Understanding how the effects of conditional conservatism measurement bias vary with the research context

Mostafa Harakeh, Edward Lee, Martin Walker·

ABSTRACT: While the asymmetric timeliness (AT) measure of Basu (1997) underpins a large body of empirical research on conditional conservatism (CC), prior studies have demonstrated that the AT construct is biased and that such bias is likely to lead to Type 1 error. To assess how this bias could affect inferences from prior literature, we replicate previous CC studies that apply the AT measure, and we compare the outcomes with those based on the asymmetric conditional variance (ACV) measure of Dutta & Patatoukas (2017) confirmed to be less influenced by similar bias. We draw two primary conclusions. First, the AT and ACV measures yield similar inferences in interrupted time-series settings that examine the impact of exogenous accounting policy changes on CC. Second, the inferences drawn from applying the AT measure are not supported by the ACV measure in seminal studies that model the determinants of CC in cross-sectional settings. Our findings have implications for both past and future empirical studies of CC.

Keywords: conditional conservatism; asymmetric timeliness; measurement bias; Type 1 error

This version: December 2017

· Mostafa Harakeh is from Lebanese American University, and Edward Lee and Martin Walker are from University of Manchester. We are grateful for helpful comments of Colin Clubb, James Ohlson, Panos Patatoukas, Andrew Stark, Konstantinos Stathopoulos, Norman Strong, as well as participants of presentations at Fudan University (Shanghai, China), Lebanese American University (Beirut, Lebanon), and the University of Manchester (Manchester, UK).

Corresponding author:martin.walker@manchester.ac.uk

1. Introduction

We examine to what extent the inferences of prior empirical research on conditional conservatism (CC) based on the Basu (1997) asymmetric timeliness (AT) measure could be affected by the concerns raised by recent critiques (Dutta & Patatoukas, 2017; Patatoukas & Thomas, 2011, 2016) regarding its underlying bias and the potential for Type I error. The large and still growing accounting literature on CC (Beyer, Cohen, Lys, & Beverly, 2010; Mora & Walker, 2015; Ruch & Taylor, 2015; Wang, Hógartaigh, & Zijl, 2009; Watts, 2003b) has widely adopted the AT construct and drawn important conclusions from its empirical findings about the role of accounting conservatism in the capital market. Examples include the reduction of lender-shareholder conflict (Ahmed, Billings, Morton, & Stanford-Harris, 2002), the strengthening of debt contracting efficiency (Ball, Robin, & Sadka, 2008), the moderation of agency problems (LaFond & Roychowdhury, 2008), and the decrease of information asymmetry (LaFond & Watts, 2008).[1]

Nevertheless, in parallel to this development, there is also an on-going debate on the validity of the AT measure. On the one hand, some studies highlight various sources of bias in the AT measure that could induce Type I error (Dietrich, Muller, & Riedl, 2007; Gigler & Hemmer, 2001; Patatoukas & Thomas, 2011, 2016). On the other hand, other studies suggest reasons or adjustments to mitigate such concerns over the AT construct (Ball, Kothari, & Nikolaev, 2013a, 2013b; Collins, Hribar, & Tian, 2014). In response to this controversy, we heed the call of Ball (2016) to revisit previous empirical studies and verify their conclusions using new methodologies and data. Our findings have particular implications for accounting

[1] To help motivate our study, we conducted a search of all original research articles, since Basu (1997) to date, that applied the AT construct and are published in the five leading accounting journals. We counted a total of 101 articles, with the time trend as well as topic and journal distribution provided in Appendix 1.

researchers seeking to evaluate the costs and benefits of CC or its determinants and consequences.

As part of our evaluation of the AT measure, we adopt the asymmetric conditional variance (ACV) measure recently proposed by Dutta & Patatoukas (2017) as a benchmark throughout this study. Specifically, the ACV measure captures CC as the spread between the variance of bad news accruals and the variance of good news accruals. Dutta & Patatoukas (2017) observe that the AT measure depends not only on CC but also on non-accounting related and firm-specific economic factors, such as expected returns, cash-flow persistence, and asymmetric returns distributions. In contrast, the ACV measure they develop is associated more with empirically observable accounting conservatism proxies such as asset write-downs and impairments. In other words, Type I error is more likely under the AT measure than the ACV measure.

For our main empirical analysis, we re-examine the inference of two sets of previous studies based on the AT construct. Our first main analysis reconsiders previous studies that exploit an interrupted time-series setting to examine how exogenous accounting policy changes affect CC. Within this context, the Sarbanes-Oxley Act (SOX) and the widespread adoption of International Financial Reporting Standards (IFRS) are widely recognized in the accounting literature as two of the most far reaching regulatory changes in recent times (Leuz & Wysocki, 2016). As such, the two relevant studies that we reassess are Lobo & Zhou (2006), who examine the impact of SOX on CC in a U.S. sample, and André, Filip, & Paugam (2015), who evaluate the effect of IFRS on CC in an international sample. Our second main analysis reevaluates previous studies that model the determinants of CC in cross-sectional settings. Within this context, outsider equity investors and debt-holders (particularly public debt-holders) are commonly assumed to be the two most important sources of demand for CC (Watts, 2003a, 2003b). Therefore, the two relevant studies we reconsider are LaFond & Watts (2008), who compare high against low information asymmetry firms in a U.S. sample, and Ball et al. (2008), who compare the role of the debt market as a driver of CC against the equity market based on an international cross-section.

We observe a marked contrast in the effects of AT measurement bias between the interrupted time series and the cross-sectional settings. In both of our interrupted time-series analyses, we observe that the ACV measure generates similar inferences to the AT measure. Furthermore, this is corroborated by changes in empirically observable accounting properties associated with conservatism, and not driven by economic factors unrelated to accounting decisions. In stark contrast, in both of our cross-sectional analyses, we find that the ACV measure does not support the inferences yielded by the AT measure. Moreover, we confirm that the results based on the AT measure correlate with cross-sectional variations in economic factors unrelated to accounting decisions, and bear no link with variations in observable accounting properties associated with conservatism. In other words, the AT measure is more likely to induce Type I error in cross-sectional than in interrupted time-series settings. To the extent the AT measure captures both CC and bias effects (Dietrich et al., 2007), exogenous changes in accounting policy are expected to exert greater influence on the former component (i.e., the CC component). However, cross-sectional variations in firm characteristics are more likely to pick up non-accounting factors that drive the latter component (i.e., the bias component).

As additional analyses, we examine the properties of the AT construct in two ways. First, we validate the role of ACV as a benchmark measure against the AT measure. We begin by confirming that these two CC measures are correlated on an unconditional basis. This helps establish a level playing field between the two measures in a sense that they both capture CC when CC exists. We also confirm that the opening stock price, or the deflator of the dependent variable in the Basu (1997) piece-wise linear regression, is inversely related to the AT measure but not to the ACV measure. This suggests that the ACV measure is less sensitive than the AT measure to the sources of bias driven by the price-scale effects documented by Patatoukas & Thomas (2011). These findings jointly support our application of ACV as the benchmark for AT to re-examine the inferences of previous studies. This is because both measures are at least comparable on an unconditional basis to the extent that they capture a certain degree of underlying CC, and at the same time, the ACV measure is less likely to suffer from Type 1 error driven by non-accounting factors.

In our second set of additional analyses, we reassess the C_Score measure, which has been adopted by many CC studies. The C_Score was developed by Khan & Watts (2009) to enable the estimation of CC at the firm-year level through the AT construct. It is based on the predicted relationship between the AT measure and three components, i.e., market-to-book, size, and leverage. Consistent with Khan & Watts (2009), we confirm that the positive relationship between the C_Score and AT exists in our samples. However, in stark contrast, we observe no such relationship between the C_Score and ACV. This implies that the C_Score inherits the potential sources of bias that affect the AT measure. Turning to the components of the C_Score, we confirm the prediction in Khan & Watts (2009) that the AT measure is negatively related to market-to-book and size and positively related to leverage. Nevertheless, again in contrast, none of these three components bear any clear association with the ACV measure.

Our study contributes to the CC literature in two ways. First, unlike previous studies involved in the debate either for (Ball, Kothari, & Nikolaev, 2013a, 2013b; Collins, Hribar, & Tian, 2014) or against (Dietrich, Muller, & Riedl, 2007; Gigler & Hemmer, 2001; Patatoukas &

Thomas, 2011, 2016) the AT construct of Basu (1997), we provide more direct empirical evidence by revisiting the settings and replicating the analyses applied in previous CC studies. Our research design essentially provides a more vivid depiction of the extent to which measurement bias could affect empirical studies of CC across different settings. Second, also unlike previous studies associated with the debate on the AT measure, our study draws a multifaceted rather than a single-sided conclusion to the issue. We reveal that the impact of the measurement error associated with the AT construct, and any potential concerns over the inferences drawn from previous studies, depend on the research setting involved. In sum, our study has implications for the interpretation of past evidence and the implementation of future research on CC and, therefore, is potentially of interest to academics and practitioners interested in both the theoretical and empirical dimensions of CC and timely loss recognition.

The rest of this paper is organized as follows. Section 2 reviews the relevant literature and develops our testable hypotheses. Section 3 discusses sample selection and presents the findings from our empirical analyses. Section 4 presents the study's conclusions.

2. Literature Review and Hypothesis Development

2.1. Conditional conservatism and its measurement

Conservatism is one of the oldest concepts in accounting (Bliss, 1924; Sterling, 1976). Theoretically, CC is an accounting practice that anticipates economic losses before being realized and recognizes economic gains only when realized (Beaver & Ryan, 2005; Watts, 2003a). Empirical studies define conservatism as the requirement of a lower degree of verification to recognize economic losses compared to the degree of verification required to recognize economic gains (Basu, 1997; Pope & Walker, 1999). Accounting conservatism can be either unconditional or conditional, with the former associated with the understatement of the book value of net assets regardless of news, while the latter refers to the timelier recognition of bad news than good news in earnings (Ryan, 2006). Examples of unconditional conservatism include the expensing of investment in intangible assets and setting depreciation rates for property plant and equipment above the expected economic rate of depreciation (Beaver & Ryan, 2005), whereas conditional conservatism is achieved through asset impairments in response to bad news about the value of assets in place (Dutta & Patatoukas, 2017).

Basu (1997) developed a construct for measuring the level of CC by regressing earnings on stock returns. This construct has become the most widely adopted CC construct in the literature (Mora & Walker, 2015; Ruch & Taylor, 2015; Wang, Hógartaigh, & Zijl, 2009; Watts, 2003b). The AT construct of Basu (1997) is depicted in the earnings-returns piecewise linear regression below:

$$X_{it} = \beta_0 + \beta_1 RD_{it} + \beta_2 RET_{it} + \beta_3 RD_{it} \times RET_{it} + \varepsilon_{it} \tag{1}$$

where for firm *i* in year *t*, X_{it} is current year earnings per share deflated by price per share at the end of previous year, *RD* is a dummy variable that equals 1 if *RET* is negative, and 0 otherwise, and *RET* is the abnormal stock return over the fiscal year. The coefficient on the interaction term (β_3) captures the incremental timeliness with which reported earnings reflect bad news relative to good news. We refer to this as the AT measure of CC.

To facilitate the implementation of the AT construct on a firm-year basis, Khan & Watts (2009) develop the C_Score measure based on the following two-step process:

$$
\begin{aligned}
X_{it} = {} & \beta_0 + \beta_1 RD_{it} + \beta_2 RET_{it} + \beta_3 RD_{it} \times RET_{it} \\
& + \beta_4 MTB_{it} + \beta_5 MTB_{it} \times RD_{it} + \beta_6 MTB_{it} \times RET_{it} + \beta_7 MTB_{it} \times RD_{it} \times RET_{it} \\
& + \beta_8 SIZE_{it} + \beta_9 SIZE_{it} \times RD_{it} + \beta_{10} SIZE_{it} \times RET_{it} + \beta_{11} SIZE_{it} \times RD_{it} \times RET_{it} \\
& + \beta_{12} LEV_{it} + \beta_{13} LEV_{it} \times RD_{it} + \beta_{14} LEV_{it} \times RET_{it} + \beta_{15} LEV_{it} \times RD_{it} \times RET_{it} \\
& + \varepsilon_{it}
\end{aligned} \tag{2}
$$

$$
C\_Score_{it} = \hat{\beta}_3 + \hat{\beta}_7 \times MTB_{it} + \hat{\beta}_{11} \times SIZE_{it} + \hat{\beta}_{15} \times LEV_{it} \tag{3}
$$

where for firm *i* in year *t*, MTB_{it} is the market-to-book value of equity, $SIZE_{it}$ is the natural logarithm of the market value of equity, and LEV_{it} is the total debt to market value of equity. These three components have been used to model the four driving factors of CC as suggested by (Watts, 2003a, 2003b), including contracting, litigation, taxation, and regulation. When the first step regression, Equation (2), is estimated annually through the sample period, coefficients $\hat{\beta}_3$, $\hat{\beta}_7$, $\hat{\beta}_{11}$, and $\hat{\beta}_{15}$ are assumed to be constant across firms but are allowed to vary over time. The *C_Score*, calculated using Equation (3), provides a measure of CC at the firm-year level. By comparing Equations (1) with (2), one can see that the C_Score is based on the AT measure.

Appendix 1 provides a summary of CC studies that applied the AT construct, published in five leading accounting journals. The summary includes original research articles that utilize the AT or the C_Score measures either for empirical tests or theoretical model development, and excludes discussion articles. A total of 101 articles were published over the period 1997 to early 2017, which amounts to nearly five articles per year. Panel A provides a time trend analysis that indicates a large and steadily growing literature. Panel B provides a topic analysis, which indicates that most of the papers are related to equity rather than debt markets. Panel C provides a journal distribution analysis, showing that the highest number of articles were published in the *Journal of Accounting and Economics*, where Basu (1997) was originally published. Overall,

Appendix 1 helps motivate our study by confirming the importance of the AT construct in the accounting literature on CC.

2.2. Debate on the validity of the AT measure

Despite the widespread adoption of the AT construct in the CC literature, there is a parallel on-going debate within this literature on the validity of this construct. On the one hand, various studies highlight potential sources of bias that increase the likelihood of Type I error. Gigler & Hemmer (2001) develop a theoretical model showing that the AT measure might be significantly positive in the absence of CC because researchers fail to control for firms' voluntary disclosure, which jointly affects stock returns and accounting conservatism. Dietrich et al. (2007) demonstrate that the AT measure may indicate CC even in the absence of CC due to sample-variance-ratio bias and truncation bias. These biases arise from the fact that earnings cause returns and not vice versa and that the returns variable is also determined by other news unrelated to earnings. Patatoukas & Thomas (2011) identify two sources of bias driven by the use of stock price as the deflator for the dependent variable in the Basu (1997) piece-wise linear regression to estimate the AT measure. First, firms with a low stock price tend to report losses more frequently, and this leads them to have more negative values for the dependent variable in the earnings-returns piecewise regression (i.e., the loss effect). Second, firms with low stock prices have higher fluctuations in their stock prices, which results in a higher variance in stock returns (i.e., the return variance effect). These two effects jointly lead to an upward bias in the AT measure, particularly among firms with low stock prices.

