

# Detecting Earnings Management in Annual Reports: An Evolutionary Approach

Marco Maggiorani<sup>1</sup> & Riccardo Cimini<sup>1</sup>

<sup>1</sup> Department of Economics, Engineering, Society and Business, University of Tuscia, Italy

Correspondence: Marco Maggiorani, Department of Economics, Engineering, Society and Business, University of Tuscia, Viterbo, Italy. E-mail: [maggiorani@unitus.it](mailto:maggiorani@unitus.it)

Received: September 12, 2024

Accepted: December 20, 2024

Online Published: January 18, 2025

doi:10.5539/ijbm.v20n1p106

URL: <https://doi.org/10.5539/ijbm.v20n1p106>

## Abstract

The primary aim of this paper is to systematize the models used in literature to assess earnings management. By using an evolutionary approach, the paper discusses different model specifications useful to detect opportunistic behaviours in annual reports. A critical review of the literature leads us to conclude that earnings management is detectable by using specifications that can be conveniently classified in “first-generation” and “second-generation” models.

The paper contributes to the literature and has implications in the practice, especially for investors. For scholars, it might be useful because it synthesizes the possible approaches that have been used over time by academics to detect earnings management. Investors might be aware of the models useful to detect the possible forms of misrepresentations in annual reports.

**Keywords:** literature review, misrepresentation, earnings management

## 1. Introduction

The paper focuses on earnings management that is a phenomenon widely investigated by academics and practitioners. It classifies the models used to detect it into the so-called “first-generation” and “second-generation” models. By using an evolutionary approach, the paper answers the following questions. How do “first-generation” and “second-generation” models differ in their approach to detect earnings management? What are the strengths and limitations of each generation of models?

We can recognize that the annual reports issued under national law and complying with a set of generally accepted accounting standards are tools that satisfy the information needs of internal and external parties. In particular, with the publication of this document, people who are outside the firm enter into a sort of conversation with the producers of the documents (i.e., the managers), making the conversation a descriptive metaphor for the interaction between the interests of insiders and those of outsiders. The conversation starts with the outsiders’ need for information about company performance, which managers provide in compliance with the law, formalizing in the annual report all the required information related to firm performance. The conversation does not stop there but continues because outsiders react to insiders’ answers with a decision related to the convenience of continuing to invest in the entity for which financial information has been produced.

Less metaphorically, the International Accounting Standards Board’s (IASB’s) with the Conceptual Framework for financial reporting (2010) validates the notion of a strong interaction between the two interest groups, in the paragraph dedicated to the objective of financial reporting. Indeed, in paragraph OB 3 of this document, we read that “the objective of general purpose financial reporting is to provide financial information about the reporting entity that is useful to existing and potential investors, lenders and other creditors in making decisions about providing resources to the entity” (IASB, 2010). This paragraph not only helps us to understand the content of the conversation between managers and outsiders but also suggests the possible answers that outsiders may give to the managers. While the content of the conversation concerns the financial information of the reporting entity, the outsiders’ potential answers are related to the possibility of starting, continuing or stopping the provision of resources to the entity. Thus, to maximize the probability that outsiders start or continue to negotiate with the entity, managers could behave opportunistically by misrepresenting the firm’s performance to influence the outsiders’ behaviour.

Financial information is misrepresented when the fundamental qualitative characteristics of financial information are violated, so, according to the Conceptual Framework for financial reporting (2010), financial information is neither relevant nor faithful. Recall that relevance relates to financial information's capacity to make a difference in its users' decisions (Conceptual Framework for Financial Reporting, 2010; QC 6). Instead, financial information is faithful when financial reports represent economic phenomena in words and numbers (Conceptual Framework for Financial Reporting, 2010; QC 12). With regards to the faithfulness of financial information, we hypothesize that with the expression "words" and "numbers" the Conceptual Framework for financial reporting (2010) refers to the form (words) and the substance (numbers) of the annual report. While the form (word) involves the reporting choices, the substance involves the accounting numbers, suggesting that both the decisions related to the form and those related to the substance can be used to misrepresent financial information and mislead outsiders' perceptions of the company's performance. We can argue that whenever the form of the annual report (words) is manipulated, insiders carry out impression management; in contrast, whenever the substance of the annual report (numbers) is manipulated, the values in the annual report lose their relevance and faithfulness and earnings management occurs (Cimini, 2015).

