.6018	-0.0330	0.1607	-0.0731	1							
	<.0001	<.0001	<.0001	<.0001	<.0001								
(7) LEV(1)	0.0014	-0.0958	-0.1366	0.1983	-0.2350	0.0204	1						
	0.7755	<.0001	<.0001	<.0001	<.0001	<.0001							
(8) SPD(1)	0.1248	0.7411	0.9019	0.5699	0.2295	0.0393	-0.1418	1					
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001						
(9) Sg(1)	0.0595	0.1925	0.2050	0.1519	0.0515	0.0478	0.0132	0.1781	1				
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0055	<.0001					
(10) TAg(1)	0.0712	0.2324	0.2379	0.1793	0.0261	0.0665	-0.0071	0.2189	0.5847	1			
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.1374	<.0001	<.0001				
(11) CSEg(1)	0.1494	0.4158	0.3288	0.2499	0.0520	0.2411	-0.0103	0.2943	0.4310	0.6210	1		
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0307	<.0001	<.0001	<.0001			
(12) RET(0)	-0.0536	0.2493	0.2350	0.1549	0.0746	0.0972	-0.0335	0.2097	0.2637	0.2620	0.3036	1	
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		
(13) CSE(0)	-0.0174	0.0948	0.0675	0.1566	-0.1278	0.0638	0.0275	0.0714	-0.0742	-0.0575	-0.0591	-0.0127	1
	0.0003	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0076	
(14) BM(0)	0.0737	-0.3089	-0.2688	-0.1921	-0.0476	-0.1570	0.0645	-0.2700	-0.1823	-0.2161	-0.2420	-0.2293	-0.1076
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Table 2 notes: Variable definitions and constructions are in Section 3. The perfect foresight date 1 value is shown as variable(1); the known date 0 value is shown as variable(0).

Table 3: Value of Information Estimated with Pooled Time Series and Cross Section

Panel A

Variables being Valued	N	R-squared from Full Model	R-squared from Reduced Model	R-squared Ratio	MSE Ratio	Φ-Retail	Φ-Institutional
ROE(1), given ROE(0) and controls	44,146	0.0740	0.0073	10.1260	1.0720	2,515	188,680
Sales Growth Sg(1), given Sg(0) and controls	34,255	0.0230	0.0108	2.1211	1.0124	447	33,540
ROE(1), RNOA(1), OPM(1), OAT(1), LEV(1), SPD(1), Sg(1), CSEg(1) and TAg(1), given their date 0 values and controls	34,255	0.1050	0.0147	7.1403	1.1009	3,478	260,975

Panel B

Detailed Analysis of Growth Variables	N	R-squared from Full Model	R-squared from Reduced Model	R-squared Ratio	MSE Ratio	Φ-Retail	Φ-Institutional
Sales Growth Sg(1), given Sg(0) and controls	34,255	0.0230	0.0108	2.1211	1.0124	447	33,540
Total Asset growth TAg(1), given TAg(0) and controls	34,255	0.0247	0.0122	2.0208	1.0128	461	34,564
Common Shareholders' Equity Growth CSEg(1), given CSEg(0) and controls	34,255	0.0571	0.0117	4.8777	1.0482	1,702	127,703
Sales Growth Sg(1), given controls (no lagged values)	44,146	0.0141	0.0070	2.0262	1.0073	261	19,616
Total Asset growth TAg(1), given controls (no lagged values)	44,146	0.0169	0.0070	2.4236	1.0101	363	27,251
Common Shareholders' Equity Growth CSEg(1), given controls (no lagged value)	44,146	0.0425	0.0070	6.1043	1.0372	1,320	99,003

Panel C

Detailed Analysis of Operating versus Financial contributions to ROE (ROE = RNOA + FIN)	N	R-squared from Full Model	R-squared from Reduced Model	R-squared Ratio	MSE Ratio	Φ-Retail	Φ-Institutional
RNOA(1), given RNOA(0) and controls	44,146	0.0584	0.0071	8.2467	1.0545	1,920	144,020
FIN(1), given FIN(0) and controls	44,146	0.0250	0.0072	3.4747	1.0182	653	48,987
RNOA(1), given RNOA(0) and FIN(0) and controls	44,146	0.0584	0.0074	7.9267	1.0542	1,909	143,248
FIN(1), given FIN(0), RNOA(0) and controls	44,146	0.0251	0.0074	3.3996	1.0181	650	48,732

