FASB was right: Earnings beat cash flows when predicting future cash flows *

Ray Ball† Valeri Nikolaev‡

December 2020

Abstract

Do accruals-based earnings provide better information about future operating cash flows than operating cash flows themselves, as predicted by FASB's conceptual framework? The most recent evidence (Nallareddy et al., 2020) is that operating cash flows, measured correctly using cash flow statement data, consistently outperform earnings by a large margin. This evidence is largely based on comparing operating cash flows with a version of "bottom line" earnings, handicapping earnings by including non-operating and transitory components with no corresponding operating cash flow. Testing the tenet underlying FASB's statement requires comparing the same earnings variable calculated on cash and accruals bases. *Operating* earnings consistently dominate operating cash flow's predictive ability in a battery of tests, especially after addressing cross-sectional differences among firms.

Keywords: Operating Cash Flow, forecasts, predictive ability, earnings, accruals, FASB, informativeness
JEL: G11, G12, G17, M41, M48

*We thank Patty Dechow, Wyane Landsman, Suresh Nallareddy, Stephen Penman, K. Ramesh, Mani Sethuraman, David Veenman, Mohan Venkatachalam and Stephen Zeff, the referee and editor, as well as workshop participants at Rice University and in the Accounting Design Project of Columbia University, for their helpful comments. We also thank Gabriel Varela for research assistance. We both gratefully acknowledge financial support for this research from the University of Chicago Booth School of Business.

† **The University of Chicago**; Postal Address: Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637, USA; E-mail: ray.ball@chicagobooth.edu.

‡ **The University of Chicago**; Postal Address: Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637, USA; E-mail: valeri.nikolaev@chicagobooth.edu.

1. Introduction

A foundational principle in accounting is that earnings are more informative than cash flows. As expressed in the academic literature, the principle is that accruals ameliorate the problem that "realized cash flows have timing and matching problems that cause them to be a 'noisy' measure of firm performance" (Dechow 1994, p.4). Indeed, analytical accruals models conclude that earnings are better predictors of future operating cash flows than current operating cash flows (Dechow et al., 1998). The practitioner literature includes the famous Financial Accounting Standards Board (FASB, 1978) proposition: "[Users'] interest in an enterprise's future cash flows and its ability to generate favorable cash flows leads primarily to an interest in information about its earnings rather than information directly about its cash flows."

In light of the foundational nature of this issue and the notoriety of the FASB proposition, it is surprising that evidence remains mixed as to whether earnings in fact are superior to current-period cash flows as an indicator of future cash-generating ability. A recent study by Nallareddy et al. (2020) takes an important step toward reconciling the conflicting evidence. Among other things, the authors show that mixed results are largely attributable to some prior studies estimating operating cash flows from balance sheet data, without using actual numbers from cash flow statements. Balance sheet estimates are known to measure operating cash flows with error when the accounting entity changes from one balance sheet to the next (Drtina and Largay III, 1985), most notably due to mergers, acquisitions and divestitures (Greenberg et al., 1986; Hribar and Collins, 2002). The measurement error then attenuates the ability of current operating cash flows to predict future operating cash flows.[1]

The most striking result in Nallareddy et al. (2020) is that operating cash flows, when measured correctly using statement of cash flows data, consistently outperform earnings in predicting future operating cash flows. This result can be seen in pooled estimation, cross-

[1]Firms were not required to report cash flow statements until after SFAS 95 was passed in 1987, so earlier studies were particularly exposed to this problem.

sectional regressions, and firm-level time-series regressions (as well as in an international replication). In terms of economic magnitudes, the authors conclude that that the predictive ability of cash flows is an astonishing 1.56 times that of earnings. Further, cash flows outperform earnings in every single year, so the result undoubtedly is statistically significant. From this analysis, Nallareddy et al. (2020) draw the startling conclusion (p.3): "Accounting standard setters such as the FASB and the IASB would find our evidence relevant as it challenges an important tenet of financial reporting, i.e., accounting earnings, as a summary metric, provides a better basis for predicting future cash flows."

This conclusion challenges not only the famous FASB proposition; more fundamentally, it challenges the economic resources devoted to collecting, calculating, reporting and auditing accruals and accrual-based earnings information, the objective of which is to make earnings more informative than cash flows (Dechow, 1994), not less. It challenges the attention Wall Street gives to earnings, as well as the use of earnings in compensation, debt, supply, royalty, M&A and other contracts. This result is of foundational interest to firms reporting and contracting on the basis of their financials, to all financial statement users, to accounting standard setters and regulators, and even to accounting instructors.

A priori, this negative conclusion could be caused by either of two institutional breakdowns. There could be substantial deficiencies in GAAP or in its implementation (such as poor estimation of uncollectible accounts receivable). Alternatively, there could be widespread accruals manipulation by managers ("earnings management"). Both of these forces undoubtedly exist to some degree, but do they rise to the level that standard setters and regulators should re-evaluate the need for calculating and reporting earnings, as Nallareddy et al. (2020) (p.3) seem to imply? The evolution of economic institutions is by no means guaranteed to generate an economically efficient institutional structure (see the studies in Dixit et al. (2011), for example). At the same time, the emergence and survival of accruals-based accounting over centuries and in multiple jurisdictions worldwide is challenging to reconcile with the seeming superior information contained in operating cash flows relative to earnings.

Is it possible that the substantial costs of reporting accruals-based earnings outweigh the benefits? Why not merely classify and report cash flows?

The essence of FASB's famous proposition is that accruals-based accounting is more informative about firms' future outcomes than cash-based accounting. In our view, the only valid way to test the essence of the proposition is to compare the predictive ability of an earnings variable calculated on an accruals basis with that of the same variable calculated on a cash basis, and using a specification that does not unduly constrain the role of accruals in a way that is largely at odds with their function. The research design in Nallareddy et al. (2020) does not satisfy either of the above criteria. First, the study's base case compares near "bottom line" earnings (income before extraordinary items and discontinued operations) with operating cash flows, which is akin to comparing apples and oranges. Second, the main specifications involve cross-sectional regressions, whereas the role of accruals is primarily to address over-time fluctuation (timing issues) in cash flows on a firm by firm basis. Cross-sectional regressions implicitly assume that the objective of accrual accounting is to measure performance (by adjusting cash flows) as a function of where the firm stands relative to other firms, most of which are in different industries and have business models. This research design can hardly be viewed as a test of the *essence* of the FASB (1978) proposition.[2]

We employ a different research design to address the above issues. We note that income before extraordinary items has components that do not map into (articulate with) *operating* cash flows and are not intended to. Rather, many earnings components map to investing or financing cash flows. For example, depreciation is an accruals-generated expense deducted from earnings, but its cash equivalent is an investment cash flow that does not affect operating cash flows at all. Another example is non-recurring items, which affect earnings more than cash flows because they primarily originate in non-current accruals. In any year the relative

[2] Nallareddy et al. (2020) indirectly acknowledge the first of these issues when choosing earnings before extraordinary items and discontinued operations, not "bottom line" Net Income or Comprehensive Income, as the base case earnings variable. Our view is that, having taken one or two steps up from the "bottom line," it would have been preferable to proceed further up the Income Statement to extract the information in earnings that most closely corresponds with operating cash flow.

magnitudes of non-operating earnings components vary considerably across firms, so in cross-sectional regressions their inclusion in an earnings predictor adds noise that is not present in an operating cash flows predictor. This attenuates the apparent predictive ability of earnings relative to that of operating cash flows.

Further, because accrual accounting shifts cash flows over time within a given firm, and not across firms, the most meaningful comparison of cash flows vs. accrual-based measures is at the firm level. We use research designs that focus on the firm rather than the cross-section and hence allow the relation between earnings and future operating cash flows to vary by firm. We expect this to be important due to firm differences in characteristics that affect accruals, such as the persistence of operating profits, the relative magnitude of working capital, operating cycle length and volatility of exogenous shocks to future operating cash flows.

We address these issues by performing a comprehensive set of tests. We begin with cross-sectional analysis, and subsequently isolate cross-sectional firm heterogeneity using pooled panel data estimation with firm fixed effects, industry-level estimation that allows all regression parameters to vary by industry and, finally, firm-level time-series estimation. Across these tests, we find that when earnings components that do not map into the operating section of cash flow statements are excluded, and especially when the assumption of cross-section homogeneity is relaxed, earnings consistently outperform cash flows, both in and out of sample. Further, earnings then largely subsume the information in operating cash flows when both predictors are used, consistent with operating cash flows being a garbled or noisy earnings measure, and with the noise being reduced by accruals (Dechow, 1994). In industry-level regressions with firm fixed effects, thereby allowing for both firm- and industry-level differences in business models and accounting methods, all earnings variables we study out-predict operating cash flows. Indeed, the earnings variable that corresponds most closely to an accruals-based version of operating cash flows comfortably subsumes operating cash flows in a "horse race" regression. These results hold for one, two and three year prediction horizons. We find similar results when replacing operating cash flows with free cash flows, especially when adding current-year capital expenditures as a predictor. It transpires that FASB was right: Earnings beat cash flows as a predictor of future cash flows.

We also confirm the result in Kim and Kross (2005) and Nallareddy et al. (2020) that the ability of earnings to predict future operating cash flows has increased in recent decades. This change is consistent with the increase over approximately the same period in the share price reaction to earnings announcements, as reported by Ball and Shivakumar (2008) and subsequently updated using similar metrics by Beaver et al. (2018) and Beaver et al. (2020). The increased ability of earnings to predict future outcomes suggests that the increased market reaction to earnings announcements is, at least in part, a rational response to a "real" increase in the information earnings contain, and is not simply a market artifact. Further, operating cash flows exhibit a similar increase in ability to predict their future values, which suggests that the increased market reaction to earnings announcements is not due to changes in GAAP, but most likely reflects changes in underlying firm characteristics.

Our findings have implications for research using accounting information generally. One generalizable implication is that many accounting variables such as earnings, cash flows and book values are comprised of components that differ economically. We demonstrate this in the context of using earnings to predict operating cash flows, but it also is the case when predicting stock returns (Ball et al., 2015, 2020) and possibly in many other contexts. Another implication is the importance of addressing firm heterogeneity in contexts where accounting rules interact with firm characteristics to produce meaningfully different firm-level relationships. Here too we demonstrate this in only one context, but we have found it also is the case in estimating Basu (1997) conditional conservatism (Ball et al., 2013), which suggests that it also is a structural issue. Finally, a practical implication from our analysis is that an operating earnings variable based on Dechow and Dichev (2002) is the most robust and effective predictor of future operating cash flows.

2. Background: Why this is a foundational issue

From the user-focused revolution in the academic and practitioner literatures during the 1960s and 1970s there emerged two major tests of whether accrual accounting methods and practices add information relative to simply counting and reporting cash flows. Both tested this fundamental issue by benchmarking accruals-based earnings against cash flows. Earnings presumably were selected for attention largely because they are important to many users in their own right. In addition, transactions recorded in Income Statements flow through onto Balance Sheets, so the properties of earnings directly affect other accounting information as well. Cash flows presumably were selected as the benchmark in order to measure the extent to which accrual accounting methods and practices add useful information.

One test was whether earnings exhibit a higher correlation with stock returns than do cash flows (e.g., Ball and Brown (1968), Dechow (1994)), a benchmark that morphed into "value relevance." A subsequently developed test was whether earnings exhibit a higher correlation with future cash flows (e.g., Bowen et al. (1986), Greenberg et al. (1986), Finger (1994)).[3] The tests are related but differ in horizons; cash flow prediction is over short horizons such as one year, whereas stock returns reflect revisions in expectations of all future cash distributions to owners.[4] More qualitative evidence comes from the use of accounting information in debt, compensation, supply, licensing and other contracts. For example, many debt and compensation contracts are written on EBITDA (operating profit). It is rare for contracts to use operating cash flow itself. In these contexts, users reveal preferences for accrual-based measurement.

Using informativeness about future cash flows as a criterion for assessing earnings reflected the decision-usefulness theory of financial reporting that emerged in the literature, and that since has morphed into "real effects." The major proponent of this theory was Staubus (1961), who stressed (p.15) investor demand for information about "the prospects

[3]Nallareddy et al. (2020) provide a comprehensive survey.

[4]We investigate cash flow prediction horizons of one, two and three years.

for cash receipts." Widespread attention to the role of earnings in estimating future cash flows was stimulated by the influential Trueblood Report, which stated (AICPA 1973, p.22): "Users need to know about probable cash movements of an enterprise to estimate cash flows to them. The periodic measurement of earnings by enterprises becomes a basis for these estimates." This prompted FASB to assert in its first conceptual framework that the fundamental objective of financial statements is providing users with information that is useful in assessing a firm's ability to generate future cash flows (FASB 1978, p.17). Mirroring Trueblood, FASB then offered its famous proposition: "[Users'] interest in an enterprise's future cash flows and its ability to generate favorable cash flows leads primarily to an interest in information about its earnings rather than information directly about its cash flows. ... Information about enterprise earnings and its components measured by accrual accounting generally provides a better indication of enterprise performance than does information about current cash receipts and payments." (FASB 1978, p.19).[5]

The essence of this FASB proposition is that accruals-based accounting is more informative to users than cash-based accounting. It is worth noting that FASB defined neither earnings nor cash flows; it was merely making a conceptual statement about accruals-based versus cash-based measures. Indeed, FASB was careful to refer to "earnings and its components providing information about future cash flows – a specific acknowledgement that it was not referring simply to "bottom line" earnings as the source of information about future cash flows.

"General purpose" financial statements are required under GAAP to report separate line items for a variety of earnings components. Elaborating on its carefully qualified proposition (cited above) that "earnings and its components" are more informative than cash flows, FASB (1984, paras. 20-22) stated: "Analysis aimed at objectives such as predicting amounts,

[5] The issue is of worldwide interest. Zeff (2016) documents how the Trueblood Report's adoption of a cash flow prediction objective for accounting information was taken up by standard setters internationally, including the Australian, Canadian and U.K. national bodies, and the International Accounting Standards Committee (the precursor to the International Accounting Standards Board). It is reflected in the current joint conceptual framework of the FASB and the IASB.

timing and uncertainty of future cash flows requires financial information segregated into reasonably homogeneous groups. For example, components of financial statements that consist of items that have similar characteristics in one or more respects, such as continuity or recurrence, stability, risk, and reliability, are likely to have more predictive value than if their characteristics are dissimilar. ... the Board believes it is important to avoid focusing attention on the 'bottom line'."

Consequently, Income Statements under GAAP provide information about gross profit, operating profit, operating profit before interest and taxes (EBIT), net income before or after income taxes, before and after excluding discontinued operations, special items such as asset write-downs or restructuring charges.[6] GAAP thereby recognizes that users select earnings variables that meet their particular purposes, and requires firms to report a menu of line items from which users can select.[7] This richness of earnings variables available to users of GAAP-compliant financials is part of the context in which FASB referred to "information about enterprise earnings and its components" in its famous proposition.[8]

3. Selecting the earnings variable

The principles we follow in testing this important FASB proposition are simple. First, the difference between the earnings and operating cash flow variables selected for comparison should be due – as much as possible, given data constraints – only to accrual versus cash accounting. Second, as far as possible the research design should not impose restrictions on the tested relation between earnings and future cash flows that the financial statement user does not encounter, such as across-firm homogeneity.

[6] In addition, GAAP requires firms to report Comprehensive Income, which is Net Income with a variety of additional unrealized gains and losses incorporated.

[7] From this perspective, it is somewhat misleading to use the term "GAAP earnings" to describe the "bottom line" earnings variable that GAAP labels as Net Income. Under GAAP, Income Statements are required to report information about a variety of earnings measures. Earnings under GAAP is a multi-dimensional concept.

[8] Equally, GAAP requires "bottom line" cash flow (the difference between the total cash balances on hand at the beginning and at the end of the accounting period) to be classified into operating, investing and financing flows. Here too, GAAP recognizes that users select cash flow variables that meet their particular purposes.

If the first of these principles is not followed, and the earnings variable studied is not an accruals-based version of the cash flow variable studied, the results conflate two effects. One effect is the information that accrual accounting adds to – or subtracts from – the studied earnings variable. The confounding effect is due to the difference between two cash flow variables: the cash flow variable studied and the cash flow variable that corresponds to the earnings variable studied.[9]

The earnings variable used by Nallareddy et al. (2020) in the body of their study is income before extraordinary items and discontinued operations. The cash flow variable is cash flows from operations as reported in cash flow statements, also net of extraordinary items and discontinued operations. Comparing the predictive abilities of these variables leads to the conclusion (p.6): "The findings using both cross-sectional and time-series approaches are in sharp contrast to FASB assertions and recent evidence supporting the superiority of earnings as a summary metric for predicting future cash flows." In an appendix, the authors summarize untabulated results using several additional earnings variables, namely: operating income before depreciation, operating income after depreciation, and pretax income and income excluding special items. From these analyses, the authors conclude (p.7): "we find that cash flows consistently outperform earnings in predicting future cash flows."[10,11]

3.1. Excluding non-operating earnings components

We believe the most informative comparison is between operating cash flows and an *operating* earnings variable. The latter is the component of "bottom line" earnings that maps to the

[9] The meaning of the results stemming from such a test is unclear. What do we learn from comparing an accruals version of earnings variable A and a cash version of earnings variable B? What does the difference *mean*?

[10] Nallareddy et al. (2020) report another result in footnote 11: 'Among the alternative measures of earnings, we find (results not tabulated) that earnings defined as operating income before depreciation exhibit superior predictive ability for future cash flows." The footnote adds the observation: "However, this is not surprising as operating income before depreciation is, by construction, closer to operating cash flows than the other measures of earnings outlined above." This in our view comes closest to an apples-to-apples comparison of the information in cash versus accrual accounting, absent confounding effects.

[11] We demonstrate below that the apparently inferior predictive ability of earnings variables disappears when the regression specification allows for heterogeneity among firms.

operating section of the cash flow statement. Many components of "bottom line" earnings relate to operating activities and, as such, they pass through the operating section of the cash flow statement at some point. For example, accrued revenues and expenses enter operating cash flows after they enter earnings; similarly, deferred revenues and expenses enter operating cash flows before they enter earnings. However, "bottom line" earnings also contain accruals and cash flows related to the financing and investing components of the cash flow statement, and as such they do not correspond to and never enter cash flow from *operations*. An obvious example is depreciation and amortization expense, which is a long-term accrual related to capital investment that has no operating cash flow equivalent. Rather, depreciation corresponds to (is a lagged function of) investing cash flows in multiple prior periods. Other examples of non-operating components of "bottom line" earnings include gains and losses on asset dispositions, realized gains and losses related to sales of marketable securities, impairments of intangible assets such as goodwill, special items (which include other kinds of gains and losses usually related to non-operating transactions), amortization of premiums and discounts on long-term debt, a portion of equity method earnings, and non-cash stock-based compensation.[12] Because these items have no corresponding operating cash flows, "bottom line" earnings is simply not the right comparison with cash flows from operations for testing the FASB proposition, or for assessing the information added to earnings by accrual accounting.