This paper aims to provide a theoretical framework for the misrepresentation of financial information disclosed in annual reports due to earnings management, systematising most conveniently the hundreds of papers that scholars published in academic journals in the last decades. In recent years, earnings management has been the topic of thousands of research papers. Scholars have been interested in the detection of the so-called accrual accounting earnings management or real activities earnings management. While in the first case, they detect the manipulation of those components of earnings that in the literature are well-known as accruals, in the second case scholars focus on real activities, i.e., reductions in expenditures on research and development (IAS 38), price discounts to increase sales or overproduction to disclose the lower cost of goods sold that, differently from the manipulation of accruals has a direct impact on cash flow (Roychowdhury, 2006). In addition, scholars who investigated accrual-accounting earnings management focused on specific accruals or assessed abnormal accruals as a whole to detect earnings management. Despite the relevant number of research papers that deal with specific accruals, in this paper, we are interested in the so-called accrual-accounting earnings management investigated as a whole.

The contribution of the paper is for both the theory and the practice. On the one hand, scholars could find "the state of the art" of the earnings management literature useful to detect this form of misrepresentation in annual reports. On the other hand, investors could learn the possible techniques to identify those annual reports that with higher probability might be affected by this form of misrepresentation of financial information.

The paper is composed of four sections. After this introduction, the paper presents the first models introduced by academics to detect earnings management, that is, the so-called "first-generation models". In the third section, the paper presents the more recent "second-generation models". The final section concludes the paper with the possible future line of research in this field of study.

## 2. The First-Generation Models to Detect Earnings Management

As to the possible manners to detect misrepresentation, scholars used several models over time to detect accrual-accounting earnings management. In this section, by using an evolutionary approach, we distinguish the "first-generation" models from the "second-generation" models.

The first of these studies is the seminal work of Paul M. Healy, who in 1985 published a paper in the *Journal of Accounting and Economics* titled "The Effect of bonus schemes on Accounting Decisions". Even though he never used the locutions "earnings management" or "window dressing" to qualify the attitude of insiders regarding the manipulation of accounting numbers, his paper is regarded as a seminal work in the earnings management literature. Healy (1985) found that bonus schemes create incentives for managers to select accounting procedures and accruals that maximize the value of their bonus awards (Healy, 1985, p. 106). As such bonus schemes become a function of the reported earnings, Healy (1985) explained how managers can manipulate them to maximize the expected utility of their bonus awards. In detail, he decomposed earnings into cash flows from operations and total accruals.

Healy (1985) assumed that the components of total accruals ( $TA_{it}$ ) are non-discretionary accruals ( $NDA_{it}$ ) and discretionary accruals ( $DA_{it}$ ). Because of difficulties in observing and measuring them, he surmised that by exercising discretion over the aforementioned (discretionary) accruals, managers can affect only the timing of reported earnings, so discretionary accruals sum to zero over the manager's employment horizon with the firms (Healy, 1985, p. 89).

The idea behind his research is that insiders — after observing the reported cash flow from operation plus non-discretionary accruals — manipulate discretionary accruals to achieve their earnings targets, i.e., the payment of

awards, in case the (real) reported earnings do not exceed the required boundary. The main concern regarding the Healy (1985) framework is that he cannot measure either discretionary or non-discretionary accruals. Discretionary accruals are very difficult to measure and thus, over the period analysed, are assumed to amount to zero. Yet, he considered non-discretionary accruals unobservable. This is the motivation that leads Healy (1985) to assume total accruals as a proxy of earnings management. Therefore, by analysing a sample of 94 entities listed in the 1980 Fortune Directory of the 250 largest US industrial corporations, Healy (1985) used a t-test for difference in means to compare the mean total accruals with several bonus plan parameters. By doing so, Healy (1985) found that managers are more likely to choose income-decreasing accruals when their bonus plans' upper or lower boundaries are binding, and income-increasing accruals when these boundaries are not binding (Healy, 1985, p. 106). This result is quite interesting, but the methodology used suffers from limitations that led several scholars to criticize him. The main concerns regarding the Healy (1985) methodology are the use of the mean total accruals as a proxy of the earnings management, for the reasons explained above, and the use of cash flow from operations as a proxy of cash flow plus non-discretionary accruals. To overcome these limitations, in 1986 Linda Elizabeth DeAngelo published in *The Accounting Review* a paper titled "Accounting numbers as market valuation substitutes: A study of management buyouts of public stockholders". Her methodology, improving upon the one of Healy (1985) contributed to the literature at least twofold. First, she provided a deeper qualification of total accruals through the analysis of the management buyout operations; then, she introduced the abnormal accruals concept as a proxy to detect earnings management.