On the other hand, other studies suggest arguments or adjustments to mitigate concerns over the AT measure. For instance, Ryan (2006) questions the severity of the bias identified by Dietrich et al. (2007) by arguing that such concerns have a trivial impact given the low R^2

observed empirically from the returns-earnings regressions. Ball et al. (2013a) further argue that the sample-variance-ratio bias suggested by Dietrich et al. (2007) is irrelevant because the covariance between returns and earnings, conditional on returns, is equal for good and bad news when CC is absent. Furthermore, the price-deflator related sources of bias in the AT measure highlighted by Patatoukas & Thomas (2011) have motivated other studies to propose modifications to the AT measure to correct the problem. Ball et al. (2013b) argue that AT bias is driven by the expected component of earnings and returns, which are correlated with firm-specific variables and result in an association between the error term and the AT coefficient. They suggest various ways to remove such components and alleviate the AT measure from the bias suggested by Patatoukas & Thomas (2011). Collins et al. (2014) propose that the replacement of the earnings with accruals as the dependent variable in the Basu (1997) regression corrects for the bias raised by Patatoukas & Thomas (2011). Because the persistence of the cash-flow component in earnings increases with the expected component of returns, Collins et al. (2014) suggest that the removal of both components through the use of accruals and unexpected returns in the regression mitigates the spurious asymmetric timeliness in the AT measure.

The debate on the validity, or otherwise, of the AT still continues. For example, in response to Ball et al. (2013b) and Collins et al. (2014), Patatoukas & Thomas (2016) demonstrate that their revised AT constructs continue to exhibit upward bias in placebo tests based on asymmetry in the conditional relations between the inverse of lagged share price and positive and negative return news. Furthermore, Dutta & Patatoukas (2017) provide evidence that the revised AT constructs are sensitive to and can be driven by three non-accounting related and firm-specific economic factors: (i) expected returns (ii) asymmetry in the conditional variances of positive and negative unexpected returns, and (iii) cash-flow persistence. Furthermore, they argue that the revised AT constructs become statistically and economically insignificant in the presence of these non-accounting factors as controls.

In conclusion, the current state of the debate raises many questions in the minds of researchers related to the validity of the AT measure and to the reliability of the inferences that have been drawn by past papers that relied on this measure. The present paper seeks to address these issues.

2.3. Alternative measure of conditional conservatism

As an alternative approach to the measurement of CC, Dutta & Patatoukas (2017) propose the ACV measure that is less affected by the sources of bias that drive the AT measure. According to Dutta & Patatoukas (2017), CC can be empirically estimated by calculating the difference between the variance of bad news accruals and the variance of good news accruals. They use the sign of unexpected returns as a proxy for good/bad news, deflate accruals with the lagged stock price, and estimate the measure as follows:

$$ACV = Variance(ACC_{it} \mid RET_{it} < 0) - Variance(ACC_{it} \mid RET_{it} > 0) \quad (4)$$

where for firm i and year t, ACC_{it} is deflated accruals and RET_{it} is unexpected returns.[2]

Theoretically, Dutta & Patatoukas (2017) argue that the ACV measure is driven only by variations in CC. Moreover, they claim that, unlike the AT construct, their proposed measure is unaffected by the asymmetric distribution of returns and does not rely on the market efficiency

[2] The ACV measure is designated as the spread in conditional variance in Dutta & Patatoukas (2017) since they compute it as the difference between the variance of good and bad news accruals. Throughout our study, we report empirical findings based on the ACV measure computed as the ratio of instead of the differences between the conditional variances, because the use of the ratio strengthens the comparability of the measure across the different sub-samples involved in our study. However, we have also replicated our tests with the ACV measure calculated as the spread, rather than the ratio, of the conditional variances for all empirical analyses, and untabulated results suggest that our inferences remain similar.

assumption where investors incorporate all information in a timely and efficient manner in stock prices. They also demonstrate that the ACV measure is not affected by non-accounting related and firm-specific economic factors that influence the AT construct, even under the modifications of Ball et al. (2013b) and Collins et al. (2014).

However, while appealing in terms of theoretical rationale and empirical properties, the CC literature awaits further research that applies this newly proposed measure. In particular, the circumstances under which the ACV measure confirms or negates the inferences of prior CC studies that relied on the AT measure deserves to be evaluated. Our paper addresses this issue.

2.4. Hypothesis development

As we discussed in Section 2.2, existing studies on the validity of the AT measure largely focus either on the identification of potential sources of bias (Dietrich, Muller, & Riedl, 2007; Gigler & Hemmer, 2001; Patatoukas & Thomas, 2011, 2016) or on proposing reasons and solutions to mitigate such bias (Ball, Kothari, & Nikolaev, 2013a, 2013b; Collins, Hribar, & Tian, 2014). However, despite the well documented concerns about the AT construct, limited attention has been paid to assessing the consequences of these concerns for prior applications of the AT construct. Nevertheless, this is an important issue, particularly for accounting researchers seeking to rely on past empirical evidence or to develop research designs for future empirical studies. As the limitations of the AT measure become increasingly highlighted in the literature, and as potential alternative solutions to estimate CC, such as the ACV measure, emerge, it becomes both necessary and possible to re-examine the inference of previous CC studies that were based on the AT construct.

In the accounting literature, there are two commonly adopted research settings in which the AT construct is utilized to empirically examine the determinants of CC. In the first setting, researchers adopt an interrupted time-series design to observe whether and how exogenous changes in accounting policies affect CC. In the second setting, researchers examine cross-sectional variations to observe whether and how differences in firm characteristics influence CC. We expect the impact of the measurement error associated with the AT construct to differ across these two research settings for the following reason. As Dietrich et al. (2007) show in their Equations 1.7a and 1.7b, the AT measure is biased for good and bad news, respectively, where each equation comprises two components, i.e., the CC component and the bias component. The CC component is affected by accounting decisions related to timely loss recognition, while the bias component is driven by economic factors unrelated to accounting (Dutta & Patatoukas, 2017; Patatoukas & Thomas, 2011, 2016). In an interrupted time-series setting, the exogenous change in accounting policy is expected to influence the CC component but not the bias component. As such, changes in the sample AT coefficient estimated over a short period of time are, arguably, more likely to be driven by systematic changes in the underlying CC component, while the changes in the bias component are likely to offset and wash out. In this case, we would expect inferences based on AT and ACV measures to be broadly consistent, to the extent the two measures both capture certain level of CC, despite the fact that the AT measure is more sensitive to the bias component. In a cross-sectional setting, however, firm characteristics assumed to affect the CC component could also simultaneously correlate with the bias component. As such, the difference in the AT coefficients estimated on a cross-sectional basis are more likely to be driven by both the CC and bias component, and particularly the latter if the difference in economic factors across firms outweighs that of CC. In this case, we would expect a greater likelihood for the ACV measure to negate the inferences based on the AT measure, which is more affected by the bias component. Given these arguments, we formulate the following testable hypotheses:

H1: In interrupted time-series settings that examine exogenous changes in CC, the AT and ACV measures are likely to lead to similar inferences.

H2: In cross-sectional settings that examine determinants of CC, the inferences from the ACV and AT constructs are more likely to be inconsistent.

3. Empirical Analyses

3.1. Main analysis

Our main empirical analyses are based on tests of hypotheses H1 and H2. To test hypothesis H1, we re-examine the inferences of two previous studies associated with interrupted time-series settings, including Lobo & Zhou (2006) for the SOX effect in the U.S. and André et al. (2015) for the IFRS effect across the European Union. To test hypothesis H2, we reassess the inferences of two previous studies associated with cross-sectional settings, specifically LaFond & Watts (2008) for the information asymmetry effect in a U.S. sample and Ball et al. (2008) for the debt market effect in an international sample. These four studies are chosen primarily due to the importance of their inferences for the CC literature. However, they also represent different ways in which the AT construct can be implemented in empirical research. For instance, while Lobo & Zhou (2006) and LaFond & Watts (2008) both directly employ the Basu (1997) AT regression approach, André et al. (2015) apply the C_Score, which enables the estimation of AT measure on a firm-year basis, and Ball et al. (2008) apply an aggregated AT measure estimated on a country level.

In each case, we first replicate the original findings and inferences based on the AT construct, following the sample construction of that study. We then replace the AT measure with the ACV measure and look to see whether the inferences implied by the AT measure still hold under the ACV measure. To accommodate our research design, our estimation of the ACV measure differs from Dutta & Patatoukas (2017) in two ways. First, we use earnings (i.e., *X*) instead of accruals to avoid the problem of missing data that can be more severe among non-U.S. firms.[3] Second, we use the ratio of, rather than the difference between, the conditional variances of bad news and good news earnings to enhance the comparability of this measure across the different sub-samples involved in our research design.[4] In addition to the ACV measure, we also introduce two additional measures to help distinguish between accounting and non-accounting factors that drive the results. Following Dutta & Patatoukas (2017), we use negative special items (*SI*) to capture accounting decisions that reflect CC, and asymmetric returns distributions (*ARetDist*) to capture economic factors that could induce measurement errors in the AT construct. Table 1 provides detailed definitions of all the variables we use to re-examine the four previous studies in our main analysis. Since each set of re-examinations requires a different sample in accordance with the original study, we provide the description of the sample construction for each case separately in the corresponding sub-section.

[Insert Table 1 here]

[3] Dutta & Patatoukas (2017) suggest that the ACV measure can also be constructed using earnings instead of accruals. Our consistent use of the earnings variable *X* in both the AT and ACV measures also renders the results from both more comparable. When we replace earnings with accruals in our U.S. sample, the inferences we draw remain unchanged.

[4] Since the ACV is a non-linear parameter, we test the statistical significance of its differences across sub-samples through a non-linear combination of estimators with the delta method (Casella & Berger, 2002; Feiveson, 1999).

3.1.1. Lobo & Zhou (2006) (interrupted time-series setting)

Lobo & Zhou (2006) examine the change in the level of CC following the Sarbanes-Oxley (SOX) Act in 2002 for U.S. firms. The main purpose of SOX is to protect investors by improving the accuracy and reliability of corporate disclosures and to restore shareholders' and lenders' confidence in the reliability of financial reporting among U.S. firms (see the survey of Coates and Srinivasan, 2014). Lobo & Zhou (2006) apply the Basu (1997) regression and introduce an interaction with the post-SOX period indicator variable (*SOX*). They document that the SOX enactment leads to an increase in the degree of CC estimated by the AT measure, and they argue that this finding provides important and early empirical evidence that policy makers were able to achieve one of their main goals by improving timely loss recognition by firms. Furthermore, their paper has informed the debate within a large accounting literature over the economic consequences of SOX (Coates & Srinivasan, 2014).

Consistent with Lobo & Zhou (2006), we use the Compustat fundamentals annual file for accounting data and CRSP for stock return data. We also follow their sample construction procedure to retrieve accounting and return data between 2000 and 2004, and exclude firms with stock prices below $1 and observations with negative book value of equity. We then delete the top and bottom percentiles of earnings and returns distributions, and we require an equal number of observations per firm pre- and post-SOX. The final sample comprises 5,622 (5,622) firm-year observations in the pre-SOX (post-SOX) period. Table 1 defines all the variables applied in this set of analyses.

Table 2 reports our replication and re-examination of Lobo & Zhou (2006). Panel A provides summary statistics for the main variables and is comparable with Table 4 in their study. The first column in Panel B of our Table 2 reports the replication results of model (6b) in Table 4 of Lobo & Zhou (2006). The coefficient on the interaction term *SOX*×*RD*×*RET* is 0.0436 and is significant at the 1% level, which indicates an increase in the degree of CC following the passage of the SOX Act.[5] We then replace the dependent variable with *SI* in order to test the change in accounting choices around the SOX enactment. As shown in the second column in Panel B of our Table 2, the coefficient on the interaction term *SOX*×*RD*×*RET* is also significantly positive at the 1% level, and this suggests a post-SOX increase in incremental timeliness of *SI* to bad return news. The results we obtain in Panel B through *X* and *SI* as dependent variables consistently support the inference of an increase in CC after SOX. Panel C of our Table 2 reports the change in ACV values around SOX, showing a statistically significant increase from 1.16 in the pre-SOX period to 1.34 in the post-SOX period. In other words, the ACV benchmark measure essentially corroborates the inferences based on the AT measure generated by Lobo & Zhou (2006). Finally, Panel D of our Table 2 reports no significant changes in *ARetDist* following SOX, where *ARetDist* is meant to capture non-accounting economic factors that bias the AT measure upwards (Dutta & Patatoukas, 2017). Taken together, our analyses in Table 2 confirm that the original inference in Lobo & Zhou (2006) based on the AT measure is consistent with that based on its ACV counterpart and is more driven by accounting than economic factors. As such, we provide a set of findings that is consistent with hypothesis H1, which predicts that ACV and AT measures are likely to generate similar inferences for CC under interrupted time-series settings involving exogenous accounting policy changes.

[Insert Table 2 Here]

[5] The magnitude of the coefficient on the interaction term is smaller than that in Lobo & Zhou (2006) because we use abnormal returns, whereas Lobo & Zhou (2006) use raw returns. We obtain a coefficient of a similar magnitude to theirs when using raw returns.

3.1.2. Andre, Filip, & Paugam (2015) (interrupted time-series setting)

André et al. (2015) examine the change in CC following the mandatory adoption of IFRS in the European Union. The mandatory adoption of IFRS is an exogenous change in accounting policy that is meant to affect various aspects of the financial reporting system (see the survey of De George, Li, & Shivakumar, 2016). Using a sample of 16 European countries over the 2000-2010 period and a modified version of the C_Score measure (*C_Score**), André et al. (2015) find a significant reduction in the degree of CC following IFRS adoption.[6] Their findings suggest that the adoption of IFRS by EU countries materially reduced the average level of CC for EU listed firms. This is an important finding for the International Accounting Standards Board (IASB), which is engaged in a policy debate on whether to keep conservatism in its conceptual framework in light of investors' demands for a conservative financial reporting (Cooper, 2015).