Panel D

Detailed Analysis of Return on Net Operating Assets and its Components (RNOA = OPM * OAT)	N	R-squared from Full Model	R-squared from Reduced Model	R-squared Ratio	MSE Ratio	Φ-Retail	Φ-Institutional
Operating Profit Margin OPM(1), given OPM(0) and controls	44,146	0.0380	0.0074	5.1638	1.0319	1,135	85,129
Operating Asset Turnover OAT(1), given OAT(0) and controls	44,146	0.0153	0.0074	2.0683	1.0080	288	21,575
Operating Profit Margin OPM(1), given OPM(0), OAT(0) and controls	44,146	0.0389	0.0076	5.0883	1.0325	1,158	86,857
Operating Asset Turnover OAT(1), given OAT(0), OPM(0) and controls	44,146	0.0153	0.0076	1.9988	1.0077	279	20,918

Panel E

Detailed Analysis of FIN = LEV * SPD and its components	N	R-squared from Full Model	R-squared from Reduced Model	R-squared Ratio	MSE Ratio	Φ-Retail	Φ-Institutional
Spread SPD(1), given SPD(0) and controls	44,146	0.0483	0.0070	6.8686	1.0433	1,534	115,071
Leverage LEV(1), given LEV(0) and controls	44,146	0.0117	0.0077	1.5144	1.0040	144	10,831
Spread SPD(1), given its SPD(0), LEV(0) and controls	44,146	0.0493	0.0077	6.3660	1.0437	1,546	116,022
Leverage LEV(1), given LEV(0), SPD(0) and controls	44,146	0.0118	0.0077	1.5246	1.0041	148	11,072

Table 3 notes: The perfect foresight date 1 value is shown as variable(1); the known date 0 value is shown as variable(0). Variable definitions and constructions are in Section 3. All models have three control variables measured at the beginning of the year: Common Shareholders Equity (CSE(0)), Book-to-Market (BM(0)), and the previous year's Momentum (RET(0)). Φ-Retail is the value of information to a hypothetical investor with \$1M in wealth; Φ-Institutional is the value of information to a hypothetical investor with \$100M in wealth. The MSE Ratio is the ratio of the maximum likelihood estimator (MLE) of the posterior variance of the reduced model to the MLE of the posterior variance of the full model.

Table 4: Interaction Effects in the Value of Information

Panel A

Variables being Valued	N	R-squared from Full Model	R-squared from Reduced Model	R-squared Ratio	MSE Ratio	Φ-Retail	Φ-Institutional
Sales growth Sg(1), given its Sg(0) and controls							
1st Quintile of ROE(0)	6,851	0.0202	0.0135	1.4996	1.0069	248	18,594
2nd Quintile of ROE(0)	6,851	0.0228	0.0099	2.2975	1.0132	475	35,610
3rh Quintile of ROE(0)	6,851	0.0241	0.0082	2.9396	1.0163	584	43,798
4th Quintile of ROE(0)	6,851	0.0298	0.0136	2.1917	1.0167	599	44,908
5th Quintile of ROE(0)	6,851	0.0334	0.0090	3.7279	1.0253	903	67,746

Panel B

Operating Asset Turnover OAT(1), given OAT(0) and controls	N	R-squared from Full Model	R-squared from Reduced Model	R-squared Ratio	MSE Ratio	Φ-Retail	Φ-Institutional
1st Quintile of OPM(0)	8,830	0.0166	0.0114	1.4514	1.0053	190	14,218
2nd Quintile of OPM(0)	8,829	0.0107	0.0039	2.7062	1.0068	245	18,392
3rh Quintile of OPM(0)	8,829	0.0192	0.0047	4.1037	1.0148	532	39,936
4th Quintile of OPM(0)	8,829	0.0171	0.0025	6.8042	1.0148	532	39,913
5th Quintile of OPM(0)	8,829	0.0107	0.0032	3.3059	1.0075	271	20,358

Panel C

Leverage LEV(1) given LEV(0) and controls	N	R-squared from Full Model	R-squared from Reduced Model	R-squared Ratio	MSE Ratio	Φ-Retail	Φ-Institutional
1st Quintile of SPD(0)	8,830	0.0166	0.0114	1.4514	1.0053	190	14,218
2nd Quintile of SPD(0)	8,829	0.0086	0.0068	1.2579	1.0018	64	4,811
3rh Quintile of SPD(0)	8,829	0.0136	0.0064	2.1096	1.0073	261	19,595
4th Quintile of SPD(0)	8,829	0.0152	0.0084	1.8176	1.0070	251	18,822
5th Quintile of SPD(0)	8,829	0.0113	0.0032	3.5509	1.0082	295	22,163

