

# The ebbing of accrual accounting

Catalin Starica\* and Pierluigi Giosi†

First version 21.11.2017

Current version 21.03.2019

## Abstract

This paper investigates the evolution of accruals ability to improve the information content of earnings over cash flows, as reflected in market prices. We documented a pronounced decline in the price association of earnings relative to cash flows. We find that, after year 2000, cash flows have informed prices at least as much as earnings. We explain the decline in earnings price association through the decrease of the timing role of accruals related to the growth in the frequency and the magnitude of non-timing accruals (Bushman et al (2016)). Our results imply that this evolution of accruals has severely limited the information content of earnings and, consequently, significantly weakened the pertinence of accrual accounting to prices.

**Keywords:** Expectation formation pertinence; earnings; cash flows; accrual accounting; non-linear association; non-parametric regression; stochastic dominance

*JEL classification:* G10, G30, M41

---

\*School of Economics and Business, University of Neuchâtel, Switzerland.

†School of Economics and Business, University of Neuchâtel, Switzerland.

We wish to thank Mary Barth, Mark Bradshaw, Richard Sloan and the seminar participants at Stanford University for helpful questions and suggestions on an earlier version of the manuscript, and Jian Kang for for the many discussions that furthered our understanding of numerous issues related to the paper.

# 1 Introduction

Previous accounting literature has argued that accruals mitigate the timing and matching problems inherent to cash flows by being negatively correlated to earnings (Dechow (1994)). By adding accruals to operating cash flows, accrual accounting produces a less noisy measure of operating performance (than operating cash flows). As a result, earnings better summarize firm performance, as reflected in stock returns (Dechow et al (1998)).

This study re-visits the issue of the superiority of earnings over cash flows in their association to prices in the light of new evidence on the changing landscape of accrual accounting (Bushman et al (2016)). We document a pronounced and continuous decline of earnings association with prices relative to cash flows. While in the beginning of the sample, earnings showed a significantly stronger association with prices relative to cash flows, after year 2000, cash flows have become at least as relevant to price formation (if not more) as earnings.

We relate the evolution of price association of earnings relative to cash flows with changes in the relative smoothness of the two accounting variables. More precisely, we document a strong increasing relation between the *ratio of earnings to cash flow volatilities* and the *relative price association* of the two accounting measures. Earnings of firms with small (large) ratios between earnings and cash flow volatilities are significantly more (less) informative than the corresponding cash flows.

This finding is consistent with previous evidence on the relation between smoothness and predictability. Dichev and Tang (2009) documents that volatile earnings (cash flows) are less informative for predicting a firms future fundamentals. It is also pertinent in a prediction set-up: higher volatility means less reliable forecasts of the measure of future performance (Watts and Leftwich (1977), Gerakos and

Gramacy (2013), Nallareddy et al (2018)). As a consequence, increase in earnings volatility relative to cash flows should reduce investors' relative ability to project a firm's future performance using the current level of net income (Barton et al (2010)), reducing consequently earnings relative pertinence to price formation.

Next, we document that the proportion of firms with a ratio of volatilities in the range corresponding to cash flows being *at least as price-associated* as earnings, has been increasing steadily, mostly after 1980, from ten percent in the 60s to around sixty five percent in 2016. In other words, for two thirds of the firms in the recent cross-sections, cash flows are at least as informative as earnings. For around fifty percent of the firms in these cross-sections, cash flows are strictly more price formation pertinent than earnings.

We explore potential reasons for the surprising order reversal of volatilities at the origin of the evolution of the relative price association. We identify three mechanisms of change. First is the growing prominence of firms with high intangible intensity (Srivastava (2014)) which significantly alters the characteristics of earnings and their components, cash flow and accruals. (Kothari et al (2002)). A second possible reason could be an increase in the volatility of accruals due primarily to the rising incidence of one-time items and non-operating accruals (Givoly and Hayn (2000), Bradshaw and Sloan (2002)). A third competing explanation is the documented reduction of the timing role of accruals (Bushman et al (2016)). To perform their smoothing or noise reduction role, accruals need to be negatively correlated to cash flows (Dechow (1994)). Or, the magnitude of the correlation between cash flows and accruals has been increasing steadily over the years. As a result, the difference between the cash volatility and the earnings volatility should diminish and, possibly, become negative.

We perform a horse race between three competing explanatory firm character-

istics associated with the three possible reasons for the changes in earnings price association relative to cash flows (intangible intensity, accrual volatility, and the level of the correlation between cash flows and accruals). We document that the intangible intensity and accruals volatility associations are not significant when controlling with the level of correlation of cash flows and accruals, while the association of the later remains strongly statistically significant.

In the light of the findings in Bushman et al (2016) which associates the dramatic decline in the relative prominence of the timing role of accruals to economic and reporting developments, our results seem to indicate that “the growth in the frequency and the magnitude of non-timing accruals” over the last five decades have severely limited the relative information content of earnings and, consequently, significantly weakened the relevance of accrual accounting to price formation.

The comprehensive analysis we conduct is made possible by an alternative information research design based on the methodological developments in Kang and Starica (2016). Our approach addresses the lack in the market-based accounting literature of a standard design which takes into account the non-linear relationship between the market value of equity and its determining variables (earnings, book value, cash flows, etc.). Despite a growing body of empirical evidence suggesting that equity prices bear a highly non-linear relationship with their determining variables, simple linear models continue to be the standard in the literature. Misspecified linear models are beset by omitted variable bias. As such, the estimated coefficients and predicted values are possibly meaningless and hence lead to erroneous inferences.

The alternative design measures the price association of a given determining variable by the size of investors’ adjustment to a valuation incorporating expecta-

tions of firm's future income stream informed *only by the level of the given variable*. The adjustment (or valuation error), which is firm-year specific, reflects the contribution of other information (than the determining variable) to the expectation formation process. It is consistently estimated through a non-linear regression of prices on the variable under investigation.

To compare the level of price association of two competing determining variables, the alternative research design contrasts the size of the errors corresponding to the valuations informed by each of the variables. We argue that a notion close to *stochastic dominance* gives the natural frame for the comparison. Although many researchers refer to it in areas as diverse as economics, finance, agriculture, statistics, marketing and operations research (Bawa (1982), Levy (1992) and the references therein), to the best of our knowledge, the concept of stochastic dominance has not been used in accounting research related to price association.

## **2 Why an alternative research design**

The debate on the information content of earnings relative to cash flows has been part of the market-based accounting research literature since its beginnings (Ball and Brown (1968), Lev and Ohlson (1982)). Biddle et al (1995) structured the debate around the notions of relative and incremental information content. According to it, "incremental information content comparisons assess whether one accounting measure (or set of measures) provides information content *beyond* that provided by another" while relative information content comparisons investigate "whether one measure provides greater information content than another".

Standard methodology for relative information studies consists in running two linear regressions of prices/returns on each of the predictor variables and compare

the size of the slope coefficients and/or regressions  $R^2$  (Dechow (1994), Hayn (1995)). The regression set-up is often that of the simple univariate regression (Dechow (1994), Hayn (1995), Biddle et al (1995), Biddle et al (1997), Guay and Sidhu (2001)) or a multiple regression following the levels-changes framework in Easton and Harris (1991) (Francis et al (2003)). Moreover, the regressions are run cross-sectionally. The inferences derived from these procedures, that is, testing the statistical significance of coefficients or of differences between regressions'  $R^2$  are, of course, conditional on the degree to which the regressions are correctly specified.

There is a growing body of empirical evidence suggesting that equity prices bear a highly non-linear relationship with their determining variables, that is, earnings, book value, cash flows, etc. (e.g., Burgstahler and Dichev (1997), Freeman and Tse (1992), Das and Lev (1994), Ali (1994), Subramanyam (1996), Lipe et al (1998)). This non-linearity is not reflected in the regressions used in the market-based accounting literature.

The consequences of using linear models for describing a non-linear relation are well-known in the econometric literature (see, e.g., Ramsey (1969), Greene (1990)). Functional form mis-specification is equivalent to omitting relevant variables, where the omitted variables, i.e., the higher-order monomials of the independent variable, represent the non-linearities in the relationship between the dependent and independent variables (Kmenta (1971)). The estimated coefficients of the mis-specified regression will be biased to the extent to that the higher-order polynomials of the independent variable  $X$  are correlated with  $X$ . The more nonlinear the true functional form is, the more the higher-order polynomials will matter, and hence, the larger the specification error will be.

If data exhibits a nonlinear pattern, the coefficients of a fitted (mis-specified)

linear regression might be meaningless. Taking them at face value can lead to erroneous inferences in the research design (see Achen (2005)<sup>1</sup> for a convincing example).

In the accounting literature, the consequences of using linear models for describing a non-linear relation were investigated mainly in the earnings response coefficient framework. For example, Freeman and Tse (1992) shows that forcing a linear specification on an abnormal return-earnings model biases the slope coefficient on earnings toward zero. Cheng et al (1992) comprehensively and formally evaluates the specification of the cross-sectional, ordinary least squares model that relates earnings to risk-adjusted security returns. The study tests for non-linearity, omitted variables, and inter-firm systematic and random coefficient variation and concludes that “the specification error is large enough to affect conclusions regarding economic relationships”. In a replication of a previously published study, the specification problems are shown to lead to substantial instability in inferences from the model.

Estimation on cross sections, which ignores the firm-specific effects, is also problematic and is another source of bias. Lipe et al (1998) argues that including firm-specific effects when estimating the relation between stock returns and accounting earnings provides the largest individual improvement (with significant additional improvements if non-linearity is also considered). Kothari and Shanken (2003) analyzes the time-series variation in the magnitudes of the slope coeffi-

---

<sup>1</sup>Achen (2005) constructs a data set of triplets: a dependent variable paired to an independent variable ( $x_2$ ) and a control variable ( $x_1$ ). The dependent quantity is linear in the independent variable  $x_2$  and there is a small non-linearity in the effect of the control variable  $x_1$ :

$$y = \alpha + \beta_1 f(x_1) + 0.1x_2 + u.$$

The non-linearity in the control causes the mis-specified linear model to give a wildly incorrect estimate of the coefficient of the independent variable  $x_2$ . Estimating a mis-specified linear regression  $y = \alpha + \beta_1 x_1 + \beta_2 x_2 + u$ , yields a strongly significant negative coefficient  $\hat{\beta}_2 = -2.78$ , that is, 28 times larger in magnitude than its true value and with the wrong sign.

coefficients in cross-sectional regressions of prices on financial variables (earnings and book values) and bring indirect evidence of bias due to correlated-omitted variables. It finds that coefficients on earnings are significantly negatively correlated with proxies of the aggregate growth and of expected rate of return. The result is consistent with Collins and Kothari (1989) which examines cross-sectional determinants of earnings response coefficient and finds that the regression coefficient is positively related to firm's earnings persistence and negatively related to firms systematic risk. In a related study, Teets and Wasley (1996) shows that estimating a cross-sectionally constant earnings response coefficient can produce downward biased coefficients and understate the explanatory power of the linear regression.

The alternative research design we propose addresses both econometric issues<sup>2</sup>. Next section briefly presents the theoretical foundation of the alternative research design based on results from Kang and Starica (2016) and introduces carefully its components.

### **3 Methodology**

This section gives the Expectation Formation Pertinence (EFP) design definition (section 3.1), it details the use of the EFP design for relative association studies (section 3.2) and for explaining changes in relative price association (section 3.3). Finally, it discusses the use of the 'in-levels' regression for consistently measuring the association between firms characteristics and the EFP measure (section 3.4).

---

<sup>2</sup>That is, the miss-specification of the linear regression when modeling non-linear relations and ignoring the firm-specific effects.

### 3.1 Expectation Formation Pertinence (EFP) research design

The alternative research design starts from the observation that a *non-linear* regression relation between prices ( $P$ ), on one hand, and a set of accounting variables relevant to the price formation<sup>3</sup> ( $\mathcal{X}$ ), on the other hand, *always* exists (while a linear regression most likely will not). Hence the price  $P_{i,t}$  can be decomposed as the sum of the best non-linear prediction of value based on the observed accounting variables ( $\mathbb{E}[P_{i,t} | \mathcal{X}_{i,t}]$ ) and an adjustment term which reflects all other information available to investors:

$$P_{i,t} = \mathbb{E}[P_{i,t} | \mathcal{X}_{i,t}] + \varepsilon_{i,t} := \mathbf{m}_i^{\mathcal{X}}(\mathcal{X}_{i,t}) + \varepsilon_{i,t}, \quad (1)$$

where  $\mathbb{E}[P_{i,t} | \mathcal{X}_{i,t}]$  stands for the conditional expectation<sup>4</sup> of  $P$  given  $\mathcal{X}$ . The notation  $\mathbf{m}^{\mathcal{X}}$  emphasizes that the function  $\mathbf{m}$  is set-specific: if the set of accounting variables changes, the regression function changes. From a statistical point of view, investors' adjustment is regression's error term that is orthogonal<sup>5</sup> to the predictor, that is,  $\mathbb{E}[\varepsilon_{i,t} | \mathcal{X}_{i,t}] = 0$ .

Estimating a linear relation between the price and the accounting variables pertinent to the process of price formation when the true regression relation is a *non-linear* one, will yield coefficients that do not reflect the economic relation in (1) and that, consequently, cannot be used to test hypothesis<sup>6</sup>.