André et al. (2015) use Thomson Reuters for accounting and returns data and DataStream for firm-level beta coefficients and stock price volatility. We replicate their main findings using Compustat Global, and we calculate firm-level beta coefficients and stock price volatility as described in the DataStream manual.[7] Broadly following their sample construction approach, we first download all accounting data from the Compustat Global fundamental annual file and stock return data from the Compustat Global security file, and then exclude firms that are cross-listed or belong to financial sectors, or have a negative book value of equity. We also keep only mandatory adopters and delete firms that did not adopt IFRS in 2005.[8] Finally, we keep firms

[6] In addition to the original market-to-book, size, and leverage applied in Khan & Watts (2009), the modified C_Score of André et al. (2015) also incorporates proxies for the cost of equity and unconditional conservatism.
[7] André et al. (2015) retrieve firms' beta coefficients from DataStream and use this variable as a proxy for the cost of capital. We follow the DataStream manual and estimate firms' beta coefficients over the last 36 months with a minimum of 23 consecutive monthly returns. We also calculate the stock price volatility, which is used in estimating unconditional conservatism, as the annualized variance of daily stock returns.
[8] We follow the accounting standards classification in Daske, Hail, Leuz, & Verdi (2013) in order to identify firms that adopted IFRS in 2005.

that have at least one observation before and one observation after IFRS adoption. This leaves us with 5,520 (7,211) firm-year observations in the pre-IFRS (post-IFRS) period. Table 1 defines all the variables applied in this set of analyses.

Table 3 presents our replication and re-examination of André et al. (2015). Panel A reports the summary statistics for the variables used in our replication of their main findings, and our statistics are broadly similar to those reported in their Panel C of Table 1.[9] The first column in Panel B of our Table 3 reports the regression that replicates column 1 of Table 2 in their study. The negative and significant coefficient (–0.0174 and significant at 1% level) on the IFRS adoption indicator variable (*IFRS*) suggests a significant reduction in the degree of CC when estimated using their modified C_Score (i.e., *C_Score**). In the second column in Panel B of our Table 3, we replace the original dependent variable *C_Score** with *SI*, to examine changes in the level of negative special items reported by firms after IFRS adoption. The positive and significant coefficient (0.0058 and significant at 1% level) on *IFRS* in the second column in Panel B of our Table 3 shows that the level of negative special items increases (i.e., becomes less negative), which supports the inference that CC declines after IFRS adoption. In Panel C of our Table 3, the ACV measure further corroborates this inference by showing a statistically significant decrease from 1.25 in the pre-IFRS period to 0.91 in the post-IFRS period. Finally, Panel D of our Table 3 reports a significant increase in the economic factor *ARetDist* from the pre- and post-IFRS adoption period. To the extent this measure is supposed to induce an upward bias in CC as Dutta & Patatoukas (2017) suggest, the finding in our Panel D implies that this non-accounting confounding effect is unlikely to drive the inference of a reduction in CC. Collectively, our findings in Table 3 confirm the conclusion of André et al. (2015) and are

[9] Obtaining differences in summary statistics and regression results is inevitable as we use a different database, yet we reach the same conclusion from our replication.

consistent with hypothesis H1 that the ACV and AT measures are likely to draw similar inferences under interrupted time-series settings that examine how CC is affected by exogenous changes in accounting policies. The corroboration of findings in our Tables 2 and 3 based on replications and re-examinations of two studies with vastly different samples serves as a powerful mutual robustness check.

[Insert Table 3 Here]

3.1.3. LaFond & Watts (2008) (cross-sectional setting)

LaFond & Watts (2008) examine whether information asymmetry in the equity market generates a demand for CC in financial reporting. They measure information asymmetry through the probability of information-based trading (PIN) developed by Easley & O'Hara (1992). The PIN score captures the difference in information asymmetries between informed and uninformed investors through abnormal order flow. The inference of LaFond & Watts (2008) is important in the CC literature because it demonstrates that, in addition to the protection of debt-holders and the promotion of debt contracting efficiency (Ahmed, Billings, Morton, & Stanford-Harris, 2002; Ball, Robin, & Sadka, 2008; Zhang, 2008), outside investors in the equity market also contribute an important source of demand for timely loss recognition to the extent such accounting approach contributes to the mitigation of agency problems. Their evidence has important implications for standard setters that are in doubt or opposed to accounting conservatism (e.g., FASB, 1980, 2005)

LaFond & Watts (2008) conduct their study based on a U.S. sample that includes firms listed on the NYSE and AMEX exchanges over the period of 1983 to 2001. To construct their sample, they require firms to have December fiscal year-ends as well as data for the PIN score and sufficient CRSP and Compustat data for the empirical analyses. Firms are required to have

December fiscal year-ends because the PIN scores are estimated using calendar year trading data. We follow their sampling procedure and acquire data for the PIN score from the same source.[10] Our final sample comprises 19,831 firm-year observations over the same sample period. Table 1 defines all variables used in this set of analyses.

Table 4 presents our replication and re-examination of LaFond & Watts (2008). Panel A reports summary statistics for the variables we apply. Note that the number of observations and the relevant statistics for *X* are generally close to those reported for $NI_{one\text{-}year}$ in Table 1 Panel A of their study. The first column of Panel B in our Table 4 reports our replication of their main analyses on the conditioning effect of PIN on CC measured through the AT measure. Note that the coefficients estimated for *RD×RET* (0.1208) and *PIN×RD×RET* (1.0444) are both positive and statistically significant at the 1% level. This yields an inference consistent with Table 2 Panel C of their study and confirms that firms with larger PIN scores (i.e., greater information asymmetry) are associated with higher levels of CC. However, in the second column of our Table 4 Panel B, the coefficient on *PIN×RD×RET* is no longer statistically significant when we apply *SI* as the dependent variable of the regression. This suggests that firms with higher PIN scores are not incrementally more associated with an accounting measure that is directly associated with conservative reporting. Turning to Panel C of our Table 4, we apply ACV as our benchmark measure for the AT construct. We split the sample into high and low information asymmetry using a dummy variable (*HighPIN*) that takes the value 1 if the observations' PIN score is higher than the median value of the sample. Note that the ACV value estimated for the higher and lower PIN groups are 1.69 and 1.76, respectively, and the difference between these

[10] We are grateful to Soeren Hvidkjaer for providing the PIN measure dataset on his website (https://sites.google.com/site/hvidkjaer/). LaFond & Watts' (2008) footnote 10 indicates that they also acquired the PIN score from his earlier website (http://www.smith.umd.edu/faculty/hvidkjaer).

groups is statistically insignificant. In other words, the ACV measure suggests that firms with higher PIN scores are not associated with higher levels of CC, which negates the findings based on the AT measure. Panel D of our Table 4 compares the *ARetDist* economic factor between higher and lower PIN score samples, and confirms that the former group is indeed significantly higher than the latter group, i.e., 3.58 and 2.06, respectively. The finding that higher PIN firms are associated with higher AT and *ARetDist* measures, but not with *SI* and ACV measures, suggests that the inference from previous studies of such firms having more pronounced CC is likely to be confounded by differences in non-accounting economic factors. Overall, the findings across Table 4 are consistent with our prediction in hypothesis H2 that the ACV measure can conflict with inferences drawn from the AT construct in cross-sectional settings. Our evidence suggests that the conclusion of higher demand for CC among firms with greater information asymmetry needs to be interpreted with caution.

[Insert Table 4 here]

3.1.4. Ball, Robin, & Sadka (2008) (cross-sectional setting)

Ball et al. (2008) examine whether debt markets or equity markets constitute the primary source of demand for timely loss recognition. Given that timely loss recognition is costly, and the supply of this activity is dependent on demand, the authors argue that debt markets are a more important driver of the demand for CC than equity markets. The findings of Ball et al. (2008) are important because they empirically confirm that CC primarily caters for the information demands of debt investors rather than equity investors, and this in turn addresses fundamental issues in the accounting literature regarding the objectives of financial statement information (Beyer et al., 2010; Kothari, 2001).

Ball et al. (2008) generate country-level AT measures by running the Basu (1997) piece-wise linear regression separately for 22 countries over the years 1992-2003. Next, these researchers run a cross-sectional regression of their 22 AT estimates on country-level proxies for the importance of the debt market and the equity market as their main explanatory variables, along with controls for other country-level characteristics. Their measures of the importance of debt and equity markets are based on La Porta, Lopez-De-Silanes, Shleifer, & Vishny (1997, 1998), who capture the value of these markets in each country relative to the Gross National Product (GNP). To replicate their analyses, we use accounting data from the Compustat Global fundamental annual file and stock return data from the Compustat Global security file for the selected 22 countries.[11] We delete firms that are cross-listed or belong to the financial and utility sectors as well as observations in the top and bottom 1% of the deflated earnings and returns distributions. The final sample comprises 96,298 firm-year observations.

Table 5 reports our summary statistics for this set of analyses, with Panel A showing those for *X*, *RET*, and *SI* across the full international sample and Panel B reporting country-level variables, including the AT and ACV measures as well as the control variables. Specifically, variables *B2* and *B3* are country-level Basu (1997) regression estimates that captures asymmetric timely gains and losses respectively. All other variables are defined in Table 1.

[Insert Table 5 here]

Table 6 presents our replication and re-examination results of Table 5 in Ball et al. (2008). In Panel A, we apply the *B3* measure as the dependent variable, which is the country-level AT measure. Note that in all columns, the coefficient on *DEBT* is significantly positive while the coefficient on *EQUITY* is negative and mostly significant. These findings are consistent with

[11] Ball et al. (2008) use the Global Vantage database, which has been succeeded by Compustat Global.

those reported in Table 5 of Ball et al. (2008), which they interpret to imply that the demand for CC is driven more by the needs of debt than equity investors. In Panel B of our Table 6, we replace *B3* with *SI* as the dependent variable. Note that the coefficient on *DEBT* is no longer statistically significant in any of the columns. To the extent *SI* captures accounting factors that are more directly related to CC than the AT measure, the contrast between our findings in Panels A and B suggests that the inference drawn in Ball et al. (2008) could be specific only to the use of the AT construct to measure CC. Panel C of our Table 6 presents the comparison of ACV values estimated for countries with higher and lower importance of debt markets. We use the variable *HighDEBT* to split the sample into high and low importance of debt markets, where *HighDEBT* is a dummy variable that takes the value 1 if the country's importance of debt is higher than the median value of the sample. The difference in ACV measure between the higher and lower debt market importance groups is small, i.e., 1.23 and 1.25, respectively, and statistically insignificant. In other words, based on the ACV measure, it is not possible to draw the inference that CC is more pronounced in countries with higher debt market importance. As such, the ACV measure negates the inference based on the AT measure in the original Ball et al. (2008) analyses. This finding is consistent with our prediction in hypothesis H2 that ACV is more likely to negate the inference of the AT measure in cross-sectional settings. Furthermore, Panel D of our Table 6 shows that the *ARetDist* economic factor is significantly higher among countries where debt markets are more important. As Dutta & Patatoukas (2017) note, the *ARetDist* economic factor, measured as the asymmetric return variance, is likely to induce upward bias in the AT measure and, accordingly, leads to Type I error. Overall, the findings in Table 6 are consistent with our hypothesis H2, and they suggest that the inference that debt markets drive the demand for CC deserves to be interpreted with caution. The consistency in inferences obtained from our re-examination of Ball et al. (2008) and LaFond & Watts (2008) provides a powerful mutual robustness check across two independent studies.

[Insert Table 6 here]

3.2. Additional analyses

Our additional analyses comprise two parts. The first part seeks to further validate the use of ACV as the benchmark for comparison with AT. The second part provides an analysis of *C_Score* and its three components (i.e., *MTB*, *SIZE*, and *LEV*). To ensure comparability across these two parts, we construct and implement them on the same samples of U.S. and non-U.S. firms. Our U.S. sample comprises 7,004 firms and 70,033 firm-year observations, and our international sample comprises 22,254 firms and 215,903 firm-year observations from 22 countries with the U.S. included among them. For the U.S. and Canadian firms, we acquire data from Compustat, and for the international sample, we use Compustat Global. We exclude cross-listed or financial firms, and exclude observations with a negative book value of equity.

For the basic AT regression described in Equation 1, we define *X* as income before extraordinary items divided by the lagged market value of equity, and *RET* as abnormal returns at the end of fiscal year.[12] To reduce the effect of outliers, we drop the top and bottom percentiles of *X* and *RET*. For the estimation of *C_Score* presented in Equations 2 and 3, we follow Khan & Watts (2009) to define *MTB* as the ratio of market value of equity to book value of equity, and *SIZE* as the natural logarithm of market value of equity. We define *LEV* as the ratio of total liabilities to the sum of market value of equity and total liabilities following Fama & French

[12] This is intended to remove the effect of annual earnings announcement on stock prices, which occurs approximately three months later (García Lara, García Osma, & Penalva, 2009). Inferences from our analyses are unchanged when we apply stock returns calculated three months after the closing date.

(2002).[13] We winsorize *MTB*, *SIZE*, and *LEV* at the top and bottom 1%. Table 1 provides a detailed definition of variables used in these empirical analyses.

Table 7 reports the observations and summary statistics for the variables used in these analyses. Panel A reports the countries we select along with their corresponding number of observations.[14] Panels B and C of Table 7 provide summary statistics of variables for the U.S. and international samples, respectively. On average, U.S. firms have higher *MTB*, non-U.S. firms have higher *LEV*, and *RET* in both samples is close to zero. Our summary statistics are also comparable to relevant previous studies. For instance, the mean and standard deviation of *C_Score* for the U.S. sample are 0.0946 and 0.1074, respectively, while the corresponding statistics reported in Khan & Watts (2009) are 0.105 and 0.139.

[Insert Table 7 here]

3.2.1. ACV benchmark validation

To help validate the ACV as a benchmark measure for comparing with the AT construct, we conduct two analyses. Table 8 reports the first analysis, which examines to what extent the ACV and AT measures are correlated on an unconditional basis. This enables us to evaluate whether the two measures are comparable in a sense that they both capture CC to some extent and, thereby, establish a level playing field between both measures in our study. We first calculate the AT coefficient estimate for each industry-year based on the Fama and French twelve-industry classification.[15] We then sort the AT coefficient estimates into deciles and

[13] Defining *LEV* following Khan & Watts (2009) as the ratio of the sum of long-term debt and short-term debt to market value of equity leads to 15% of observations with zero value, which is not suitable for our research design, which involves decile ranking of observations. Nevertheless, our results are qualitatively similar when using the Khan & Watts (2009) definition.

[14] The countries we select in our additional analysis are broadly consistent with Ball et al. (2008) and André et al. (2015).

[15] Using a more specific industry classification (such as SIC two-digit) will leave many industry-year groups with a small number of observations, which affects the accuracy of regression estimates.

calculate the corresponding ACV measure for each decile. Finally, we plot the values of the ACV measure across the deciles of the AT measure for the U.S. and the international samples, separately, as shown in Table 8. This graphical evidence confirms a positive correlation between the AT and ACV measures on an unconditional basis. This suggests that both measures indeed empirically capture some common underlying effect, which is presumably related to CC, and renders them comparable on that basis in our tests.