Table 4 notes: The perfect foresight date 1 value is shown as variable(1); the known date 0 value is shown as variable(0). All models have three control variables measured at the beginning of the year: Common Shareholders Equity (CSE(0)), Book-to-Market (BM(0)), and the previous year's Momentum (RET(0)). Φ-Retail is the value of information to a hypothetical investor with $1M in wealth; Φ-Institutional is the value of information to a hypothetical investor with $100M in wealth. The MSE Ratio is the ratio of the maximum likelihood estimator (MLE) of the posterior variance of the reduced model to the MLE of the posterior variance of the full model. Variable definitions and constructions are in Section 3.

Table 5: Distribution of Value Estimates over Fama/French 38 Industries

Panel A: ROE(1) given ROE(0) and controls

	N	Mean	Minimum	25th	50th	75th	Maximum	Pooled estimates from Table 3
R-squared Ratio	36	8.5945	1.0870	3.2475	5.6626	10.7035	47.2629	10.1260
Φ-Retail	36	3,091	132	2,101	2,968	4,338	6,764	2,515
Φ-Institutional	36	231,897	9,878	157,617	222,659	325,468	507,506	188,680

Panel B: Sg(1) given Sg(0) and controls

	N	Mean	Minimum	25th	50th	75th	Maximum	Pooled estimates from Table 3
R-squared Ratio	36	2.2569	1.0000	1.1854	1.5337	2.6516	7.9884	2.1211
Φ-Retail	36	702	0	227	379	811	4,036	447
Φ-Institutional	36	52,652	18	17,065	28,431	60,837	302,777	33,540

Panel C: RNOA(1) given RNOA(0) and controls

	N	Mean	Minimum	25th	50th	75th	Maximum	Pooled estimates from Table 3
R-squared Ratio	36	7.0160	1.0012	2.7094	4.9753	8.1265	39.7274	8.2467
Φ-Retail	36	2,462	1	1,592	2,220	3,287	5,467	1,920
Φ-Institutional	36	184,712	99	119,470	166,590	246,622	410,144	144,020

Panel D: FIN(1) given FIN(0) and controls

	N	Mean	Minimum	25th	50th	75th	Maximum	Pooled estimates from Table 3
R-squared Ratio	36	4.1500	1.0065	1.3961	3.2578	7.2011	14.3183	3.4747
Φ-Retail	36	1,460	3	486	1,495	2,036	3,548	653
Φ-Institutional	36	109,561	207	36,483	112,134	152,778	266,170	48,987

Panel E: OPM(1) given OPM(0) and controls

	N	Mean	Minimum	25th	50th	75th	Maximum	Pooled estimates from Table 3
R-squared Ratio	36	4.5939	1.1298	2.1577	3.2440	5.9443	17.8202	5.1638
Φ-Retail	36	1,589	508	800	1,536	2,092	4,306	1,135
Φ-Institutional	36	119,228	38,109	60,031	115,208	156,924	323,086	85,129

Panel F: LEV(1) given LEV(0) and controls

	N	Mean	Minimum	25th	50th	75th	Maximum	Pooled estimates from Table 3
R-squared Ratio	36	1.9451	1.0011	1.0949	1.3759	1.9797	11.9209	1.5144
Φ-Retail	36	380	4	81	207	443	3,548	144
Φ-Institutional	36	28,525	310	6,074	15,522	33,253	266,191	10,831

Table 5 notes: This table presents the distribution of value estimates over 36 industries of the Fama-French 38 (two industries have insufficient data). The perfect foresight date 1 value is shown as variable(1); the known date 0 value is shown as variable(0). All models have three control variables measured at the beginning of the year: Common Shareholders Equity (CSE(0)), Book-to-Market (BM(0)), and the previous year's Momentum (RET(0)). Φ-Retail is the value of information to a hypothetical investor with $1M in wealth; Φ-Institutional is the value of information to a hypothetical investor with $100M in wealth. The MSE Ratio is the ratio of the maximum likelihood estimator (MLE) of the posterior variance of the reduced model to the MLE of the posterior variance of the full model. Variable definitions and constructions are in Section 3.