The function  $\mathbf{m}$  can be *consistently* estimated on cross-sections of firms by employing proven inference techniques from the field of non-parametric regres-

---

<sup>3</sup>  $\mathcal{X}$  can be one variable, like *NI* or *CFO*, or a set of accounting variables, like  $\{CFO, TACC\}$ .

<sup>4</sup>It is a well-known fact that  $\mathbb{E}[P | \mathcal{X}]$  is a (generally) *non-linear* function of  $\mathcal{X}$ . It is also the best predictor (in the  $L^2$ -norm) of  $P$  when one knows  $\mathcal{X}$ .

<sup>5</sup>This condition guarantees the consistent estimation of  $\mathbf{m}$  and is violated if a linear regression is estimated instead.

<sup>6</sup>This is, fundamentally, the explanation of the well-known bias issue that affects the price-levels or the returns-earnings regressions (Lo (2005), Barth and Clinch (2009)). Estimating a *non-linear* regression resolves the issue.

sion. Its estimate,  $\widehat{\mathbf{m}}$ , is theoretically guaranteed to be *unbiased* (see Kang and Starica (2016)).

**A non-linear valuation.** Assuming that prices follow the RI valuation relation (Preinreich (1936) and (1938), Edwards and Bell (1961), Peasnell (1982))

$$P_{i,t} = B_{i,t} + \sum_{u=1}^{\infty} \frac{\mathbb{E}_t[NI_{i,t+u} - r_{i,t} \times B_{i,t+u-1}]}{(1+r_{i,t})^u} = B_{i,t} + \sum_{u=1}^{\infty} \frac{\mathbb{E}_t[RI_{i,t+u}]}{(1+r_{i,t})^u}, \quad (2)$$

( $r_t$  denotes the price of equity risk at time  $t$ ,  $\mathbb{E}_t$  stands for market's expectation conditional on all information available at time  $t$ , while  $RI$  stands for the abnormal earnings  $NI - r \times B_{-1}$ ), Kang and Starica (2016) show that the regression function  $\mathbf{m}$  in the decomposition (1) can be expressed as:

$$\mathbf{m}_i^{\mathcal{X}}(\mathcal{X}_{i,t}) = B_{i,t} + \sum_{u=1}^{\infty} \frac{\mathbb{E}[RI_{i,t+u} | \mathcal{X}_{i,t}]}{(1+r_{i,t})^u}. \quad (3)$$

By comparing the expression in (3) to the RI representation (2), one notes that the regression function  $\mathbf{m}$  is a *valuation where the expectations of future abnormal earnings are informed only by the level of the accounting variables  $\mathcal{X}_{i,t}$* .

**Empirical specification.** The main determinants of the future evolution of abnormal earnings are firm's cost of equity, its growth profile and, possibly, its level of accounting conservatism<sup>7</sup> (Kang and Starica (2016)). Hence, for the firm  $i$  in a cross-section  $t$  (and belonging to industry I), the empirical specification of the RI economic relation (2) is:

$$P_{i,t} := \mathbf{m}_{I,t}^{\mathcal{X}}(\mathcal{X}_{i,t}; r_{i,t}, g_{i,t}, C_{i,t}) + \varepsilon_{i,t}, \quad (4)$$

where  $r_{i,t}$  stands for firm's cost of equity,  $g_i$  denotes firm's income growth and

---

<sup>7</sup>How current abnormal earnings forecast future earnings might depend on how the earnings are constructed.

$C_i$ , its conservative accounting level, and where the error term  $\varepsilon_{i,t}$  satisfies the *orthogonality condition*

$$\mathbb{E}_t[\varepsilon_i | \mathcal{X}_i] = 0.$$

The regression in (4) is estimated cross-sectionally, for each industry<sup>8</sup>.

**Firm-specific measure of association.** Since the regression function can be interpreted as a valuation, the (relative) size of the estimated regression error

$$\left| \frac{\widehat{\varepsilon}_{i,t}}{P_{i,t}} \right| = \left| \frac{P_{i,t} - \widehat{\mathbf{m}}_{I,t}^{\mathcal{X}}}{P_{i,t}} \right|$$

quantifies the extent to which future earnings expectations are shaped by information *other than* the level of the accounting variables  $\mathcal{X}$ . This error is defined to be the *firm-year-specific measure of association* (Expectation Formation Pertinence) for the data entry  $(i,t)$ . It amounts to the proportion of the price corresponding to earnings expectations shaped by information other than the current value of the accounting variables  $\mathcal{X}$ .

### 3.2 EFP design for relative association studies

The research design introduced in section 3.1 is instrumental to comparing the information content (pertinent to the expectation formation process impounded in prices) of two sets of predictor variables,  $\mathcal{X}$  vs.  $\mathcal{Y}$ <sup>9</sup>.

To use it towards this goal, we take the following steps. First, for the firm  $i$  in the cross-section  $t$  belonging to the industry  $I$ , we separate the price impact of each of the two predictors  $\mathcal{X}$  and  $\mathcal{Y}$  from the contribution of other information

---

<sup>8</sup>The underlying assumption is that firms sharing the same industry code have similar non-linear valuation functions  $\mathbf{m}$  once we control for risk, growth, and level of conservatism.

<sup>9</sup>In the case of our study,  $\mathcal{X} := \{NI\}$  and  $\mathcal{Y} := \{CFO\}$ .

by inferring (in cross-sections) the non-linear valuation regression (4) for each of the two sets of accounting variables in the comparison. This step yields two valuations:

$$\widehat{P}_{i,t}^{\mathcal{X}} = \widehat{\mathbf{m}}_{i,t}^{\mathcal{X}}(\mathcal{X}_{i,t}; r_{i,t}, g_{i,t}, C_{i,t})$$

and

$$\widehat{P}_{i,t}^{\mathcal{Y}} = \widehat{\mathbf{m}}_{i,t}^{\mathcal{Y}}(\mathcal{Y}_{i,t}; r_{i,t}, g_{i,t}, C_{i,t}),$$

respectively, and, hence, two (relative) valuation errors:

$$|P_{i,t} - \widehat{P}_{i,t}^{\mathcal{X}}|/P_{i,t} \quad \text{and} \quad |P_{i,t} - \widehat{P}_{i,t}^{\mathcal{Y}}|/P_{i,t}. \quad (5)$$

Second, we statistically compare the distributions of the valuation errors in each cross-section  $t$ . A distribution of the error corresponding to the set  $\mathcal{X}$  more concentrated towards 0 (than that corresponding to the predictor set  $\mathcal{Y}$ ) is evidence of smaller size errors for the valuation informed by the set  $\mathcal{X}$ . Since the investors need to adjust the former valuation less than the later, we would conclude that the predictor set  $\mathcal{X}$  is more *expectation formation pertinent* than the set of predictor variables  $\mathcal{Y}$ .

In this paper, we use two criteria for measuring the relative pertinence of two sets of accounting variables,  $\mathcal{X} = NI$  and  $\mathcal{Y} = CFO$ . First, when documenting that earnings superiority in informing prices (relative to cash flows) has disappeared over time, we compare the cross-sectional *cumulative distribution functions (cdfs)* of the two valuation error terms:

$$F_{\mathfrak{X},t}(x) := \mathbb{P}\left(|P_t - \widehat{P}_t^{\mathfrak{X}}|/P_t \leq x\right), \quad x \in \mathbb{R},$$

where the symbol  $\mathfrak{X}$  stands for  $NI$  and  $CFO$ , respectively. We address this element of the research design in sections 3.2.1, 3.2.2, and 3.2.3.

Second, when exploring potential reasons for this disappearance, we focus on the (log) *ratio between the EFP measures* of the two sets, defined in (5):

$$Rel\ EFP_{i,t} := |P_{i,t} - \widehat{P}_{i,t}^{CFO}| / |P_{i,t} - \widehat{P}_{i,t}^{NI}|. \quad (6)$$

In order to consistently infer the individual association of various correlated firm characteristics with this ratio, we use all year-firm observations and regress it *jointly* on competing firm characteristics hypothesized to explain it. We give more details about this part of the investigation in sections 3.3 and 3.4.

### 3.2.1 Stochastic dominance - an exhaustive criterion for comparing EFP

The concept of *stochastic dominance* gives the natural set-up for comparing the pertinence of competing predictor sets to the expectation formation process. While stochastic dominance has been extensively used in finance research, especially for the comparison of portfolios' performance, to our knowledge, it has not yet been applied for the analysis of valuation accuracy.

The stochastic dominance criterion is both intuitively appealing and comprehensive. We say that valuation  $\mathcal{X}$  (informed by the predictors set  $\mathcal{X}$ ) dominates valuation  $\mathcal{Y}$  (informed by the competing predictors set  $\mathcal{Y}$ ) if the proportion of firms valued within a specified error is higher for valuation  $\mathcal{X}$  than for valuation  $\mathcal{Y}$ , and that, for all error levels<sup>10</sup>.

This definition can be stated formally making use of the cumulative distribution functions (cdf)  $F_{\mathcal{X}}$  and  $F_{\mathcal{Y}}$  of the distributions of absolute relative errors in

---

<sup>10</sup>To preserve the intuition, recall that in the set-up of this investigation  $\mathcal{X} = \{CFO, NI\}$  while  $\mathcal{Y} = \{CFO\}$ .

the cross-section  $t$ :

$$|P_t - \widehat{P}_t^{\mathcal{X}}|/P_t \quad \text{and} \quad |P_t - \widehat{P}_t^{\mathcal{Y}}|/P_t,$$

corresponding to the two valuations informed only by the level of the predictors  $\mathcal{X}$  and  $\mathcal{Y}$ , respectively. The valuation  $\mathcal{X}$  dominates<sup>11</sup> valuation  $\mathcal{Y}$  (we write  $\mathcal{X} \geq \mathcal{Y}$ ), in the cross-section  $t$ <sup>12</sup>, if

$$F_{\mathcal{X}}(e) \geq F_{\mathcal{Y}}(e), \quad \text{i.e.} \quad (7)$$

$$\mathbb{P} \left( \frac{|P_t - \widehat{P}_t^{\mathcal{X}}|}{P_t} \leq e \right) \geq \mathbb{P} \left( \frac{|P_t - \widehat{P}_t^{\mathcal{Y}}|}{P_t} \leq e \right),$$

for all error sizes  $e$ . In this case, the predictors set  $\mathcal{X}$  is more expectation formation pertinent than the predictors set  $\mathcal{Y}$ .

If  $e$  denotes a specific absolute error, say 20%, then the inequality  $F_{\mathcal{X}}(e) \geq F_{\mathcal{Y}}(e)$  in the definition (7) means that the percentage of firms whose values based on expectations informed by the predictors set  $\mathcal{X}$  are within 20% of the actual price is greater than, or equal to, the percentage of firms for which the value informed by the predictors set  $\mathcal{Y}$  lies within 20% of the actual price. If method  $\mathcal{X}$  dominates method  $\mathcal{Y}$ , then whatever error level we may choose, there is always more precision delivered by the predictors set  $\mathcal{X}$  than by the set  $\mathcal{Y}$ .

Inequality (7) implies a clear relationship between the corresponding percentiles of the two error distributions. Since for a  $p \in [0, 1]$ , the  $(p * 100)\%$  percentile of the distribution  $\mathcal{X}$  is defined as  $F_{\mathcal{X}}^{-1}(p)$ , if method  $\mathcal{X}$  dominates

---

<sup>11</sup>For our purpose, we had to adjust the common definition of *stochastic dominance* which states that *distribution Y dominates distribution X stochastically* at first order if, for any argument  $e$ ,  $F_{\mathcal{X}}(e) \geq F_{\mathcal{Y}}(e)$  where  $F_{\mathcal{X}}$  stands for the cumulative distribution function of the distribution  $\mathcal{X}$ . This definition fits the case where smaller probabilities of low values are desirable, like in the study of poverty. In our case small values mean higher precision and are hence desirable.

<sup>12</sup>For the simplicity of notation, in this section, we will not emphasize the cross-section script  $t$ .

method  $\mathcal{Y}$ , then

$$(p * 100)\% \text{ percentile of } \frac{|P - \hat{P}^{\mathcal{X}}|}{P} \leq (p * 100)\% \text{ percentile of } \frac{|P - \hat{P}^{\mathcal{Y}}|}{P},$$

for all  $p \in [0, 1]$ . In words, all percentiles of the distribution of the error  $|P - \hat{P}^{\mathcal{X}}|/P$  are smaller than the corresponding  $|P - \hat{P}^{\mathcal{Y}}|/P$ -percentiles. In particular, the median absolute error (and even the mean, if the distributions are symmetric) of a valuation informed by the predictors set  $\mathcal{X}$  is smaller than that of a valuation informed by the competing predictors set  $\mathcal{Y}$ .

The notion of dominance yields the most exhaustive criterion for comparing the valuation accuracy, and hence, the expectation formation pertinence (EFP), of competing predictor sets. The relation between the two concepts is summarized in Table 1. In particular, if valuation informed by the predictors set  $\mathcal{X}$  (stochastically) dominates pricing method informed by the competing predictors set  $\mathcal{Y}$ , we will conclude that the predictor set  $\mathcal{X}$  is more pertinent to the expectation formation process than the predictor set  $\mathcal{Y}$ .