As to total accruals, she similarly calculated them as Healy (1985) and explained that total accruals are negative because of the non-monetary depreciation expenses that are the most important components of total accruals. As to the research design, DeAngelo (1986) used a methodology that she considered roughly similar to the one typical of the event study to assess abnormal accruals. In this regard, she calculated abnormal accruals as the difference between the total accruals reported over the period in which earnings management is expected and the ones reported during the prior year.

In DeAngelo (1986), compared with Healy (1985), the proxy used to detect earnings management are abnormal accruals that are not simply total accruals but the first difference ( $AC_1 - AC_0$ ) of the reported total accruals reported at the end of different periods. Despite the improvements in comparison to the methodology followed by Healy (1985), the difficulties in assessing non-discretionary accruals remain. Thus, DeAngelo (1986) continued to assume that the average change in non-discretionary accruals ( $NDA_1 - NDA_0$ ) is approximately zero, so that a significant average decrease in total accruals primarily reflects a significant average decrease in discretionary accruals ( $DA_1 - DA_0$ ) (DeAngelo, 1986, p. 409). Once she calculated abnormal accruals, DeAngelo (1986) tested whether both the quarterly and the annual change measures of abnormal accruals, earnings and cash flow were statistically different from zero. By using a parametric test, that is, the t-test and the non-parametric Wilcoxon (1945) signed-rank test, her findings did not validate the research hypothesis that managers carry out earnings management before management buyouts. DeAngelo (1986) was aware of the weakness in the adopted methodology and argued that "one possible explanation for both sets of insignificant results is that the accrual methodology is simply not powerful enough to detect systematic income manipulation" (DeAngelo, 1986, p. 418).

At the end of the 1980s, the methodology followed to detect earnings management radically changed, thanks to a more rigorous use of event studies. This method allowed scholars to measure the magnitude of discretionary accruals that, assumed to be a measure of abnormal accruals, have become a measure of earnings management.

The paper of McNichols and Wilson (1988) was the first to use this new methodology. It is titled "Evidence of earnings management from the provisions for bad debts" authored by Maureen McNichols and G. Peter Wilson in the *Journal of Accounting Research* in 1988. Unlike the aforementioned studies of Healy (1985) and DeAngelo (1986), they detected earnings management due to the manipulation of a specific accrual which is a provision for bad debts, instead of total accruals.

McNichols and Wilson (1988) adopted an event study statistical model similar to that introduced by Izan (1978) and then used by Binder (1998). The peculiarity of the model is that they used an accrual estimation model that includes, between the explicative variables, a dummy (called PART) that is equal to 1 if the vector of accounting amounts that are assumed to be associated with non-discretionary accruals are observed over the event period, and 0 if accounting amounts are observed over the estimation period, with the assumption that insiders, over the estimation period, have not carried out any manipulation that significantly affected the magnitude of the abnormal accruals' random walk. In the model of McNichols and Wilson (1988), the statistical significance of the coefficient of the PART variable ( $\beta$ ) is very important to determine whether insiders have manipulated accounting amounts. That is, the regression coefficient  $\beta$  is used to test the presence of earnings management. In this regard, from a statistical point of view, two possibilities exist.

The former is that the regression coefficient  $\beta$  is statistically significant; the latter is that the same coefficient is not statistically significant. In the first case ( $\beta$  significant), there are two regression lines with different intercepts and the same slope. One interpolates the observations over the estimation period, whose intercept is  $\alpha$ , being PART = 0, and one interpolates the observations over the event period, whose intercept is  $\alpha + \beta$ , being PART = 1. The economic meaning of this first case ( $\beta$  significant) is that after the event, there is a structural break in the magnitude of discretionary accruals in comparison to the period before the event; insiders, because of the event, manipulated earnings enough to produce a structural break in the discretionary accruals' random walk. In the second case ( $\beta$  not significant), there is a single regression line whose intercept is  $\alpha$ , that interpolates observations over the estimation period and the event period. In this case, the economic meaning is that before and after the event, discretionary accruals are statistically the same.

This innovative approach adopted to detect earnings management suffers from several limitations that may bias the inference on the  $\beta$  regression coefficient. Dechow *et al.* (1995) or Dechow *et al.* (2012), describing the model of McNichols and Wilson (1988), argue that both measurement errors in discretionary accruals and the problem of omitted variables bias the regression coefficient, its standard errors, and ultimately the statistical inference in tests for earnings management.