Table 6: Value of ROE(1) Given ROE(0) and Controls by Fama/French 12 industry

FF12 Code	FF12 Industry Name	N	R-squared from Full Model	R-squared from Reduced Model	R-squared Ratio	MSE Ratio	Φ-Retail	Φ-Institutional
1	Consumer Nondurables	3,778	0.1069	0.0075	14.1942	1.1112	3,815	286,237
11	Finance	2,499	0.0943	0.0062	15.3147	1.0973	3,358	251,976
9	Wholesale/Retail	5,985	0.0921	0.0105	8.7560	1.0899	3,114	233,650
7	Telecom	1,109	0.2230	0.1688	1.3211	1.0698	2,439	183,023
12	Other	6,617	0.0659	0.0016	41.1004	1.0689	2,409	180,763
8	Utilities	2,280	0.1365	0.0771	1.7706	1.0688	2,408	180,640
3	Manufacturing	7,653	0.0740	0.0105	7.0446	1.0685	2,398	179,947
6	Business Equipment	6,076	0.0738	0.0110	6.6981	1.0678	2,374	178,091
5	Chemicals	1,574	0.0752	0.0143	5.2467	1.0658	2,306	172,993
10	Healthcare	2,965	0.0650	0.0046	14.2798	1.0647	2,267	170,079
4	Energy	2,073	0.0667	0.0069	9.7147	1.0641	2,248	168,665
2	Consumer Durables	1,537	0.0994	0.0500	1.9883	1.0548	1,931	144,908

Table 6 notes: The perfect foresight date 1 value is shown as variable(1); the known date 0 value is shown as variable(0). All models have three control variables measured at the beginning of the year: Common Shareholders Equity (CSE(0)), Book-to-Market (BM(0)), and the previous year's Momentum (RET(0)). Φ-Retail is the value of information to a hypothetical investor with \$1M in wealth; Φ-Institutional is the value of information to a hypothetical investor with \$100M in wealth. The MSE Ratio is the ratio of the maximum likelihood estimator (MLE) of the posterior variance of the reduced model to the MLE of the posterior variance of the full model. Variable definitions and constructions are in Section 3.

Table 7: Value of ROE(1) Given ROE(0) and Controls by Fama-French 12 Industry on ETF Sample

		Two-Asset Investment Set					One-Asset Investment Set				
FF12 Code	FF12 Industry Name	N	Determinant from Full Models	Determinant from Reduced Models	Phi-Retail	Phi-Institutional	N	MSE Ratio	Phi-Retail	Phi-Institutional	Phi-Retail Two-Asset minus Phi-Retail One-Asset
7	Telecom	251	0.0064	0.0075	6,158	461,980	251	1.09	3,124	234,370	3,034
11	Finance	558	0.0101	0.0121	6,345	476,021	558	1.148	5,004	375,458	1,341
6	Business Equipment	1,530	0.0162	0.0174	2,685	201,450	1,530	1.046	1,626	121,973	1,059
9	Wholesale/Retail	938	0.0049	0.0053	2,903	217,832	938	1.068	2,389	179,217	514
4	Energy	589	0.0555	0.0572	1,070	80,245	589	1.017	592	44,446	478
1	Consumer Nondurables	545	0.0023	0.0025	2,569	192,760	545	1.067	2,329	174,772	240
12	Other	1,106	0.0154	0.0157	636	47,697	1,106	1.011	399	29,916	237
5	Chemicals	405	0.0100	0.0111	3,676	275,834	405	1.1	3,451	258,903	225
3	Manufacturing	1,224	0.0120	0.0123	857	64,317	1,224	1.022	775	58,123	83
2	Consumer Durables	199	0.0238	0.0260	3,247	243,602	199	1.093	3,217	241,364	30
10	Healthcare	841	0.0099	0.0105	1,897	142,309	841	1.056	1,956	146,756	-59
8	Utilities	528	0.0019	0.0020	2,961	222,134	528	1.087	3,027	227,133	-66

The perfect foresight date 1 value is shown as variable(1); the known date 0 value is shown as variable(0). All models have three control variables measured at the beginning of the year: Common Shareholders Equity (CSE(0)), Book-to-Market (BM(0)), and the previous year's Momentum (RET(0)). Φ-Retail is the value of information to a hypothetical investor with $1M in wealth; Φ-Institutional is the value of information to a hypothetical investor with $100M in wealth. The MSE Ratio is the ratio of the maximum likelihood estimator (MLE) of the posterior variance of the reduced model to the MLE of the posterior variance of the full model. Variable definitions and constructions are in Section 3. Estimated on the industry ETF sample. The last column gives the value from two-asset estimation less the value from the one-asset estimation on the industry ETF sample.