The relative magnitudes of non-operating items that are in included in "bottom line" earnings, but not in operating cash flows, vary across firms. For example, depreciation and amortization expense has a larger effect on "bottom line" earnings in capital-intensive firms. Consequently, in cross-sectional regressions of future operating cash flows on earnings, these items make the earnings variable studied a noisy measure of the accruals-based earnings equivalent to operating cash flows, thereby attenuating its estimated predictive power.[13]

[12]Stock option expense does not map to *any* cash flows. If and when options are exercised, there is an *increase* in financing cash flows, which bears no mapping to the decreases in "bottom line" earnings in the prior periods over which the option grants earlier were expensed.

[13]This is why we believe the untabulated test reported briefly in footnote 11 of Nallareddy et al. (2020) is

3.2. Excluding non-recurring earnings components

An additional reason for not using "bottom line" earnings to predict future operating cash flows is that, relative to operating earnings, they contain more transitory components. By definition, transitory items do not recur in the future and have no predictive power. Transitory components tend to be large and on average negative (Basu, 1997), and are a major reason for the high frequency of negative earnings (Collins et al., 1997) in firms with positive prices.

Because these "bottom line" earnings components are transitory, security analysts regularly exclude them when constructing a "Street" version of actual earnings and when issuing earnings forecasts. Gu and Chen (2004) show that Street earnings are more persistent and have higher valuation multiples than the excluded components. Li (2010) demonstrates that long term debt contracts are more likely to base covenants on current earnings exclusive of transitory components than are short term contracts. Bradshaw and Sloan (2002) p.41) observe that "managers, security analysts, investors, and the press rely increasingly on modified definitions of GAAP net income, known by such names as 'operating' and 'pro forma' earnings." They show that stock prices respond primarily to Street earnings, and conjecture (p.42) that managers and analysts could be trying to inflate share prices, or they could be stripping transitory components from earnings to obtain "an improved measure for determining future cash flows and hence firm value." Their evidence does not allow them to distinguish between these competing explanations. Our evidence contributes to this important debate and is rather consistent with the second explanation.

Large transitory components of earnings such as asset impairment charges or loss provisions tend to be accruals. Consequently, they affect "bottom line" earnings more than they affect cash flows, so they hobble the apparent predictive ability of earnings relative to that of cash flows. Here also, the relative magnitude of transitory components in earnings varies across firms, further attenuating the explanatory power of "bottom line" earnings in

the study's most meaningful earnings and cash flows comparison.

cross-sectional regressions.

3.3. Cross-sectional heterogeneity

Cross-sectional regressions in this literature typically assume the relation between current earnings, current cash flows and future cash flows is homogeneous across firms. This assumption seems unlikely to be valid because accrual accounting does not to adjust (nor it should) for cross-sectional differences in cash flows. One issue is that firms differ in how profitably they have invested in the past. Consequently, even when variables are scaled by book value of total assets, the magnitudes of firms' scaled earnings and operating cash flows are positively correlated in cross-section. This will be the case independently of the extent to which year-to-year variation in earnings or in operating cash flows is informative about future operating cash flows. Failure to take cross-sectional heterogeneity into consideration in this context thus can result in misleading inferences. Another reason for cross-sectional heterogeneity to be problematic is the presence of earnings components without a corresponding mapping to operating section of the cash flow statement that are not homogeneous across firms.

3.4. Addressing the three research design issues

In each of the above three ways, cross-sectional regressions of future cash flows on "bottom line" earnings constrain the use of earnings information when predicting cash flows in ways that the user of earnings information in practice does not encounter. Users can choose an operating earnings variable that corresponds most closely to operating cash flows. They can choose to ignore or to incorporate into their predictions any information in the transitory accruals components of earnings. In addition, based on their knowledge of firm characteristics such as industry membership and operating cycle length, users can use accounting information in different ways for different firms, or in different ways over time.[14]

[14] Similarly, in offering its famous proposition, FASB (1978) did not constrain the relation it envisaged between earnings and future cash flows to conform to a linear prediction model. Nor did it constrain the

In our analysis that follows, we are able to partially relax the constraints that cross-sectional regressions impose on the real world by making our own choices. We choose an operating earnings measure, avoid transitory earnings components, and employ a variety of techniques to address heterogeneity across firms. These statistical techniques cannot fully replicate the ability of users to base cash flow predictions on earnings. For example, they do not incorporate all of the knowledge users possess about firm characteristics. Nevertheless, even a partial relaxation of the the above constraints has a profound effect on the empirical results, as we show next.

4. Methodology and Variables

We explore several proxies for accruals-based operating earnings. They are described below in increasing order of what we assess to be their validity for testing the essence of the Trueblood/FASB proposition. We expect this to also be the order in which they are able to predict future operating cash flows. Table 1 summarises how each variable is constructed.

We use income before extraordinary items (IBC), as reported in cash flow statements, to represent "bottom line" earnings.[15] For reasons discussed above, we expect it to perform the worst of the earnings variables in predicting operating cash flows. We then compute "adjusted income before extraordinary items" (IBC_A), removing items from earnings that have no future operating cash flow equivalent: depreciation and amortization, extraordinary items and discontinued operations, realized gains and losses on sale of fixed assets and investments, equity method income and minority interest, special items, and other unrealised gains and losses. We expect IBC_A to predict better than IBC, but not as well as the following operating earnings variables.

As one representation of accruals-based operating earnings, we use Compustat's operating

parameters of any prediction to be constant across firms, or across time. In its newly-articulated user orientation, the standard setter presumably was aware that users may differ in how they use accounting information, and may use it in different ways for different firms and in different time periods.

[15] With the exception of gains and losses classified as "extraordinary," which are infrequent, IBC does not exclude non-operating or transitory earnings components.

profit (OP) variable, which is measured before depreciation and amortization. For what we regard as the legitimate operating earnings variable for comparison with cash flows from operations, we add working capital accruals (i.e., operating accruals) to operating cash flows, following Dechow and Dichev (2002), and label it "operating earnings" (OE). By definition, working capital assets and liabilities are those with a cycle length of one year or less. The accruals that create working capital assets (other than cash itself) and liabilities therefore are designed to bring into current-year earnings those cash flows that current-year operating transactions generated in the previous year or are expected to generate in the following year. Because operating cash flows by definition arise from transactions with a cycle length of one year or less, whereas longer-cycle transactions are classified as investing or financing cash flows, they naturally are additive to working capital accruals to create operating profit. OE excludes long-term (non-operating) accruals and cash flows that are included in earnings but are unrelated to operating activities, such as realized gains on asset sales. These are the "bottom line" earnings components that do not naturally map to next-period operating cash flows and that consequently add noise to the independent variable in predictive regressions.

Conceptually, OE is the version of earnings we study that is best aligned with cash flows from operating activities. It is the earnings variable that we expect to exclude the most cross-sectional variation in non-operating items that create noise in predicting operating cash flows, so we expect it to show the most predictive ability.

We use several techniques to address heterogeneity across firms, an issue that plagues cross-sectional estimation. We employ firm-level fixed effects in pooled regressions, conduct within-industry analyses with firm-fixed effects, and also estimate predictive ability at the individual-firm level using time series data.

With these measures at hand, we perform a comprehensive set of tests to explore how the predictive ability of operating cash flows compares to that of more appropriately measured (for the purpose of this comparison) accrual-based earnings. First, we employ variations on univariate OLS regressions to predict operating cash flows that take these forms:

$$\text{Operating Cash Flow}_{i,t+1} = \alpha + \beta \text{Predictor}_{i,t} + \tilde{\epsilon}_{i,t+1} \tag{1}$$

where i indexes firms, t indexes years and *Predictor* is one of operating cash flows (CF), income before extraordinary items (IBC), income adjusted for non-operating items (IBC_{A}), operating profit (OP), and the Dechow-Dichev (2002) operating earnings (OE). See Section (4.2) and Table 1 for variable definitions.

We use equation (1) to compare the predictive abilities of earnings and operating cash flows. We examine cross-sectional, pooled, and firm-level estimation to generate forecasts. We also perform analyses both in sample and out of sample. We focus on R^2s but also examine the slope coefficients, which should be inversely related to the degree of noise in a particular proxy for expected operating cash flows.[16]

Second, we explore the predictive ability of accrual-based measures incrementally to operating cash flows. This test is motivated by the models of accruals (Dechow and Dichev, 2002; Nikolaev, 2018) suggesting that operating earnings and cash flows can be viewed as noisy measures of the underlying economic performance. Specifically, Nikolaev (2018) model states:

$$CF_{i,t+1} = \pi_{i,t} + \Delta w_{i,t}, \qquad OE_{i,t+1} = \pi_{i,t} + \Delta v_{i,t}$$

where $\pi_{i,t}$ is the observable economic performance that persists over time, $w_{i,t}$ and $v_{i,t}$ are timing and estimation errors in operating cash flows and earnings, respectively (i.e., transitory components). The model implies that both cash flows and accruals will carry noisy information about $\pi_{i,t}$ and hence, generally, both should be incrementally informative. However, if the estimation error in earnings is relatively small as compared to timing errors in cash flows, i.e., if accruals are highly effective at reducing the noise in cash flows, earnings will dominate and potentially even subsume the information in cash flows. To test this, we

[16] Note that a proxy for expected future operating cash flows measured without expectational error is expected to have a slope coefficient of one.

use the following model:

$$\text{(2)} \quad \text{Operating Cash Flow}_{i,t+1} = \alpha + \beta_1 \text{Operating Cash Flow}_{i,t} + \beta_2 \text{Predictor}_{i,t} + \tilde{\epsilon}_{i,t+1}$$

where the variables are defined as previously and now *Predictor* is one of the four earnings measures: income before extraordinary items (IBC), income adjusted for non-operating items (IBC_{A}), operating income (OP), and the Dechow and Dichev (2002) operating earnings variable (OE). This model allows us to test whether earnings and operating cash flows add incremental information about future operating cash flows, including whether one of the variables subsumes the other. For example, if operating cash flow (earnings) is a noisier measure of future operating cash flows, we would expect it to have an insignificant amount of incremental information relative to the less noisy variable.

Given the findings in Nallareddy et al. (2020), we formulate the null hypothesis in the following form:

H_0: *Operating cash flows are a superior predictor of future operating cash flows compared to accruals-based earnings.*

4.1. Out-of-sample performance

To evaluate the out-of-sample predictive ability of operating cash flows versus earnings, we perform two types of analysis: (1) cross-sectional predictive regressions; and (2) rolling window firm-level predictive regressions. To measure the out-of-sample predictability of operating cash flows, we use the out-of-sample R^2, as is standard in the literature (Campbell and Thompson (2008); Welch and Goyal (2008); Kelly and Pruitt (2013); Lu and Nikolaev (2019)).

For the cross-sectional predictive regressions, we estimate equation (1) for every year T in the sample and use the estimated parameters to construct the out-of-sample forecasts based data up to time $T-1$. Based on such forecasts, we calculate the out-of-sample R^2 as follows:

$$R_T^2 = 1 - \frac{\sum_j (y_{j,T} - \hat{y}_{j,T|T-1})^2}{\sum_j (y_{j,T} - \bar{y}_T)^2}$$

where j indexes firms and T indexes years, $y_{j,T}$ denotes actual future operating cash flows, $\hat{y}_{j,T|T-1}$ denotes predicted operating cash flows conditional on data up to time $T-1$, and $\bar{y}_T$ denotes the unconditional mean across firms based on the information up to time $T-1$. For example, when predicting 2002 operating cash flow for a given firm, we use observations prior to 2002 to estimate $\bar{y}_{j,T|T-1}$. We do not use cash flows for 2002 or subsequent years in this calculation. The same process is repeated for 2003 and so on.

For the firm-level forecasts, we use a rolling window to estimate equation (1), such that $t \in \{1, ..., T\}$. For each firm, we start the sample in the first year with available data and end at year T. Subsequently, we forecast operating cash flows CF_{T+1}. We require 15 years of data with non-missing observations prior to the forecasted year, i.e., for years $T-14$ to T. The firm-level out-of-sample R^2 is computed based on these forecasts, in an analogous manner:

$$R_j^2 = 1 - \frac{\sum_t (y_{j,T} - \hat{y}_{j,T|T-1})^2}{\sum_t (y_{j,T} - \bar{y}_{j,T})^2}$$

where $\hat{y}_{j,T|T-1}$ is predicted operating cash flows based on the rolling window ending at time T, and $\bar{y}_{j,T}$ denotes the (unconditional) mean estimated over the same rolling window. Out-of-sample R^2s can be negative when the unconditional mean is a more accurate predictor.

4.2. Sample construction and data

The data are from Compustat's Fundamentals Annual file. The sample starts in 1988 as this is the first fiscal year for which cash flow statement data are available, following the passage of SFAS 95. Our data requirements follow those of Nallareddy et al. (2020). We require that a company has total assets (at) and sales (sale) above ten million US dollars and a fiscal year-end closing price (prcc_f) above one US dollar. We drop financial firms, defined as those with a one-digit SIC code (sich) of 6. Firm/year observations with zero values are not treated as missing because all variables we use are routinely reported by firms whenever their values are materially non-zero. However, the results are little changed when firm/years with a zero value for any of the variables are excluded.

Operating cash flows (CF) are measured as cash flows from operating activities as reported in the Statement of Cash Flows (oancf). Income before extraordinary items (IBC) as reported on the Statement of Cash Flows (ibc) is used as the measure of near "bottom line" earnings.

The three earnings variables that exclude items without mapping to operating cash flows are measured as follows. Adjusted income before extraordinary items (IBC_A) starts with IBC and adds back depreciation and amortization (dpc), extraordinary items and discontinued operations (xidoc), realized gains and losses on sale of fixed assets or investments (sppiv), equity method income (esubc), minority interest, special items, and other unrealised gains and losses (fopo).[17] The second variable is Compustat's operating profitability (OP), which also is measured before depreciation and amortization (oidbp). The third earnings variable is operating profitability (OE) defined as the sum of working capital accruals and operating cash flows, following Dechow and Dichev (2002), which does not include long term accruals (notably, depreciation and amortization).[18]

All variables are scaled by average total assets. Subsequently, we truncate the scaled variables at the top and bottom 1% of their distributions (pooled across all firms and years).

[17] Compustat item fopo aggregates a range of heterogenous accrual components that are effectively unrelated to working capital accruals and are largely transitory, e.g., fair value adjustments related to financial assets and liabilities (Dechow et al., 2020).

[18] Regulation S-X Rule 5-02 requires registrants to report totals for Current Assets and Current Liabilities "when appropriate." While ASC 210-10-05-4 notes that most firms comply, some firms report "non-classified" balance sheets, without clearly distinguishing between current and non-current items, apparently because the distinction is not always clear. For example, cigarette companies typically classify leaf tobacco inventory as current, even though a portion is aged more than one year. Some receivables originate in complex transactions, hence the Compustat accounts receivable variable recch includes "Long-term receivables included by the company in the Operating Activities section." The effect would be to include a small amount of noise in our OE earnings variable.

4.3. Summary statistics

Table 2 Panel A provides summary statistics for the five variables. Income before extraordinary items (IBC) shows a substantial frequency of losses, consistent with the evidence in Collins et al. (1997). Its mean and median values (as a proportion of total assets) of 0.021 and 0.038 are approximately only one-fifth and two-fifths of the equivalent OE values, respectively. This is intuitive as IBC includes "one time" charges, which are predominantly negative (Basu, 1997) and are not treated under GAAP as extraordinary. Consequently, IBC is more left-skewed than the three other earnings variables. IBC also deducts (substantial) depreciation and amortization expenses, unlike operating cash flows and the other earnings measures. As expected, IBC exhibits the largest cross-sectional standard deviation of the variables, reflecting its inclusion of non-operating items that vary in cross section and add noise to the predictive power of "bottom line" earnings. The table shows that IBC_A, which adjusts for these non-operating components, has an average (median) level of earnings much closer to the average (median) level of operating cash flows.

The last two columns of the table indicate substantial cross-sectional heterogeneity, which is known to cause specification issues in earnings time series (e.g., Ball et al. (2013)). These columns are constructed after removing firm-specific and year-specific fixed effects from each variable and indicate a considerable reduction in the within-firm standard deviations as a result. The reduction for CF is 36% ($1-0.065/0.101$) and for IBC the equivalent reduction after removing firm fixed effects is a similar 35%. For the other earnings variables it is 39-44%, indicating that they are more affected by cross-sectional heterogeneity.

Table 2 Panel B provides a correlation matrix. IBC_A and IBC are almost perfectly positively correlated. OE and OP are highly correlated. A key result is that the correlation between our preferred operating earnings variable OE and operating cash flow CF is 0.753, indicating that they share only a little over one third of the variation in common, and promising a discriminating "horse race" between their predictive abilities.

5. Predictive Ability Results

We begin our analysis by evaluating the predictive ability of operating cash flows vs. earnings based on cross-sectional predictive regressions. Subsequently, we introduce the time series dimension by moving to pooled panel data estimation, which allows us to explore the confounding effect of cross-sectional heterogeneity. We then proceed with industry level and ultimately firm level estimation, to address cross-sectional heterogeneity more fully.

5.1. Cross-sectional estimation

The cross-sectional analysis is performed in-sample and then out-of-sample. Our focus is on two complementary statistics: regression slope coefficients and R^2s. We estimate the single-predictor equation (1) each year and report time-series means and medians of the slope coefficients and R^2s. We use the model based on current period operating cash flow (CF) as a benchmark to evaluate the predictive ability of earnings. For each year, we examine the slope coefficients and R^2s associated with a given predictor (one of: IBC, IBC_A, OP, and OE) and compute differences from the corresponding slope coefficient and R^2 associated with CF. Based on the yearly differences, we evaluate whether the predictive ability of operating cash flows exceeds that of the earnings variable by examining the statistical significance of the mean (median) difference, based on a standard t-test (Wilcoxon rank test).

In regressions of future cash flow from operations on current cash flow from operations, both the slope coefficients and R^2s reflect firm but not accounting characteristics (assuming accountants correctly count and classify operating cash flows). For example, firms whose businesses generate more transitory shocks to operating cash flows will exhibit smaller slope coefficients and smaller R^2s. Similar effects will occur in years that generate more transitory cash flow volatility, such as recessions and crises. However, in regressions of future cash flow from operations on current *earnings*, both statistics will reflect accounting as well as firm characteristics. The distinguishing feature of accruals-based earnings is that it reflects (captures) *expected* operating cash flows associated with the current period operations, i.e., accruals incorporate in the current period those cash flows that the current-year operating transactions are expected to generate in the following year, such as credit sales; it also excludes a portion of current cash flows expected to be earned in the future, such as deferred revenue. It follows that, given the firm characteristics that determine the extent of exogenous cash flow volatility, and hence the operating cash flow regression statistics, accruals-based earnings is expected to exhibit larger slope coefficients and larger R^2s, simply because the predictor contains relatively more information about future operating cash flows.[19] The accuracy with which accrual accounting technology estimates the next-year expected cash flow consequences of current-year operating transactions will affect the magnitude of the increase.[20] For example, less accurate estimates of the proportion of credit sales that turn out to be uncollectible will reduce the additional predictive information in earnings and will attenuate the increase in regression statistics relative to the cash flow predictor. If earnings measure expected operating cash flows without estimation error, the slope coefficient on earnings is expected to be one. This need not be the case for R^2, however. It is thus also logically possible that accruals-based earnings only adds noise to operating cash flows, in which unlikely case the earnings predictor will exhibit smaller regression slopes and R^2s than the cash flows predictor.