Due to the limitations of the dummy variable approach of Izan (1978), after the research of McNichols and Wilson (1988), scholars used the original market model to detect earnings management with regression models estimated both over the estimation period and the event period.

One of this research is the paper published by Jennifer J. Jones in 1991 in the *Journal of Accounting Research*, titled "Earnings management during import relief investigations". Like McNichols and Wilson (1988), she used the locution earnings management to identify the attitude of insiders toward manipulating earnings. However, while the former used the discretionary components of a specific accrual (e.g., bad debts), the latter avoided distinctions by studying the aggregate accruals as a whole, similar to the aforementioned research of Healy (1985) and DeAngelo (1986).

Jones (1991) tested whether insiders of domestic producers, who would like to benefit from import protection, carried out earnings management by manipulating total accruals one year before and one year after the completion of the investigation by the US International Trade Commission (ITC). Using a sample of 23 firms, she used two different methodologies to detect earnings management. The first is a replication of the one used by DeAngelo (1986) which led them to conclude that accruals decreased, but this result does not allow for a firm conclusion that earnings management decreased because also earnings and cash flow decreased. The second methodology required four steps and relied on a classic event study adopted by using a market model.

With particular reference to the first step, Jones (1991) qualifies as the event the year when the ITC completed its import relief investigation. Then she identified an event period (p) and an estimation period (t). Jones (1991) assumed that the event period was three years, that is, the year of the event, the one before and the one after the event. The estimation period ranges from 14 to 32 years before the event period, according to the availability of data.

With particular reference to the second step, Jones (1991) moved from Kaplan's (1985) statement that changes in working capital account and thus accruals depend on the economic circumstances of the firm. Instead, Jones (1991) built a model to assess normal accruals during the estimation period, with total accruals as the dependent variable and the determinants of non-discretionary accruals as regressors that are  $\Delta REV_{it}$  (the change in revenues from t-1 and t of the firm i) and  $PPE_{it}$  (the gross property, plant and equipment in year t of the firm i that control the portion of total accruals related to non-discretionary depreciation expense). All the variables of the model were deflated by the lagged total asset to control for the scale effect. Jones (1991) justified the presence of the abovementioned variables as regressors of her accrual estimation model, arguing that: "Revenues are used to control for the economic environment of the firm because they are an object measure of the firms' operations before managers' manipulation, but they are not completely exogenous. Gross property, plant and equipment are included to control for the portion of total accruals related to nondiscretionary depreciation expense. Gross property, plant, and equipment are included in the expectations model rather than changes in this account because total depreciation expense (versus the change in depreciation expense) is included in the total accruals measure" (Jones, 1991, pp. 211-212).

Regarding the third step, Jones (1991) calculated abnormal accruals as the difference between the reported total accruals and the non-discretionary accruals predicted, by using the regression coefficients estimated over the estimation period, under the strong hypothesis of stationarity of non-discretionary accruals (Jones, 1991, p. 210). Abnormal accruals are assumed to be discretionary accruals, because once she controlled for the effect of the

economic circumstances on non-discretionary accruals, Jones (1991) could assume that the difference between current and prior-year accruals is due solely to changes in discretionary accruals. Algebraically, such a difference is calculated as follows:

For the fourth step, Jones (1991) used the Wilcoxon (1945) signed-rank test to verify her research hypothesis that during the import relief investigations by the ITC, accounting accruals are underestimated by insiders to maximize the benefit from increased import protection and/or the amount of protection granted. With this non-parametric test, she found that discretionary accruals for year 0 are significantly less than zero ( $p\text{-value} < 0.001$ ), validating her research hypothesis.

In 1995 the Jones (1991) model was improved by a model specification proposed by Dechow *et al.* (1995). They published the results, “Detecting Earnings Management” in *The Accounting Review*.