[Insert Table 8 here]

Table 9 presents the second of our benchmark validation analysis, which examines to what extent the ACV and AT measures are correlated with the opening stock price that is used as the deflator of the dependent variable in the Basu (1997) piece-wise linear regression. As Patatoukas & Thomas (2011) suggest, the price deflator introduces loss and return variance effects that increase the upward bias in the AT measure, particularly among low price stocks. We confirm in Table 9 that deciles sorted on opening stock prices are inversely related with the AT measure in both the U.S. and international samples, consistent with the prediction of Patatoukas & Thomas (2011). In contrast, the association between the price deciles and the ACV measure appears to be more random in both the U.S. and international samples, and broadly in the opposite direction relative to the AT measure. This suggests that the sources of bias driven by the price deflator of the AT construct does not affect the ACV measure. The findings of the initial analyses reported in Tables 8 and 9 jointly suggest that, while the ACV and AT measures are both able to capture a common underlying accounting factor associated with CC, the former is less prone to Type I error than the latter in the presence of confounding effects unrelated to accounting decisions. As such, these findings help validate the ACV as a benchmark measure against the AT construct throughout our main analysis and hypotheses testing.

[Insert Table 9 here]

3.2.2. C_Score analyses

Table 10 presents our comparison of the ACV and AT measures in terms of their correlation with the C_Score measure. We first estimate *C_Score* on a firm-year basis for our U.S. and international samples, and then sort observations into deciles accordingly. Consistent with Khan & Watts (2009), we provide evidence that the AT measure increases with the *C_Score* deciles. The difference in the AT measure between the tenth and first deciles is 0.2640 and 0.1800 in the U.S. and international samples, respectively, where both differences are significant at the 1% level. In stark contrast to these findings, ACV and *C_Score* appear to have no clear association. In the U.S. sample, the level of ACV values across *C_Score* deciles is nearly flat, and the difference in ACV values between the tenth and first deciles is –0.0937, which is in the opposite direction to the AT measure. In the international sample, although the difference in ACV values between the extreme *C_Score* deciles is consistent in direction with that of the AT measure, there is, however, no interpretable pattern in ACV values across the *C_Score* deciles. The observation that *C_Score* correlates more closely with AT than ACV strongly suggests that the cross-sectional variation in the C_Score measure is likely to be largely driven by the bias effect, which confounds the AT measure than by the CC effect represented more clearly by the ACV measure. As such, Table 10 essentially urges caution against the application of the C_Score measure in empirical studies of CC in cross-sectional settings. The observation that changes in ACV yield broadly similar inferences in comparison to changes in *C_Score** in our re-examination of André et al. (2015) is attributable to the interrupted time series research design of that study.

[Insert Table 10 here]

As depicted in Equations 2 and 3, the C_Score measure is essentially a firm-year AT measure predicted by market-to-book, size, and leverage. Khan & Watts (2009) motivate these three components on the basis that they potentially capture important drivers of CC such as contracting, litigation, taxation, and regulation (Watts, 2003a, 2003b). Nevertheless, our test in Table 10 reveals that *C_Score* bears positive correlation only with the AT measure but not with its ACV counterpart. To gain further insights into this issue, we examine the association between each of the individual C_Score components separately with AT and ACV measures.

Tables 11, 12, and 13 report our findings for market-to-book, size, and leverage, respectively. Note that across these analyses, based on both our U.S. and international samples, the AT measure is indeed negatively associated with market-to-book and size, and positively related to leverage. The changes in AT value across deciles sorted on these three C_Score component variables are almost monotonic in the predicted directions. In stark contrast, there are no discernible and interpretable patterns for the ACV measure across these deciles. In general, the findings across Tables 11, 12, and 13 suggest that these firm characteristics bear limited empirical association with the CC effect (that is more clearly captured by the ACV measure) and are much more likely to be due to the bias effect that confounds the AT measure. This explains why the incorporation of these three conditioning effects does little to strengthen the ability of the C_Score measure to capture the CC effect, as we show in Table 10. More broadly, the findings across Tables 11, 12, and 13 imply that the intuition in the existing literature that market-to-book, size, and leverage are associated with CC also deserves re-examination.

[Insert Tables 11, 12, and 13]

4. Conclusions

Empirical evidence on CC holds an important place in the accounting literature, since it informs debates among academics and standard setters and helps to enrich our understanding of the role of accounting information in capital markets (Beyer, Cohen, Lys, & Beverly, 2010; Mora & Walker, 2015; Ruch & Taylor, 2015; Wang, Hógartaigh, & Zijl, 2009; Watts, 2003b). Since Basu (1997), the accounting literature has largely adopted the AT measure to draw inferences on the sources of costs and benefits of CC as well as the determinants and consequences of CC.

Nevertheless, as a parallel development in the CC literature, there is an on-going debate over the validity the AT measure. On the one hand, concerns have been raised about the tendency of the AT measure to be driven by non-accounting related factors and, accordingly, induce Type I error (Dietrich, Muller, & Riedl, 2007; Dutta & Patatoukas, 2017; Gigler & Hemmer, 2001; Patatoukas & Thomas, 2011, 2016). On the other hand, arguments to mitigate such concerns as well as solutions to adjust the AT construct have been suggested (Ball, Kothari, & Nikolaev, 2013a, 2013b; Collins, Hribar, & Tian, 2014).

Given the importance of CC as an issue in the accounting literature, and the on-going debate over the validity of the widely adopted Basu (1997) AT measure, we believe that it is necessary to heed the call of Ball (2016) and revisit inferences drawn from previous studies using more recently developed research methodologies. As such, unlike previous studies involved in the debate over the validity of the AT measure, we contribute to the CC literature by providing more direct evidence on whether and how bias in the AT measure varies across frequently adopted research settings. On the one hand, we provide evidence through the re-examination of Lobo & Zhou (2006) and André et al. (2015) that inferences based on the AT measure are likely to hold for interrupted time-series settings that examine the impact of exogenous changes in accounting policies. On the other hand, we provide evidence through the re-assessment of LaFond & Watts (2008) and Ball et al. (2008) that the AT measure is likely to generate biased inferences for cross-sectional settings that seek to identify the determinants of CC.

On a slightly more positive note, we do find that ACV and AT are correlated unconditionally, which is consistent with AT being capable of indicating the presence of CC where it exists. Thus, the problem with AT is that it is biased in favor of concluding CC even when it does not exist, i.e., Type 1 error. Our study has important implications for both past and future applications of the AT construct. For past applications, we suggest that prior claims based on cross-sectional applications that rely on the AT measure need to be reworked using alternative measures of CC that are less likely to suffer from the biases in AT highlighted in the literature. Prior results produced from interrupted time series settings are likely to be more robust. Future attempts to model the determinants and consequences of CC in cross-sectional settings should avoid use of the AT construct.

References:

Ahmed, A., Billings, B., Morton, R., & Stanford-Harris, M. (2002). The role of accounting conservatism in mitigating bondholder-shareholder conflicts over dividend policy and in reducing debt costs. *The Accounting Review*, *77*(4), 867–890.

André, P., Filip, A., & Paugam, L. (2015). The effect of mandatory IFRS adoption on conditional conservatism in Europe. *Journal of Business Finance & Accounting*, *42*((3) & (4)), 482–514.

Ball, R. (2016). Why we do international accounting research. *Journal of International Accounting Research*, *15*(2), 1–6.

Ball, R., Kothari, S. P., & Nikolaev, V. (2013a). Econometrics of the Basu asymmetric timeliness coefficient and accounting conservatism. *Journal of Accounting Research*, *51*(5), 1071–1097.

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

Ball, R., Robin, A., & Sadka, G. (2008). Is financial reporting shaped by equity markets? An international study of timeliness and conservatism. *Review of Accounting Studiesi*, *13*, 168–205.

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

Beaver, W., & Ryan, S. (2005). Conditional and Unconditional Conservatism:Concepts and Modeling. *Review of Accounting Studies*, *10*(2–3), 269–309.

Beyer, A., Cohen, D., Lys, T., & Beverly, W. (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics*, *50*, 296–343.

Bliss, J. H. (1924). *Management through accounts*. New York, NY: The Ronald Press Co.

Casella, G., & Berger, R. (2002). *Statistical Inference* (2nd ed.). Duxbury.

Coates, J., & Srinivasan, S. (2014). SOX after Ten Years: A Multidisciplinary Review. *Accounting Horizons*, *28*(3), 627–671.

Collins, D., Hribar, P., & Tian, X. (2014). Cash flow asymmetry: Causes and implications for conditional conservatism research. *Journal of Accounting and Economics*, *58*(2), 173–200.

Cooper, S. (2015). A tale of "prudence." Retrieved from http://www.ifrs.org/Investor-resources/Investor-perspectives-2/Documents/Prudence_Investor-Perspective_Conceptual-FW.PDF

Daske, H., Hail, L., Leuz, C., & Verdi, R. (2013). Adopting a label: Heterogeneity in the economic consequences around IAS/IFRS adoptions. *Journal of Accounting Research*, *51*(3), 495–547.

De George, E., Li, X., & Shivakumar, L. (2016). A review of the IFRS adoption literature. *Review of Accounting Studies*, *21*(3), 898–1004.

Dietrich, D., Muller, K., & Riedl, E. (2007). Asymmetric timeliness tests of accounting conservatism. *Review of Accounting Studies*, *12*(1), 95–124.

Dutta, S., & Patatoukas, P. (2017). Identifying conditional conservatism in financial accounting data: theory and evidence. *The Accounting Review*, (Forthcoming).

Easley, D., Hvidkjaer, S., & O'Hara, M. (2002). Is Information Risk a Determinant of Asset Returns? *The Journal of Finance*, *57*(5), 2185–2221.

Easley, D., & O'Hara, M. (1992). Time and process of security price adjustment. *Journal of Finance*, *47*, 577–605.

Fama, E., & French, K. (2002). Testing Tradeoff and Pecking Order Predictions About

Dividends and Debt. *The Review of Financial Studies*, *15*(1), 1–33.
Fama, E., & MacBeth, J. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636.
Feiveson, A. (1999). FAQ: What is the delta method and how is it used to estimate the standard error of a transformed parameter? Retrieved from http://www.stata.com/support/faqs/stat/deltam.html
Financial Reporting Standards Board (FASB). (1980). *Qualitative Characteristics of Accounting Information*. Statement of Financial Accounting Concepts No. 2. FASB, Norwalk, CT.
Financial Reporting Standards Board (FASB). (2005). *Board meeting Handout: Conceptual framework*. July 27. Norwalk, CT: FASB.
García Lara, J. M., García Osma, B., & Penalva, F. (2009). Accounting conservatism and corporate governance. *Review of Accounting Studies*, *14*, 161–201.
Gigler, F., & Hemmer, T. (2001). Conservatism, optimal disclosure policy, and the timeliness of financial reports. *The Accounting Review*, *76*(4), 471–493.
Khan, M., & Watts, R. (2009). Estimation and empirical properties of a firm-year measure of accounting conservatism. *Journal of Accounting and Economics*, *48*(2–3), 132–150.
Kothari, S. P. (2001). Capital accounting research in accounting. *Journal of Accounting and Economics*, *31*, 105–231.
La Porta, R., Lopez-De-Silanes, F., Shleifer, A., & Vishny, R. (1997). Legal Determinants of External Financing. *The Journal of Finance*, *LII*(3).
La Porta, R., Lopez-De-Silanes, F., Shleifer, A., & Vishny, R. (1998). Law and Finance. *Journal of Political Economy*, *106*(6), 1113–1155.
LaFond, R., & Roychowdhury, S. (2008). Managerial Ownership and Accounting Conservatism. *Journal of Accounting Research*, *46*(1), 101–135.
LaFond, R., & Watts, R. (2008). The information role of conservatism. *The Accounting Review*, *83*(2), 447–478.
Leuz, C., & Wysocki, P. (2016). The Economics of Disclosure and Financial Reporting Regulation: Evidence and Suggestions for Future Research. *Journal of Accounting Research*, *54*(2), 525–622.
Lobo, G., & Zhou, J. (2006). Did Conservatism in Financial Reporting Increase after the Sarbanes-Oxley Act and CEO/CFO Certification of Financial Statements? *Accounting Horizons*, *20*(1), 57–73.
Mora, A., & Walker, M. (2015). The implications of research on accounting conservatism for accounting standard setting. *Accounting and Business Research*, *45*(5), 620–650.
Patatoukas, P., & Thomas, J. (2011). More Evidence of Bias in the Differential Timeliness Measure of Conditional Conservatism. *The Accounting Review*, *86*(5), 1765–1793.
Patatoukas, P., & Thomas, J. (2016). Placebo tests of conditional conservatism. *The Accounting Review*, *91*(2), 625–648.
Pope, P., & Walker, M. (1999). International differences in the timeliness, conservatism, and classification of earnings. *Journal of Accounting Research*, *37*, 53–87.
Ruch, G., & Taylor, G. (2015). Accounting conservatism: A review of the literature. *Journal of Accounting Literature*, *34*, 17–38.
Ryan, S. (2006). Identifying Conditional Conservatism. *European Accounting Review*, *15*(4), 511–525.
Sterling, R. (1976). Conservatism: The fundamental principle of valuation in traditional accounting. *ABACUS*, *3*(2), 109–122.

Wang, R. Z., Hógartaigh, C. Ó., & Zijl, T. V. (2009). Measures of accounting conservatism: A construct validity perspective. *Journal of Accounting Literature*, *28*, 165–203.

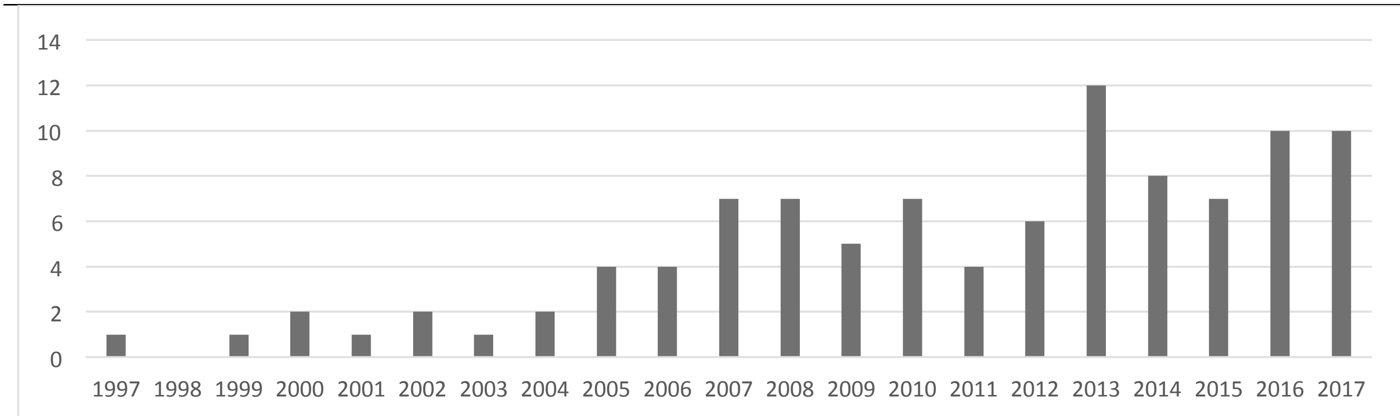
Watts, R. (2003a). Conservatism in accounting, Part I: explanations and implications. *Accounting Horizons*, *7*(3), 207–221.