In the set-up of our study, documenting that the valuation informed by the union of cash flows and earnings (stochastically) dominates the valuation informed only by cash flows will establish an incremental expectation formation pertinence of earnings with respect to cash flows. Contrary, if the distributions of the two valuation errors are statistically indistinguishable, we will conclude that accruals do not contain incremental information relevant to the expectation formation process. From a valuation point of view, accruals do not enhance the information content of cash flows.

In practice, before performing a comparison of the accuracy of two valuation methods  $\mathcal{X}$  and  $\mathcal{Y}$ , we need to estimate  $F_{\mathcal{X}}$  and  $F_{\mathcal{Y}}$ , the cdfs of the corresponding absolute errors. When comparing the EF pertinence of competing predictor sets,

Stochastic dominance	Expectation formation
$\mathcal{X} > \mathcal{Y}$ or $F_{\mathcal{X}}(e) > F_{\mathcal{Y}}(e), \forall e$	Predictors set $\mathcal{X}$ is more pertinent than predictors set $\mathcal{Y}$
$\mathcal{X} < \mathcal{Y}$ or $F_{\mathcal{X}}(e) < F_{\mathcal{Y}}(e), \forall e$	Predictors set $\mathcal{Y}$ is more pertinent than predictors set $\mathcal{X}$
$\mathcal{X} = \mathcal{Y}$ or $F_{\mathcal{X}}(e) = F_{\mathcal{Y}}(e), \forall e$	Predictors set $\mathcal{X}$ is as pertinent as predictors set $\mathcal{Y}$
Neither cdf dominates the other	The 2 predictors sets cannot be compared

Table 1: **Stochastic dominance and EFP of predictors sets  $\mathcal{X}$  and  $\mathcal{Y}$ .** The table establishes the connection between stochastic dominance relation between the cdfs  $F_{\mathcal{X}}$  and  $F_{\mathcal{Y}}$  of errors  $|P_i - \widehat{P}_i^{\mathcal{X}}|/P_i$  and  $|P_i - \widehat{P}_i^{\mathcal{Y}}|/P_i$  of valuations informed by the predictors set  $\mathcal{X}$  and  $\mathcal{Y}$ , on one hand, and the EFP of the two predictors sets  $\mathcal{X}$  and  $\mathcal{Y}$ .

the statistical error of this estimation needs to be taken into account. For details about how this important issue can be addressed rigorously, see next section as well as Sections A and B in the Appendices.

### 3.2.2 Statistical tests of relative EFP - hypotheses

In this section we discuss statistical issues related to the tests we use to establish the stochastic dominance in cross-sections.

We will test (in cross-sections) the hypothesis of equal EFP of two sets of predictors  $\mathcal{X}$  and  $\mathcal{Y}$  (in the case of this study,  $\mathcal{X} = \{NI\}$  and  $\mathcal{Y} = \{CFO\}$ ) against

the alternative of one set,  $\mathcal{X}$ , being more expectation formation pertinent than the other,  $\mathcal{Y}$ , through *evaluating the stochastic dominance relationship between the errors of the valuations informed by the competing predictor sets*. To this end we will use the comprehensive *Kolmogorov-Smirnov statistics*.

We begin by formally stating the hypotheses (null and alternative). The null hypothesis is that, in cross-section  $t$ ,

$$\begin{aligned}
 H_0 : & \text{ The predictor sets } \mathcal{X} \text{ and } \mathcal{Y} \text{ are equally EF pertinent} \\
 & \text{ or, equivalently,} \\
 H_0 : & \text{ The cdf of } (|P_t - \widehat{P}_t^{\mathcal{X}}|/P_t) = \text{ the cdf of } (|P_t - \widehat{P}_t^{\mathcal{Y}}|/P_t).
 \end{aligned} \tag{8}$$

The alternative hypothesis is that, in cross-section  $t$ ,

$$\begin{aligned}
 H_1 : & \text{ The predictor set } \mathcal{X} \text{ is more EF pertinent than } \mathcal{Y} \\
 & \text{ or, equivalently,} \\
 H_1 : & \text{ The cdf of } (|P_t - \widehat{P}_t^{\mathcal{X}}|/P_t) > \text{ the cdf of } (|P_t - \widehat{P}_t^{\mathcal{Y}}|/P_t).
 \end{aligned} \tag{9}$$

This hypothesis will be tested using a sample of  $n$  paired (strongly dependent) valuation errors, available at time  $t$ :

$$\begin{aligned}
 & \left( |P_{1,t} - \widehat{P}_{1,t}^{\mathcal{X}}|/P_{1,t}, |P_{1,t} - \widehat{P}_{1,t}^{\mathcal{Y}}|/P_{1,t} \right), \\
 & \left( |P_{2,t} - \widehat{P}_{2,t}^{\mathcal{X}}|/P_{2,t}, |P_{2,t} - \widehat{P}_{2,t}^{\mathcal{Y}}|/P_{2,t} \right), \\
 & \dots, \\
 & \left( |P_{n,t} - \widehat{P}_{n,t}^{\mathcal{X}}|/P_{n,t}, |P_{n,t} - \widehat{P}_{n,t}^{\mathcal{Y}}|/P_{n,t} \right),
 \end{aligned} \tag{10}$$

corresponding to the firms  $1, 2, \dots, n$ .

### 3.2.3 Statistical tests of relative EFP - Kolmogorov-Smirnov statistics

To test the null hypothesis (8), we calculate two sample Kolmogorov-Smirnov statistics by starting from the sample (10) and taking the following steps.

First, in the cross-section  $t$ , we estimate the cumulative distribution functions of the valuation errors corresponding to the two predictors  $\mathcal{X}$  and  $\mathcal{Y}$  by  $\widehat{F}_{\mathcal{X},t}$  and  $\widehat{F}_{\mathcal{Y},t}$ , respectively (see equation (18) in the Appendices).

Second, we calculate the maximum (minimum) of the positive (negative) differences between the two estimated cdfs:

$$\overline{M}_t(\mathcal{X}, \mathcal{Y}) = \sup_e \{\widehat{F}_{\mathcal{X},t}(e) - \widehat{F}_{\mathcal{Y},t}(e) : \widehat{F}_{\mathcal{X},t}(e) - \widehat{F}_{\mathcal{Y},t}(e) \geq 0\}$$

and

$$\underline{M}_t(\mathcal{X}, \mathcal{Y}) = \inf_e \{\widehat{F}_{\mathcal{X},t}(e) - \widehat{F}_{\mathcal{Y},t}(e) : \widehat{F}_{\mathcal{X},t}(e) - \widehat{F}_{\mathcal{Y},t}(e) < 0\},$$

respectively, and scale them to obtain the corresponding Kolmogorov-Smirnov statistics:

$$\overline{KS}_t(\mathcal{X}, \mathcal{Y}) = \sqrt{\frac{n_t}{2}} \overline{M}_t(\mathcal{X}, \mathcal{Y}) \quad (11)$$

and

$$\underline{KS}_t(\mathcal{X}, \mathcal{Y}) = \sqrt{\frac{n_t}{2}} \underline{M}_t(\mathcal{X}, \mathcal{Y}), \quad (12)$$

respectively, where  $n_t$  is the sample size of the cross-section  $t$ .

The two statistics have intuitive definitions. They measure the distance between the two cdf distributions and are useful in testing the statistical hypotheses of stochastic dominance.

The distributions of the statistics  $\overline{KS}_t(\mathcal{X}, \mathcal{Y})$  and  $\underline{KS}_t(\mathcal{X}, \mathcal{Y})$  depend on the joint distribution of the error pairs (10) and can be approximated via bootstrap (see Section B in the Appendices for details).

In the sequel analysis, we display the statistics  $\overline{KS}_t$  and  $\underline{KS}_t$  as functions of time  $t$ , as well as their upper (lower) 95% confidence bounds<sup>13</sup>. For a given cross-section  $t$ , values of the  $\overline{KS}_t$  ( $\underline{KS}_t$ ) statistic above (below) the confidence bound indicate that, in cross-section  $t$ , earnings (cash flows) are significantly more informative for price setting. Values of the  $\overline{KS}_t$  ( $\underline{KS}_t$ ) statistic between 0 and the confidence bounds indicate that, in cross-section  $t$ , earnings and cash flows inform equally much the price formation process.

We document a pronounced and continuous decline of earnings association with prices relative to cash flows. While in the beginning of the sample, earnings showed a significantly stronger association with prices relative to cash flows, in recent years, cash flows have become at least as pertinent to price formation (if not more) as earnings.

### 3.3 EFP design for explaining changes in price association

The EFP research design is particularly useful when exploring potential reasons for the decline of the relative pertinence of earnings. For this purpose, we quantify the relevance of earnings relative to cash flows as the *ratio between the EFP measures of the two accounting variables*, defined in (6).

In a first step, we perform a horse race between three competing explanatory firm characteristics (intangible intensity, accrual volatility, and the level of the correlation between cash flows and accruals) by regressing the ratio (6) on these characteristics. This step answers the question:

*'Which of these three firm characteristics is associated (in a non-biased analysis) with the level of earnings price pertinence relative to cash flows?'*

---

<sup>13</sup>Note that the other bounds of these statistics are at 0.

The horse race is made possible by the fact that the measure (6) is *year-firm specific* and not cross-section specific (as the  $R^2$  measure in the value relevance design). It allows for *unbiased estimation* of the individual association of otherwise correlated firm characteristics with the ratio of valuation errors.

We document that the intangible intensity and accruals volatility associations with the ratio are not significant when controlling with the level of correlation of cash flows and accruals, while the association of the later remains strongly statistically significant.

In a second step, we investigate if the evolution of the measure of accruals' smoothing role can explain the changes in the relative price association of two income accounting variables. This step addresses the question:

*'Is the evolution of earnings price association relative to cash flows explained by the pronounced and continuous decline of the timing role of accruals (Bushman et al (2016))?'*

We identify a time trend in the median of the firm-year EFP measure (6). We, then, control for the level of the correlation of cash flows and accruals. If the evolution of the relative pertinence measure is associated with the dynamics of the firm characters tic, controlling for the level of the later in cross-sections should remove the evidence of trend.

We document that controlling for the decrease in the timing role of accruals removes the strong downwards trend in the median ratio (6) after 1980, i.e. during the period when the smoothing role of the accruals declined significantly.

Next section discusses some methodological aspects related to the regressions in the two steps of the analysis outlined above.

### 3.4 'In-levels' regression - consistent estimation of association of firms characteristics to EFP measure

This section presents methodological aspects related to the estimation of the relation between the relative EFP measure (*Rel EFP*) defined in (6) and firm characteristics. Our first goal is to disentangle the individual association of firm characteristics<sup>14</sup> to the EFP measure. Multiple regression is the natural tool for attaining it. It allows to quantify the impact of a firm characteristic while holding the level of the other constant.

However, since there is no reason to assume a linear<sup>15</sup> relation between the level of the relative EFP measure and firm characteristics, the use of the plain vanilla linear regression might not be suitable. To avoid the assumption of linearity in the relation between the regressor and the independent variables, we use the linear regression as a tool for *testing differences in mean* as follows.

It is well known that, when the independent variable in a linear regression is an indicator variable, *testing the null hypothesis of zero regression coefficient is equivalent to testing the equality of the mean of the dependent variable conditional on the presence or absence of the character modeled by the binary variable* (see, e.g., Stock and Watson (2012)). In particular, if the indicator variable models contrasting values of a given character (e.g. high vs. low), a significant non-zero regression coefficient of the binary variable indicates statistically distinct means of the dependent variable conditional on the level of the character modeled by the indicator variables.

---

<sup>14</sup>Such as intangible intensity, correlation between cash flows and accruals, cash flow or accrual volatility.

<sup>15</sup>Contrary, it is very plausible that the change in the association with a EFP measure corresponding to an incremental change in the value of a firm characteristic will depend on the level of firm's characteristic.

This observation motivates our use of the linear regression: instead of plugging the numerical values of firms characteristics directly in a linear multivariate regression, we convert them in level indicators and test if firms with values under/above the reference level have lower/higher associated EFP measures (for an extensive discussion on this particular approach to linear regression, see Hardy (1993)).

Next we make precise the notation we use in presenting our analysis. We split the range of values of the firm character  $Ch$  in  $l + 1$  intervals of equal probability<sup>16</sup>,  $(a_{min}, a_1], (a_1, a_2], \dots, (a_{l-1}, a_l], (a_l, a_{max})$ , and construct  $l$  indicator functions

$$D^{(Ch,k)}, \quad k \in \{1, 2, \dots, l\},$$

modeling the different levels of character  $Ch$ . The left-out level (in this case, the smaller values of the character) will be the *reference level*. We then replace the classical linear regression

$$Rel\ EFP_{i,t} = \gamma_{0,t} + \gamma Ch_{i,t} + \varepsilon_{i,t}$$

by its corresponding version in levels:

$$Rel\ EFP_{i,t} = \gamma_{0,t} + \gamma_{1,t} D_{i,t}^{(Ch,1)} + \gamma_{2,t} D_{i,t}^{(Ch,2)} + \dots + \gamma_{l,t} D_{i,t}^{(Ch,l)} + \varepsilon_{i,t}.$$

In this set-up<sup>17</sup>, a significant non-zero regression coefficient of the  $k$ -th indicator variable  $\gamma_k$  means that the mean relative *EFP* measure *conditional on the character  $Ch$  taking values in the  $k$ -th bracket of its range,  $(a_{k-1}, a_k]$* , is statistically different from *EFP*'s mean value when conditioning with character's reference

---

<sup>16</sup>In this case  $a_i$  is the  $i/(l+1)$ -quantile.