The results from the cross-sectional analysis are presented in Table 3, Panel A. The first row of the table reports yearly mean (median) slope coefficients and R^2s associated with cash flows. These results serve as a benchmark: the mean (median) slope coefficient on CF is 0.593 (0.577) and the mean (median) model's R^2 is 0.363 (0.333). The second row in the table shows the performance of "bottom line" earnings, measured by income before extraordinary items (IBC). Relative to the CF benchmark, this row exhibits a considerable decline in both the slope coefficients and R^2s. Specifically, the mean (median) slope coefficient on

[19] The qualification "given firm characteristics" is a major reason for specifying fixed effects in regressions below.

[20] This would seem an important dimension of the hackneyed term "accounting quality."

IBC is 0.430 (0.425) and the mean (median) R^2 is 0.217 (0.230). These results closely line up with the conclusion in Nallareddy et al. (2020) that cash flows have roughly 1.56 times the predictive power of earnings. In Appendix Table 1, we also replicate their results using alternative earnings variables that they considered.

The next three rows of Panel A present results for the three earnings variables that exclude non-operating transactions: IBC_A, OP, and OE. For each of these measures, there is a considerable increase in slope coefficients relative to that of IBC. Indeed, two of the three variables show larger mean (median) slope coefficients than for operating cash flows. While the slopes are not necessarily an indication of predictive ability (we view the slope as a proxy for noise in the measurement error in the proxy for expected operating cash flows) we observe a similar increase in R^2s. The mean R^2s associated with IBC_A, OP, and OE are, respectively, 0.361, 0.353, and 0.378, which are similar to or larger than the 0.363 value for CF. The differences become more obvious with the median R^2s: 0.386 (IBC_A), 0.389 (OP), and 0.406 (OE), versus 0.333 (CF). In the case of OE, the levels of statistical significance are sufficient to reject the null hypothesis that cash flows are superior predictors. Thus, even the cross-sectional tests already show an early indication that conclusions in Nallareddy et al. (2020) of remarkable superiority of operating cash flows over earnings do not hold up to measuring earnings related to operating transactions.

We next analyze incremental informativeness and estimate the two-predictor equation (2) in cross-section: future operating cash flows are regressed on both current period operating cash flows and one earnings variable (a similar approach is used in Ball et al. (2016)). The results are presented in Table 3, Panel B. The results imply that both earnings and cash flows contain useful information in predicting future operating cash flows, irrespective of how earnings are measured. Based on the weights assigned by the regression to each predictor, it appears that cash flows dominate “bottom line” earnings IBC in terms of the information they contribute (coefficients 0.495 vs. 0.163). However, this result changes and even reverses when we switch to earnings variables that exclude non-operating items. Notably, we observe that the average coefficient on OE (0.403) exceeds that on CF (0.310).

While the analysis in Table 3 is performed in-sample, we also conduct an out-of-sample analysis, presented in Table 4. The analysis is based on equation (1) to compare R^2s across measures, and focusing on the out-of-sample R^2. We observe a very similar pattern. Cash flows have a considerably higher predictive ability relative to "bottom line" earnings, however, this disappears and even reverses when we use earnings that are devoid of non-operating components.

Overall, the cross-sectional analyses indicate that, once one chooses accruals-based earnings that are devoid of non-operating components that add noise when predicting operating cash flows, and hence put earnings on the same footing as operating cash flows, prior conclusions that operating cash flows provide a dominant predictor do not hold. In the cross-section, both variables have similar predictive ability, and indeed our preferred operating earnings variable exhibits slight dominance over operating cash flows. It then is logical to explore how cross-sectional heterogeneity affects these conclusions, which is the focus of our subsequent tests.

5.2. Pooled estimation

We now introduce the time dimension by moving from cross-sectional to pooled estimation. Recall that accruals adjust the timing of cash flows (as opposed to cross-sectional heterogeneity) and that Table 2 reveals the presence of significant cross-sectional heterogeneity in both earnings and operating cash flows, which partly confounds interpretation of the predictability of cash flows (e.g., Hsiao (1985)). Controlling for this heterogeneity is important for at least two reasons. First, it is intuitive that some firms have systematically lower cash flows and earnings than others even after scaling these numbers by total assets. These systematic cross-sectional differences can generate a misleading mechanical association between the LHS and RHS variables even if earnings carried no information about future cash flows (e.g., as would be the case if earnings were a firm-level constant plus white noise). Pooled estimation enables us to isolate this heterogeneity. A second and equally important reason for isolating firm-specific heterogeneity is to control for systematic differences across firms in the non-operating components of earnings devoid of any mapping into operating cash flows. For example, some firms may have systematically higher levels of depreciation and/or other non-current accruals than the others. In the cross-sectional regressions, these differences will act as noise in the predictor variable (earnings) and will attenuate its predictive ability.

The results from pooled estimation are presented in Table 5. Panel A shows the slope coefficients and R^2s for equation (1), whereas Panel B corresponds to equation (2). Each panel shows the specifications without firm fixed effects (left side) and with firm fixed effects (right side). Because (total) R^2 does not discriminate between the explanatory power of earnings/cash flows vs. fixed effects, we also present within R-squared, R^2_{within}. This metric captures explanatory power after fixed effects are effectively purged from the data.

Overall, pooled estimation without fixed effects mimics our cross-sectional findings above. In particular, Panel A indicates that operating cash flows outperform "bottom line" earnings in forecasting future operating cash flows. Once we remove non-operating components of earnings, the three measures of earnings perform as well as operating cash flows, with OE being slightly superior to CF.

Panel A also reveals the extent of firm-level heterogeneity in the data and its substantial effect on cross-sectional estimation. For the CF predictor, the R^2 of the model with firm fixed effects is 1.50 times that of the model without fixed effects (0.519/0.347). For the IBC, IBC_A, OP, and OE predictors, the equivalent ratios are even larger: 2.60, 1.57, 1.65 and 1.53, respectively. These large increases in explanatory power are consistent with our expectation that the manner in which the accruals component of earnings converts into future operating cash flows varies substantially across firms (and that accruals are not meant to undo these cross-sectional differences). This has to be addressed in the cross-sectional models.

Unadjusted "bottom line" earnings IBC is the version of earnings that is most enhanced in predictive power by including firm fixed effects into the model. It exhibits a striking recovery in predictive ability. Operating cash flows ($R^2 = 0.519$) no longer outperform it ($R^2 = 0.518$). We also find that the three earnings variables devoid of non-operating components all exhibit noticeably higher R^2s than operating cash flows, ranging between 0.541 and 0.546. What's more, R^2_{within} indicates that after removing cross-sectional variance explained by firm fixed effects (i.e., zooming in on timing of cash flows addressed by the accruals accounting), we observe that accrual-based measures comfortably dominate cash flows. For example, in cases of OP and OE (0.112 and 0.113), the corresponding adjusted R-squared is almost double as compared to CF (0.061).

These results suggest that the seemingly superior predictive ability of operating cash flows relative to some earnings variables is driven by cross-sectional differences and not due to accruals being ineffective at performing their primary function.

Another insight obtained from fixed effects estimation can be seen in Panel B, which reports the model with operating cash flow and an earnings predictor included simultaneously. While the model without fixed effects tracks the cross-sectional results we have seen earlier, we find that controlling for firm heterogeneity considerably dents the predictive power of operating cash flows when the regression is also conditioned on earnings. In particular, the slope coefficient on CF declines from 0.490 to 0.168 in the case of IBC, from 0.342 to 0.084 in the case of IBC_A, from 0.366 to 0.091 in the case of OP, and from 0.319 to 0.064 in the case of OE. In contrast, we observe the coefficients on the earnings variables fall by a considerably lower magnitude when firm fixed effects are specified, and (with the exception of "bottom line" earnings IBC) substantially exceed the coefficient on CF.

We provide an additional reconciliation between our results and those in Nallareddy et al. (2020), presented in Appendix Table 1. Besides income before extraordinary items IBC, they considered pretax income, PI, operating income after depreciation, $OPAD$, and operating income before depreciation, OP. While these variables are dominated in the cross-sectional regression (Panel A, left side), we show that after controlling for firm heterogeneity

(Panel A, right side), the "bottom line" earnings does equally well as compared to CF and three alternative predictors used in their analysis dominate the cash flow variable.[21]

Overall, pooled estimation reveals that once cross-sectional heterogeneity is taken into account by specifying a simple firm fixed effects model, earnings dominate operating cash flows in predictive ability, both in univariate comparisons and, particularly, in terms of incremental information.

5.3. Pooled estimation by industry

Given the importance of firm-level heterogeneity documented above, we further relax the assumption of homogeneity across firms by estimating separately by industry. Estimation on industry-level pooled data with firm fixed effects allows the intercepts to vary at the firm level and also allows the intercepts and slope coefficients to vary at the industry level.

The results of industry-level tests are presented in Table 6. We present the mean slope coefficients and R^2s aggregated across 2-digit SIC industries (at industry level, the medians closely mimic the means and to preserve space we do not tabulate them) . The test statistics are based on variation across industries, which is analogous to the procedure in Fama and MacBeth (1973) and which results in more conservative standard errors as compared to pooled estimation. As previously, we estimate equation (1), based on one predictor at a time, and equation (2), which runs horse races between earnings and operating cash flow predictors.

Panel A indicates that, compared to pooled estimation, allowing for industry-level heterogeneity generates even more aligned mean slope coefficients for CF and unadjusted IBC: 0.240 versus 0.225. Further, the mean R^2s for these two variables are identical: 0.488. This implies that even "bottom line" earnings before extraordinary items do as well as cash flows when allowing for firm- and industry-level differences in business models and accounting methods. Panel A also shows that all earnings variables generate comfortably higher R^2s

[21] Their study does not tabulate alternative predictors but summarises the results verbally in the Appendix.

relative to operating cash flows, and that the levels of statistical significance allow rejecting the null of dominant predictability of CF in favor of the superior predictive ability of accrual-based earnings. For example, the R^2 associated with OE, which is our preferred proxy, is 0.520 which compares favorably with 0.488 in the case of cash flows. Further, as previously, the within firm R-squared indicates that, after removing cross-sectional variance explained by firm fixed effects, the explanatory power added by earnings is almost double: R^2_{within} for OE is 0.137 vs. 0.084 in case of CF

Panel B of Table 6 yields another telling result: the incremental coefficient on CF is even closer to zero and becomes statistically insignificant when it competes with OE, the accruals-based earnings equivalent of operating cash flows. Namely, the coefficient on CF is 0.029 ($t = 1.498$), whereas the coefficient on the earnings variable is 0.370 and is highly significant ($t = 14.6$). Similar patterns are observed for the other earnings variables.

In Appendix Table 3, we perform an alternative estimation where, instead of including firm fixed effects, we use instrumental variable estimation that isolates cross-sectional heterogeneity. Specifically, we use changes from period $t-1$ to t in the right hand side variables to eliminate firm fixed effects and subsequently use these changes as instruments.[22] The conclusions from this analysis are largely the same as from fixed effect models.

The upshot from recognizing industry differences (which analysts using accounting information undoubtedly do) is that not only do earnings dominate operating cash flows in predictive ability, they also effectively subsume the information carried by the cash flow variable. This means that operating cash flows CF can be viewed as a garbled or noisy version of earnings when it comes to predicting future operating cash flows, consistent with Dechow (1994); Dechow and Dichev (2002).

[22] The reason for this estimation is that the presence of unobserved heterogeneity in time series models may not be fully resolved via the inclusion of fixed effects and they are auto-correlated with lagged regressors. This is a known issue and is typically addressed via instruments based on differenced time series (Hsiao, 2014). This is also a reason we next perform the analysis at the firm level, which is not subject to this potential problem.

5.4. Firm-level estimation

Our final set of tests fully eliminate the cross-sectional dimension and perform time series estimation by firm. Doing so imposes additional data requirements, selecting firms with longer lifespans. The resulting sample is likely to be over-represented by more mature companies, where operating cash flows are less likely to be affected by growth. Operating cash flows are then expected to be relatively more informative than in the population; conversely, accruals are expected to play a lesser role in stripping out changes in working capital from earnings. Specifically, we require that a firm has at least 15 time series observations to estimate the models given by equations (1), and (2) separately for each firm.

The results of this analysis are presented in Table 7, Panels A (equation 1) and B (equation 2). Consistent with our findings above that take cross-sectional differences into consideration, Panel A indicates that the predictive abilities of operating cash flows (CF) and "bottom line" earnings (IBC) are similar. Mean (median) slope coefficients on CF and IBC are 0.327 (0.344) and 0.311 (0.297), respectively. Mean (median) R^2s are 0.205 (0.135) and 0.181 (0.117), respectively. We also continue to find both economically and statistically larger slope coefficients and R^2s on the accrual-based earnings measures that exclude non-operating components than those on operating cash flows. In particular, the mean (median) slope coefficients on IBC_A, OP, and OE are 0.419 (0.462), 0.354 (0.386), and 0.428 (0.479). The equivalent mean (median) coefficients on the CF variable are only 0.327 (0.344). Recall that the slope coefficients can be viewed as a measure of noise in a proxy for expected operating cash flows and this analysis indicates that accruals-based measures score better in this respect, up to a factor of approximately 1.4.

Similar results are observed for R^2s. The mean (median) R^2s associated with IBC_A, OP, and OE are 0.232 (0.173), 0.243 (0.186), and 0.246 (0.187), versus 0.205 (0.135) for CF. The differences are significant both economically and statistically, allowing us to firmly reject the null of superior predictability of operating cash flows. Similar to the slope coefficients, the median R^2s associated with these accruals-based earnings measures exceed that of operating cash flows by approximately a factor of 1.4. The corresponding factor associated with the mean R^2s is approximately 1.2, which is still economically large.

Panel B of Table 7 presents firm-level horse races between accruals and cash-based predictors. As previously, we observe that both operating cash flows and "bottom line" earnings are important in forecasting future operating cash flows, and their regression weights are roughly the same: 0.213 (CF) vs. 0.191 (IBC). Here too, after removing non-operating components from earnings, the importance of operating cash flows dissipates, whereas earnings clearly comes out as a dominant variable. In particular, the regression weights given to IBC_A, OP, and OE are 0.330 (t-value, 34.0), 0.290 (t-value, 36.2), and 0.380 (t-value, 36.8), respectively, which are multiples of the corresponding weights on CF: 0.111 (t-value, 14.9), 0.100 (t-value, 14.5), and 0.062 (t-value, 7.8). Interestingly, these values as somewhat larger compared to their equivalents in the industry-level tests, and all are statistically significant. A possible explanation is that the sample is limited to more mature firms, whose operating cash flows are expected to carry more information on average. Nevertheless, the economic magnitudes point to earnings being the dominant predictor, whereas cash flows carry little incremental value.

Firm-level analysis carried out in-sample is subject to over-fitting of the data, given that the number of observations at the firm level is relatively small (as low as 15), and two degrees of freedom are lost. For this reason, it is important to examine whether the results hold up out-of-sample, which is the purpose of the following test.

5.5. Out-of-sample firm-level prediction

To perform these out-of-sample tests, we use a rolling window approach, in which for each firm and each year T we estimate equation (1) based on at least 15 years leading up to and including that year. We then use the estimates to construct forecasts of cash flows in year $T + 1$. This procedure is carried out recursively every year to construct forecasts that do not use any information from the future (unlike the in-sample design). Given the small number of observations, in some instances this procedure generates non-meaningful forecasts resulting in some 'extreme' negative R^2s.[23] Accordingly, we delete 1% of observations at the bottom of the R^2 distribution. We also explore the exclusion of all negative R^2 values, as discussed further below.

The results of this analysis are reported in Table 8. The slope coefficients are presented for descriptive convenience and are comparable to the prior tests, albeit somewhat lower. Not surprisingly, this analysis generates considerably lower R^2s and in fact suggests that CF outperforms "bottom line" earnings IBC both in terms of mean and median values. However, the earnings variables that exclude non-operating items exhibit a different pattern. Based on the median R^2 values, all three measures have considerably larger R^2s relative to operating cash flows. In particular, the median R^2s associated with IBC_A, OP, and OE are 0.146, 0.161, and 0.175, vs. 0.116 for CF. Recall that our preferred measure is OE, which exhibits a median R^2 1.5 times that of CF. For both OP and OE, the levels of statistical significance associated with the median differences in R^2 allow rejecting the null that operating cash flow is a superior predictor. When we examine the mean R^2s the results are not as strong although OE still dominates CF: the mean R^2s associated with IBC_A, OP, and OE are 0.117, 0.147, and 0.157, vs. 0.149 associated with CF. The reason for these somewhat different patterns is that in a number of cases, the out-of-sample R^2s are negative (which is common for firm-level out-of-sample forecasts). This is not very surprising given that few observations are being used in this firm-level estimation and hence influential observations (noise) are present in the data.

One could also reasonably argue in favor of excluding all negative values of R^2 since those point to poor model estimation. To preserve space, we investigate this choice in the Internet Appendix, Table 4 which excludes negative R^2s. As expected, this table exhibits higher slopes and R^2s. It indicates similar performance for both CF and IBC. More importantly, it also shows that all three earnings variables devoid of non-operating components outperform

[23] Out-of-sample R^2 can become negative when forecasts are particularly noisy, i.e., when the unconditional mean is a superior predictor (e.g., Kelly and Pruitt (2013)).

cash flows: IBC_A, OP, and OE, comfortably dominate CF in terms of both means and median R^2 values, as well as stronger statistical significance of the corresponding differences as compared to Table 8.

Overall, the results of firm-level analysis, both in and out of sample, confirm the previous findings that that firm-level heterogeneity is an important confounding factor in the cross-sectional regressions and that accruals-based earnings dominate cash flows in predicting future cash flows.

6. Time variation in predictive ability of earnings vs. operating cash flow

Prior evidence suggests that the predictive abilities of earnings and of operating cash flows have been increasing over time (Kim and Kross, 2005; Nallareddy et al., 2020).[24] These results are based on cross-sectional forecasting without addressing heterogeneity, which we find to be inferior and generally inconclusive. In particular, trends in predictability could be partly explained by changes over time in the relative importance of within-firm versus across-firm differences in earnings and operating cash flows. To investigate this possibility, we follow the approach we adopted earlier to address firm-level heterogeneity before estimating the cross-sectional models. To remove firm-fixed effects, we demean all earnings and operating cash flow variables at the firm level then estimate models (1) and (2) each year. The yearly estimates can be found in Appendix Table 5 and are summarized graphically in Figures 1, 2, and 3.