Watts, R. (2003b). Conservatism in accounting, Part II: evidence and research opportunities. *Accounting Horizons*, *7*(4), 287–301.

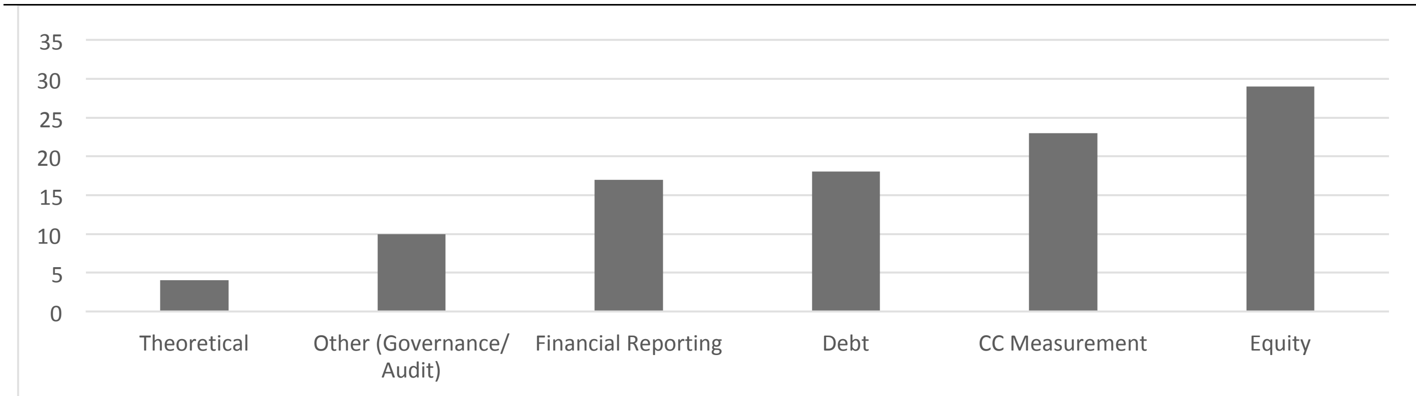
Zhang, J. (2008). The contracting benefits of accounting conservatism to lenders and borrowers. *Journal of Accounting and Economics*, *45*(1), 27–54.

Appendix 1: Literature analysis

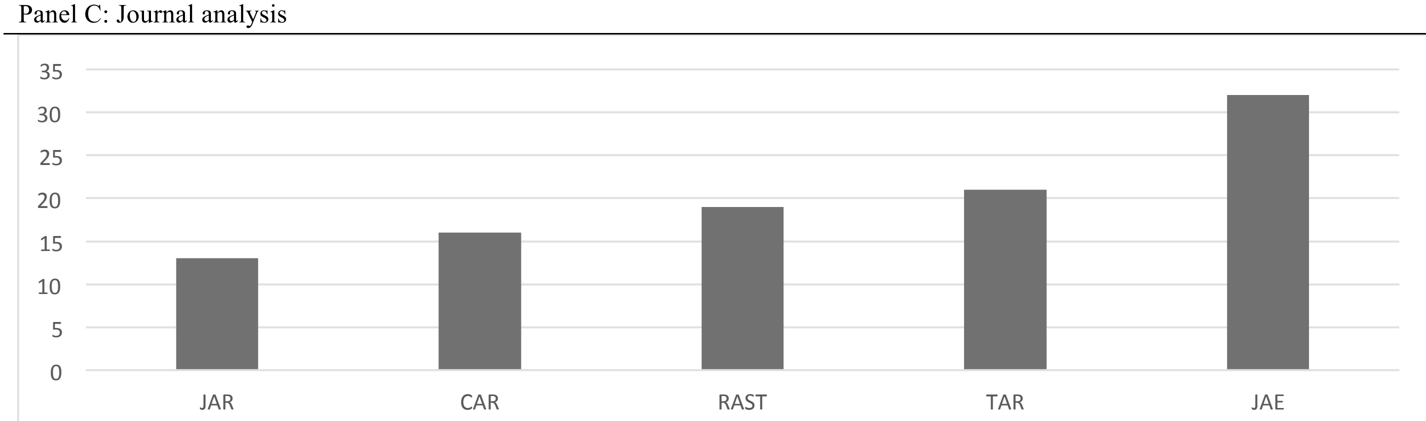
Panel A: Time trend analysis



Panel B: Topic analysis



Panel C: Journal analysis



This appendix presents the literature analysis of conditional conservatism (CC) studies that applied the Basu (1997) asymmetric timeliness (AT) approach (including the C_Score of Khan & Watts, 2009) either in their empirical analyses or theoretical models. We cover studies published over the period of 1997 to 2017 across five leading accounting journals, including (in alphabetical order) *Contemporary Accounting Research* (CAR), *Journal of Accounting and Economics* (JAE), *Journal of Accounting Research* (JAR), *Review of Accounting Studies* (RAST), and *The Accounting Review* (TAR). We exclude discussion papers associated with some of the studies. The vertical axis in each panel indicates the number of publications. Panels A, B, and C depict the number of publications each year, topic, and journal respectively.

Table 1: Variable definitions

Variable	Definition
ACV measure and other factors (Dutta & Patatoukas, 2017)	
ACV	ACV measure of CC estimated as the variance of bad news earnings divided by variance of good news earnings. ACV = (X_{it} | *RD*=1) / (X_{it} | *RD*=0).
ARetDist	Economic factor confounding effect based on asymmetry in returns distributions, estimated as the variance of positive abnormal returns divided by the variance of negative abnormal returns.
SI	Accounting factor associated with conservatism based on negative special items scaled by lagged market value of equity (missing and positive values are set to 0).
Lobo & Zhou (2006) inference re-examination	
RD	Bad news dummy variable that equals 1 if *RET* is negative, and 0 otherwise.
RET	Abnormal stock return calculated on an annual basis at the end of the fiscal year and adjusted for the country-year average of returns.
SOX	SOX indicator that equals 1 for fiscal years ending in August 2002 or beyond, and 0 otherwise.
X	Income before extraordinary items deflated by lagged market value of equity.
Andre, Filip, & Paugam (2015) inference re-examination	
BETA	CAPM beta estimated over the last 36 months and estimated from the time-series regressions of monthly stock returns corresponding period market returns.
*C_Score**	Modified version of the C_Score measure estimated with the addition of *BETA* and *UCC* in addition to *MTB*, *SIZE* and *LEV* as prediction components.
IFRS	IFRS indicator that equals 1 for fiscal years ending 2005 or beyond, and 0 otherwise.
*LEV**	Leverage calculated as the sum of long-term and short-term debt divided by market value of equity.
MTB	Market-to-book ratio calculated as market value of equity to book value of equity.
RET	Abnormal stock return calculated on an annual basis at the end of the fiscal year and adjusted for the country-year average of returns.
SIZE	Firm size calculated as the natural logarithm of market value of equity.
UCC	Unconditional conservatism estimated using the residual of annual cross-sectional regressions of *MTB* on raw returns, intangibles assets, property plant and equipment, capital expenditures, percentage change in sales, return on equity, price volatility, leverage ratio and firm size.
LaFond & Watts (2008) inference re-examination	
HighPIN	High PIN indicator that equals 1 for observations with *PIN* score above sample median, and 0 otherwise.
PIN	the probability of an information-based trade derived from a structural market microstructure model (see Easley, Hvidkjaer, & O'Hara, 2002).
RD	Bad news dummy variable that equals 1 if *RET* is negative, and 0 otherwise.
RET	Abnormal stock return calculated on an annual basis at the end of the fiscal year and adjusted for the country-year average of returns.
X	Income before extraordinary items deflated by lagged market value of equity.

(continue next page)

Table 1: *(continued from previous page)*

Ball, Robin, & Sadka (2008) inference re-examination	
B0	Country-level constant term in the Basu (1997) AT regression.
B1	Country-level coefficient on *RD* in the Basu (1997) AT regression.
B2	Country-level coefficient on *RET* in the Basu (1997) AT regression.
B3	Country-level coefficient on *RD*×*RET* in the Basu (1997) AT regression.
BTM	Country-level book-to-market calculated as the median for all firm and years in each country.
CORRP	Country-level corruption index based on La Porta, Lopez-De-Silanes, Shleifer, and Vishny (1997, 1998).
CRED	Country-level creditor rights index based on La Porta, Lopez-De-Silanes, Shleifer, and Vishny (1997, 1998).
DEBT	Country-level debt market importance, calculated as the sum of bank debt of the private sector and outstanding non-financial bonds divided by GNP in 1994, based on La Porta, Lopez-De-Silanes, Shleifer, and Vishny (1997, 1998).
ENGLISH	English legal origin indicator that equals 1 for such countries, and 0 otherwise.
EQUITY	Country-level equity market importance, calculated as stock market capitalization held by minorities divided by GNP in 1994, based on La Porta, Lopez-De-Silanes, Shleifer, and Vishny (1997, 1998).
FRENCH	French legal origin indicator that equals 1 for such countries, and 0 otherwise.
HighDEBT	High debt market importance indicator that equals 1 for countries with *DEBT* value above sample median, and 0 otherwise.
LAW	Country-level law and order index based on La Porta, Lopez-De-Silanes, Shleifer, and Vishny (1997, 1998).
SCAND	Scandinavian legal origin indicator that equals 1 for such countries, and 0 otherwise.
C_Score and components (Khan & Watts, 2009)	
RD	Bad news dummy variable that equals 1 if *RET* is negative, and 0 otherwise.
RET	Abnormal stock return calculated on an annual basis at the end of the fiscal year and adjusted for the country-year average of returns.
X	Income before extraordinary items deflated by lagged market value of equity.
C_Score	C_Score measure of CC estimated as the predicted value of AT regression conditional on *MTB*, *SIZE*, and *LEV*.
LEV	Leverage calculated as total liabilities divided by the sum of market value of equity and total liabilities.
MTB	Market-to-book ratio calculated as the ratio of market value of equity to book value of equity.
SIZE	Firm size measured as the natural logarithm of market value of equity.

This table presents the definition of all variables applied in our main and additional analyses. We follow the original definition from the designated previous studies whenever possible, and make adjustments whenever necessary and appropriate in order to adapt to our available sample and data.

Table 2: Lobo & Zhou (2006) inference re-examination (main analysis to test hypothesis H1)

Panel A: Summary statistics

	Observations	Mean	Stdev.	Q1	Median	Q3
RET	11,244	0.0128	0.6170	−0.3398	−0.0738	0.2119
X	11,244	0.0150	0.1602	−0.0054	0.0472	0.0856
SI	11,244	−0.0211	0.0656	−0.0086	0.0000	0.0000

Panel B: SOX effect through AT regression based on *X* or *SI* as dependent variable

	X	*SI*
RD	0.0176***	0.0086***
	(2.94)	(3.34)
RET	−0.0043	−0.0032*
	(−0.94)	(−1.65)
RD×*RET*	0.2289***	0.0567***
	(20.75)	(11.94)
SOX	−0.0115**	0.0045*
	(−2.11)	(1.91)
SOX×*RD*	−0.0014	−0.0011
	(−0.17)	(−0.30)
SOX×*RET*	0.0028	−0.0053*
	(0.41)	(−1.79)
SOX×*RD*×*RET*	0.0436***	0.0208***
	(2.63)	(2.92)
Intercept	0.0605***	−0.0147***
	(15.42)	(−8.74)
Adjusted R^2	12.93%	3.83%
Observations	11,244	11,244

Panel C: Comparison of ACV measure in pre- and post-SOX periods

Groups	Observations	Average *X*	Stdev.
SOX = 0, *RD* = 1	2,380	0.0582	0.1509
SOX = 0, *RD* = 0	3,242	−0.0050	0.1628
SOX = 1, *RD* = 1	2,459	0.0483	0.1424
SOX = 1, *RD* = 0	3,163	−0.0230	0.1647
ACV | (*SOX* = 0) = 1.16	Chi^2 = 6.54		
ACV | (*SOX* = 1) = 1.34	*p*-value = 0.010		

Panel D: Comparison of *ARetDist* in pre- and post-SOX periods

Groups	Observations	Average *RET*	Stdev.
SOX = 0, *RD* = 1	2,380	0.5366	0.6728
SOX = 0, *RD* = 0	3,242	−0.3702	0.2615
SOX = 1, *RD* = 1	2,459	0.4458	0.5883
SOX = 1, *RD* = 0	3,163	−0.3254	0.2373
ARetDist | (*SOX* = 0) = 6.62	Chi^2 = 1.91		
ARetDist | (*SOX* = 1) = 6.15	*p*-value = 0.167		

This table presents replication and re-examination of the inference from Lobo & Zhou (2006) for our main analysis to test hypothesis H1 based on an interrupted time-series setting. Panel A reports summary statistics for the variables used in the analysis based on a U.S. sample over the period of 2000-2004. Panel B reports the conditioning effect of SOX through Basu (1997) AT regression separately with *X* and *SI* as dependent variables, with *t*-statistics in parentheses calculated based on clustered standard errors at the firm level. Panel C reports the summary statistics in terms of *X* across four observation groups sorted on *SOX* and *RD*, the ACV measure across pre- and post-SOX periods, and statistical significance tests of the difference between these two periods. Panel D reports the summary statistics in terms of *RET* across four observation groups sorted on *SOX* and *RD*, the *ARetDist* values across pre- and post-SOX periods, and statistical significance tests of the difference between these two periods. All variables are defined in Table 1. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 3: Andre, Filip, & Paugam (2015) inference re-examination (main analysis to test hypothesis H1)

Panel A: Summary statistics

	Observations	Mean	Stdev.	Q1	Median	Q3
RET	12,731	0.0252	0.4474	−0.1925	0.0429	0.2786
*C_Score**	12,731	0.0500	0.0669	0.0084	0.0477	0.0914
SI	12,731	−0.0198	0.0782	−0.0029	0.0000	0.0000
SIZE	12,731	6.4201	2.0122	4.9541	6.2554	7.6563
MB	12,731	2.3350	2.8101	0.9536	1.6106	2.6521
*LEV**	12,731	0.1207	0.1239	0.0146	0.0874	0.1844
BETA	12,731	0.9967	0.8531	0.4488	0.9071	1.4557
UCC	12,731	−0.5921	2.6226	−2.1991	−0.7703	0.7368