<sup>17</sup>The set-up is particularly relevant to the static analysis in Section 5.4.1.

values,  $(a_{min}, a_1]$ . Significant non-zero coefficients are evidence of association between the  $EFP$  measure and firm's characteristic  $Ch$ .

The following short-hand notation simplifies the expression for the 'in-levels' regressions above:

$$Rel EFP_{i,t} = \gamma_{0,t} + \gamma_t \times D_{i,t}^{(Ch)} + \varepsilon_{i,t}. \quad (13)$$

As mentioned in the previous section, we will employ the 'in-levels' regressions<sup>18</sup> in two set-ups. First, in section 5.4 we will use its multivariate version:

$$Rel EFP_{i,t} = \gamma_{0,t} + \gamma_{1,t} \times D_{i,t}^{(Ch_1)} + \dots + \gamma_{m,t} \times D_{i,t}^{(Ch_m)} + \varepsilon_{i,t}, \quad (14)$$

( $Ch_i, i = 1, 2, \dots, m$  are  $m$  different firm characteristics) to run a horse race (on the overall sample) between variables competing to explain the relative association measure: intangible intensity, correlation between cash flows and accruals, cash flow or accrual volatility.

Second, when investigating if the decline in the timing role of accruals might explain the evolution of the price association of earnings relative to cash flows (section 5.4.2), we will estimate *cross-sectionally* the regression (13) for  $Ch = cor(CFO, TACC)$  and look at the coefficient  $\gamma_{0,t}$  as a function of time. A statistically constant (through time)  $\gamma$  is evidence that the decline of earnings pertinence to price formation relative to cash flows is associated with the decline of the magnitude of the negative correlation of cash flows and accruals.

---

<sup>18</sup>For robustness, we estimate and report results for the quantile regression version of the 'in-level' regressions.

## 4 Variables definition and sample

### 4.1 Variables definition

We define earnings ( $NI$ ) as income before extraordinary items (Compustat  $IB$ ). Following Bushman et al (2016) we use the balance sheet approach to estimate total accruals before 1988<sup>19</sup>. Specifically, we defined current accruals ( $CACC$ ) as  $CACC_{i,t} = (\Delta ACT_{i,t} - \Delta LCT_{i,t}) - (\Delta CHE_{i,t} - \Delta DLC_{i,t})$ , where the four terms are changes in current assets (Compustat  $ACT$ ), current liabilities (Compustat  $LCT$ ), cash and short-term investments (Compustat  $CHE$ ), and total debt in current liabilities (Compustat  $DLC$ ) in the fiscal year ending at time  $t$ , respectively. Total accruals ( $TACC$ ) are equal to current accruals minus depreciation and amortization (Compustat  $DP$ ). Operating cash flows ( $CFO$ ) are equal to earnings minus total accruals.

To address the fact that the balance sheetbased accruals suffer from measurement errors (Hribar and Collins (2002)), after 1988, we construct the total accruals from the statement of cash flows as follows. Cash flows ( $CFO$ ) are cash flows from operations as disclosed on the statement of cash flows. Total accruals ( $TACC$ ) are defined as earnings minus cash flows.

The number of common shares outstanding are equal to the Compustat variable  $CSHO$ . The share price is equal to the monthly closing price (Compustat  $PRCCM$ ) collected at the beginning of the fourth month after the fiscal year end. The key variables are defined as follows.

**SG&A intensity.** Following Dichev and Tang (2008) and Srivastava (2014), we compute total expenses by subtracting income before extraordinary items (Com-

---

<sup>19</sup>Firms were not required to disclose the statement of cash flows until the promulgation of SFAS No. 95 in 1987.

pustat *IB*) from revenues (Compustat *SALE*). We measure the *SG&A* expenses by the Compustat data item *XSGA*, and define the *SG&A* intensity as the proportion the *SG&A* expenses represent in the firm-year's total expenses, that is, as the ratio between *XSGA* and total expenses.

**Earnings, cash flow and accruals volatility.** Following Givoly and Hayn (2000) and Dichev and Tang (2009) we scale earnings and their components by the average of the beginning and ending of the year total assets and estimate the standard deviations of these variables for each firm-year using at least four values<sup>20</sup> of the eight most recent annual observations ( $t - 7$  through  $t$ ).

**The timing role of accruals.** We estimate the correlation coefficient:

$$\text{corr}(CFO_{i,t}, TACC_{i,t}),$$

for each firm-year using at least four values<sup>21</sup> of the eight most recent annual observations ( $t - 7$  through  $t$ ). (All variables are scaled by the average of the beginning and ending of the year total assets.)

**Earnings Expectation Formation Pertinence.** To construct the firm-year measure of pertinence we first consistently estimate the regression in (4) for each industry  $I$  and year  $t$  in the sample. We use firm's size (as measured by market value) as a second proxy of risk (industry membership being the other), as well as the  $P/B$  ratio as measures of the level of conservatism. The growth proxy is mean earnings growth (both calculated on firm's most recent 8-year history,  $t - 7$  through  $t$ ). The *absolute relative valuation errors*<sup>22</sup> defined in (5) are our measures of the expectation formation pertinence of earnings and cash flows, respec-

<sup>20</sup>The sample contains firm-year with available data for at least 4 of the last 8 most recent years.

<sup>21</sup>The sample contains firm-year with available data for at least 4 of the last 8 most recent years.

<sup>22</sup>Our results are robust to alternative proxying of size by sales or total assets, of growth by mean sales growth, or of the level of conservatism by the conservatism index in Penman and Zhang (2002).

tively. The non-linear valuation regression is estimated on sub-samples defined by the intersection of year and industry (Fama-French 12 industries). The results reported in the paper were obtained using the Random Forest estimation approach (Breiman (2001), James et al (2014)).

## 4.2 Sample

The sample, obtained from Compustat (accounting information) and CRSP (prices) data bases, covers the 55-year period between 1962 and 2016 and contains all firm-year observations with non-missing values of the variables needed to define the valuation errors (5) as well as the three fundamental firm's characteristics (intangible intensity, accrual volatility, and the level of the correlation between cash flows and accruals).

Variable	Mean	SD	10%	25%	50%	75%	90%
<b>Panel A: EFP measures</b>							
$EFP^{NI}$	0.49	1.04	0.04	0.10	0.23	0.47	1.08
$EFP^{CFO}$	0.50	1.00	0.05	0.11	0.24	0.49	1.11
$EFP^{CFO}/EFP^{NI}$	1.06	3.14	0.27	0.58	1.04	1.94	4.35
<b>Panel B: Firm's explanatory characteristics</b>							
$\sigma(NI)$	0.09	0.12	0.01	0.02	0.04	0.09	0.20
$\sigma(CFO)$	0.10	0.29	0.02	0.04	0.06	0.10	0.18
$\sigma(NI)/\sigma(CFO)$	0.76	2.03	0.30	0.47	0.78	1.22	1.89
$SG\&A$	0.27	0.18	0.08	0.14	0.23	0.36	0.53
$\sigma(TACC)$	0.10	0.43	0.02	0.04	0.06	0.11	0.19
$\rho(CFO, TACC)$	-0.56	0.47	-0.97	-0.92	-0.75	-0.34	0.18

Table 2: **Descriptive statistics.** *Panel A:* EFP measures. *Panel B:* Firm characteristics.

It consists of 189,107 firm-year observations and contains 16,834 distinct firms. For calculating volatilities and correlations, we required that the firm-year

had at least four non-missing values in the eight most recent years. Following the prior literature, we exclude financial firms (SIC two-digit code from 60 through 69).

Table 2 displays the summary statistics of the EFP measures (Panel A) and the explanatory firm characteristics (Panel B). The variables were winsorized at the 1% level.

	$EFP^{CFO}/EFP^{NI}$	$\sigma(NI)/\sigma(CFO)$	SG&A	$\sigma(TACC)$	$\rho(CFO, TACC)$
$EFP^{CFO}/EFP^{NI}$	1.00	-0.12	-0.06	-0.04	-0.13
$\sigma(NI)/\sigma(CFO)$	-0.11	1.00	0.20	0.18	0.90
SG&A	-0.05	0.21	1.00	0.23	0.22
$\sigma(TACC)$	-0.04	0.25	0.25	1.00	0.07
$\rho(CFO, TACC)$	-0.10	0.75	0.27	0.19	1.00

Table 3: **Correlation between regression variables (EFP relative measure and explanatory characteristics).** *Lower triangle:* Pearson correlation, *Upper triangle:* Spearman correlation.

Table 3 displays the values of Spearman (upper triangle) and Pearson (lower triangle) correlation between the relevant variables (EFP relative measure, relative volatility, and explanatory firm characteristics).

## 5 Empirical results

This section presents the empirical findings of the analysis of the EFP of earnings relative to cash flows. The conclusions are based on an in-depths analysis of the distributions of the valuation errors

$$|P_i - \widehat{P}_{i,t}^{NI}|/P_i \quad \text{and} \quad |P_i - \widehat{P}_{i,t}^{CFO}|/P_i \quad (15)$$

obtained from the consistent (yearly and industry-specific) inference of the non-linear regression specification of the RI model :

$$P_{i,t} = \mathbf{m}_I^{\mathfrak{X}}(\mathfrak{X}; (SALES.g)_{i,t}, Mktv_{i,t-1}, (P/B)_{i,t-1}) + \varepsilon_{i,t}, \quad (16)$$

where the function  $\mathbf{m}_I^{\mathfrak{X}}$  is specific to each predictor set and is estimated cross-sectionally and for each industry  $I$ . The symbol  $\mathfrak{X}$  stands for  $\{NI\}$  and  $\{CFO\}$ , respectively. The sales growth ( $SALES.g$ ) is a proxy for future growth of income, market value ( $Mktv$ ) (together with industry) proxies for risk while the price to book ratio serves as a measure of the level of conservative accounting  $C_i$  (Pae et al (2005))<sup>23</sup>.

## 5.1 Illustration of the stochastic dominance concept

Since the notion of stochastic dominance is relatively new to the accounting literature we decided to illustrate it with a few graphs which we hope will help the reader further her intuitive grip on the notion. This section displays the estimated cdfs of the errors of the two competing valuations in (15) conditional on year and the level of correlation between CFO and TACC. Statistical test are performed in the next section.

Figure 1 displays the cdfs of the valuation errors in (15) for the predictors  $CFO$  (dotted, blue) and  $NI$  (full, black), respectively, for the years 1968, 1980, 1990, and 2009 (logarithmic scale). The graphs illustrate the notion of stochastic dominance with one cdf being larger than the other. The first and the last graphs show that the order of the curves has inverted between the beginning and the end of the sample. While in 1968 (as well as 1980), the errors corresponding to

---

<sup>23</sup>In an unreported analysis the later was proxied by . The results are qualitatively equivalent.

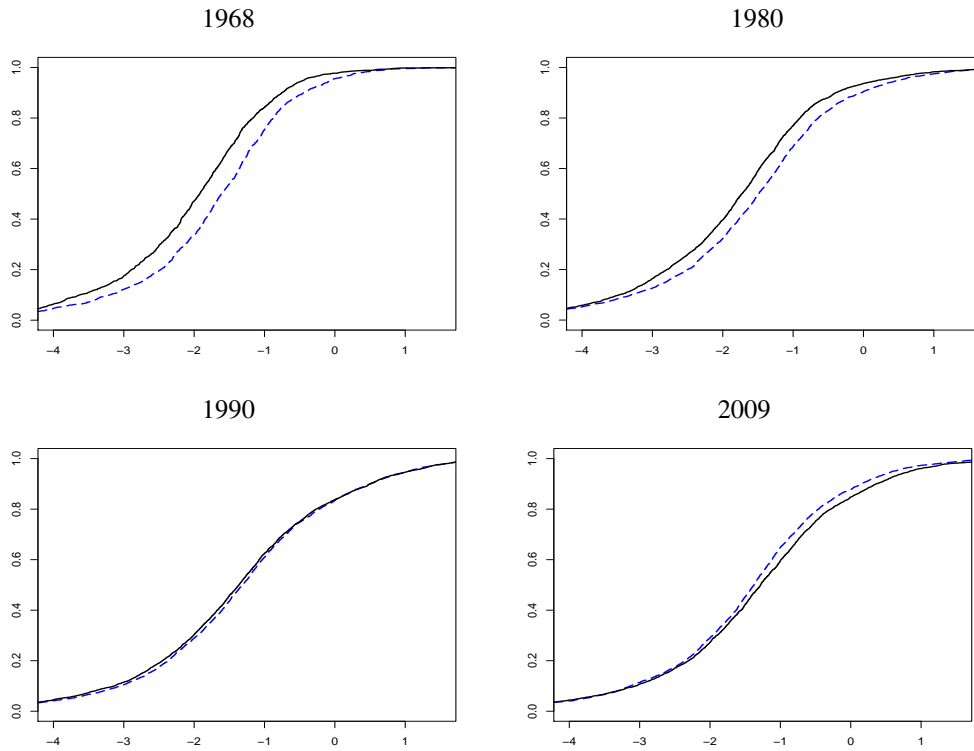


Figure 1: **EFP of NI relative to CFO conditional of year - cdfs.** The plots display the cdfs of the valuation errors in (15) for the predictor set  $\{CFO\}$  dotted (blue) and for  $\{NI\}$  full (black), respectively, conditional on the years specified above each graph (logarithmic scale). In the first two graphs, the valuation informed by earnings stochastically dominates the one informed by cash flows. Cash flows are more informative than earnings in the last graph while the two measures are equally informative in the third graph. For the statistical significance of the differences, see Figure 3.

a valuation informed by earnings stochastically dominated those of a valuation informed by earnings, the relation reversed in 2009. In this year, cash flows seem to inform the price formation more than earnings. 1990 is an example of an year when the two cdfs were practically identical. Results of statistical tests of the significance of the differences between the curves in the graphs are presented in Figure 3.