This evidence is generally consistent with the previously documented upward trend in the predictability of operating cash flows, and also confirms our predictive ability findings in the prior section. As Figure 1 indicates, the slope coefficient is increasing over time for each of the four earnings variables. The source of these trends is unclear, though likely contributors are the secular decline over this period in the relative importance of non-cash

[24] Bushman et al. (2016) report evidence that "one-time" and non-operating items have increased in importance over time. Nallareddy et al. (2020) observe that this could increase noise that reduces the predictive ability of earnings, but do not follow up on the implication that Income before extraordinary items and discontinued operations might not be the appropriate earnings variable for predicting operating cash flows.

working capital and increased importance of expenditures on intangibles (Bushman et al., 2016; Nallareddy et al., 2020).[25] The former is evident in our sample (Appendix Figure 4). The fact that the operating cash flow predictor also performs better over time would seem to rule out changes in GAAP as an explanation for the improvement in earnings-based predictors. Consequently, the explanation most likely lies in some type of secular change or changes in firm characteristics generally, or in the characteristics of firms that choose to be listed.

Whatever the origins of the upward trend in the ability of earnings to predict operating cash flows, it could help explain, at least in part, an upward trend over approximately the same period in the average share price reaction to earnings announcements. Ball and Shivakumar (2008) report "a sharp increase during recent years in the proportion of annual price revision that occurs in the four [quarterly] earnings-event windows ... a similar increase is observed in event-window abnormal volume." Using similar metrics, Beaver et al. (2018) and Beaver et al. (2020) confirm that these results continue in subsequent years. Ball and Shivakumar (2008) canvass several possible explanations for increased event-window reactions, including investors finding earnings to be more informative, which the evidence of increased ability of earnings to predict cash flows would seem to support.

The upward trend in the ability of earnings to predict operating cash flows also implies that the parallel increase in market reaction to earnings announcements is, at least in part, a rational response to a "real" increase in the information that earnings contain about future cash flows, and is not a mere market artifact.

Figure 1 also indicates that, in the case of "bottom line" earnings (solid line), the slope coefficients closely track those for operating cash flows (dotted line), suggesting similar predictive power. However, in the case of earnings devoid of non-operating components, the slope coefficients almost uniformly exceed those on operating cash flows. Figure 2 indicates similar patterns in R^2s, although these exhibit higher variance. We generally observe an

[25] Bates et al. (2009) observe reduced holdings of inventory and receivables over time, presumably due to advances in inventory and receivables management.

upward trend, with operating earnings showing consistently greater predictive ability than operating cash flows.

Finally, Figure 3 shows slope coefficients for the model with two predictors. While the upper left segment indicates roughly equal importance of operating cash flows and "bottom line" earnings, both of which trend upward, the slopes for the three earnings variables that are devoid of non-operating components exhibit much greater magnitudes, which trend upward over time. At the same time, the regression weight placed on operating cash flows fluctuates around zero, with virtually no trend. The results confirm our evidence above that accrual-based measures dominate operating cash flows and effectively subsume the latter's predictive ability. They also show that our conclusions are not specific to any particular time period.

7. Longer-horizon predictive ability

In this section, we extend our prediction horizon from one year to two and three years. The ability of earnings to predict cash flows beyond one year is of interest in its own right. Besides being a robustness test, a two-year horizon helps our understanding of the one-year results. Specifically, the relation between current-year and following-year operating cash flows incorporates two offsetting effects. One effect is that cash flows reflect the firm's underlying economic performance, which introduces a positively auto-correlated component. The other effect is that timing shocks to cash flows reverse (Dechow, 1994; Dechow and Dichev, 2002), which implies a negative auto-correlation component. Horizons longer than one year are not expected to incorporate the reversal effect, because working capital (i.e., operating) accruals generally reverse within one year, by definition.

To preserve space, we focus on R^2s estimated at industry and firm levels. Two-year horizon results are presented in Table 9 (the corresponding one-year results are in Tables 7 and 6). The loss of one observation drops some firms below our minimum time-series length, so the firm-level sample size is almost 10 per cent smaller. Consequently, the results are closely but not precisely comparable with those in Tables 7 and 6, and involve a slightly greater survivorship selection bias.

Panel A presents the results of univariate estimation. The evidence indicates that our findings with respect to one-year ahead horizons extend to two year ahead horizons. Specifically, both at the industry level with firm fixed effects and the firm level, the explanatory powers of CF and bottom line IBC are virtually identical. At the same time, IBC_A, OP, and OE consistently outperform CF, with the differences being particularly noticeable at the firm level.

In the bivariate estimation reported in Panel B, CF continues to predict as well as – or nearly as well as – IBC, but its predictive ability is substantially weakened when conditioning on one of the other three earnings variables IBC_A, OP, and OE, consistent with the one-year results. This effect is particularly noticeable at the industry level where it is important to recognize that we do condition on relatively long firm history.

Three-year horizon results are reported in Appendix Table 2. As expected, explanatory power is a little diminished relative to the two-year horizon results. The R^2s in Panel A univariate regressions are only marginally different than the two-year results in industry estimation, and are diminished by approximately one quarter at the firm level. In the Panel B bivariate horse races, CF is insignificant when controlling for IBC_A, OP, and OE in industry-level estimation, and is outperformed by those earnings variables in firm-level estimation.

8. Operating cash flow before interest and taxes on investment capital gains

It is worth pointing out that the GAAP definition of operating cash flows is not without some controversies. Many investors view interest payments as investing cash flows. Similarly, dividend and interest income, as well as taxes paid in connection to sale of investments, also are viewed as non-operating cash flows. These issues may interfere with some of our predictors. Most notably, operating profit reported on Compustat, OP, is measured before interest revenue or expense and does not include taxes. This introduces an inconsistency between the LHS (CF) and RHS (OP) variables. However OE, which is our preferred measure, does not suffer to such a degree from such inconsistency because it adds working capital accruals to operating cash flows; nevertheless it does reflect interest payments and non-operating taxes.

In this section we adjust the operating cash flow definition to alleviate these issues. Specifically, CF_A excludes net interest paid and also taxes paid in connection to the sale of assets. These adjustments should largely eliminate any first order issues with the controversial GAAP definition of cash flows.

The results are reported in Table 10. Again, we report R^2s for industry and firm level analyses. The results generally are very similar to, and in fact are somewhat more pronounced than, those reported earlier using the GAAP definition of operating cash flows (see Tables 6 and 7). For example, Panel A shows that the mean industry level R^2 associated with CF_A is 0.474 and that of OE is 0.503. The firm-level mean (median) counterparts are 0.202 and 0.242 (0.135 and 0.184), respectively. Panel B indicates that OE continues to subsume (at the industry level) or dramatically reduce the incremental information contained in (adjusted) operating cash flows.

9. Predicting Free Cash Flows

Our final analysis investigates free cash flows ($Free\,CF$) as both the predicted and the predictor variable. Free cash flows are defined as operating cash flows less net investment cash outflows, and are the cash analog of IBC, our only earnings variable that is net of depreciation. Because depreciation is accrued as an average of lagged investment cash outflows over the asset life[26], the relative abilities of $Free\,CF$ and IBC to predict future $Free\,CF$ depends on the time-series behavior of investing cash outflows and on the forecasting horizon.

As we discuss in Section 2, FASB's position is that prediction of future cash flows requires

[26] Depreciation is an unweighted average under the Straight Line method, and a weighted average with exponentially decaying weights under the Accelerated method.

segregating information into reasonably homogeneous groups. Consequently, GAAP requires cash flow reporting to distinguish between operating and investing cash flows. Indeed, the information contained in capital expenditures is of a very different nature than the information in operating cash flows or operating earnings, so we include $Capx$ as a standalone predictor in all regressions.

Results are summarized in Table 11. As previously, we focus on R^2s estimated by industry and by firm. Panel A shows that, when the analysis is carried out at the industry level, all variables exhibit similar explanatory power. Free cash flows on a stand along basis carry approximately as much information as "bottom line" earnings. As is the case with operating cash flows, here IBC_A, OP, and OE perform better than $Free\,CF$ in predicting free cash flows, with the differences being economically more appreciable at the firm level.

The most revealing results are in Panel B, where the regressions include both $Free\,CF$ and an earnings variable. When current capital expenditures are added to the regressions alongside those earnings variables that are not net of depreciation (IBC_A, OP, and OE), current-year free cash flows are either largely or wholly subsumed as a predictor of next-period free cash flows. Overall, the results for free cash flows are broadly consistent with the above findings for operating cash flows. [27]

10. Conclusions

We contribute new evidence to a long-standing debate revolving around the value of accounting earnings for predicting an enterprise's future operating cash flows. Despite the central importance of this issue to practitioners, academics, and standard setters, the debate has not previously been settled. In particular, recent evidence presented in Nallareddy et al. (2020) suggests that operating cash flows consistently dominate earnings in this context.

[27] In all Panel B regressions (i.e., for all earnings variables), the coefficient on $Capx$ is both significantly greater than zero and significantly less than one. The coefficient is smallest in the case of IBC, the only variable that is already net of depreciation. Together, these results imply that capital expenditures have elements of both AR(0) and AR(1) processes, in which both the past average and the most recent observation contain information about the next observation.

Their evidence questions the FASB's famous proposition about the usefulness of earnings in informing investors about future cash flows (FASB (1978)). It also questions the allocation of substantial economic resources to the collection, computation, reporting and auditing of accruals-based earnings, as well as the attention given to earnings by Wall Street and the use of earnings in a variety of contracting contexts. It implies that most of GAAP *reduces* information content. If operating cash flows are superior to earnings in terms of FASB's predictive ability criterion, why incur the costs of collecting, calculating, auditing, reporting and analyzing earnings? Why not simply count and classify operating cash flows?

Our evidence is that *operating* earnings clearly dominate operating cash flows in predicting *operating* cash flows over horizons of one to three years. Testing the essence of the notorious FASB proposition requires putting earnings on an equal footing with operating cash flows. Once non-operating components that are unrelated to operating cash flows are excluded from earnings, leaving working capital accruals to correct the timing limitations of cash flows from operations, earnings dominate operating cash flows in predictive ability. We observe this across multiple forecasting approaches, including cross-sectional, pooled, industry level, and firm level estimation, in-sample as well as out-of-sample, carried out as univariate comparisons as well as in horse races between operating earnings and operating cash flows. Our evidence is consistent with operating cash flows being a garbled or noisy measure of operating earnings (Dechow, 1994), with the noise being reduced by accruals.

Our evidence also demonstrates the importance of addressing firm-level and industry-level heterogeneity in cross-sectional research designs. The mapping from current period operating earnings to future-period operating cash flows varies across firms and industries, which differ in business models and accounting methods. Failure to address this issue in cross-sectional estimation penalizes the estimated predictive ability of operating earnings relative to operating cash flows. We believe this is a common problem in cross-sectional research designs.

References

AICPA. (1973). Objectives of financial statements. New York: American Institute of Certified Public Accountants, Study Group on the Objectives of Financial Statements.

Ball, R. and P. Brown (1968). An empirical evaluation of accounting income numbers. Journal of Accounting Research 6, 159–178.

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

Ball, R., J. Gerakos, J. T. Linnainmaa, and V. Nikolaev (2020). Earnings, retained earnings, and book-to-market in the cross section of expected returns. Journal of Financial Economics 135(1), 231–254.

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

Ball, R., S. Kothari, and V. V. Nikolaev (2013). On estimating conditional conservatism. The Accounting Review 88(3), 755–787.

Ball, R. and L. Shivakumar (2008). How much new information is there in earnings? Journal of Accounting Research 46(5), 975–1016.

Basu, S. (1997). The conservatism principle and the asymmetric timeliness of earnings. Journal of Accounting and Economics 24(1), 3–37.

Bates, T. W., K. M. Kahle, and R. M. Stulz (2009). Why do us firms hold so much more cash than they used to? The journal of finance 64(5), 1985–2021.

Beaver, W. H., M. F. McNichols, and Z. Z. Wang (2018). The information content of earnings announcements: new insights from intertemporal and cross-sectional behavior. Review of Accounting Studies 23(1), 95–135.

Beaver, W. H., M. F. McNichols, and Z. Z. Wang (2020). Increased market response to earnings announcements in the 21st century: An empirical investigation. Journal of Accounting and Economics, 101244.

Bowen, R. M., D. Burgstahler, and L. A. Daley (1986). Evidence on the relationships between earnings and various measures of cash flow. Accounting Review, 713–725.

Bradshaw, M. T. and R. G. Sloan (2002). Gaap versus the street: An empirical assessment of two alternative definitions of earnings. Journal of Accounting Research 40(1), 41–66.

Bushman, R. M., A. Lerman, and X. F. Zhang (2016). The changing landscape of accrual accounting. Journal of Accounting Research 54(1), 41–78.

Campbell, J. Y. and S. B. Thompson (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? The Review of Financial Studies 21(4), 1509–1531.

Collins, D. W., E. L. Maydew, and I. S. Weiss (1997). Changes in the value-relevance of earnings and book values over the past forty years. Journal of Accounting and Economics 24(1), 39–67.

Dechow, P., C. R. Larson, and R. J. Resutek (2020). The effect of accrual heterogeneity on accrual quality inferences. Working paper.

Dechow, P. M. (1994). Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. Journal of accounting and economics 18(1), 3–42.

Dechow, P. M. and I. D. Dichev (2002). The quality of accruals and earnings: The role of accrual estimation errors. The Accounting Review 77(s-1), 35–59.

Dechow, P. M., S. Kothari, and R. L. Watts (1998). The relation between earnings and cash flows. Journal of accounting and Economics 25(2), 133–168.

Dixit, A. K., E. M. M. Milgrom, and P. R. Milgrom (2011). Dynamics of social, political, and economic institutions. Proceedings of the National Academy of Sciences 108(Supplement 4), 21283–21284.

Drtina, R. E. and J. A. Largay III (1985). Pitfalls in calculating cash flow from operations. Accounting Review, 314–326.

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

FASB (1978). Statement of financial accounting concepts no. 1, objectives of financial reporting by business enterprises. Norwalk, CT: Financial Accounting Standards Board.

FASB (1984). Statement of financial accounting concepts no. 5, recognition and measurement in financial statements of business enterprises. Norwalk, CT: Financial Accounting Standards Board.

Finger, C. A. (1994). The ability of earnings to predict future earnings and cash flow. Journal of Accounting Research 32(2), 210–223.

Greenberg, R. R., G. L. Johnson, and K. Ramesh (1986). Earnings versus cash flow as a predictor of future cash flow measures. Journal of Accounting, Auditing & Finance 1(4), 266–277.

Gu, Z. and T. Chen (2004). Analysts' treatment of nonrecurring items in street earnings. Journal of Accounting and Economics 38, 129–170.

Hribar, P. and D. W. Collins (2002). Errors in estimating accruals: Implications for empirical research. Journal of Accounting Research 40(1), 105–134.

Hsiao, C. (1985). Benefits and limitations of panel data. Econometric Reviews 4(1), 121–174.

Hsiao, C. (2014). Analysis of panel data. Number 54. Cambridge university press.

Kelly, B. and S. Pruitt (2013). Market expectations in the cross-section of present values. The Journal of Finance 68(5), 1721–1756.

Kim, M. and W. Kross (2005). The ability of earnings to predict future operating cash flows has been increasing—not decreasing. Journal of Accounting research 43(5), 753–780.

Li, N. (2010). Negotiated measurement rules in debt contracts. Journal of Accounting Research 48(5), 1103–1144.

Lu, Y. and V. V. Nikolaev (2019). Expected loan loss provisioning: An empirical model. Chicago Booth Research Paper (19-11).

Nallareddy, S., M. Sethuraman, and M. Venkatachalam (2020). Changes in accrual properties and operating environment: Implications for cash flow predictability. Journal of Accounting and Economics, 101313.

Nikolaev, V. V. (2018). Identifying accounting quality. Chicago Booth Research Paper (14-28).

Staubus, G. J. (1961). A theory of accounting to investors. University of California Press, Berkeley.

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. The Review of Financial Studies 21(4), 1455–1508.

Zeff, S. A. (2016). The trueblood study group on the objectives of financial statements (1971–73): A historical study. Journal of Accounting and Public policy 35(2), 134–161.

Table 1: Variable definitions

Variable	Definition
CF	Cash flow from operating activities (oancf)
IBC	Income before extraordinary items reported in the cash flow statement (ibc)
IBC_A	Income adjusted for non-operating items, calculated as the sum of income before extraordinary items (ibc); depreciation and amortization (dpc); extraordinary items and discontinued operations (xidoc); sale of property, plant and equipment and investments gain (sppiv); equity in net loss earnings (esubc); and other items involved in the calculation of funds from operations (fopo)
OE	Dechow-Dichev operating earnings, calculated as cash flow (oancf) minus the sum of accounts receivable decrease (recch); inventory decrease (invch); accounts payable and accrued liabilities increase (apalch); income taxes accrued increase (txach); and net change in other assets and liabilities (aoloch)
OP	Operating income before depreciation and amortization (oibdp)
CF_A	Cash flow from operations as reported under GAAP (oancf) plus net interest paid (intpn) plus the difference between gains (losses) on sale of assets before tax (glp) and after tax (gla)
Free CF	Cash flow from operations (oancf) less capital expenditures (capx) plus proceeds from sale of property plant and equipment (sppe)
Capx	Capital expenditures (capx) less proceeds from sale of property plant and equipment (sppe)

All variables are from Compustat's Fundamentals Annual database. All variables are scaled by the average of the beginning and end of the year total assets (at).

Table 2: Summary statistics

Panel A: Summary Statistics

Variable	Min	1st Quartile	Median	3rd Quartile	Max	Mean	SD	Firm SD	Year SD
CF	-0.392	0.031	0.082	0.134	0.382	0.077	0.101	0.065	0.100
IBC	-0.616	-0.006	0.038	0.079	0.312	0.021	0.112	0.073	0.110
IBC_A	-0.379	0.053	0.094	0.145	0.401	0.092	0.095	0.058	0.094
OP	-0.423	0.069	0.121	0.178	0.468	0.116	0.114	0.064	0.111
OE	-0.378	0.052	0.094	0.144	0.390	0.092	0.095	0.057	0.094
CF_A	-0.380	0.049	0.101	0.151	0.398	0.093	0.102	0.066	0.100
Free CF	-0.491	-0.026	0.030	0.080	0.322	0.019	0.104	0.073	0.102
Capx	-0.021	0.019	0.040	0.075	0.373	0.057	0.057	0.030	0.055

Panel B: Correlation matrix

	CF	IBC	IBC_A	OP	OE	CF_A	Free CF	Capx
CF	1.000							
IBC	0.602	1.000						
IBC_A	0.740	0.798	1.000					
OP	0.700	0.808	0.876	1.000				
OE	0.753	0.800	0.978	0.891	1.000			
CF_A	0.980	0.575	0.719	0.711	0.731	1.000		
Free CF	0.847	0.524	0.586	0.551	0.591	0.826	1.000	
Capx	0.230	0.111	0.245	0.238	0.257	0.232	-0.323	1.000

Panel A presents summary statistics for the variables used in the analysis. The statistics are aggregated over all firm years in the sample. SD denotes standard deviation; Firm SD is calculated for each firm, subsequently averaged across all firms; Year SD is calculated for each year, subsequently averaged across all years.
Panel B presents the correlation matrix between the variables.
The sample is over the period 1988-2019. We require total assets and net sales to exceed \$10 million; we also require the closing fiscal year stock price to exceed \$1. We exclude financial firms. To remove outliers and data errors, for each variable, we drop 1% of extreme observations at each tail. See Table 1 for variable definitions.