Panel B: IFRS effect though AT regressions based on *C_Score** or *SI* as dependent variable

	*C_Score**	*SI*
IFRS	-0.0174^{***}	0.0058^{***}
	(−22.58)	(3.66)
SIZE	-0.0238^{***}	-0.0038^{***}
	(−94.61)	(−4.46)
MB	0.0099^{***}	0.0047^{***}
	(38.51)	(7.17)
*LEV**	0.1631^{***}	−0.0073
	(35.79)	(−0.58)
BETA	-0.0142^{***}	-0.0060^{***}
	(−26.12)	(−5.52)
UCC	-0.0215^{***}	-0.0036^{***}
	(−84.67)	(−5.53)
Intercept	0.1709^{***}	−0.0045
	(144.44)	(−1.09)
Adjusted R^2	54.85%	2.82%
Observations	12,731	12,731

Panel C: Comparison of ACV measure in pre- and post-IFRS periods

Groups	Observations	Average *X*	Stdev.
IFRS = 0, *RD* = 0	3,089	0.0921	0.2064
IFRS = 0, *RD* = 1	2,431	−0.0006	0.2309
IFRS = 1, *RD* = 0	3,929	0.0897	0.1866
IFRS = 1, *RD* = 1	3,282	0.0104	0.1774
ACV | (*IFRS* = 0) = 1.25	Chi^2= 35.2		
ACV | (*IFRS* = 1) = 0.91	*p*-value = 0.000		

Panel D: Comparison of *ARetDist* in pre- and post-IFRS periods

Groups	Observations	Average *RET*	Stdev.
IFRS = 0, *RD* = 0	3,089	0.3414	0.2922
IFRS = 0, *RD* = 1	2,431	−0.3843	0.4170
IFRS = 1, *RD* = 0	3,929	0.2959	0.2539
IFRS = 1, *RD* = 1	3,282	−0.2930	0.2875
ARetDist | (*IFRS* = 0) = 0.49	Chi^2 = 80.5		
ARetDist | (*IFRS* = 1) = 0.78	*p*-value = 0.000		

This table presents replication and re-examination of the inference from Andre, Filip, & Paugnam (2015) for our main analysis to test hypothesis H1 based on an interrupted time-series setting. Panel A reports summary statistics for the variables used in the analysis based on an international sample of 16 European Union countries over the period of 2000-2010. Panel B reports the conditioning effect of IFRS through regression separately with *C_Score** and *SI* as dependent variables, with *t*-statistics in parentheses calculated based on clustered standard errors at the firm level. Panel C reports the summary statistics in terms of *X* across four observation groups sorted on *IFRS* and *RD*, the ACV measure across pre- and post-IFRS periods, and statistical significance tests of the difference between these two periods. Panel D reports the summary statistics in terms of *RET* across four observation groups sorted on *IFRS* and *RD*, the *ARetDist* values across pre- and post-IFRS periods, and statistical significance tests of the difference between these two periods. All variables are defined in Table 1. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 4: LaFond & Watts (2008) inference re-examination (main analysis to test hypothesis H2)

Panel A: Summary statistics

	Observations	Mean	Stdev	Q1	Median	Q3
X	19,831	0.0353	0.1632	0.0224	0.0610	0.0965
RET	19,831	0.0000	0.3873	−0.2291	−0.0346	0.1754
PIN	19,831	0.1955	0.0786	0.1395	0.1825	0.2354
SI	19,831	−0.0172	0.0608	−0.0024	0.0000	0.0000

Panel B: PIN effect through AT regression based on *X* or *SI* as dependent variable

	X	*SI*
RD	0.0063	0.0035
	(0.72)	(1.19)
RET	0.0181	−0.0061
	(0.79)	(−0.69)
RD×RET	0.1208***	0.0610***
	(3.18)	(3.73)
PIN	−0.0223	−0.0109
	(−1.11)	(−1.42)
PIN×RD	0.0867*	−0.0018
	(2.10)	(−0.15)
PIN×RET	0.0126	0.0148
	(0.13)	(0.52)
PIN×RD×RET	1.0444***	0.0127
	(5.33)	(0.23)
Intercept	0.0738***	−0.0077***
	(11.72)	(−5.47)
Adjusted R^2	21.48%	4.50%
Observations	19,831	19,831

Panel C: Comparison of ACV measure across high and low PIN groups

Groups	Observations	Average *X*	Stdev.
HighPIN = 0, *RD* = 0	4,573	0.0758	0.0947
HighPIN = 0, *RD* = 1	5,346	0.0313	0.1231
HighPIN = 1, *RD* = 0	4,409	0.0723	0.1625
HighPIN = 1, *RD* = 1	5,503	−0.0240	0.2159
ACV | (*HighPIN* = 0) = 1.69	Chi^2 = 1.18		
ACV | (*HighPIN* = 1) = 1.76	*p*-value = 0.276		

Panel D: Comparison of *ARetDist* across high and low PIN groups

Groups	Observations	Average *RET*	Stdev.
HighPIN = 0, *RD* = 0	4,573	0.2449	0.2511
HighPIN = 0, *RD* = 1	5,346	−0.2199	0.1748
HighPIN = 1, *RD* = 0	4,409	0.3691	0.3979
HighPIN = 1, *RD* = 1	5,503	−0.2857	0.2101
ARetDist | (*HighPIN* = 0) = 2.06	Chi^2 = 166.03		
ARetDist | (*HighPIN* = 1) = 3.58	*p*-value = 0.000		

This table presents replication and re-examination of the inference from LaFond & Watts (2008) for our main analysis to test hypothesis H2 based on a cross-sectional setting. Panel A reports summary statistics for the variables used in the analysis based on a U.S. sample over the period of 1983−2001. Panel B reports the conditioning effect of *PIN* through Basu (1997) AT regression separately with *X* and *SI* as dependent variables, with *t*-statistics in parentheses calculated based on Fama & MacBeth (1973) standard errors. Panel C reports the summary statistics in terms of *X* across four observation groups sorted on *HighPIN* and *RD*, the ACV measure across higher and lower *HighPIN* groups, and statistical significance tests of the difference between these two groups. Panel D reports the summary statistics in terms of *RET* across four observation groups sorted on *HighPIN* and *RD*, the *ARetDist* values across high and low *HighPIN* groups, and statistical significance tests of the difference between these two groups. All variables are defined in Table 1. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 5: Summary statistics for Ball, Robin, & Sadka (2008) inference re-examination

Panel A: Summary statistics

	Observations	Mean	Stdev.	Q1	Median	Q3
X	96,298	−0.0050	0.2518	−0.0091	0.0322	0.0729
RET	96,298	−0.0075	0.5523	−0.3018	−0.0281	0.2327
SI	96,298	−0.0145	0.0830	0.0000	0.0000	0.0000

Panel B: Country-level variables

Country	*B0*	*B1*	*B2*	*B3*	ACV	*SI*	*DEBT*	*EQUITY*	*LAW*	*CORRP*	*CRED*	*BTM*
Australia	0.017	−0.027	0.004	0.272	1.90	−0.0343	0.76	0.49	10.00	8.52	1	0.646
Canada	0.051	0.002	−0.001	0.293	1.64	−0.0280	0.72	0.39	10.00	10.00	1	0.657
Malaysia	−0.012	−0.010	−0.023	0.160	1.19	−0.0149	0.84	1.48	6.78	7.38	4	0.831
Singapore	0.016	0.004	0.087	0.013	1.00	−0.0071	0.60	1.18	8.57	8.22	3	0.861
South Africa	0.101	−0.001	0.147	−0.017	0.70	−0.0149	0.93	1.45	4.42	8.92	4	0.722
Thailand	0.030	−0.016	0.003	0.365	0.86	−0.0021	0.93	0.56	6.25	5.18	3	1.003
UK	0.041	−0.018	−0.026	0.193	1.48	−0.0183	1.13	1.00	8.57	9.10	4	0.515
USA	0.037	0.010	−0.023	0.203	1.30	−0.0196	0.81	0.58	10.00	8.63	1	0.461
Brazil	0.043	−0.061	−0.019	0.027	0.61	−0.0001	0.39	0.18	6.32	6.32	1	0.003
Chile	0.061	0.002	0.098	0.116	1.29	−0.0023	0.63	0.80	7.02	5.30	2	0.848
France	0.043	−0.007	0.022	0.216	2.33	−0.0176	0.96	0.23	8.98	9.05	0	0.690
Indonesia	−0.021	−0.006	0.045	−0.025	1.09	−0.0016	0.42	0.15	3.98	2.15	4	0.767
Italy	0.054	0.000	−0.019	0.129	1.04	−0.0061	0.55	0.08	8.33	6.13	2	0.990
Netherlands	0.079	−0.005	−0.036	0.221	1.82	−0.0058	1.08	0.52	10.00	10.00	2	0.565
Spain	0.119	−0.018	−0.046	0.132	0.61	−0.0843	0.75	0.17	7.80	7.38	2	0.769
Germany	0.012	−0.039	0.023	0.212	1.72	−0.0188	1.12	0.13	9.23	8.93	3	0.693
Japan	0.009	−0.010	0.045	0.081	2.00	−0.0001	1.22	0.62	8.98	8.52	2	0.844
South Korea	0.056	−0.039	0.239	0.032	1.25	−0.0001	0.74	0.44	5.35	5.30	3	2.000
Denmark	0.088	−0.028	0.048	0.127	1.36	−0.0058	0.34	0.21	10.00	10.00	3	0.853
Finland	0.093	−0.024	0.075	0.071	0.78	−0.0092	0.75	0.25	10.00	10.00	1	0.829
Norway	0.052	−0.011	−0.016	0.230	1.81	−0.0168	0.64	0.22	10.00	10.00	2	0.651
Sweden	0.022	0.011	0.078	0.270	4.16	−0.0237	0.55	0.51	10.00	10.00	2	0.672
Mean	0.051	0.005	0.064	0.168	1.452	−0.015	0.766	0.529	8.208	7.956	2.273	0.767
Median	0.047	0.003	0.047	0.177	1.295	−0.012	0.750	0.465	8.775	8.575	2.000	0.744
Stdev.	0.031	0.005	0.067	0.096	0.769	0.018	0.245	0.416	1.947	2.102	1.162	0.346

This table presents the summary statistics for the replication and re-examination of Ball, Robin, and Sadka (2008) for our main analysis to test hypothesis H2 based on a cross-sectional setting. Panel A reports summary statistics for *X*, *RET*, and *SI* variables for the international sample of 22 countries over the period of 1992-2003. Panel B reports the country-level variables, including the Basu (1997) coefficients, ACV and *SI* values, debt and equity market importance, and control variables. All variables are defined in Table 1.

Table 6: Ball, Robin, and Sadka (2008) inference re-examination (main analysis to test hypothesis H2)

Panel A: Debt vs. equity market effects based on AT measure (*B3*) as dependent variable

DEBT	0.2652**	0.2013*	0.2601*	0.2531**	0.2561**	0.2000*	0.2685*	0.2569*	0.2571*
	(2.79)	(1.87)	(2.12)	(2.52)	(2.22)	(1.80)	(2.12)	(2.13)	(2.06)
EQUITY	−0.1837***	−0.1451*	−0.1832**	−0.1648**	−0.1162	−0.1491*	−0.1617*	−0.1148	−0.1198
	(−3.00)	(−2.13)	(−2.88)	(−2.26)	(−1.63)	(−2.01)	(−2.10)	(−1.45)	(−1.45)
ENGLISH	0.2173***	0.1828**	0.2155**	0.2053**	0.1902**	0.1844**	0.2088**	0.1899**	0.2047**
	(3.18)	(2.50)	(2.86)	(2.77)	(2.63)	(2.42)	(2.67)	(2.52)	(2.38)
FRENCH	0.0825	0.0697	0.0817	0.0718	0.0752	0.0729	0.0719	0.0743	0.0922
	(1.24)	(1.05)	(1.17)	(1.01)	(1.15)	(1.03)	(0.98)	(1.06)	(1.09)
SCAND	0.1691**	0.1101	0.1644	0.1587*	0.1581	0.1086	0.1726	0.1589	0.1656
	(2.19)	(1.22)	(1.59)	(1.94)	(1.62)	(1.16)	(1.61)	(1.55)	(1.55)
LAW		0.0156			0.032	0.0172		0.0317	0.0314
		(1.21)			(1.72)	(1.08)		(1.57)	(1.50)
CORRP			0.001		−0.0228		−0.0035	−0.023	−0.0203
			(0.07)		(−1.21)		(−0.21)	(−1.15)	(−0.94)
CRED				−0.0101		0.0042	−0.0127	−0.0012	−0.001
				(−0.50)		(0.18)	(−0.53)	(−0.05)	(−0.04)
BTM									0.0301
									(0.41)
Intercept	−0.0911	−0.1635	−0.0933	−0.0592	−0.1865	−0.184	−0.0428	−0.1811	−0.2338
	(−0.85)	(−1.34)	(−0.81)	(−0.46)	(−1.54)	(−1.07)	(−0.28)	(−1.07)	(−1.08)
Adjusted R^2	52.66%	56.90%	52.68%	53.45%	60.97%	57.00%	53.59%	60.97%	61.52%
Observations	22	22	22	22	22	22	22	22	22

Panel B: Debt vs. equity market effects based on *SI* as dependent variable

DEBT	−0.0202	−0.0139	−0.0029	−0.0196	−0.0031	−0.0136	−0.0012	−0.0016	−0.0016
	(−0.92)	(−0.54)	(−0.11)	(−0.84)	(−0.11)	(−0.51)	(−0.04)	(−0.05)	(−0.05)
EQUITY	0.0164	0.0126	0.0148	0.0154	0.0183	0.0137	0.0193	0.0209	0.0224
	(1.17)	(0.78)	(1.05)	(0.91)	(1.06)	(0.77)	(1.13)	(1.09)	(1.13)
ENGLISH	−0.0229	−0.0196	−0.0168	−0.0223	−0.0181	−0.02	−0.0182	−0.0188	−0.0232
	(−1.46)	(−1.12)	(−1.01)	(−1.30)	(−1.03)	(−1.10)	(−1.05)	(−1.03)	(−1.13)
FRENCH	−0.0159	−0.0147	−0.0132	−0.0154	−0.0136	−0.0155	−0.0153	−0.0152	−0.0205
	(−1.04)	(−0.93)	(−0.86)	(−0.93)	(−0.86)	(−0.92)	(−0.94)	(−0.90)	(−1.01)
SCAND	−0.0151	−0.0094	0.0004	−0.0146	0.0001	−0.009	0.0021	0.0016	−0.0003
	(−0.85)	(−0.43)	(0.02)	(−0.77)	(0.00)	(−0.40)	(0.09)	(0.07)	(−0.01)
LAW		−0.0015			0.0017	−0.0019		0.0011	0.0012
		(−0.50)			(0.37)	(−0.51)		(0.23)	(0.23)
CORRP			−0.0032		−0.0045		−0.0042	−0.0048	−0.0056
			(−1.06)		(−0.98)		(−1.14)	(−1.00)	(−1.08)
CRED				0.0005		−0.0011	−0.0027	−0.0022	−0.0023
				(0.11)		(−0.19)	(−0.49)	(−0.38)	(−0.38)
BTM									−0.0089
									(−0.51)
Intercept	0.0079	0.015	0.0153	0.0063	0.0104	0.0204	0.0258	0.021	0.0367
	(0.32)	(0.51)	(0.60)	(0.21)	(0.35)	(0.50)	(0.76)	(0.51)	(0.70)
Adjusted R^2	14.37%	15.77%	20.37%	14.44%	21.15%	16.00%	21.73%	22.04%	23.69%
Observations	22	22	22	22	22	22	22	22	22