Figure 2 displays the cdfs of the valuation errors in (15) for the predictors  $CFO$  (dotted, blue) and  $NI$  (full, black), respectively, conditional on the extreme values of the correlation between earnings components. On the left, the correlation is

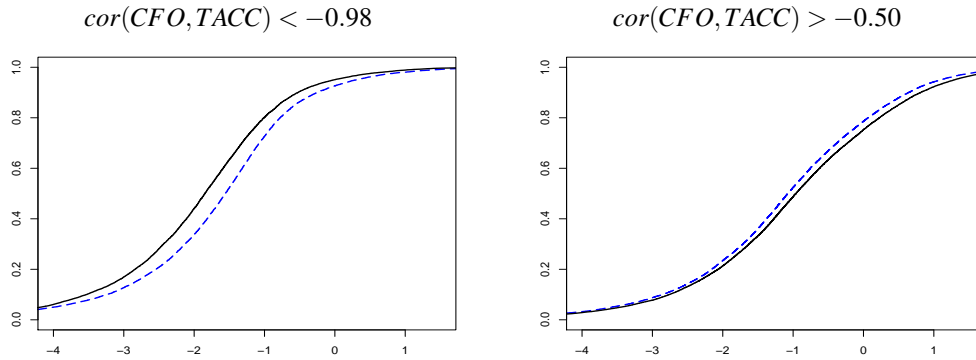


Figure 2: **EFP of NI relative to CFO conditional of the level of  $cor(CFO, TACC)$  - cdfs.** The plots display the cdfs of the valuation errors in (15) for the predictor set  $\{CFO\}$  (*dotted, blue*), and for  $\{NI\}$  (*full, black*), respectively, conditional on the level of  $cor(CFO, TACC)$  specified above each graph (logarithmic scale). The valuation informed by earnings stochastically dominates the one informed by cash flows for more negative correlation. Cash flows become at least as informative as earnings (if not more) for less negative correlation. For the statistical significance of the differences, see Figure 5.

smaller than the first quartile (0.98) (strongly negative), while on the right it is larger than the third quartile (-0.5) (logarithmic scale). While for strongly negative correlation of CFO with TACC the information content of earnings seems larger than that of cash flows, the order reverses for less negative correlations. Results of statistical tests of the significance of the differences between the curves in the graphs are presented in Figure 5.

## 5.2 Evolution of earnings price association relative to cash flows

Figure 3 shows the time evolution of the  $KS$  statistics between 1962 and 2016. The plot displays the  $\overline{KS}$  (*full, black*) and  $\underline{KS}$  (*dashed, blue*) statistics in (11) and (12), respectively, for the pair  $(NI, CFO)$ , on the y-axis, and time, on the x-axis. It also shows the two 95%-confidence bounds: the upper one (*dotted, red*) for the  $\overline{KS}$  and the lower one (*dotted, magenta*) for the  $\underline{KS}$ .

The figure presents the results of a large number of hypothesis tests (one for each cross-section). A yearly value of the  $\overline{KS}_t$  ( $\underline{KS}_t$ ) above (below) the upper

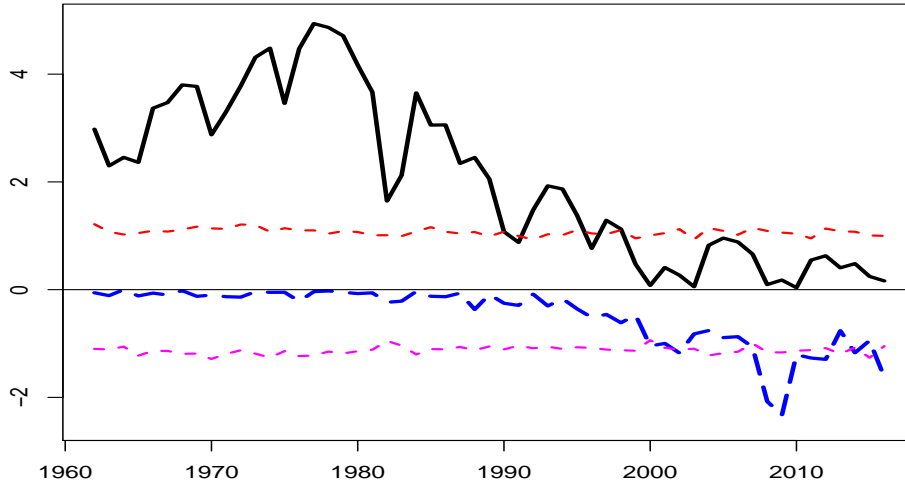


Figure 3: **The time evolution of EFP -  $KS$  statistics.** The plot displays the  $\overline{KS}$  (full, black) and  $\underline{KS}$  (dashed, blue) statistics in (11) and (12), respectively, for the pair  $(NI, CFO)$  as a function of time. Values of  $\overline{KS}$  above the upper 95%-confidence bound (dotted, red) indicate that the valuation informed by earnings levels stochastically dominates the one informed only by cash flows. Values of  $\underline{KS}$  smaller than the lower 95%-confidence bound (dotted, magenta) are significant evidence that cash flows inform the price setting more than earnings. Values of the two statistics inside the confidence band mark the years when the two accounting variables inform the price formation process to the same extent.

(lower) confidence bound rejects the hypothesis (8) of equal pertinence of NI and CFO (for the cross-section  $t$ ) for an alternative (9) that NI (CFO) are more pertinent (to the price formation) than CFO (NI). Values the  $\overline{KS}_t$  and  $\underline{KS}_t$  statistics between the two confidence bounds do not reject the null hypothesis (8) (for the cross-section  $t$ ).

Up to (around) year 2000,  $\overline{KS}$  lies above the upper 95%-confidence bound. This is (statically-significant) evidence that, in the beginning of the sample, valuations informed by earnings levels stochastically dominated the ones informed only by cash flows (although the significance diminishes strongly from mid-70s). Between 2000 and 2007, the two statistics are (mostly) within the confidence band indicating that, during this period, the two accounting variables have informed the

price formation process to the same extent. Between 2008 and 2011, the  $KS$  is (mostly) significantly smaller than the lower 95%-confidence bound. This is statistically-significant evidence that during and immediately after the subprime financial crisis the cash flows informed the price setting more than earnings.

### **5.3 Possible explanations**

Having documented the pronounced decline in the overall price association of earnings relative to cash flows, we explore potential reasons for this attenuation in two steps. In the next section we postulate and document that the pertinence reduction might be associated with changes in the relation between the smoothness of the two accounting variables.

In Section 5.3.2 we investigate different economic and accounting developments that could be responsible for the pronounced increase in the volatility of earnings relative to that of cash flows we document in section 5.3.1.

#### **5.3.1 Earnings volatility vs. cash flow volatility**

In this section we first theorize and then document that changes in the relation between the smoothness of the two accounting variables might be responsible for the pertinence reduction illustrated in the previous section.

We begin by documenting a consequential change in the magnitude of the volatility of earnings relative to that of cash flows. The graph on the left-hand side of Figure 4 displays the volatility of earnings (full line) and of cash flow (dotted line) as functions of time. It highlights a very different evolution of the variability of the two accounting measures. After a slight increase towards the end of the 80s, the volatility of cash flows decreased to roughly the same level it had in

the beginning of the sample. In contrast to this, cross-sectional median earnings volatility increased three fold between the the 60s and year 2000 and went down only slightly afterwards. As a result of these different dynamics, after year 2000, the cross-sectional median volatility of cash flows has been equal or smaller than that of earnings.

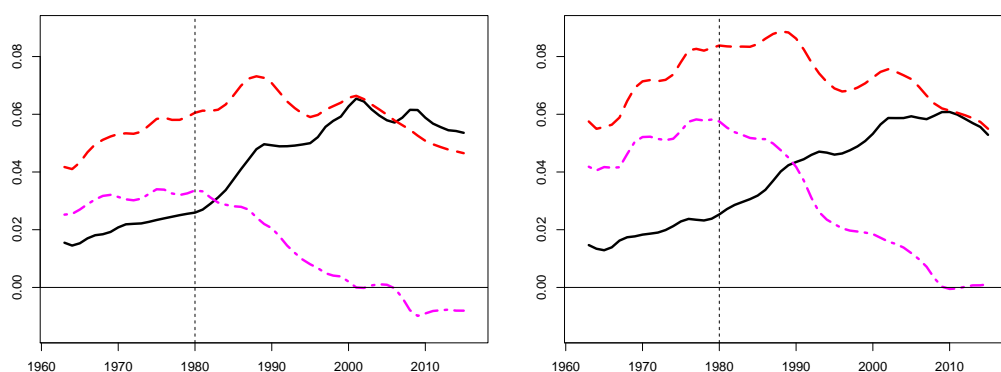


Figure 4: **Evolution of earnings and cash flow volatilities.** The graphs display the cross-sectional median volatility of earnings (left graph) and change in earnings (right graph), respectively (*full, black*), of cash flows and change in cash flows, respectively (*long dashed, red*), and of their difference (*dashed dotted, magenta*). After 1980, the median difference has decreased steadily. The cash flows have been smoother than earnings (for more than half of the firms in the cross-sample) for the last ten years of the sample.

To better see this evolution, the dash-dotted (magenta) curve shows the difference between the volatilities of cash flow and earnings. After a plateau until around 1980, the difference diminishes swiftly becoming 0 around year 2000 and strictly negative in the last ten years of the sample.

The graph in the right-hand side of the figure displays the volatility of changes in earnings, that of changes in cash flows and their difference. The evolution of the two graphs are very similar.

Next, we postulate that the pertinence reduction documented in section 5.2 is, possibly, associated with changes in the relation between the smoothness of the two accounting variables illustrated by Figure 4.

Our explanation is based on the extensive literature on earnings and cash flow forecasting and is consistent with previous evidence on the relation between smoothness and predictability (Dichev and Tang (2009)). Recent accounting research shows that cross-sectional ordinary least squares regressions provide relatively well-behaved forecasts of both the level of earnings and cash flows (Hou et al (2012), Gerakos and Gramacy (2013), Nallareddy et al (2018)). In particular, Gerakos and Gramacy (2013) provides a thorough examination of regression-based out-of-sample forecasts of net income using both linear and non-linear approaches. It documents that 'predictors generated using ordinary least squares and small sets of explanatory variables (often just lagged net income) lead to more accurate forecasts' than other more sophisticated approaches and large predictor sets. Their findings can be summarized by one sentence: 'Random walk and AR(1) are hard to beat'.

Brown (1993) overviews the accounting literature taking a time series perspective on forecasting earnings and cash flows and concludes that 'the random walk with drift, IMA(1,1) and AR(1) models are reasonable characterizations of the annual earnings time-series process'. In a out-of-sample forecasting comparison between random walk and an AR(2) prediction, Finger (1994)<sup>24</sup> reports that 'random walk models outperform individually estimated earnings models for one-year out-of-sample forecasts' while the AR(2) model is slightly better for four- and eight-year forecasting horizon. Moreover, 'cash flow is a better short-term predictor of cash flow than are earnings, both in and out of sample, and the two are approximately equivalent long-term'.

To summarize, a certain consensus seems to exist in the literature respecting the fact that random walks or very simple linear models (say, with one or two

---

<sup>24</sup>The study uses linear models with up to eight lags of the explanatory variables.

lags) provide reasonable forecast mechanisms<sup>25</sup> for both earnings and cash flows. Or, the prediction accuracy of such models is fundamentally determined by the volatility of the variable to forecast, in the case of the linear model, or of the period-to-period changes in this variable, for the random walk, respectively.

More precisely, in the setting of linear model with two lags and assuming local stationary<sup>26</sup>:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t,$$

the mean square one-step-ahead prediction error is given by

$$MSE_2^{(1)} = \left( 1 - \frac{1 + \phi_2}{1 - \phi_2} \phi_1^2 + \phi_2^2 \right) \times Var(Y).$$

The expression becomes:

$$MSE_1^{(1)} = (1 - \phi_1^2) \times Var(Y),$$

for an linear model with one lagged predictor. More generally, for such a model, the mean square  $h$ -step-ahead prediction error ( $h > 1$ ) is given by:

$$MSE_1^{(h)} = (1 - \phi_1^2) (1 + \phi_1^2 + \phi_1^4 + \dots + \phi_1^{2(h-1)}) \times Var(Y).$$

For a random walk,

$$Y_t = \phi_0 + Y_{t-1} + \varepsilon_t,$$

the mean square one-step-ahead prediction error is the variance of the period-to-

---

<sup>25</sup>One of main findings of Gerakos and Gramacy (2013) is that historically variables can show high (and statistically significant) correlations with net income. However, such correlations do not necessarily imply that the historical relationships (of these variables with net income) contributes to more accurate forecasts of future net income.