Table 3: Cross-sectional regressions aggregated across years

Panel A: One predictor

	Coefficient		Δ Coef.		R^2		ΔR^2		
Predictor	Mean	Median	Mean	Median	Mean	Median	Mean	Median	N
CF	0.593	0.577			0.363	0.333			31
IBC	0.430	0.425	-0.163	-0.169	0.217	0.230	-0.146	-0.144	31
			(-13.509)	(-4.840)			(-18.679)	(-4.860)	
IBC_A	0.639	0.646	0.046***	0.041***	0.361	0.386	-0.001	-0.003	31
			(6.429)	(4.448)			(-0.162)	(-0.157)	
OP	0.533	0.528	-0.060	-0.050	0.353	0.389	-0.009	-0.012	31
			(-7.228)	(-4.664)			(-1.251)	(-1.156)	
OE	0.654	0.669	0.061***	0.063***	0.378	0.406	0.015**	0.020**	31
			(10.072)	(4.860)			(2.149)	(2.077)	

Panel B: Two predictors

Predictor	CF Coef.	Predictor Coef.	R^2	N
IBC	0.495***	0.163***	0.383	31
	(36.624)	(20.026)		
IBC_A	0.340***	0.369***	0.416	31
	(24.920)	(32.205)		
OP	0.358***	0.307***	0.422	31
	(26.809)	(31.975)		
OE	0.310***	0.403***	0.422	31
	(24.158)	(28.488)		

This table presents the results of cross-sectional regressions that forecast future operating cash flows in-sample. Panel A is based on yearly cross-sectional OLS regressions of this form:

$$\text{CF}_{t+1} = \alpha + \beta \text{Predictor}_t + \tilde{\epsilon}_{t+1}$$

The columns report, respectively, the means and medians of the following statistics aggregated across years: the regression coefficient, β; the difference between the coefficient on a given predictor and the coefficient on CF, which is shown in the first row (benchmark specification); the coefficient of determination, R^2; the difference between the R^2 for a given predictor and its equivalent based on CF as the predictor (ΔR^2); and the number of yearly observations, N.
The difference between the coefficients or R^2s is tested for (positive) statistical significance based on a one-sided test, given that we are testing to reject the null hypothesis of cash flows' superior predictive power. For means, we report the standard t-test and for medians, we report a Wilcoxon signed-rank test, with respective t- or z-statistics shown in parentheses.
Panel B uses the following cross-sectional OLS regression, estimated by year:

$$\text{CF}_{t+1} = \alpha + \beta_1 \text{CF}_t + \beta_2 \text{Predictor}_t + \tilde{\epsilon}_{t+1}$$

The columns report, respectively, the means of the following statistics aggregated across years: the two regression slope coefficients, β_1 and β_2; the coefficient of determination, R^2; and the number of yearly observations, N.
See Table 1 for variable definitions and Table 2 for sample construction. For both panels, ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Table 4: Out-of-sample cross-sectional forecasts aggregated accross years

	Coefficient		Δ Coef.		R^2		ΔR^2		
Predictor	Mean	Median	Mean	Median	Mean	Median	Mean	Median	N
CF	0.591	0.573			0.359	0.392			29
IBC	0.428	0.421	-0.162	-0.166	0.200	0.222	-0.158	-0.167	29
			(-13.051)	(-4.762)			(-15.689)	(-4.703)	
IBC_{A}	0.636	0.643	0.046***	0.041***	0.355	0.388	-0.004	-0.004	29
			(6.198)	(4.350)			(-0.501)	(-0.465)	
OP	0.532	0.527	-0.059	-0.047	0.346	0.389	-0.012	-0.013	29
			(-6.929)	(-4.576)			(-1.564)	(-1.330)	
OE	0.652	0.662	0.061***	0.064***	0.371	0.419	0.013*	0.019**	29
			(9.747)	(4.782)			(1.608)	(1.676)	

This table presents the results for out-of-sample cross-sectional forecasts carried out by year. Specifically, for a given year T, we estimate the following regression:

$$\text{CF}_T = \alpha_T + \beta_T \text{Predictor}_{T-1} + \tilde{\epsilon}_T$$

We use the coefficients β_T and earnings/cash flow data for year T to forecast CF_{T+1}. We subsequently compute the out-of-sample R_T^2s as discussed in section 4.1.

The columns report, respectively, the means and medians of the following statistics aggregated across years: the regression coefficient, β; the difference between the coefficient on a given predictor and the coefficient on CF, which is shown in the first row (benchmark specification); the coefficient of determination, R^2; the difference between the R^2 for a given predictor and its equivalent based on CF as the predictor (ΔR^2); and the number of observations, N.

The difference between the coefficients or R^2s is tested for (positive) statistical significance based on a one-sided test, given that we are testing to reject the null hypothesis of cash flows' superior predictive power. For means, we report the standard t-test and for medians, we report a Wilcoxon signed-rank test, with respective t- or z-statistics shown in parentheses.

See Table 1 for variable definitions and Table 2 for sample construction. ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

Table 5: Pooled estimation and firm heterogeneity

Panel A: One predictor

	No fixed effects			Firm fixed effects			
Predictor	Coef.	R^2	N	Coef.	R^2	R^2_{within}	N
CF	0.583***	0.347	109041	0.242***	0.519	0.061	109041
	(28.724)			(13.528)			
IBC	0.415***	0.199	109041	0.215***	0.518	0.060	109041
	(16.942)			(21.494)			
IBC_A	0.630***	0.344	109041	0.352***	0.541	0.105	109041
	(28.496)			(24.994)			
OP	0.516***	0.330	109041	0.321***	0.545	0.112	109041
	(27.191)			(22.344)			
OE	0.645***	0.358	109041	0.371***	0.546	0.113	109041
	(30.763)			(24.720)			

Panel B: Two predictors

	No fixed effects				Firm fixed effects				
Predictor	CF	Predictor	R^2	N	CF	Predictor	R^2	R^2_{within}	N
IBC	0.490***	0.157***	0.367	109041	0.168***	0.148***	0.530	0.084	109041
	(85.434)	(34.818)			(26.375)	(30.117)			
IBC_A	0.342***	0.360***	0.400	109041	0.084***	0.297***	0.544	0.109	109041
	(50.517)	(52.069)			(12.619)	(40.780)			
OP	0.366***	0.291***	0.404	109041	0.091***	0.273***	0.548	0.118	109041
	(57.191)	(55.772)			(14.028)	(44.047)			
OE	0.319***	0.387***	0.405	109041	0.064***	0.327***	0.547	0.116	109041
	(45.378)	(53.141)			(9.479)	(42.502)			

The table presents the results for the pooled in-sample forecasts of future operating cash flows. Panel A is based on the following regression estimated over the entire sample:

$$\text{CF}_{i,t+1} = \alpha + \beta \text{Predictor}_{i,t} + \tilde{\epsilon}_{i,t+1}$$

while Panel B is based on the following regression, also estimates over the entire sample:

$$\text{CF}_{i,t+1} = \alpha + \beta_1 \text{CF}_{i,t} + \beta_2 \text{Predictor}_{i,t} + \tilde{\epsilon}_{i,t+1}$$

Each regression format is run twice; with and without firm fixed effects. Panel A's regressions have standard errors clustered by industry. We cluster firms with missing industry data together. Panel B's regressions have standard errors clustered by firm.
The columns in Panel A report, respectively, the coefficients, β; the coefficient of determination, R^2; the within-firm coefficient of determination, R^2_{within}; and the number of firm-year observations used in each regression, N.
Panel B reports both coefficients, β_1 and β_2; the coefficient of determination, R^2; the within-firm coefficient of determination, R^2_{within}; and the number of firm-year observations, N. t-statistics are in parentheses under the coefficient value. The standard errors are clustered by industry.
Each panel shows the specifications estimated without (left side) and with (right side) firm fixed effects. See Table 1 for variable definitions and Table 2 for sample construction. ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

Table 6: Estimation by industry

Panel A: One predictor

Predictor	Coefficient	Δ Coef.	R^2	ΔR^2	R^2_{within}	$\Delta R^2_{\text{within}}$	N
CF	0.240		0.488		0.084		63
IBC	0.225	-0.015	0.488	-0.000	0.078	-0.006	63
		(-0.780)		(-0.060)		(-0.767)	
IBC_A	0.373	0.132***	0.514	0.026***	0.127	0.043***	63
		(6.512)		(4.778)		(5.948)	
OP	0.332	0.092***	0.516	0.028***	0.134	0.050***	63
		(6.354)		(6.485)		(6.480)	
OE	0.387	0.147***	0.520	0.032***	0.137	0.053***	63
		(7.033)		(4.689)		(6.303)	
Firm FEs:	Yes, for all						

Panel B: Two predictors

Predictor	CF Coef.	Predictor Coef.	R^2	R^2_{within}	N
IBC	0.163***	0.157***	0.509	0.118	63
	(8.006)	(10.116)			
IBC_A	0.061***	0.331***	0.523	0.144	63
	(3.306)	(14.549)			
OP	0.057***	0.303***	0.525	0.151	63
	(2.736)	(13.961)			
OE	0.029	0.370***	0.529	0.153	63
	(1.498)	(14.617)			
Firm FEs:	Yes, for all				

This table presents the results for pooled industry-level in-sample forecasts of future operating cash flows. Panel A is based on the following OLS regressions estimated by (two-digit SIC) industry:

$$\text{CF}_{t+1} = \alpha + \beta \text{Predictor}_t + \tilde{\epsilon}_{t+1}$$

The columns report, respectively, the means of the following statistics aggregated across industries: the regression coefficient, β; the industry-level difference between the coefficient on a given predictor and the coefficient on CF, which is shown in the first row (benchmark specification); the coefficient of determination, R^2; the industry-level difference between the R^2 for a given predictor and its equivalent based on CF as the predictor (ΔR^2); the within-firm coefficient of determination R^2_{within}; the industry-level difference between the R^2_{within} for a given predictor and its equivalent based on CF as the predictor ($\Delta R^2_{\text{within}}$); and the number of yearly observations, N.
The difference in coefficients or R^2s is tested for (positive) statistical significance based on a one-sided t-test, given that we are testing to reject the null hypothesis of cash flows' superior predictive power. We report the t-statistics in parentheses.
Panel B is based on the following pooled OLS regression, also estimated by (two-digit SIC) industry:

$$\text{CF}_{t+1} = \alpha + \beta_1 \text{CF}_t + \beta_2 \text{Predictor}_t + \tilde{\epsilon}_t$$

The columns report, respectively, the means of the following statistics aggregated across industries: the two regression slope coefficients, β_1 and β_2; the coefficient of determination, R^2; the R^2_{within}; and the number of yearly observations, N.
See Table 1 for variable definitions and Table 2 for sample construction. For both panels, ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Table 7: Firm-level estimation

Panel A: One predictor

	Coefficient		Δ Coef.		R^2		ΔR^2		
Predictor	Mean	Median	Mean	Median	Mean	Median	Mean	Median	N
CF	0.327	0.344			0.205	0.135			2709
IBC	0.311	0.297	-0.016	-0.034	0.181	0.117	-0.025	-0.006	2709
			(-2.056)	(-4.223)			(-7.495)	(-5.572)	
IBC_A	0.419	0.462	0.093***	0.081***	0.232	0.173	0.027***	0.020***	2709
			(13.934)	(16.485)			(9.810)	(10.771)	
OP	0.354	0.386	0.028***	0.018***	0.243	0.186	0.038***	0.029***	2709
			(4.414)	(4.607)			(12.892)	(13.630)	
OE	0.428	0.479	0.101***	0.090***	0.246	0.187	0.041***	0.031***	2709
			(15.602)	(19.377)			(15.326)	(16.134)	

Panel B: Two predictors

Predictor	CF Coef.	Predictor Coef.	R^2	N
IBC	0.213***	0.191***	0.284	2709
	(31.830)	(22.735)		
IBC_A	0.111***	0.330***	0.300	2709
	(14.912)	(34.041)		
OP	0.100***	0.290***	0.311	2709
	(14.536)	(36.195)		
OE	0.062***	0.380***	0.308	2709
	(7.756)	(36.796)		

This table presents the results of firm-level forecasting of future cash flows, carride out in-sample. Panel A is based on the following time-series regression, estimated separately for each firm:

$$\text{CF}_{i,t+1} = \alpha + \beta \text{Predictor}_{i,t} + \tilde{\epsilon}_{i,t+1}$$

The columns report, respectively, the means and medians of the following statistics aggregated across firms: the regression coefficient, β; the firm-level difference between the coefficient on a given predictor and the coefficient on CF, which is shown in the first row (benchmark specification); the coefficient of determination, R^2; the firm-level difference between the R^2 for a given predictor and its equivalent based on CF as the predictor (ΔR^2); and the number of yearly observations, N.
The difference in coefficients or R^2s is tested for (positive) statistical significance based on a one-sided test, given that we are testing to reject the null hypothesis of cash flows' superior predictive power. For means, we report the standard t-test and for medians, we report a Wilcoxon signed-rank test, with respective t- or z-statistics shown in parentheses.
Panel B is based on the following time-series regression, also estimated by firm:

$$\text{CF}_{i,t+1} = \alpha + \beta_1 \text{CF}_{i,t} + \beta_2 \text{Predictor}_{i,t} + \tilde{\epsilon}_{i,t+1}$$

The columns report, respectively, the means of the following statistics aggregated across firms: the two regression slope coefficients, β_1 and β_2; the coefficient of determination, R^2; and the number of yearly observations, N.
See Table 1 for variable definitions and Table 2 for sample construction. For both panels, ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Table 8: Out-of-sample firm-level forecasts

	Coefficient		R^2		ΔR^2		
Predictor	Mean	Median	Mean	Median	Mean	Median	N
CF	0.290	0.293	0.149	0.116			2453
IBC	0.287	0.276	0.022	0.044	-0.130	-0.056	2441
					(-10.833)	(-10.772)	
IBC_{A}	0.388	0.420	0.117	0.146	-0.035	0.009	2448
					(-3.367)	(0.440)	
OP	0.322	0.352	0.147	0.161	-0.007	0.030***	2443
					(-0.692)	(4.364)	
OE	0.387	0.434	0.157	0.175	0.003	0.034***	2452
					(0.333)	(6.421)	

This table presents results of firm-level out-of-sample forecasts performed based on rolling window time-series regressions of this form, estimated separately by firm:

$$\text{CF}_{i,t+1} = \alpha_T + \beta_T \text{Predictor}_{i,t} + \tilde{\epsilon}_{i,t+1}, \forall t \in \{1, ..., T\}$$

Each rolling window starts in the first year each firm is present and ends in year T. We use the estimated parameters and the information at time T to forecast cash flow for the period $T + 1$.
After obtaining out-of-sample forecasts based on above regression, we calculate the out-of-sample R^2 for each firm as described in section 4.1. We drop firm-years in the lowest percentile of the R^2 distribution to mitigate the effect of not economically meaningful observations. Regression statistics, β and R^2, are aggregated across all firms.
The columns report, respectively, the means and medians of the following statistics aggregated across firms: the regression coefficient, β; the coefficient of determination, R^2; the difference in R^2 between each regression and its equivalent using CF as the predictor (ΔR^2); and the number of observations, N.
The difference between coefficients or R^2s is tested for (positive) statistical significance based on a one-sided test, given that we are testing to reject the null hypothesis of cash flows' superior predictive power. For means, we report the standard t-test and for medians, we report a Wilcoxon signed-rank test, with respective t- or z-statistics shown in parentheses.
See Table 1 for variable definitions and Table 2 for sample construction. ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

Table 9: Two-year-ahead operating cash flows

Panel A: One predictor

Predictor	Industry-level estimation R^2 Mean	R^2 Median	ΔR^2 Mean	ΔR^2 Median	N	Firm-level estimation R^2 Mean	R^2 Median	ΔR^2 Mean	ΔR^2 Median	N
CF	0.455	0.461			63	0.128	0.067			2457
IBC	0.454	0.462	-0.002	0.001	63	0.124	0.065	-0.005	-0.000	2457
			(-0.967)	(-0.363)				(-1.622)	(-0.866)	
IBC_A	0.462	0.467	0.007***	0.006***	63	0.150	0.081	0.021***	0.009***	2457
			(3.713)	(4.238)				(8.925)	(8.753)	
OP	0.468	0.470	0.012***	0.013***	63	0.161	0.089	0.033***	0.016***	2457
			(4.960)	(5.128)				(12.520)	(12.467)	
OE	0.463	0.467	0.008***	0.007***	63	0.152	0.084	0.023***	0.010***	2457
			(4.241)	(4.833)				(10.215)	(10.207)	
Firm FEs:			Yes					No		

Panel B: Two predictors

Predictor	Industry-level estimation CF Coef.	Predictor Coef.	N	Firm-level estimation CF Coef.	Predictor Coef.	N
IBC	0.089***	0.082***	63	0.132***	0.131***	2457
	(6.405)	(6.645)		(18.966)	(14.088)	
IBC_A	0.039**	0.164***	63	0.066***	0.211***	2457
	(2.593)	(9.622)		(8.725)	(20.840)	
OP	0.004	0.201***	63	0.043***	0.203***	2457
	(0.180)	(8.422)		(5.901)	(24.289)	
OE	0.029*	0.175***	63	0.047***	0.230***	2457
	(1.837)	(10.021)		(5.598)	(21.400)	
Firm FEs:		Yes			No	

This table presents the results for industry-level pooled (left side) and firm-level time-series (right side) in-sample forecasts of future operating cash flows. The industry is defined based on the two-digit SIC code. Panel A is based on panel or time-series OLS regressions of this form:

$$\mathrm{CF}_{t+2} = \alpha + \beta \mathrm{Predictor}_t + \tilde{\epsilon}_{t+1}$$

The columns report, respectively, the means and medians of the following statistics aggregated across industries: the coefficient of determination, R^2, and the corresponding industry-level difference between the R^2 for a given predictor and its equivalent based on CF used as the predictor (ΔR^2); the number of yearly observations, N; and the equivalent statistics for the firm-level regressions.
The difference in R^2s is tested for (positive) statistical significance based on a one-sided test, given that we are testing to reject the null hypothesis of cash flows' superior predictive power. For means, we report the standard t-test and for medians, we report a Wilcoxon signed-rank test, with respective t- or z-statistics shown in parentheses.
Panel B uses the following OLS regression, estimated by two-digits SIC industry (left side) or by firm (right side):

$$\mathrm{CF}_{t+2} = \alpha + \beta_1 \mathrm{CF}_t + \beta_2 \mathrm{Predictor}_t + \tilde{\epsilon}_t$$