(*continue next page*)

Table 6: *(continue from previous page)*

Panel C: Comparison of ACV measure across debt and equity markets			
Groups	Observations	Average *X*	Stdev.
HighDEBT = 0, *RD* = 0	6,494	0.0656	0.3181
HighDEBT = 0, *RD* = 1	6,842	−0.0440	0.3560
HighDEBT = 1, *RD* = 0	38,781	0.0262	0.2189
HighDEBT = 1, *RD* = 1	44,181	−0.0368	0.2425
ACV | (*HighDEBT* = 0) = 1.25	$Chi^2 = 0.60$		
ACV | (*HighDEBT* = 1) = 1.23	*p*-value = 0.437		

Panel D: Comparison of *ARetDist* across debt and equity markets			
Groups	Observations	Average *RET*	Stdev.
HighDEBT = 0, *RD* = 0	6,494	0.4026	0.5517
HighDEBT = 0, *RD* = 1	6,842	−0.3942	0.4570
HighDEBT = 1, *RD* = 0	38,781	0.3839	0.4762
HighDEBT = 1, *RD* = 1	44,181	−0.3514	0.2970
ARetDist | (*HighDEBT* = 0) = 1.46	Chi2 = 647.59		
ARetDist | (*HighDEBT* = 1) = 2.57	p-value = 0.000		

This table presents replication and re-examination of the inference from Ball, Robin, and Sadka (2008) for our main analysis to test hypothesis H2 based on a cross-sectional setting. Panels A and B report results based on regressions with AT measure (*B3*) and *SI* respectively as the dependent variable, with *t*-statistics in parentheses. Panel C reports the summary statistics in terms of *X* across four observation groups sorted on *HighDebt* and *RD*, the ACV measure across high and low *HighDebt* groups, and statistical significance tests of the difference between these two groups. Panel D reports the summary statistics in terms of *RET* across four observation groups sorted on *HighDebt* and *RD*, the *ARetDist* values across high and low *HighDebt* groups, and statistical significance tests of the difference between these two groups. All variables are defined in Table 1. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 7: Samples and summary statistics for additional analyses

Panel A: Number of firm-year observations across countries

Australia	15,291	France	8,978	Netherlands	2,269	Sweden	4,145
Austria	1,026	Germany	8,868	New Zealand	1,295	Switzerland	2,855
Belgium	1,451	Greece	2,159	Norway	2,277	UK	20,986
Canada	4,574	Ireland	605	Singapore	6,655	USA	70,033
Denmark	1,809	Italy	2,898	South Africa	3,297		
Finland	1,824	Japan	50,815	Spain	1,793		

Panel B: Summary statistics for U.S. sample

	Observations	Mean	Stdev.	Q1	Median	Q3
X	70,033	0.0107	0.1346	−0.0074	0.0409	0.0714
RET	70,033	−0.0067	0.5649	−0.3392	−0.0857	0.1931
MTB	70,033	3.1855	4.0001	1.3118	2.0742	3.4985
SIZE	70,033	5.8650	2.0503	4.3644	5.8200	7.2620
LEV	70,033	0.3276	0.2181	0.1421	0.2934	0.4872
C_Score	70,033	0.0946	0.1074	0.0348	0.0842	0.1397

Panel C: Summary statistics for international sample

	Observations	Mean	Stdev.	Q1	Median	Q3
X	215,903	0.0123	0.2822	−0.0105	0.0406	0.0792
RET	215,903	0.0004	0.4450	−0.2497	−0.0105	0.2276
MTB	215,903	2.4499	3.2104	0.8991	1.5506	2.7089
SIZE	215,903	6.4456	2.9050	4.1851	6.1918	8.5714
LEV	215,903	0.4129	0.2451	0.2068	0.4009	0.6046
C_Score	215,903	0.0569	0.1320	−0.0194	0.0553	0.1271

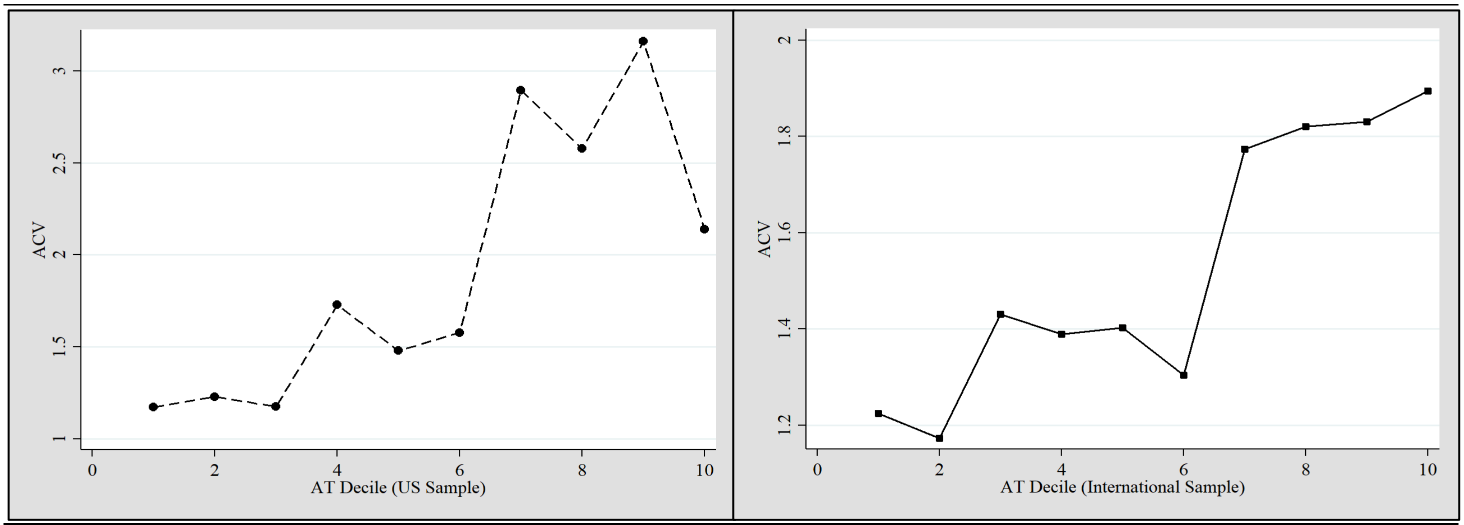
This table presents the samples and summary statistics for all additional analyses, including benchmark validation (Tables 8 and 9) and C_Score analyses (Tables 10 to 13). We construct a U.S. and an international sample of 22 countries (including the U.S.) over the period of 1990-2015. The U.S. sample comprises of 70,033 firm-year observations and the international sample comprises of 215,903 firm-year observations. Panel A reports the number of firm-year observations for each country. Panels B and C reports for the U.S. and international sample respectively their summary statistics for the variables used to estimate the AT and the C_Score measures. All variables are defined in Table 1.

Table 8: Association between AT and ACV measures on an unconditional basis (additional analysis)

Panel A: Portfolio average AT and ACV values

	U.S. sample		International sample	
AT measure sorted deciles	AT measure	ACV measure	AT measure	ACV measure
(Lowest) 1	0.0137	1.1724	−0.0187	1.2247
2	0.0847	1.2289	0.0495	1.1734
3	0.1102	1.1744	0.0734	1.4303
4	0.1321	1.7283	0.1049	1.3889
5	0.1571	1.4783	0.1343	1.4023
6	0.1842	1.5774	0.1663	1.3040
7	0.2135	2.8953	0.2007	1.7732
8	0.2484	2.5770	0.2423	1.8199
9	0.3146	3.1620	0.3112	1.8303
(Highest) 10	0.4363	2.1389	1.2566	1.8938

Panel B: Graphical depiction



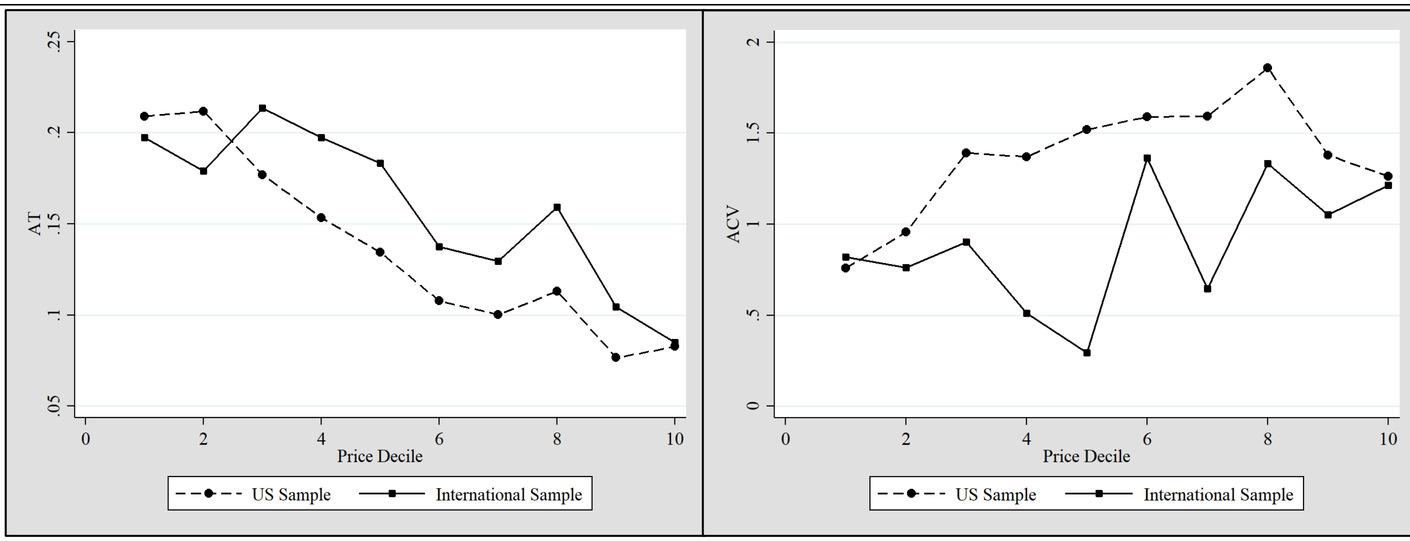
This table presents the association between AT and ACV measures on an unconditional basis for our additional analysis. We construct a U.S. and an international sample of 22 countries (including the U.S.) over the period of 1990-2015. The U.S. sample comprises of 70,033 firm-year observations, and the international sample comprises of 215,903 firm-year observations. Panel A reports across the AT measure sorted deciles their average AT and ACV values, where the AT coefficient estimates were estimated on an industry-year basis using the Fama and French 12-industry classification. Panel B provides graphical depiction of the correlation between the AT and ACV measures across the AT sorted decile portfolios. All variables are defined in Table 1.

Table 9: Stock price association with AT and ACV measures (additional analysis)

Panel A: Portfolio average AT and ACV values

	AT measure		ACV measure	
Price sorted deciles	U.S. sample	International sample	U.S. sample	International sample
(Lowest) 1	0.2090	0.1973	0.7586	0.8193
2	0.2118	0.179	0.9581	0.7616
3	0.1770	0.2135	1.3921	0.9013
4	0.1533	0.1974	1.3704	0.512
5	0.1344	0.1833	1.5202	0.2943
6	0.1077	0.1374	1.5894	1.3643
7	0.1002	0.1296	1.5939	0.6444
8	0.1131	0.1592	1.8577	1.3322
9	0.0766	0.1046	1.3795	1.0515
(Highest) 10	0.0827	0.0849	1.2628	1.2147

Panel B: Graphical depiction



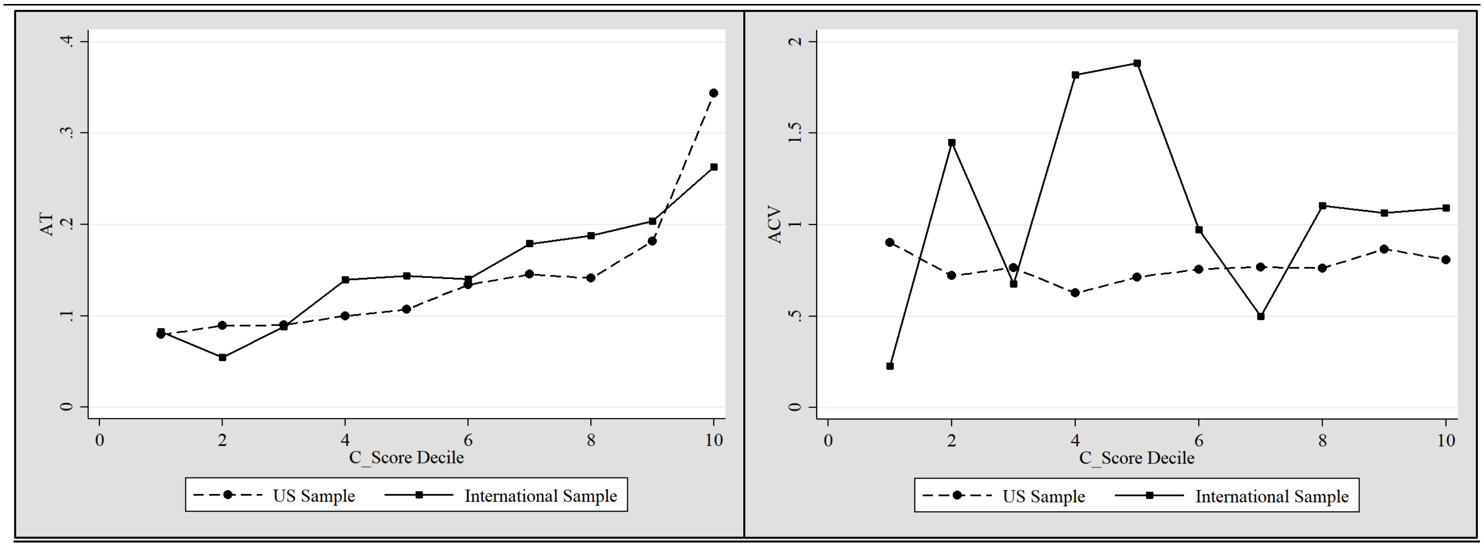
This table presents the association of opening stock price with AT and ACV measures for our additional analysis. We construct a U.S. and an international sample of 22 countries (including the U.S.) over the period of 1990-2015. The U.S. sample comprises of 70,033 firm-year observations, and the international sample comprises of 215,903 firm-year observations. Panel A reports across the opening price sorted decile portfolios their average AT and ACV values. Panel B provides graphical depiction of the association between the price sorted deciles with the AT and ACV measures separately. All variables are defined in Table 1.