<sup>26</sup>Local stationary implies that the moments of the variable  $Y$  do not change between  $t - 2$  and  $t$ .

period changes:

$$MSE_{RW} = Var(\Delta Y) = Var(\varepsilon),$$

while the mean square  $h$ -step-ahead prediction error ( $h > 1$ ) is given by:

$$MSE_{RW}^{(h)} = h^2 \times Var(\Delta Y).$$

It follows that higher earnings or cash flow (or changes in earnings and changes in cash flow) volatility translates into less precise expectations about future income and hence decreased price formation relevance of the concerned variable. In other words, the increase in earnings (change in earnings) volatility relative to cash flows (change in cash flows) illustrated in Figure 4 should reduce investors' relative ability to project a firm's future performance using the current level of net income (Barton et al (2010)), reducing consequently earnings relative pertinence to price formation.

To empirically verify the fit of our explanation, we define a relative measure of variability as the *ratio between earnings and cash flow volatility* and document a strong increasing relation between the size of this measure and the price pertinence of earnings relative to cash flows as measured by the  $KS$  statistics. Since the results when using changes in the variability of the two performance measure are similar, we do not report them.

Figure 5 brings evidence that low (high) values of the ratio of earnings volatility to cash flows volatility are associated with high (low) relative information content of earnings relative to cash flows, reflected in high (low) values of the  $\overline{KS}$  ( $KS$ ) statistic. It displays the  $KS$  statistics *conditional on the level* of the ratio of volatility of earnings on cash flow volatility. It shows that, for the largest four deciles of the ratio (ratios greater than 0.86), the cash flows are at least as associated to prices as earnings while for the largest three deciles (ratios greater than 1),

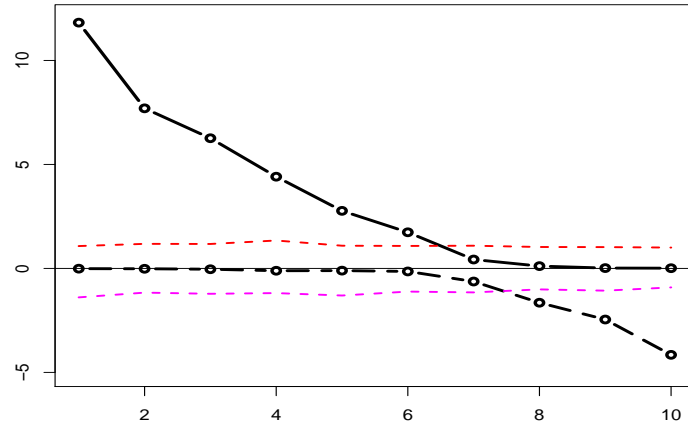


Figure 5: **EFP of NI relative to CFO conditional of the level of the ratio  $\sigma(NI)/\sigma(CFO)$ .** The plot displays the  $\overline{KS}$  (full, black) and  $\underline{KS}$  (dashed, blue) statistics in (11) and (12), respectively, for the pair  $(NI, CFO)$  as a function of the deciles of the volatility ratio. Values of  $\overline{KS}$  ( $\underline{KS}$ ) above (below) the upper (lower) 95%-confidence bound indicate that, for firms with relative volatility magnitudes in the decile indicated on the  $x$ -axis, the valuation informed by earnings levels stochastically dominates (is stochastically dominated by) the one informed only by cash flows.

they are strictly more informative than earnings.

Next, we investigate what happens to the proportion of firms in the cross-section for which the cash flows are at least as pertinent as earnings. Figure 6 displays the time evolution of the proportion of firms in the cross-section with the ratio between earnings and cash flow volatilities belonging to the deciles for which cash flows are *at least as informative as* earnings, i.e., 7 to 10, (full, black curve) and *strictly more informative* than earnings, i.e, 8 to 10 (dotted, red curve). The proportion of these firms in cross-section increased with time, particularly after 1980, possibly explaining the dynamics of the  $KS$  statistics in Figure 3. The proportion rises from 10 percent in the 60s to around 65 percent in 2016. In other words, for two thirds of the firms in the recent cross-sections, cash flows are at least as informative as earnings. For around 50 percent of the firms in these cross-sections, cash flows are *strictly* more price formation pertinent than earnings.

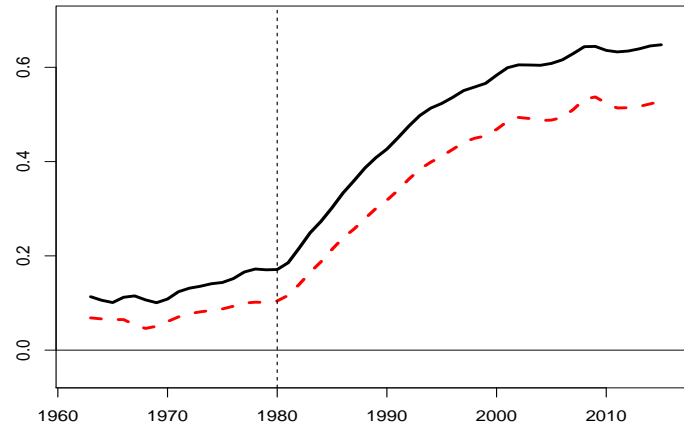


Figure 6: **Proportion of the sample with relative magnitude of NI and CFO volatilities associated with equal (or lower) pertinence of NI relative to CFO.** The plot displays the time evolution of the proportion of firms in the sample with a ratio of NI volatility to CFO volatility belonging to the deciles for which cash flows are *at least as informative as* earnings (deciles 7 to 10) (upper, full, black) and *strictly more informative than* earnings (deciles 8 to 10) (the second higher, long dashed, red), respectively). For two thirds of the firms in the recent cross-sections, cash flows are at least as informative as earnings, while, for around 50 percent of the firms, cash flows are *strictly* more price formation pertinent than earnings.

### 5.3.2 Mechanisms of change

The pronounced increase in the volatility of earnings relative to that of cash flows could be associated to a number of different economic and accounting developments.

**Growing prominence of firms with high intangible intensity.** The increase of the intangible intensity significantly alters the characteristics of earnings and their components, cash flow and accruals. For example, the future benefits of intangible investments are more uncertain than those of tangible investments (Kothari et al (2002)), lessening the informational content of current earnings relative to cash flows. Moreover, the immediate expensing of investments in (internally generated) intangibles (such as patents, trade names, human capital, and customer relations) upon the outflow of cash should increase cash flows ability to predict

future cash flows, as these investments do not generate accruals (Bushman et al (2016)). Growing investments in intangible assets should increase the volatility in expenses and reduce the matching between concurrent revenues and expenses. As a result, earnings volatility should increase.

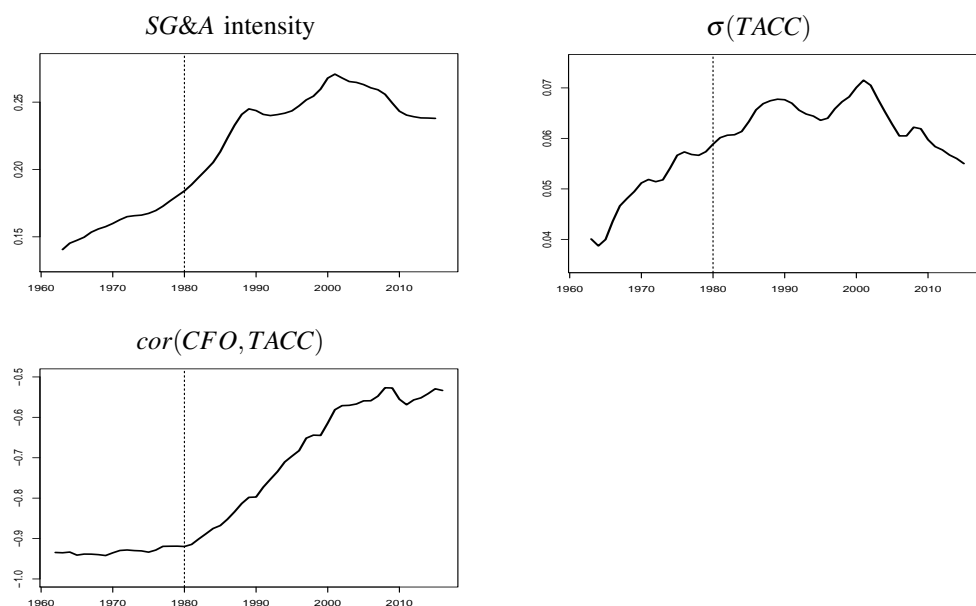


Figure 7: **Evolution of firm characteristics that potentially explain the decrease in earnings price association relative to cash flows.** *Top, left:* the cross-sectional median volatility of earnings (full, black), cash flows (long dashed, red), and total accruals (dotted, blue). *Top, right:* the cross-sectional median of the average SG&A intensity. *Bottom:* the cross-sectional median  $cor(CFO, TACC)$ .

The top-left graph in Figure 7 shows a double fold increase in the cross-sectional median mean intangible intensity from around 0.12 in the beginning of the 60s to 0.27 around year 2000.

**Increase in the volatility of accruals.** The top-right graph in Figure 7 displays the evolution of cross-sectional median of accruals volatility which, in our sample, increased from 4% in the beginning of the sample to around 6% around year 2000. Givoly and Hayn (2000) finds that the greater accrual variability is due primarily to the rising incidence of one-time items and non-operating accruals (as documented by Elliot and Shaw (1988), Collins et al. (1997) and Givoly

and Hayn (1994), Bradshaw and Sloan (2002)). The magnitudes of both of these categories have increased drastically over the time period examined. More important, the cross-sectional volatility of special (Compustat SPI) and non-operating items (Compustat NOPI) has also experienced a pronounced increase from about 1% to 15% of total assets (Bushman et al (2016)).

**Reduction of the timing role of accruals.** To perform their smoothing or noise reduction role, accruals need to be negatively correlated to cash flows (Dechow (1994)). Or, the magnitude of the correlation between cash flows and accruals has been increasing steadily over the years (Bushman et al (2016)). As a result, the positive difference between the cash volatility and the earnings volatility should diminish and, possibly, become negative. The bottom-left graph in Figure 7 displays the time evolution of the cross-sectional median correlation between cash flows and accruals and shows a dramatic increase from around -0.93 in the beginning of the sample to -0.55 in the end of the sample.

## 5.4 Idiosyncratic relative price association

The next two sections analyze which of the three mechanisms above is the most likely reason behind the decrease of earnings EFP relative to cash flows. The investigation articulates in two steps: a regression analysis of the all year-firm observations in the sample and a cross-sectional analysis. The statistical set-up is that described in Section 3.4. Most importantly, in this two sections we no longer quantify the relative pertinence through the *KS* statistics, a cross-sectional measure, but through the ratio in (6), a year-firm specific measure, suited for multivariate regression analyses.

#### 5.4.1 Static analysis: all-sample regressions

The investigation in this section shows that the association of two of the competing (explanatory) firm characteristics, that is, accrual volatility and *SG&A* intensity, with the relative pertinence measure, although significant in a regression on *only* the levels of individual variables, disappears once we control for the level of the timing effect of accruals. The significant coefficients of the individual levels in the simple regression are due to the omitted variable bias.

We pool all cross-sections in the sample and run, first, four univariate regressions (13), one for the ratio of earnings to cash flows volatilities,

$$Ch = \sigma(NI)/\sigma(CFO),$$

the variable the analysis in section 5.3.1 ties to the evolution of the relative EFP, and one for each of the three competing explanatory firm characteristics discussed in the previous section, i.e.,

$$Ch = \text{SG\&A intensity}, \quad Ch = \sigma(TACC), \quad \text{and} \quad Ch = \text{cor}(CFO, TACC). \quad (17)$$

The estimated coefficients  $\gamma^{(u)}$  (uncontrolled) reflect the association without controlling for possible bias due to omitted characteristics.

Second, we ran a multivariate regression (14) with the firm characteristics (17) as independent variables. Note that the pair  $(\sigma(TACC), \text{cor}(CFO, TACC))$  represents a decomposition<sup>27</sup> of the explanatory variable  $\sigma(NI)/\sigma(CFO)$  in section 5.3.1. The multivariate regression yields consistent estimates,  $\gamma^{(c)}$  (controlled), of the idiosyncratic association of each of the firm characteristics in (17) to the

---

<sup>27</sup>Holding constant the pair  $(\sigma(TACC), \text{cor}(CFO, TACC))$  results in the ratio  $\sigma(NI)/\sigma(CFO)$  being constant.

relative pertinence measures. It indicates if the firm characteristic is significantly associated to the pertinence measures.