The columns report, respectively, the means of the following statistics aggregated across industries or firms: the two regression slope coefficients, β_1 and β_2; and the number of observations, N.
See Table 1 for variable definitions and Table 2 for sample construction. For both panels, ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Table 10: Operating cash flow excluding interest and taxes paid on capital gains

Panel A: One predictor

	Industry-level estimation					Firm-level estimation				
Predictor	R^2		ΔR^2		N	R^2		ΔR^2		N
	Mean	Median	Mean	Median		Mean	Median	Mean	Median	
$\mathrm{CF_A}$	0.474	0.478			63	0.202	0.135			2709
IBC	0.466	0.463	-0.008	-0.001	63	0.167	0.098	-0.036	-0.010	2709
			(-1.579)	(-1.602)				(-10.258)	(-7.810)	
$\mathrm{IBC_A}$	0.490	0.492	0.015***	0.015***	63	0.214	0.153	0.011***	0.008***	2709
			(2.854)	(4.005)				(3.917)	(5.550)	
OP	0.502	0.499	0.028***	0.027***	63	0.243	0.186	0.040***	0.030***	2709
			(6.050)	(5.504)				(13.903)	(14.554)	
OE	0.503	0.502	0.029***	0.026***	63	0.242	0.184	0.040***	0.028***	2709
			(4.200)	(5.627)				(15.321)	(15.870)	
Firm FEs:			Yes					No		

Panel B: Two predictors

	Industry-level estimation			Firm-level estimation		
Predictor	CF Coef.	Predictor Coef.	N	CF Coef.	Predictor Coef.	N
IBC	0.175***	0.124***	63	0.225***	0.157***	2709
	(8.694)	(8.129)		(33.891)	(18.218)	
$\mathrm{IBC_A}$	0.095***	0.268***	63	0.143***	0.275***	2709
	(4.968)	(12.692)		(19.359)	(28.654)	
OP	0.053***	0.298***	63	0.095***	0.292***	2709
	(2.691)	(13.067)		(13.493)	(35.377)	
OE	0.030	0.359***	63	0.062***	0.380***	2709
	(1.485)	(13.965)		(7.607)	(37.424)	
Firm FEs:		Yes			No	

This table presents the results for industry-level pooled (left side) and firm-level time-series (right side) in-sample forecasts of future operating cash flows adjusted to exclude interest paid and the portion of income taxes related to realization of gains (losses) on sale of assets. The industry is defined based on the two-digit SIC code. Panel A is based on panel or time-series OLS regressions of this form:

$$\mathrm{CF}_{\mathrm{A}t+1} = \alpha + \beta \mathrm{Predictor}_t + \tilde{\epsilon}_{t+1}$$

The columns report, respectively, the means and medians of the following statistics aggregated across industries: the coefficient of determination, R^2, and the corresponding industry-level difference between the R^2 for a given predictor and its equivalent based on $\mathrm{CF_A}$ used as the predictor (ΔR^2); the number of yearly observations, N; and the equivalent statistics for the firm-level regressions.
The difference in R^2s is tested for (positive) statistical significance based on a one-sided test, given that we are testing to reject the null hypothesis of cash flows' superior predictive power. For means, we report the standard t-test and for medians, we report a Wilcoxon signed-rank test, with respective t- or z-statistics shown in parentheses.
Panel B uses the following OLS regression, estimated by two-digits SIC industry (left side) or by firm (right side):

$$\mathrm{CF}_{\mathrm{A}t+1} = \alpha + \beta_1 \mathrm{CF}_{\mathrm{A}t} + \beta_2 \mathrm{Predictor}_t + \tilde{\epsilon}_t$$

The columns report, respectively, the means of the following statistics aggregated across industries or firms: the two regression slope coefficients, β_1 and β_2; and the number of observations, N.
See Table 1 for variable definitions and Table 2 for sample construction. For both panels, ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Table 11: Free cash flow predicted by industry and by firm

Panel A: Using free cash flows or earnings as predictors

	Industry-level estimation					Firm-level estimation				
Predictor	R^2		ΔR^2		N	R^2		ΔR^2		N
	Mean	Median	Mean	Median		Mean	Median	Mean	Median	
Free CF	0.417	0.404			63	0.254	0.203			2649
IBC	0.416	0.405	-0.001	-0.000	63	0.243	0.201	-0.011	-0.004	2649
			(-0.382)	(-0.705)				(-4.530)	(-4.402)	
IBC_A	0.426	0.421	0.010	0.006	63	0.265	0.218	0.011***	0.004***	2649
			(4.235)	(4.655)				(5.586)	(5.705)	
OP	0.427	0.414	0.010	0.008	63	0.273	0.232	0.018***	0.009***	2649
			(4.297)	(4.573)				(8.418)	(8.789)	
OE	0.427	0.420	0.010	0.007	63	0.273	0.232	0.019***	0.009***	2649
			(4.474)	(4.738)				(9.559)	(9.987)	
Firm FEs:			Yes					No		

Panel B: Using both free cash flows and earnings as predictors

	Industry-level estimation				Firm-level estimation			
Predictor	Free CF Coef.	Pred. Coef.	Capx Coef.	N	Free CF Coef.	Pred. Coef.	Capx Coef.	N
IBC	0.106***	0.068***	-0.232***	63	0.152***	0.065***	-0.261***	2649
	(5.173)	(3.589)	(-5.392)		(18.366)	(5.482)	(-10.684)	
IBC_A	0.016	0.217***	-0.338***	63	0.065***	0.199***	-0.406***	2649
	(0.745)	(7.052)	(-7.583)		(6.974)	(16.327)	(-16.269)	
OP	0.016	0.193***	-0.352***	63	0.053***	0.177***	-0.406***	2649
	(0.642)	(6.366)	(-8.549)		(6.212)	(16.717)	(-14.977)	
OE	-0.003	0.244***	-0.370***	63	0.026**	0.248***	-0.450***	2649
	(-0.136)	(7.565)	(-8.508)		(2.462)	(18.507)	(-17.382)	
Firm FEs:		Yes			No			

This table presents the results for future free cash flow forecasts by two-digit SIC industry (left side) or by firm (right side), respectively. Panel A is based on OLS regressions of this form:

$$\text{Free CF}_{t+1} = \alpha + \beta_1 \text{Predictor}_t + \beta_2 \text{Capx}_t + \tilde{\epsilon}_{t+1}$$

The columns report, respectively, the means and medians of the following statistics aggregated across industries: the coefficient of determination, R^2; the industry-level difference between the R^2 for a given predictor and its equivalent based on Free CF as the predictor (ΔR^2); the number of observations per industry, N; and the equivalent statistics for the firm-level estimation. The difference in R^2s is tested for (positive) statistical significance based on a one-sided test, given that we are testing to reject the null hypothesis of cash flows' superior predictive power. For means, we report the standard t-test and for medians, we report a Wilcoxon signed-rank test, with respective t- or z-statistics shown in parentheses.
Panel B uses the following OLS regression:

$$\text{Free CF}_{t+1} = \alpha + \beta_1 \text{Free CF}_t + \beta_2 \text{Predictor}_t + \beta_3 \text{Capx}_t + \tilde{\epsilon}_t$$

The columns report, respectively, the means of the three regression slope coefficients, β_1, β_2 and β_3, aggregated across industries; the number of observations per industry, N; and the equivalent statistics for the firm-level estimation.
See Table 1 for variable definitions and Table 2 for sample construction. For both panels, ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

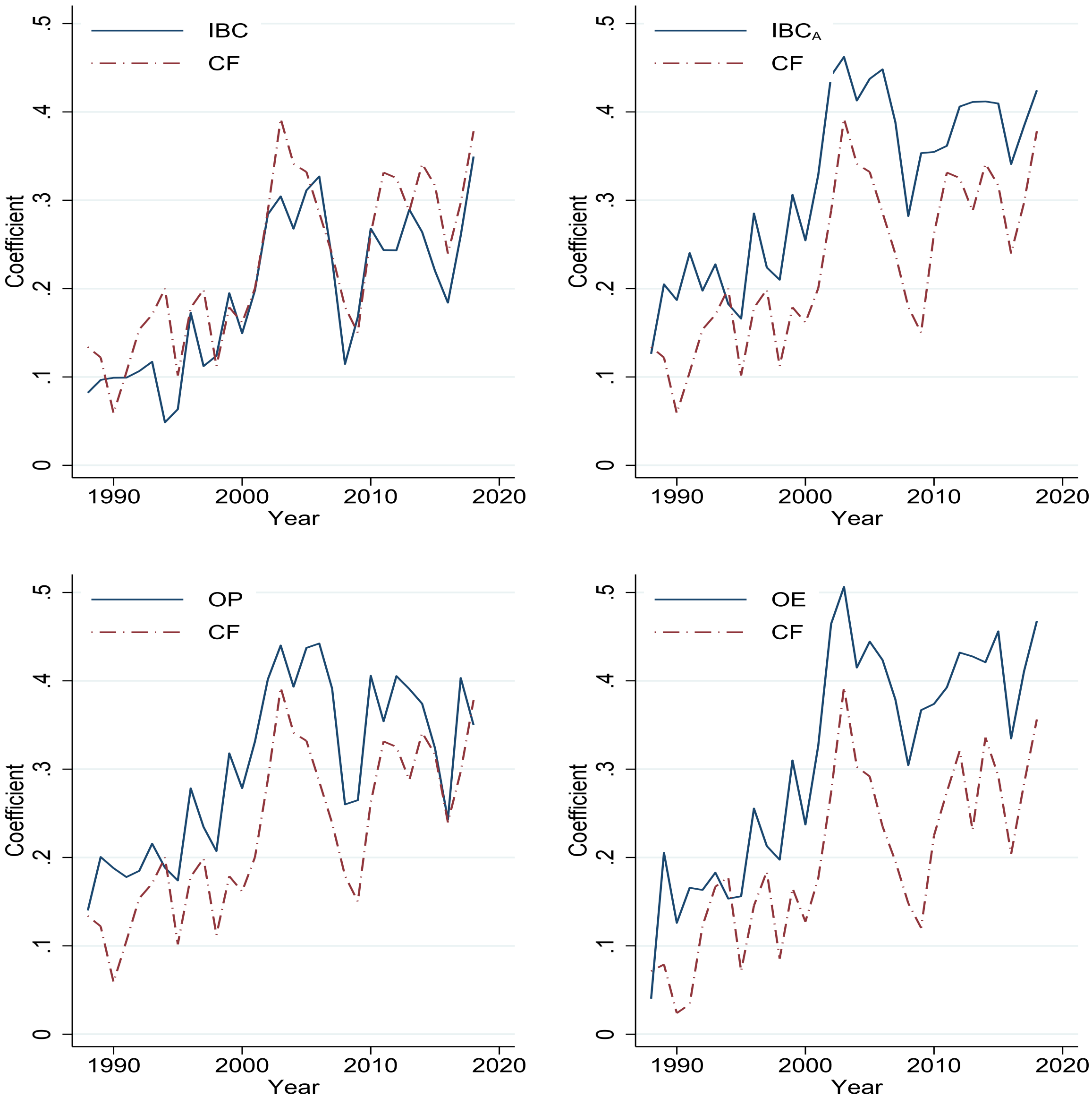


Figure 1: The figure displays the coefficients from the cross-sectional regressions of future cash flows on a single predictor. Variables in each regressions are de-meaned by firm to remove cross-sectional heterogeneity (firm fixed effects). Each plot contrasts the coefficients obtained using a non-CF predictor with that obtained using CF. Clockwise, starting at the top-left, the predictors are IBC, IBC$_A$, OE and OP. In all plots, solid lines denote the coefficients for the non-CF predictor and dashed lines denote that for CF. See Table 1 for variable definitions.

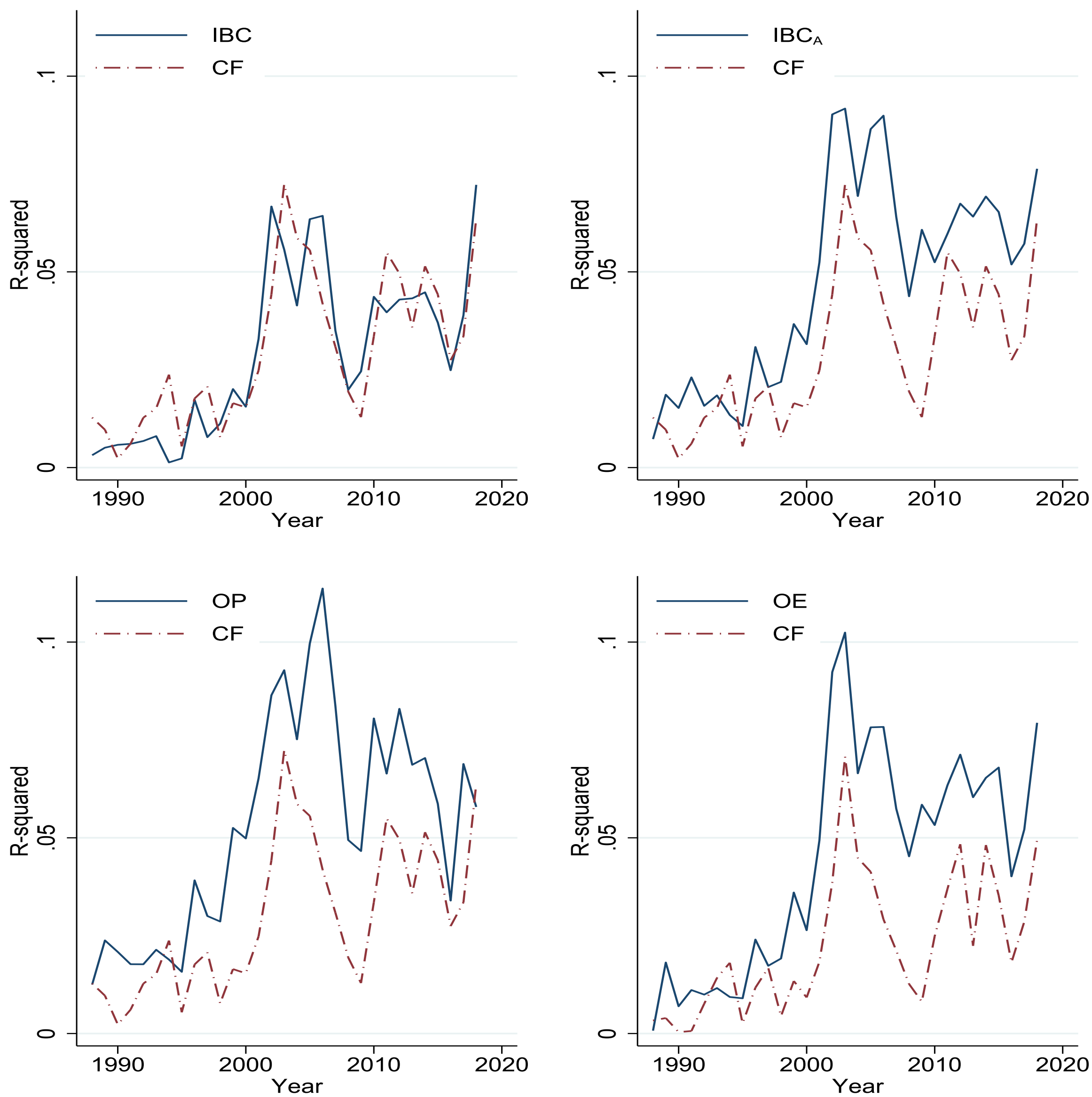


Figure 2: The figure displays the R-squareds based on the cross-sectional regressions of future cash flows on a single predictor. Variables in each regressions are de-meaned by firm to remove cross-sectional heterogeneity (firm fixed effects). Each plot contrasts the R^2 obtained using a non-CF predictor with that obtained using CF. Clockwise, starting at the top-left, the predictors are IBC, IBC_A, OE and OP. In all plots, solid lines denote the R^2 for the non-CF predictor and dashed lines denote that for CF. See Table 1 for variable definitions.

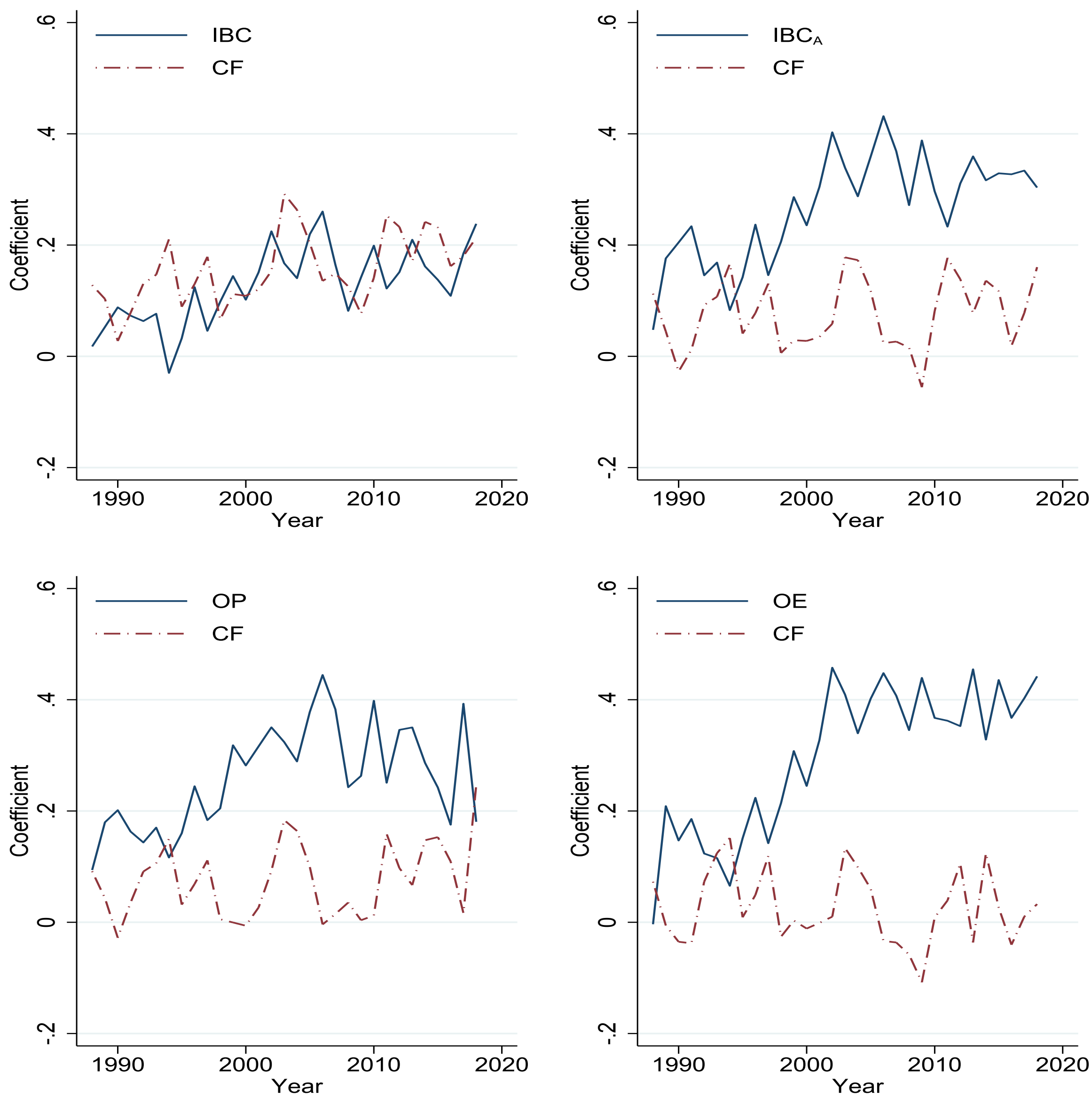


Figure 3: The figure displays the coefficients from the cross-sectional regression of future cash flows on two-predictors. Variables in each regressions are de-meaned by firm to remove cross-sectional heterogeneity (firm fixed effects). Each plot contrasts the coefficients to CF and the other predictor simultaneously present as an independent variable in each regression. Clockwise, starting at the top-left, the (non-CF) predictors are IBC, IBC_A, OE and OP. In all plots, solid lines denote the coefficient for the non-CF predictor and dashed lines denote that for CF. See Table 1 for variable definitions.