Table 10: Association of the *C_Score* with AT and ACV measures (additional analysis)

Panel A: Portfolio average AT and ACV values

	AT measure		ACV measure	
C_Score sorted deciles	U.S. sample	International sample	U.S. sample	International sample
(Lowest) 1	0.0795	0.0827	0.9024	0.2261
2	0.0892	0.0545	0.7211	1.4496
3	0.0899	0.0883	0.7648	0.6760
4	0.1001	0.1393	0.6259	1.8187
5	0.1071	0.1440	0.7139	1.8823
6	0.1342	0.1402	0.7565	0.9727
7	0.1454	0.1789	0.7683	0.4984
8	0.1413	0.1876	0.7628	1.1041
9	0.1817	0.2035	0.8669	1.0650
(Highest) 10	0.3435	0.2627	0.8087	1.0910
Decile 10 – Decile 1	0.2640	0.1800	−0.0937	0.8649
Chi^2 (*p*-value)	132.00 (0.000)	53.91 (0.000)	4.88 (0.027)	1561.41 (0.000)

Panel B: Graphical depiction



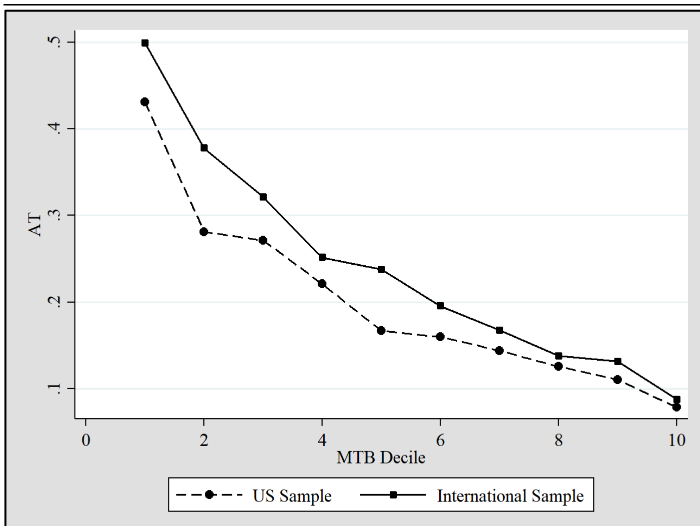
This table presents the association of *C_Score* with AT and ACV measures for our additional analysis. We construct a U.S. and an international sample of 22 countries (including the U.S.) over the period of 1990-2015. The U.S. sample comprises of 70,033 firm-year observations, and the international sample comprises of 215,903 firm-year observations. Panel A reports across the *C_Score* sorted decile portfolios their average AT and ACV values. The differences between Deciles 10 and 1 are also reported along with their statistical significance tests. Panel B provides graphical depiction of the association between *C_Score* sorted deciles with AT and ACV measures separately. All variables are defined in Table 1.

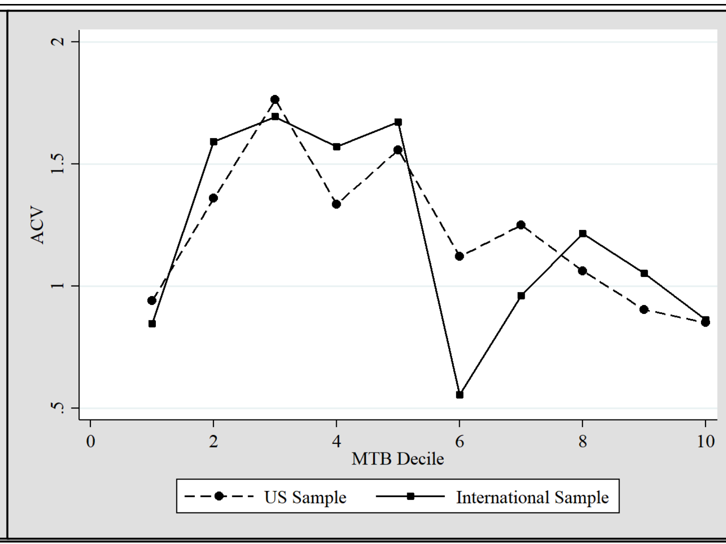
Table 11: Association of *C_Score* component *MTB* with AT and ACV measures (additional analysis)

Panel A: Portfolio average AT and ACV values

	AT measure		ACV measure	
MTB sorted deciles	U.S. sample	International sample	U.S. sample	International sample
(Lowest) 1	0.4309	0.4990	0.9403	0.8473
2	0.2813	0.3778	1.3595	1.5902
3	0.2713	0.3213	1.7630	1.6932
4	0.2210	0.2517	1.3351	1.5711
5	0.1670	0.2379	1.5568	1.6721
6	0.1601	0.1958	1.1210	0.5558
7	0.1440	0.1680	1.2502	0.9618
8	0.1257	0.1383	1.0610	1.2157
9	0.1105	0.1317	0.9039	1.0533
(Highest) 10	0.0788	0.0879	0.8503	0.8630
Decile 10 – Decile 1	−0.3521	−0.4111	−0.0900	0.0157
Chi^2 (*p*-value)	215.14 (0.000)	163.14 (0.000)	4.33 (0.037)	0.44 (0.504)

Panel B: Graphical depiction





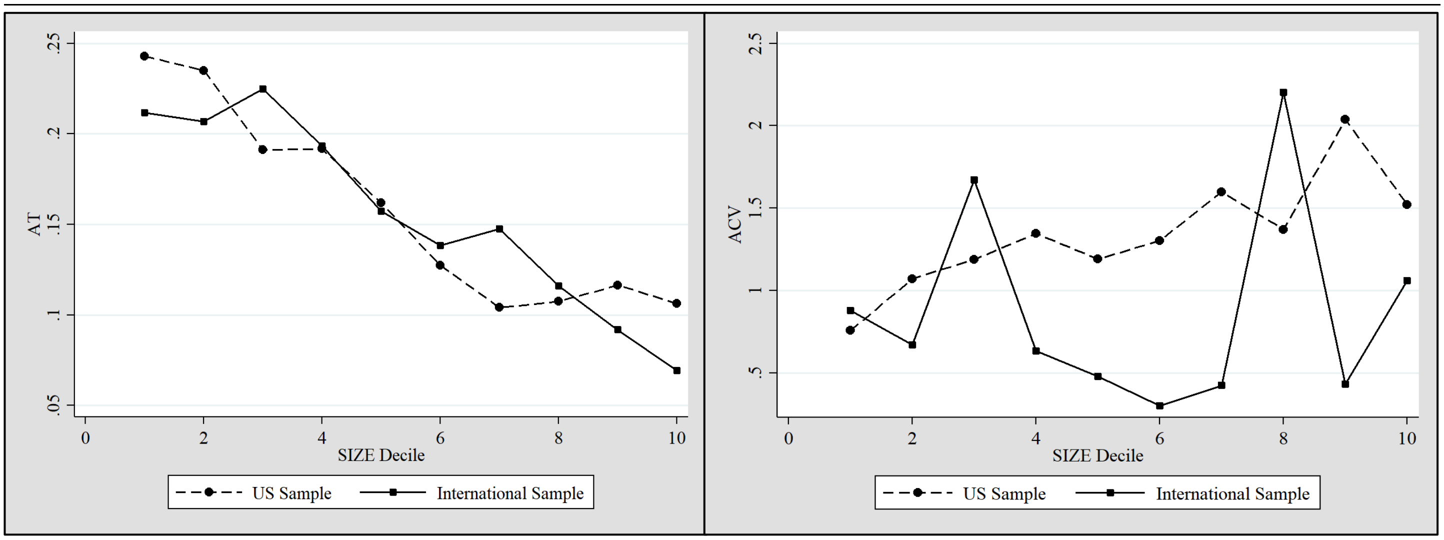
This table presents the association of *C_Score* component based on market-to-book (*MTB*) with AT and ACV measures for our additional analysis. We construct a U.S. and an international sample of 22 countries (including the U.S.) over the period of 1990-2015. The U.S. sample comprises of 70,033 firm-year observations, and the international sample comprises of 215,903 firm-year observations. Panel A reports across the *MTB* sorted decile portfolios their average AT and ACV values. The differences between Deciles 10 and 1 are also reported along with their statistical significance tests. Panel B provides graphical depiction of the association between *MTB* sorted deciles with AT and ACV measures separately. All variables are defined in Table 1.

Table 12: Association of *C_Score* component *SIZE* with AT and ACV measures (additional analysis)

Panel A: Portfolio average AT and ACV values

	AT measure		ACV measure	
SIZE sorted deciles	U.S. sample	International sample	U.S. sample	International sample
(Lowest) 1	0.2429	0.2116	0.7567	0.8795
2	0.2350	0.2068	1.0715	0.6700
3	0.1912	0.2249	1.1877	1.6717
4	0.1917	0.1935	1.3474	0.6334
5	0.1620	0.1573	1.1904	0.4786
6	0.1273	0.1384	1.3026	0.2994
7	0.1041	0.1474	1.5989	0.4232
8	0.1075	0.1162	1.3701	2.2036
9	0.1165	0.0918	2.0371	0.4302
(Highest) 10	0.1063	0.0694	1.5212	1.0588
Decile 10 – Decile 1	−0.1366	−0.1422	0.7645	0.1793
Chi^2 (*p*-value)	52.65 (0.000)	80.86 (0.000)	171.88 (0.000)	45.50 (0.000)

Panel B: Graphical depiction



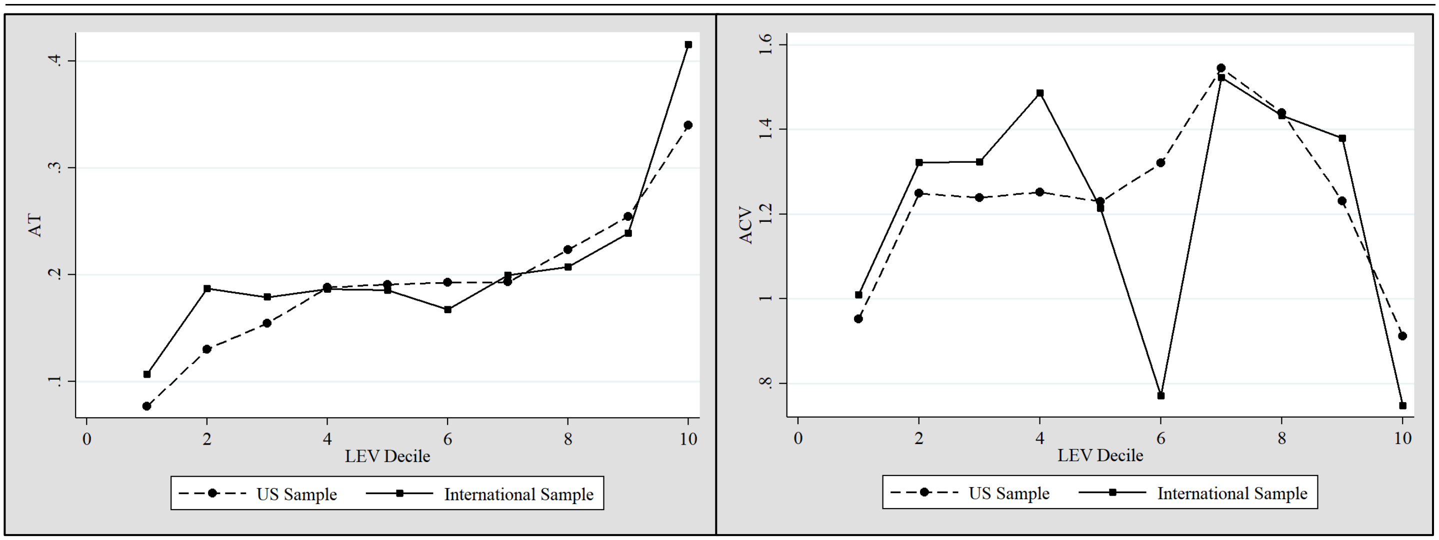
This table presents the association of *C_Score* component based on firm size (*SIZE*) with AT and ACV measures for our additional analysis. We construct a U.S. and an international sample of 22 countries (including the U.S.) over the period of 1990-2015. The U.S. sample comprises of 70,033 firm-year observations, and the international sample comprises of 215,903 firm-year observations. Panel A reports across the *SIZE* sorted decile portfolios their average AT and ACV values. The differences between Deciles 10 and 1 are also reported along with their statistical significance tests. Panel B provides graphical depiction of the association between *SIZE* sorted deciles with AT and ACV measures separately. All variables are defined in Table 1.

Table 13: Association of *C_Score* component *LEV* with AT and ACV measures (additional analysis)

Panel A: Portfolio average AT and ACV values

	AT measure		ACV measure	
LEV sorted deciles	U.S. sample	International sample	U.S. sample	International sample
(Lowest) 1	0.1068	0.0767	0.9523	1.0090
2	0.1871	0.1304	1.2489	1.3221
3	0.1792	0.1546	1.2389	1.3229
4	0.1868	0.1882	1.2520	1.4861
5	0.1858	0.1910	1.2293	1.2135
6	0.1676	0.1927	1.3208	0.7714
7	0.1997	0.1933	1.5452	1.5223
8	0.2076	0.2237	1.4394	1.4331
9	0.2390	0.2547	1.2307	1.3789
(Highest) 10	0.4159	0.3400	0.9120	0.7472
Decile 10 – Decile 1	0.3091	0.2633	−0.0403	−0.2618
Chi^2 (*p*-value)	133.47 (0.000)	92.37 (0.000)	0.78 (0.376)	113.57 (0.000)

Panel B: Graphical depiction



This table presents the association of *C_Score* component based on leverage (*LEV*) with AT and ACV measures for our additional analysis. We construct a U.S. and an international sample of 22 countries (including the U.S.) over the period of 1990-2015. The U.S. sample comprises of 70,033 firm-year observations, and the international sample comprises of 215,903 firm-year observations. Panel A reports across the *LEV* sorted decile portfolios their average AT and ACV values. The differences between Deciles 10 and 1 are also reported along with their statistical significance tests. Panel B provides graphical depiction of the association between *LEV* sorted deciles with AT and ACV measures separately. All variables are defined in Table 1.