We structured the estimating set-up as follows. First, we divided the range of each of the (continuous) independent variable in ten intervals of equal probability (deciles). Then, we created ten dummy variables indicating that the explaining characteristic belongs to each one of these intervals but one, the reference level.

A coefficient of the  $j$ -th indicator ( $1 \leq j \leq 10$ , reference level excluded) significantly different from zero provides statistical evidence on the difference between the mean EFP measures of firms with values of the firm characteristic in the  $j$ -th interval and the mean EFP measure of firms with value of the characteristic in the reference interval. Significant negative/positive coefficients indicate that the characteristic is associated with the relative EFP measure.

Given the large number of estimated parameters and performed tests, the results of the analysis are presented through graphs instead of tables. For a given firm characteristic, we display three curves. The first one traces the ten coefficients of the indicator variables modeling the levels of the firm characteristic and is shown using a full line. The other two trace the lower and the upper limits of the ten 95%-confidence intervals<sup>28</sup> and are shown as dotted lines.

Such a graph tests ten hypothesis of equal mean EFP measure (that is, association of firm's characteristic with the relative EFP measure), one for each of the levels one to ten of firm's characteristic, the reference level excepted. The tests are performed by checking if 0, the null difference between the two mean EFP

---

<sup>28</sup>The confidence intervals are based on heteroskedasticity-consistent (HC) estimation of the covariance matrix of the coefficient estimates in the regression model. Long and Ervin (2000) conduct a simulation study of HC estimators in the linear regression model (HC0 to HC3; HC0 is Whites estimator (White (1980)) while the other estimators are refinements of this), recommending to use of the so-called HC3 estimator which we used consistently through the analysis. Qualitatively identical results were obtain using alternative HC estimators.

measures, belongs to each of the ten confidence intervals. If 0 does not belong to the confidence interval of the  $j$ -th coefficient, we reject the null.

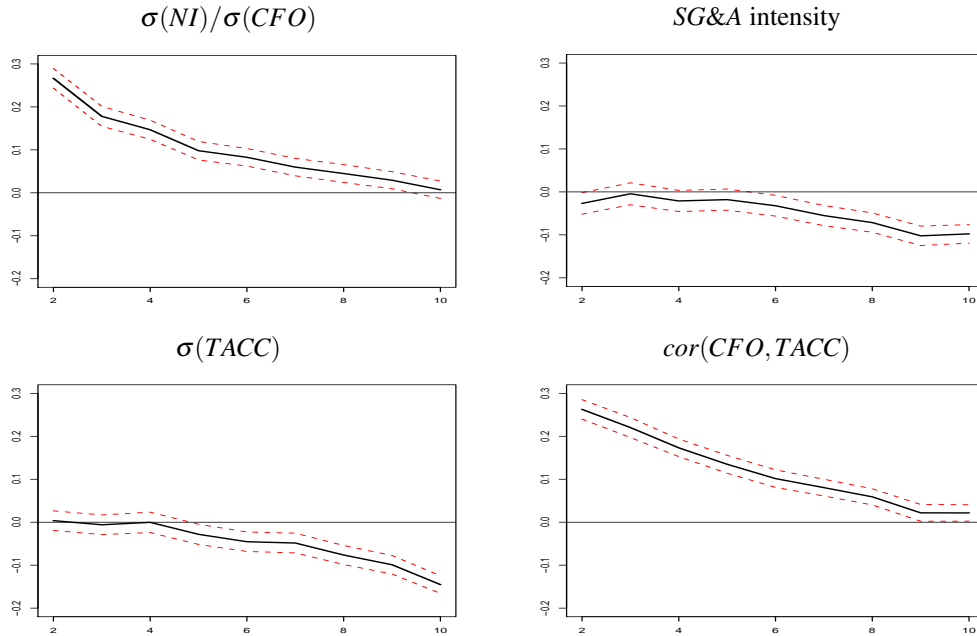


Figure 8: **Firm characteristics that explain the relative EFP measure - simple regressions.** The figure displays the coefficients of the level indicators corresponding to the characteristics named in the title of the graph in the simple regression (13) ( $\gamma^{(u)}$ ) together with their 95% confidence intervals. Each graph presents the results of 10 hypothesis tests. For any coefficient of a level indicators, the value 0 outside the confidence interval is proof of significant association of the firm characteristic to the relative EFP measure. All firm characteristics are associated with the EFP measure albeit with different strength.

Figure 8 presents the results of the four univariate regressions (13). The graphs display the estimated coefficients  $\gamma^{(u)}$  (uncontrolled), with the corresponding 95% confidence intervals for each of the four competing characteristics. The first graph confirms the analysis based on the  $KS$  statistic in section 5.3.1 and shows that the ratio of the two volatilities is also significantly and strongly monotone associated to the year-firm specific measure of price association defined in (6). Earnings of firms with earnings less volatile than cash flows yield valuations that are significantly more accurate than those informed by firm's cash flow level.

The remaining three graphs show significant association of the characteristics in (17) to the year-firm specific EFP measure (although the degree of association varies). Lower levels of accrual volatility, *SG&A* intensity and correlation of cash flows and accruals are associated with earnings that are more informative relative to cash flows. We also see that the correlation of cash flows and accruals has the strongest association to the relative EFP measure.

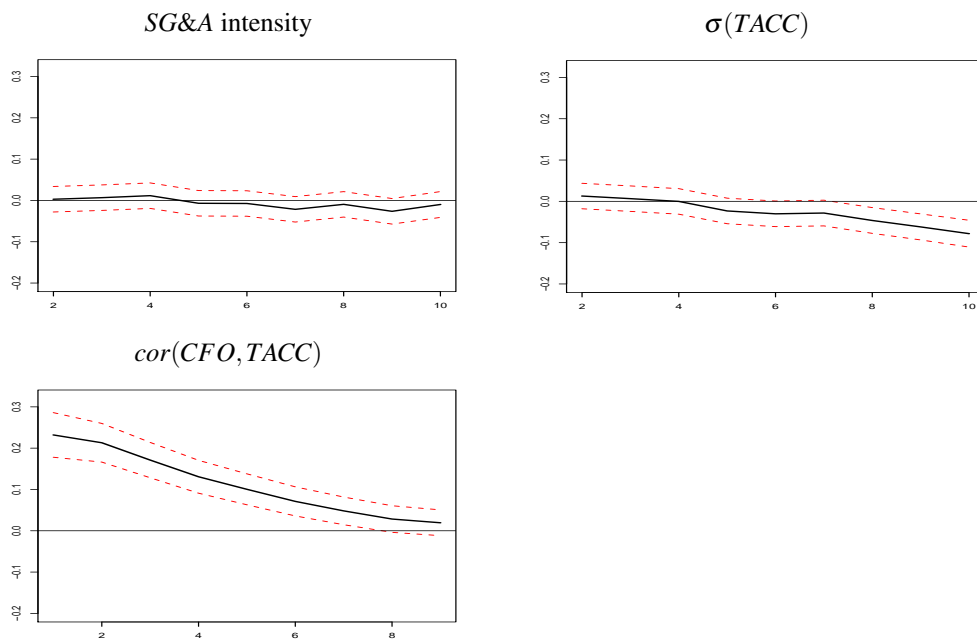


Figure 9: **Firm characteristics that explain the contemporaneous correlation between revenues and expenses - joint analysis.** The figure displays the coefficients of the level indicators corresponding to the characteristics named in the title of the graph in the multiple regression (14) ( $\gamma^{(c)}$ ) together with their 95% confidence intervals. Each graph presents the results of 10 hypothesis tests. For any coefficient of a level indicators, the value 0 outside the confidence interval is proof of significant association of the firm characteristic to the relative EFP measure. We note that the association of all firms characteristics but the measure of the timing role of the accruals are no longer significant.

Figure 9 presents the results of the multivariate regressions (14). The graphs display the idiosyncratic association of each of the competing firm characteristics (17) to the relative pertinence measures as measured by the estimated coefficients  $\gamma^{(c)}$  (controlled), together with the corresponding 95% confidence intervals. We

note that the association of all firms characteristics but the measure of the timing role of the accruals is greatly reduced ( $\sigma(TACC)$ ) or no longer significant (*SG&A* intensity).

Based on Figures 8 and 9 we conclude that a statistically-consistently association exists mainly between the relative pertinence measures and the strength of the smoothing effect of accruals. The association between the relative pertinence measures and accrual volatility or knowledge intensity in Figure 8 is largely spurious and due to omitted variable bias.

#### **5.4.2 Dynamic analysis - cross-sectional regressions**

The analysis in this section brings evidence that the decline in the timing role of accruals explains the evolution of the price association of earnings relative to cash flows. First, we identify a significant downward time trend in the median of the year-firm specific EFP measure (6). If the evolution of this measure is associated with the dynamic of change in the timing role of accruals, controlling for the level of the later, cross-section by cross-section, should remove the evidence of trend.

We document that controlling for the level of the correlation between cash flows and accruals does remove the strong trend in the median year-firm specific pertinence measure (6) during the period when the smoothing role of accruals is significantly decreasing, i.e. after 1980 (see bottom graph in Figure 7).

We want to emphasize that an analysis trying to relate directly the evolution of relative price pertinence to changes in the timing role of accruals by regressing the cross-sectionally-constant measure of relative pertinence in Figure 3 on the correlation measure in the bottom graph of Figure 7 is not feasible. Only stationary variables can be used in such time-series regressions. Granger and Newbold (1974) documents that when a non-stationary variable is regressed on another, the

bias toward rejection of no relationship is proportional to the sample size. Two non-stationary variables may hence seem to be strongly correlated even when, in fact, they are independent. Our use of a firm-year specific relative pertinence measure is essential to the statistical soundness of our conclusions.

First graph in Figure 10 displays the evolution of the cross-sectional median of the logarithm of the year-firm specific EFP measure in (6) together with its 95% confidence interval. Values above (below) 0 identify the cross-sections for which earnings are more (less) informative relative to cash flows. It shows a significant downward trend (particularly after 1980) highlighting the decrease and, finally, vanishing of earnings superiority in informing prices relative to cash flows. It complements the analysis in section 5.2<sup>29</sup> with evidence based on the firm-year specific EFP measure. The the median ratio (6) is bigger than 1 (indicating larger errors for the valuation informed by cash flows relative to one informed by earnings) until roughly 1990, it is equal to one for the cross-sections of the 90s and it is significantly smaller than 1 after year 2000.

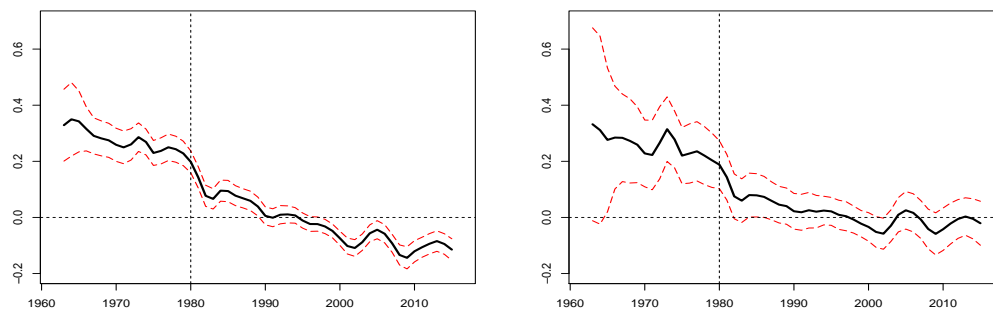


Figure 10: **Evolution of cross-sectional median of log year-firm specific EFP measure (6).** Unconditional (*left*) and conditional of the level of  $cor(CFO, TACC)$  (*right*). The dotted lines display the 95% confidence interval. Values above (below) 0 identify the cross-sections for which earnings are more (less) informative relative to cash flows. As the number of firms in the cross-section with less negative correlations increases after 1980, controlling for the level of accrual smoothing removes the strong negative trend visible in the left-hand graph.

<sup>29</sup>There the analysis focused on the *KS* statistic as a measure of the information gap between earnings and cash flows.

Second graph in Figure 10 displays the evolution of the cross-sectional median of the logarithm of the year-firm specific EFP measure (6) *when keeping the level of the correlation constant*, together with its 95% confidence interval. Formally, we estimate cross-sectionally, the regression<sup>30</sup> (13) with  $Ch = cor(CFO, TACC)$  and display the estimated constant  $\gamma_{0,t}^{(c)}$  and its confidence interval as functions of time. Before 1980, the range of the correlation is rather tight resulting in large confidence intervals. Moreover, the percentage of firms with correlation values that correspond to at least as informative cash flows does not change while the cross-sectional median relative EFP decreases (see Figure 6). Consequently, controlling for a level of accrual smoothing that does not change cannot explain the slight decrease of the dependent variable. The median ratio is significantly larger than 1. However, as the range of the correlation between earnings components expands and, more importantly, as the number of firms in the cross-section with less negative correlations increases, controlling for the level of accrual smoothing removes the strong negative trend visible in the first graph of Figure 10. The resulting function  $t \rightarrow \gamma_{0,t}^{(c)}$  is (statistically) constant for  $t$  after 1981.