Appendix

to

"FASB was right: Earnings beat cash flows when predicting future cash flows"

by

Ray Ball and Valeri Nikolaev

Appendix Table 1: Reconciliation with Nallareddy et al. (2020) baseline results

Panel A: One predictor

	No fixed effects			Firm fixed effects			
Predictor	Coef.	R^2	N	Coef.	R^2	R^2_{within}	N
CF	0.583***	0.346	108846	0.242***	0.518	0.060	108846
	(28.685)			(13.504)			
IBC	0.416***	0.197	108846	0.214***	0.517	0.060	108846
	(17.083)			(21.230)			
PI	0.442***	0.250	108846	0.270***	0.532	0.089	108846
	(19.688)			(21.755)			
OPAD	0.468***	0.258	108846	0.301***	0.538	0.099	108846
	(19.594)			(21.746)			
OP	0.519***	0.329	108846	0.323***	0.544	0.112	108846
	(27.364)			(22.528)			
OE	0.645***	0.357	108846	0.370***	0.544	0.112	108846
	(30.883)			(25.065)			

Panel B: Two predictors

	No fixed effects				Firm fixed effects				
Predictor	CF	Predictor	R^2	N	CF	Predictor	R^2	R^2_{within}	N
IBC	0.490***	0.157***	0.365	108846	0.168***	0.147***	0.529	0.083	108846
	(85.708)	(34.725)			(26.456)	(29.935)			
PI	0.447***	0.197***	0.377	108846	0.124***	0.212***	0.538	0.101	108846
	(72.781)	(42.344)			(19.128)	(38.216)			
OPAD	0.440***	0.217***	0.381	108846	0.110***	0.245***	0.543	0.109	108846
	(70.584)	(43.633)			(17.121)	(40.981)			
OP	0.366***	0.293***	0.403	108846	0.091***	0.275***	0.547	0.118	108846
	(57.081)	(55.958)			(14.033)	(44.135)			
OE	0.320***	0.387***	0.404	108846	0.065***	0.326***	0.546	0.115	108846
	(45.387)	(52.864)			(9.557)	(42.060)			

The table presents the results for pooled in-sample forecasts of future operating cash flows, carried out with (right side) and without (left side) firm fixed effects. Panel A is based on the following regression estimated over the entire sample:

$$\text{CF}_{t+1} = \alpha + \beta \text{Predictor}_{i,t} + \tilde{\epsilon}_{i,t+1}$$

while Panel B uses regressions of the following form:

$$\text{CF}_{t+1} = \alpha + \beta_1 \text{CF}_{i,t} + \beta_2 \text{Predictor}_{i,t} + \tilde{\epsilon}_{i,t+1}$$

The table compares performance across four baseline profitability measures: IBC is income before extraordinary items as reported in the cash flow statement, PI is pretax income (defined as pi - spi), OPAD is operating profit after depreciation (defined as oiadp) and OP is operating profit. The definitions of these predictors are the same is in Nallareddy et al. (2020).
Panel A's regressions have standard errors clustered by industry; panel B's regressions have standard errors clustered by firm. The columns in Panel A report, respectively, the coefficients, β; the coefficient of determination, R^2; the within-firm coefficient of determination, R^2_{within}; and the number of firm-year observations used in each regression, N.
Panel B reports both coefficients, β_1 and β_2, as well as the R^2 and the number of firm-year observations. t-statistics are in parentheses under the coefficient value. The standard errors are clustered by industry.
Each panel shows the specifications estimated without (left side) and with (right side) firm fixed effects. See Table 1 for variable definitions and Table 2 for sample construction. ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

Appendix Table 2: Three-year-ahead operating cash flows

Panel A: One predictor

	Industry-level estimation					Firm-level estimation				
Predictor	R^2		ΔR^2		N	R^2		ΔR^2		N
	Mean	Median	Mean	Median		Mean	Median	Mean	Median	
CF	0.445	0.457			63	0.096	0.043			2283
IBC	0.444	0.470	-0.001	-0.000	63	0.099	0.046	0.003	0.001	2283
			(-0.557)	(-0.815)				(1.090)	(1.266)	
IBC_A	0.449	0.478	0.005**	0.001***	63	0.113	0.053	0.016***	0.004***	2283
			(2.348)	(3.170)				(7.615)	(6.904)	
OP	0.451	0.479	0.006***	0.003***	63	0.124	0.061	0.027***	0.010***	2283
			(3.239)	(4.532)				(11.284)	(11.226)	
OE	0.449	0.476	0.004**	0.001***	63	0.112	0.053	0.016***	0.005***	2283
			(2.339)	(2.807)				(7.750)	(7.175)	
Firm FEs:			Yes					No		

Panel B: Two predictors

	Industry-level estimation			Firm-level estimation		
Predictor	CF Coef.	Predictor Coef.	N	CF Coef.	Predictor Coef.	N
IBC	0.044***	0.059***	63	0.099***	0.079***	2283
	(2.952)	(4.830)		(13.730)	(7.920)	
IBC_A	0.019	0.104***	63	0.045***	0.144***	2283
	(1.357)	(6.371)		(5.544)	(13.330)	
OP	-0.011	0.129***	63	0.031***	0.136***	2283
	(-0.538)	(5.963)		(4.109)	(15.475)	
OE	0.018	0.103***	63	0.052***	0.130***	2283
	(1.193)	(6.041)		(5.859)	(11.729)	
Firm FEs:		Yes			No	

This table presents the results for industry-level pooled (left side) and firm-level time-series (right side) in-sample forecasts of future operating cash flows. The industry is defined based on the two-digit SIC code. Panel A is based on panel or time-series OLS regressions of this form:

$$\mathrm{CF}_{t+3} = \alpha + \beta \mathrm{Predictor}_t + \tilde{\epsilon}_{t+1}$$

The columns report, respectively, the means and medians of the following statistics aggregated across industries: the coefficient of determination, R^2, and the corresponding industry-level difference between the R^2 for a given predictor and its equivalent based on CF used as the predictor (ΔR^2); the number of yearly observations, N; and the equivalent statistics for the firm-level regressions.
The difference in R^2s is tested for (positive) statistical significance based on a one-sided test, given that we are testing to reject the null hypothesis of cash flows' superior predictive power. For means, we report the standard t-test and for medians, we report a Wilcoxon signed-rank test, with respective t- or z-statistics shown in parentheses.
Panel B uses the following OLS regression, estimated by two-digits SIC industry (left side) or by firm (right side):

$$\mathrm{CF}_{t+3} = \alpha + \beta_1 \mathrm{CF}_t + \beta_2 \mathrm{Predictor}_t + \tilde{\epsilon}_t$$

The columns report, respectively, the means of the following statistics aggregated across industries or firms: the two regression slope coefficients, β_1 and β_2; and the number of observations, N.
See Table 1 for variable definitions and Table 2 for sample construction. For both panels, ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Appendix Table 3: Instrumental variable estimation by industry

Panel A: One predictor

Predictor	Coefficient		Δ Coef.		N
	Mean	Median	Mean	Median	
CF	0.198	0.168			63
IBC	0.155	0.157	-0.042	-0.036	63
			(-1.703)	(-1.376)	
IBC_A	0.339	0.304	0.141***	0.116***	63
			(4.817)	(5.073)	
OP	0.338	0.341	0.141***	0.148***	63
			(5.675)	(5.333)	
OE	0.357	0.335	0.159***	0.172***	63
			(4.171)	(5.326)	
Firm FEs:	Yes, for all				

Panel B: Two predictors

Predictor	CF Coef.	Predictor Coef.	N
IBC	0.158***	0.117***	63
	(6.908)	(6.379)	
IBC_A	0.081***	0.285***	63
	(4.279)	(8.642)	
OP	0.074***	0.310***	63
	(4.752)	(11.021)	
OE	0.048**	0.318***	63
	(2.540)	(6.971)	
Firm FEs:	Yes, for all		

As an alternative to fixed effects regressions, this table presents the results from instrumental variable estimation, carried out by two-digit SIC industry. Panel A is based on the following instrumental variable regressions:

$$\text{CF}_{t+1} = \alpha + \beta\text{Predictor}_t + \tilde{\epsilon}_{t+1}$$

where the instrument is constructed to eliminate firm fixed effects: specifically, we use the difference in the RHS variable as the instrument for its level. The columns report, respectively, the means and medians of the following statistics aggregated across industries: the regression coefficient, β; the industry-level difference between the coefficient andthe coefficient on CF, which is shown in the first row (benchmark specification); and the number of yearly observations, N. We do not report R^2 because it is not meaningful for IV estimation.
The difference in coefficients is tested for (positive) statistical significance based on a one sided-test, given that we are testing to reject the null hypothesis of cash flows' superior predictive power. For means, we report the standard t-test and for medians, we report a Wilcoxon signed-rank test, with respective t- or z-statistics shown in parentheses.
Panel B uses the following instrumental variable regression, estimated by (two-digit SIC) industry:

$$\text{CF}_{t+1} = \alpha + \beta_1\text{CF}_t + \beta_2\text{Predictor}_t + \tilde{\epsilon}_t$$

where the instruments are constructed to eliminate firm fixed effects: specifically, we use the differences in the RHS variables as instruments. The columns report, respectively, the means of the following statistics aggregated across industries: the two regression slope coefficients, β_1 and β_2; and the number of yearly observations, N.
See Table 1 for variable definitions and Table 2 for sample construction. For both panels, ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.

Appendix Table 4: Out-of-sample firm-level forecasts: excluding negative R^2s

	Coefficient		R^2		ΔR^2		
Predictor	Mean	Median	Mean	Median	Mean	Median	N
CF	0.383	0.392	0.357	0.301			1611
IBC	0.411	0.389	0.335	0.273	-0.033	-0.010	1378
					(-3.741)	(-3.203)	
IBC_A	0.490	0.517	0.390	0.350	0.045***	0.036***	1586
					(6.746)	(7.025)	
OP	0.411	0.435	0.407	0.366	0.064***	0.063***	1616
					(9.746)	(10.298)	
OE	0.492	0.514	0.404	0.378	0.064***	0.056***	1677
					(10.472)	(11.156)	

This table presents results of firm-level out-of-sample forecasts performed based on rolling window time-series regressions of this form, estimated separately by firm:

$$\text{CF}_{i,t+1} = \alpha_T + \beta_T \text{Predictor}_{i,t} + \tilde{\epsilon}_{i,t+1}, \forall t \in \{1, ..., T\}$$

Each rolling window starts in the first year each firm is present and ends in year T. We use the estimated parameters and the information at time T to forecast cash flow for the period $T+1$.
After obtaining out-of-sample forecasts based on above regression, we calculate the out-of-sample R^2 for each firm as described in section 4.1. In this table, we drop all firm-level observations with negative out-of-sample R^2s to mitigate the effect of not economically meaningful observations. Regression statistics, β and R^2, are aggregated across all firms.
The columns report, respectively, the means and medians of the following statistics aggregated across firms: the regression coefficient, β; the coefficient of determination, R^2; the difference in R^2 between each regression and its equivalent using CF as the predictor (ΔR^2); and the number of observations, N.
The difference between coefficients or R^2s is tested for (positive) statistical significance based on a one-sided test, given that we are testing to reject the null hypothesis of cash flows' superior predictive power. For means, we report the standard t-test and for medians, we report a Wilcoxon signed-rank test, with respective t- or z-statistics shown in parentheses.
See Table 1 for variable definitions and Table 2 for sample construction. ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

Appendix Table 5: Cross-sectional forecasts by year: mean-adjusted estimation

Panel A: One predictor

	CF		IBC		IBC_A		OP		OE	
Year	Coef.	R^2	Coef.	R^2	Coef.	R^2	Coef.	R^2	Coef.	R^2
1988	0.134	0.013	0.082	0.003	0.126	0.007	0.140	0.013	0.159	0.012
1989	0.122	0.010	0.097	0.005	0.205	0.019	0.201	0.024	0.214	0.021
1990	0.059	0.002	0.099	0.006	0.187	0.015	0.188	0.021	0.178	0.015
1991	0.106	0.006	0.099	0.006	0.240	0.023	0.178	0.018	0.217	0.020
1992	0.154	0.013	0.107	0.007	0.198	0.016	0.185	0.018	0.189	0.015
1993	0.171	0.015	0.117	0.008	0.228	0.018	0.216	0.021	0.220	0.018
1994	0.200	0.024	0.049	0.001	0.183	0.013	0.188	0.019	0.185	0.014
1995	0.102	0.005	0.064	0.002	0.166	0.011	0.174	0.016	0.177	0.012
1996	0.178	0.018	0.173	0.017	0.285	0.031	0.278	0.039	0.292	0.032
1997	0.199	0.021	0.112	0.008	0.224	0.021	0.235	0.030	0.222	0.020
1998	0.112	0.008	0.124	0.011	0.210	0.022	0.207	0.029	0.208	0.022
1999	0.179	0.016	0.195	0.020	0.306	0.037	0.318	0.053	0.324	0.041
2000	0.162	0.015	0.149	0.016	0.255	0.032	0.278	0.050	0.271	0.036
2001	0.200	0.025	0.198	0.033	0.329	0.052	0.331	0.065	0.339	0.057
2002	0.288	0.044	0.284	0.067	0.441	0.090	0.402	0.086	0.472	0.099
2003	0.392	0.072	0.304	0.056	0.462	0.092	0.440	0.093	0.507	0.106
2004	0.341	0.059	0.268	0.041	0.413	0.069	0.393	0.075	0.448	0.078
2005	0.332	0.056	0.311	0.063	0.438	0.086	0.437	0.100	0.492	0.102
2006	0.286	0.042	0.327	0.064	0.448	0.090	0.442	0.114	0.471	0.095
2007	0.239	0.031	0.233	0.035	0.388	0.064	0.391	0.084	0.417	0.072
2008	0.179	0.019	0.115	0.020	0.282	0.044	0.260	0.049	0.321	0.053
2009	0.150	0.013	0.168	0.025	0.353	0.061	0.265	0.047	0.376	0.064
2010	0.262	0.034	0.268	0.044	0.355	0.052	0.406	0.081	0.400	0.062
2011	0.331	0.055	0.244	0.040	0.362	0.060	0.354	0.066	0.426	0.078
2012	0.325	0.050	0.243	0.043	0.406	0.067	0.405	0.083	0.470	0.082
2013	0.288	0.036	0.289	0.043	0.411	0.064	0.391	0.069	0.476	0.078
2014	0.341	0.051	0.264	0.045	0.412	0.069	0.374	0.070	0.447	0.078
2015	0.317	0.044	0.220	0.037	0.410	0.065	0.324	0.059	0.463	0.077
2016	0.240	0.027	0.184	0.025	0.341	0.052	0.243	0.034	0.351	0.050
2017	0.297	0.033	0.260	0.039	0.384	0.057	0.403	0.069	0.446	0.067
2018	0.378	0.064	0.350	0.072	0.424	0.076	0.350	0.058	0.449	0.082
Diff.	0.129***	0.025***	0.120***	0.028***	0.150***	0.038***	0.128***	0.037***	0.186***	0.046***
	(5.958)	(5.397)	(5.355)	(4.917)	(6.535)	(6.077)	(5.162)	(4.863)	(7.441)	(6.525)

Appendix Table 5: Cross-sectional forecasts by year: mean-adjusted estimation (continued)

Panel B: Two predictors

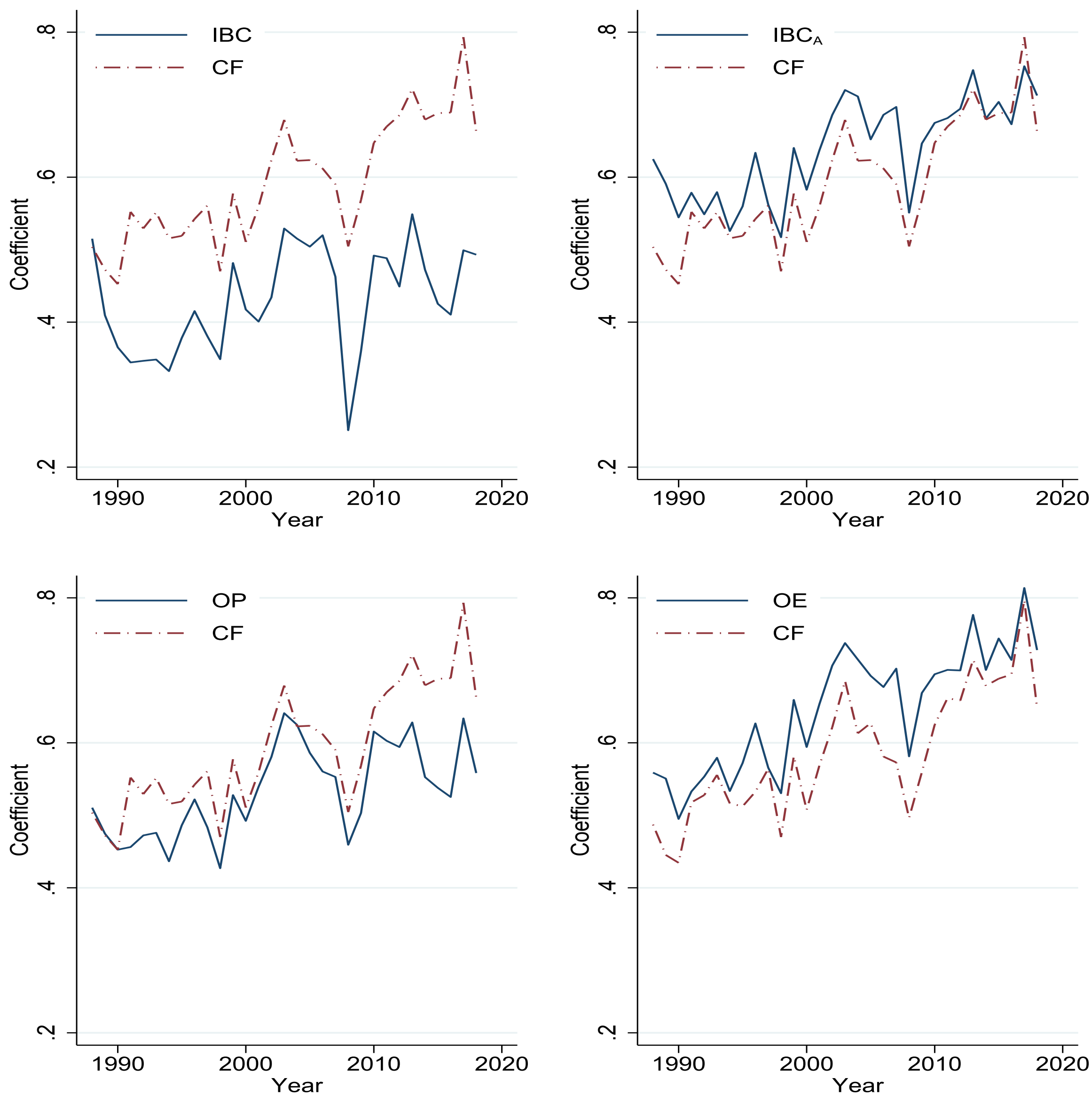
	IBC			IBC_A			OP			OE		
Year	CF	Pred.	R^2	CF	Pred.	R^2	CF	Pred.	R^2	CF	Pred.	R^2
1988	0.128	0.018	0.013	0.113	0.048	0.013	0.092	0.094	0.017	0.089	0.099	0.016
1989	0.103	0.053	0.011	0.045	0.176	0.020	0.043	0.180	0.025	0.036	0.191	0.022
1990	0.028	0.088	0.006	-0.027	0.205	0.016	-0.028	0.201	0.021	-0.025	0.194	0.015
1991	0.078	0.073	0.009	0.012	0.234	0.023	0.035	0.163	0.018	0.022	0.205	0.020
1992	0.131	0.063	0.015	0.092	0.146	0.019	0.091	0.143	0.021	0.096	0.136	0.019
1993	0.147	0.077	0.018	0.107	0.169	0.023	0.106	0.170	0.026	0.108	0.162	0.023
1994	0.211	-0.030	0.024	0.166	0.083	0.026	0.150	0.116	0.029	0.164	0.086	0.026
1995	0.089	0.032	0.006	0.041	0.142	0.011	0.032	0.160	0.016	0.034	0.157	0.012
1996	0.130	0.124	0.025	0.078	0.237	0.033	0.069	0.245	0.041	0.071	0.246	0.034
1997	0.179	0.046	0.022	0.131	0.146	0.027	0.111	0.184	0.035	0.131	0.143	0.027
1998	0.066	0.098	0.013	0.006	0.206	0.022	0.005	0.205	0.029	0.004	0.206	0.022
1999	0.112	0.144	0.025	0.029	0.286	0.037	-0.000	0.318	0.053	0.006	0.320	0.041
2000	0.109	0.102	0.021	0.028	0.236	0.032	-0.006	0.282	0.050	0.004	0.268	0.036
2001	0.120	0.151	0.040	0.035	0.305	0.053	0.027	0.316	0.065	0.021	0.324	0.057
2002	0.154	0.225	0.076	0.059	0.403	0.091	0.092	0.350	0.090	0.016	0.461	0.099
2003	0.292	0.167	0.084	0.178	0.339	0.100	0.184	0.324	0.102	0.126	0.415	0.110
2004	0.263	0.140	0.067	0.173	0.288	0.078	0.164	0.289	0.083	0.137	0.342	0.083
2005	0.203	0.219	0.079	0.118	0.359	0.091	0.099	0.378	0.103	0.067	0.444	0.103
2006	0.136	0.260	0.071	0.024	0.432	0.090	-0.004	0.445	0.114	-0.007	0.476	0.095
2007	0.149	0.165	0.044	0.027	0.369	0.064	0.015	0.383	0.084	-0.013	0.427	0.072
2008	0.126	0.082	0.028	0.016	0.272	0.044	0.036	0.243	0.050	-0.038	0.350	0.053
2009	0.076	0.142	0.027	-0.056	0.388	0.062	0.004	0.263	0.047	-0.075	0.426	0.066
2010	0.141	0.199	0.050	0.081	0.297	0.054	0.012	0.398	0.081	0.036	0.372	0.062
2011	0.253	0.122	0.062	0.177	0.233	0.068	0.159	0.251	0.073	0.099	0.349	0.080
2012	0.233	0.151	0.062	0.139	0.311	0.073	0.097	0.346	0.086	0.070	0.416	0.083
2013	0.171	0.210	0.053	0.078	0.359	0.066	0.067	0.350	0.070	0.001	0.475	0.078
2014	0.242	0.162	0.064	0.136	0.316	0.074	0.147	0.287	0.076	0.088	0.381	0.080
2015	0.232	0.138	0.056	0.117	0.329	0.069	0.153	0.242	0.065	0.050	0.424	0.077
2016	0.162	0.109	0.033	0.019	0.327	0.052	0.109	0.175	0.037	-0.006	0.356	0.050
2017	0.180	0.185	0.048	0.078	0.334	0.059	0.016	0.393	0.069	-0.030	0.470	0.067
2018	0.213	0.238	0.085	0.160	0.303	0.082	0.247	0.181	0.071	0.125	0.351	0.085
Diff.	0.066***	0.084***	0.034***	0.025	0.126***	0.039***	0.034*	0.100***	0.038***	-0.018	0.191***	0.044***
	(3.646)	(4.093)	(5.260)	(1.117)	(4.816)	(6.194)	(1.422)	(3.407)	(5.119)	(-0.841)	(6.697)	(6.528)

The table presents the results of cross-sectional forecasting of future cash flows for each year. To eliminate firm fixed effects, each variable is de-meaned at firm level. Panels A and B are based on the following regressions, respectively:

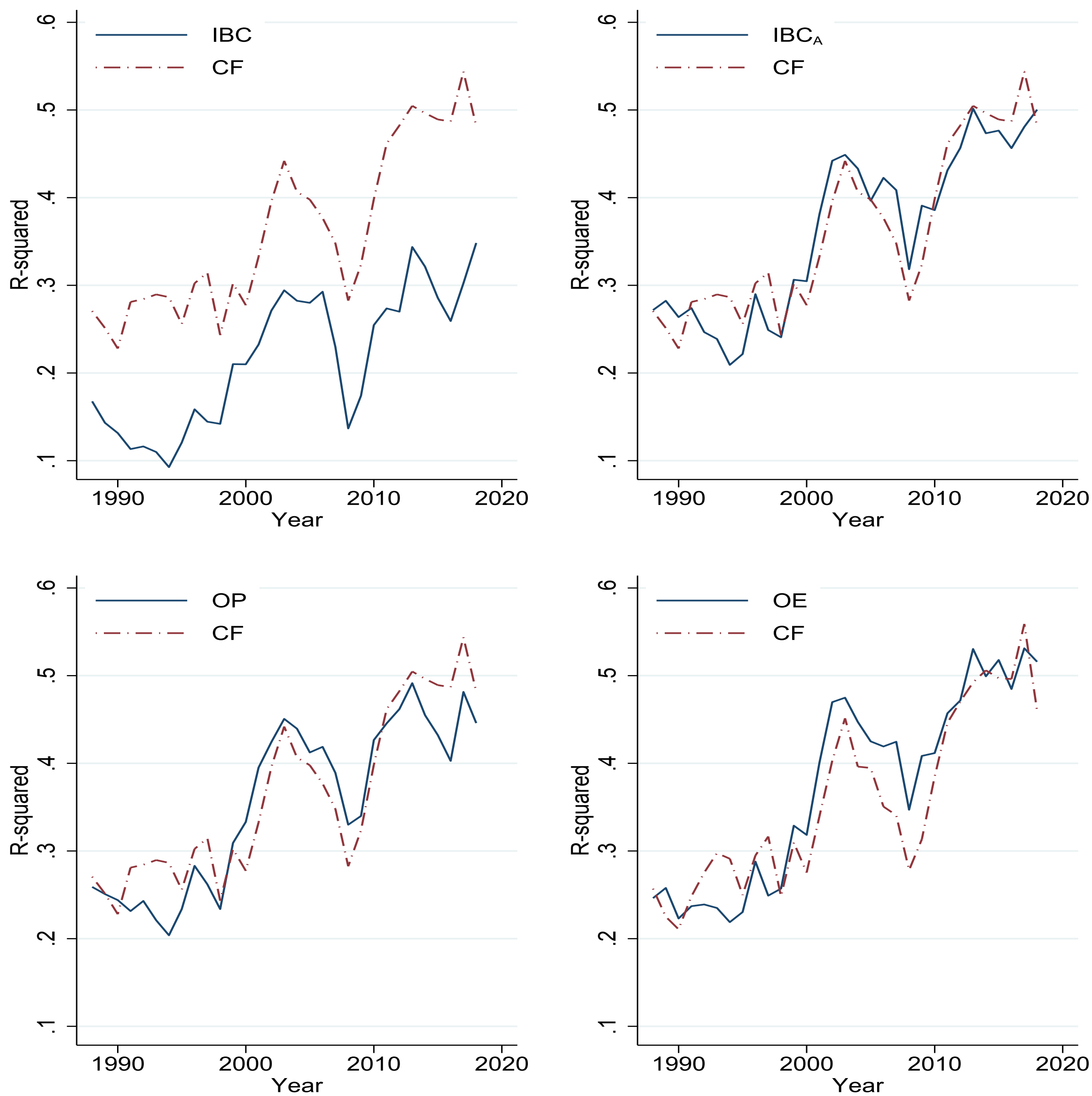
$$\text{CF}_{t+1} = \alpha + \beta \text{Predictor}_t + \tilde{\epsilon}_{t+1}$$

$$\text{CF}_{t+1} = \alpha + \beta_1 \text{CF}_t + \beta_2 \text{Predictor}_t + \tilde{\epsilon}_{t+1}$$

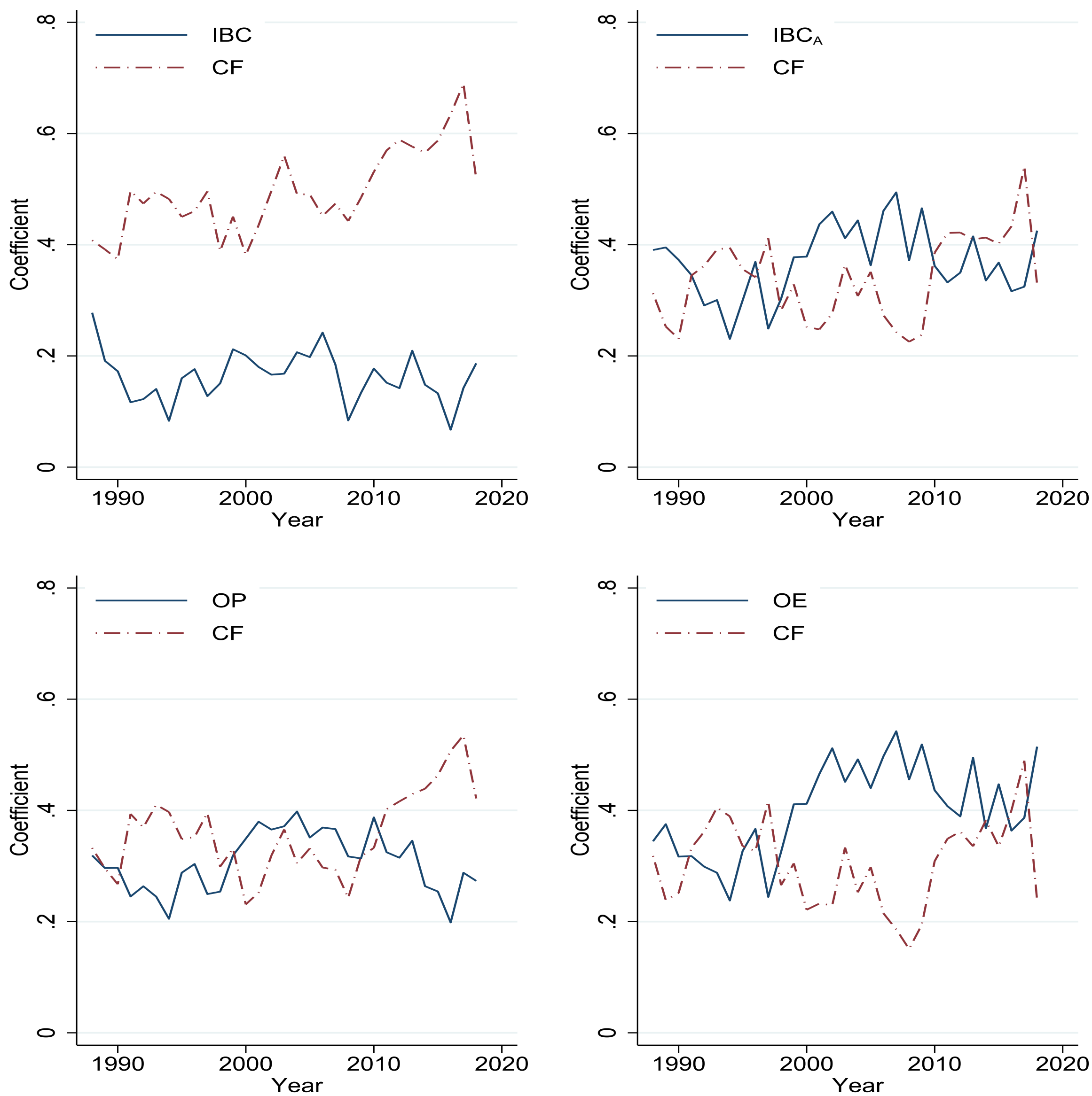
Panel A reports the coefficient, β, and the R^2 each year, whereas Panel B reports both coefficients, β_1 and β_2, and the R^2. The last two rows present and test for the difference between the mean value in the coefficients and R^2s in the latter half of the sample period (2004-19) and the former half (1988-2003). t-statistics are shown in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively.



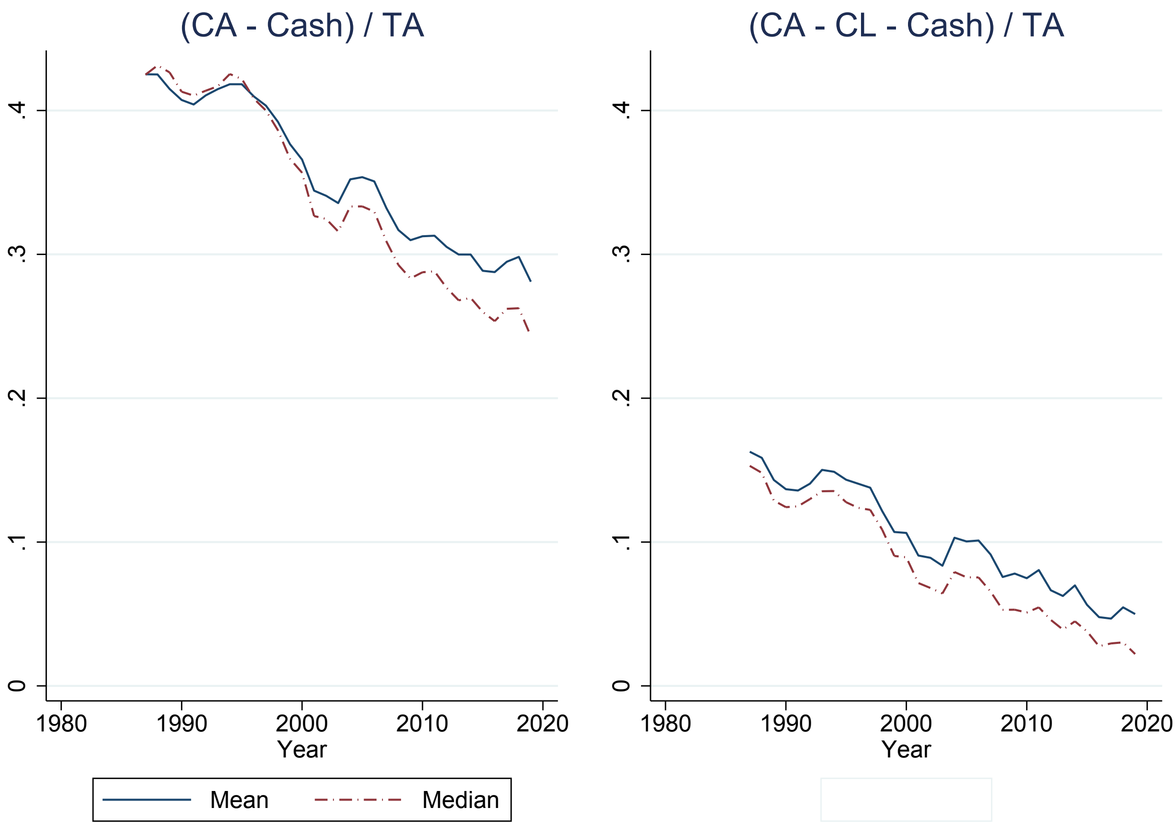
Appendix Figure 1: The figure displays the coefficients from the cross-sectional regressions of future cash flows on a single predictor (no firm fixed effects are included). Each plot contrasts the coefficients obtained using a non-CF predictor with that obtained using CF. Clockwise, starting at the top-left, the predictors are IBC, IBC_A, OE and OP. In all plots, solid lines denote the coefficients for the non-CF predictor and dashed lines denote that for CF. See Table 1 for variable definitions.



Appendix Figure 2: The figure displays the R-squareds based on the cross-sectional regressions of future cash flows on a single predictor (no firm fixed effects are included). Each plot contrasts the R^2 obtained using a non-CF predictor with that obtained using CF. Clockwise, starting at the top-left, the predictors are IBC, IBC$_{\mathrm{A}}$, OE and OP. In all plots, solid lines denote the R^2 for the non-CF predictor and dashed lines denote that for CF. See Table 1 for variable definitions.



Appendix Figure 3: The figure displays the coefficients from the cross-sectional regression of future cash flows on two-predictors (no firm fixed effects are included). Each plot contrasts the coefficients to CF and the other predictor simultaneously present as an independent variable in each regression. Clockwise, starting at the top-left, the (non-CF) predictors are IBC, IBC$_A$, OE and OP. In all plots, solid lines denote the coefficient for the non-CF predictor and dashed lines denote that for CF. See Table 1 for variable definitions.



Appendix Figure 4: On the left: non-cash current assets as a proportion of total assets. On the right: non-cash current assets minus current liabilities, also as a proportion of total assets.