## 6 Robustness

The results are robust to a number of alternative specifications. We estimated the non-linear regression yielding the EFP relative measure using other non-linear methods. We used the conservatism index in Penman and Zhang (2002) as an alternative measures of the level of conservatism to the  $P/B$  ratio. We used total assets as an alternative to the market value as a proxy for risk. We varied the

---

<sup>30</sup>For a robust estimation, we use the quantile version of the linear regression. As such, the constant  $\gamma_{0,t}$  is the median ratio for firms in cross-section  $t$  with the characteristic belonging to the reference level of the independent variable, i.e.,  $cor(CFO, TACC)$ . The graphs display the median of this constant over all reference levels and as a function of time.

number of different levels we split the ranges of the firm characteristics. We used *R&D* instead of *SG&A* intensity as a proxy of knowledge intensity. Moreover, we performed the analysis conditional on cohorts (Srivastava (2014)). Our findings are not a cohort effect. The results are qualitatively the same for all cohorts.

## 7 Conclusions

By adjusting cash flows with accruals, accrual accounting attempts to address the timing and matching issues that cause cash flows to be a noisy measure of firm performance. Consequently, accruals are able to perform their noise-dampening role and yield a smoother, more pertinent performance measure to the extent to which they are negatively correlated with cash flows. Accruals that are less negative correlation (or even a positive correlation) with cash flows will not reduce (or will even amplify) the latter's temporary fluctuations.

We document that the relation between the magnitudes of the variability of earnings and cash flows implied by the above view on accruals, i.e. earnings are smoother than cash flows, has inverted through time. While earnings were less volatile in the cross-sections up to until around year 2000, their variability has surpassed that of cash flows for most firms in the recent cross-sections.

Moreover, we argue that the pronounced and continuous decline in the correlation between accruals and cash flows (Bushman et al (2016)) is associated with a significant reduction of the pertinence of earnings to the process of price formation relative to cash flows. After year 2000, cash flows have informed prices at least as much as earnings.

To perform our analysis we introduce an alternative relative information research design which addresses the econometric issues related to miss-specification

of the linear regression when modeling non-linear relations.

The significance of our results is better understood when considered next to the findings in Bushman et al (2016)<sup>31</sup>. In the light of their conclusions which associate the dramatic decline in the relative prominence of the timing role of accruals to economic and reporting developments, our results seem to indicate that “the growth in the frequency and the magnitude of non-timing accruals” over the last five decades have severely limited the incremental information content of earnings and, consequently, significantly weakened the pertinence of accrual accounting to prices.

Hence, we give an affirmative answer to the question posed in Bushman et al (2016): “Have earnings become less meaningful than before because of the changing nature of the information content of accruals?” For the last decade, the cash flows of as much as two thirds of the firms in the cross-sections informed prices at least as much as earnings. During this period, for roughly half of the firms, cash flows association to prices has been significantly stronger than that of earnings.

## A Cdf estimation

For a given sample of absolute errors  $(err_1^{\mathcal{X}}, err_2^{\mathcal{X}}, \dots, err_{n_{\mathcal{X}}}^{\mathcal{X}})$  of a valuation informed by the predictor set  $\mathcal{X}$ , the *empirical cumulative distribution function*, the estimator of  $F_{\mathcal{X}}$ , is defined as:

$$\widehat{F}_{\mathcal{X}, n_{\mathcal{X}}}(x) := \frac{1}{n_{\mathcal{X}}} \sum_{i=1}^{n_{\mathcal{X}}} I_{(-\infty, x]}(err_i^{\mathcal{X}}) = \frac{\# \text{ of errors } \leq x}{n_{\mathcal{X}}}, \quad (18)$$

---

<sup>31</sup>The paper documents a dramatic decline in the magnitude of the correlation between accruals and cash flows over the past half century.

where  $I_A(x)$  is the indicator function:

$$I_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A. \end{cases}$$

## B Bootstrap of the Kolmogorov-Smirnov statistic for dependent samples

When the two error samples (5) are independent, the asymptotic distribution of the *KS* statistics (11) and (12) are known and do not depend on the marginal distributions of the two error terms. However, when the two error samples are dependent, as it is our case<sup>32</sup>, the two distributions depends on the (unknown) joint distribution of the error pair. Hence, we need to approximate the asymptotic distribution of the two statistics via bootstrap as follows.

To make sure that we draw samples from the joint distribution of errors with a common marginal cdf and a dependency structure that reproduces the one in the sample, from the original sample of pairs of errors,  $(err_i^{\mathcal{X}}, err_i^{\mathcal{Y}})$ , we construct the sample of pairs of ranks  $(R_i^{\mathcal{X}}, R_i^{\mathcal{Y}})$ . We then

1. draw with replacement  $n$  pairs from the sample of ranks,
2. combine the two samples  $(err_1^{\mathcal{X}}, err_2^{\mathcal{X}}, \dots, err_n^{\mathcal{X}})$  and  $(err_1^{\mathcal{Y}}, err_2^{\mathcal{Y}}, \dots, err_n^{\mathcal{Y}})$  and for each pair  $(R_i^{\mathcal{X}}, R_i^{\mathcal{Y}})$  in the simulated sample we construct the pair of errors from the union which have ranks equal to  $R_i^{\mathcal{X}}$  and  $R_i^{\mathcal{Y}}$ , respectively; in this way, we generate pairs of errors with the same dependency structure (copula) as in our sample but with a common marginal distribution;

---

<sup>32</sup>In our case the observations in the two samples arrive in pairs and they are strongly dependent. The Spearman correlation between the relative pricing errors is 0.59.

3. calculate the Kolmogorov-Smirnov statistics (11) and (12) for the simulated sample of error pairs.

By repeating a large number of times this sequence of operations, we approximate the sampling distribution of the statistics of interest (see, e.g., Efron and Tibshirani (1994)) under the null hypothesis of equal cdf of the error terms when the data arrives as dependent pairs.

## References

- Achen CH (2005) Let's put garbage-can regressions and garbage-can probits where they belong. *Conflict Management and Peace Science* 22(4):327–339
- Ali A (1994) The incremental information content of earnings, working capital from operations, and cash flows. *Journal of Accounting Research* 32(1):61–74
- Ball R, Brown P (1968) An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6(2):159–178
- Barth M, Clinch G (2009) Scale effects in capital markets-based accounting research. *Journal of Business Finance and Accounting* 36(3-4):253–288
- Barton J, Hansen TB, Pownall G (2010) Which performance measures do investors around the world value the most and why? *The Accounting Review* 85(3):753–789
- Bawa VS (1982) Stochastic dominance: A research bibliography. *Management Science* 28(6):698–712
- Biddle GC, Seow GS, Siegel AF (1995) Relative versus incremental information content. *Contemporary Accounting Research* 12(1):1–23

- Biddle GC, Bowen RM, Wallace JS (1997) Does EVA beat earnings? Evidence on associations with stock returns and firm values. *Journal of Accounting and Economics* 24(3):301 – 336
- Bradshaw MT, Sloan RG (2002) GAAP versus The Street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research* 40(1):41–66
- Breiman L (2001) Random forests. *Machine Learning* 45(1):5–32
- Brown LD (1993) Earnings forecasting research: its implications for capital markets research. *International Journal of Forecasting* 9(3):295 – 320
- Burgstahler DC, Dichev ID (1997) Earnings, adaptation and equity value. *The Accounting Review* 72(2):187–215
- Bushman RM, Lerman A, Zhang XF (2016) The changing landscape of accrual accounting. *Journal of Accounting Research* 54:41–78
- Cheng CSA, Hopwood WS, McKeown JC (1992) Non-linearity and specification problems in unexpected earnings response regression model. *The Accounting Review* 67(3):579–598
- Collins DW, Kothari S (1989) An analysis of intertemporal and cross-sectional determinants of earnings response coefficients. *Journal of Accounting and Economics* 11(23):143 – 181
- Das S, Lev B (1994) Nonlinearity in the returns-earnings relation: Tests of alternative specifications and explanations. *Contemporary Accounting Research* 11(1):353–379

- Dechow PM (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 PM, Kothari S, Watts RL (1998) The relation between earnings and cash flows. *Journal of Accounting and Economics* 25(2):133 – 168
- Dichev ID, Tang VW (2008) Matching and the changing properties of accounting earnings over the last 40 years. *The Accounting Review* 83(6):1425–1460
- Dichev ID, Tang VW (2009) Earnings volatility and earnings predictability. *Journal of Accounting and Economics* 47(1):160–181
- Easton PD, Harris TS (1991) Earnings as an explanatory variable for returns. *Journal of Accounting Research* 29(1):19–36
- Edwards E, Bell P (1961) *The Theory and Measurement of Business Income*. University of California Press
- Efron B, Tibshirani RJ (1994) *An Introduction to the Bootstrap*. CRC press
- Finger CA (1994) The ability of earnings to predict future earnings and cash flow. *Journal of Accounting Research* 32(2):210–223
- Francis J, Schipper K, Vincent L (2003) The relative and incremental explanatory power of earnings and alternative (to earnings) performance measures for returns. *Contemporary Accounting Research* 20(1):121–164
- Freeman RN, Tse SY (1992) A nonlinear model of security price responses to unexpected earnings. *Journal of Accounting Research* 30(2):185–209
- Gerakos JJ, Gramacy R (2013) Regression-based earnings forecasts. Available at SSRN: <https://ssrncom/abstract=2112137>

- Givoly D, Hayn C (2000) The changing time-series properties of earnings, cash flows and accruals: Has financial reporting become more conservative? *Journal of Accounting and Economics* 29(3):287 – 320
- Granger C, Newbold P (1974) Spurious regressions in econometrics. *Journal of Econometrics* 2(2):111 – 120
- Greene W (1990) *Econometric Analysis*. Pearson Prentice Hall
- Guay WR, Sidhu BK (2001) The usefulness of long-term accruals. *Abacus* 37(1):110–131
- Hardy MA (1993) *Regression with dummy variables*, Sage university papers series. Quantitative applications in the social sciences, vol 93. Sage Publications, Inc.
- Hayn C (1995) The information content of losses. *Journal of Accounting and Economics* 20(2):125–153
- Hou K, van Dijk MA, Zhang Y (2012) The implied cost of capital: A new approach. *Journal of Accounting and Economics* 53(3):504 – 526
- Hribar P, Collins DW (2002) Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40(1):105–134
- James G, Witten D, Hastie T, Tibshirani R (2014) *An Introduction to Statistical Learning with Applications in R*. Springer
- Kang J, Starica C (2016) On the consistent inference of the economic relation of the Residual Income valuation model. Available at SSRN: <https://ssrncom/abstract=2958733> or <http://dxdoior/102139/ssrn2958733>
- Kmenta J (1971) *Elements of econometrics*. Macmillan New York

- Kothari SP, Shanken J (2003) Time-series coefficient variation in value-relevance regressions: A discussion of Core, Guay, and Van Buskirk and new evidence. *Journal of Accounting and Economics* 34:69–87
- Kothari SP, Laguerre TE, Leone AJ (2002) Capitalization versus expensing: Evidence on the uncertainty of future earnings from capital expenditures versus r&d outlays. *Review of Accounting Studies* 7(4):355–382
- Lev B, Ohlson JA (1982) Market-based empirical research in accounting: A review, interpretation, and extension. *Journal of Accounting Research* 20:249–322
- Levy H (1992) Stochastic dominance and expected utility: Survey and analysis. *Manage Sci* 38(4):555–593
- Lipe RC, Bryant L, Widener SK (1998) Do nonlinearity, firm-specific coefficients, and losses represent distinct factors in the relation between stock returns and accounting earnings? *Journal of Accounting and Economics* 25(2):195 – 214
- Lo K (2005) The effects of scale differences on inferences in accounting research: Coefficient estimates, tests of incremental association, and relative value relevance. Available at SSRN: <http://ssrncom/abstract=555684>
- Long JS, Ervin LH (2000) Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician* 54(3):217–224
- Nallareddy S, Sethuraman M, Venkatachalam M (2018) Earnings or cash flows: Which is a better predictor of future cash flows? Available at: <https://ssrncom/abstract=3054644> or <http://dxdoiorg/102139/ssrn3054644>
- Pae J, Thornton DB, Welker M (2005) The link between earnings conservatism

- and the price-to-book ratio\*. *Contemporary Accounting Research* 22(3):693–717
- Peasnell K (1982) Some formal connections between economic values and yields and accounting numbers. *Journal of Business Finance and Accounting* 9(3):361–381
- Penman SH, Zhang XJ (2002) Accounting conservatism, the quality of earnings, and stock returns. *The Accounting Review* 77(2):237–264
- Preinreich G (1936) The fair value and yield of common stock. *The Accounting Review* 11:130–140
- Ramsey JB (1969) Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society Series B (Methodological)* 31(2):350–371
- Srivastava A (2014) Why have measures of earnings quality changed over time? *Journal of Accounting and Economics* 57(2):196 – 217
- Stock JH, Watson MW (2012) *Introduction to Econometrics*. Addison-Wesley
- Subramanyam KR (1996) Uncertain precision and price reactions to information. *The Accounting Review* 71(2):207–219
- Teets WR, Wasley CE (1996) Estimating earnings response coefficients: Pooled versus firm-specific models. *Journal of Accounting and Economics* 21(3):279 – 295
- Watts RL, Leftwich RW (1977) The time series of annual accounting earnings. *Journal of Accounting Research* 15(2):253–271

White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48(4):817–838