First, they reference all the previous models used by scholars to detect earnings management, then propose a model that they called a “modified Jones” (1991) model. Continuing to assume that the components of total accruals are non-discretionary accruals and discretionary accruals, like in the modified Jones (1991) model, there is the strong assumption that non-discretionary accruals are constant over time. The most significant difference regards the prediction of normal accruals over the event period. In the modified Jones (1991) model, Dechow *et al.* (1995) used  $\Delta REV_{it} - \Delta REC_{it}$  instead of  $\Delta REV_{it}$ , explaining that the original model implicitly assumes that discretion is not exercised over revenue in either the estimation period or the event period. The modified version of the Jones (1991) model implicitly assumes that all changes in credit sales in the event period result from earnings management. This is based on the reasoning that it is easier to manage earnings by exercising discretion over the recognition of revenue on credit sales than it is to manage earnings by exercising discretion over the recognition of revenue on cash sales (Dechow *et al.*, 1995, p. 199).

Several scholars demonstrated that the modified version of the Jones (1991) model improved the original Jones (1991) model even though it is affected by several limitations. To describe such limitations also argued by Dechow *et al.* (1995), we quote Dechow *et al.* (2012) in which the sub-section 2.3 is dedicated to the presentation of the limitations of those that we call “first-generation” models. Such limitations are:

1. all of the models generate well-specified test statistics when applied to random samples;
2. all models generate tests of low power for earnings management of economically plausible magnitudes (e.g., 1% to 5% of total assets), and
3. all models are misspecified when applied to samples of firms with extreme financial performance (Dechow *et al.*, 2012, p. 282).

An in-depth analysis of the literature leads us to integrate the concerns listed above. In our opinion, the paper that summarizes these issues is McNichols (2000), where we can find three other limits of the “first-generation” models.

They are:

1. all the models require the event study approach. In each event study, we have an assumption over the estimation period and a hypothesis to test over the event period. The assumption is that earnings management does not occur during the so-called estimation period. In other words, the presence of earnings management does not cause any structural break over the estimation period in the accruals’ random walk. In the event period, the hypothesis to be tested is the presence of earnings management. Thus, the researcher has to test whether, after the occurrence of a specific event, the abnormal accruals vary differently in comparison to the estimation period, because of a major (minor) attitude of insiders to carry out earnings management as a consequence of the occurrence of the event. To implement such a research protocol, the non-discretionary accruals are assumed to be constant over time (hypothesis of stationarity) and a lot of subjectivity is required to identify both the estimation period and the event period. Take also into account that the use of OLS requires the linearity of any relationship between the magnitude of performance and accruals, and that in event studies, controlling for reversals in the event period of earnings management which occurred in prior years is quite difficult;
2. all the models assess accruals following the balance sheet approach that according to McNichols (2000) is neither really sensitive to the hypothesized manipulation nor readily controlled;
3. all the methodologies adopted in the “first-generation” models are firm-specific and require the use of panel data of at least 10 years. This increases the probability of losing observations because of missing data. Alternatively, an industry-specific approach (e.g., estimation of accruals on a cross-sectional basis instead of with the use of time series) could limit the loss of missing data.

All these limitations led to a decline of that strand of literature that detected earnings management by using event studies, regression models and test statistics, i.e., the “first-generation” models, and we turned to a new strand of

literature that has assessed earnings management by using the “second-generation” models.

### 3. The Second-Generation Models to Detect Earnings Management

After a period during which scholars detected earnings management using ratios (Leuz et al., 2003), the “second-generation models” have been introduced to overcome the limitations of the “first-generation models”. Between them, in 2002, Maureen F. McNichols published a paper in *The Accounting Review* discussing Dechow and Dichev (2002), who proposed a measure of the quality of accruals and earnings based on the extent to which accruals map into cash flow realizations in contemporaneous and adjacent periods (McNichols, 2002, p. 61). Moving from the Jones (1991) model, she proposed a correction (a new accruals estimation model) to avoid one of the main concerns that scholars had with the “first-generation” model, i.e., their misspecification due to the presence of omitted (independent) variables (e.g., the so-called economic fundamentals) that are correlated with non-discretionary accruals’ determinants. In this regard, McNichols (2002) began by assuming that discretionary accruals estimates are correlated with earnings performance. This is why the proposal of McNichols (2002) was to add cash flow in the Jones (1991) model, to reduce the extent to which the model omits variables that are correlated with sample firms’ economic performance fundamentals. In addition, McNichols (2002) used the cash flow statement approach and overcame a limitation of the “first-generation” models, which extensively adopted the balance sheet approach to calculate total accruals. Other than using a cash flow statement approach, she did not consider total accruals as a whole, but only the working capital accruals. In other words, she avoided considering the depreciation expenses. McNichols (2002) justified this choice by arguing that this is consistent with Dechow and Dichev (2002) and because depreciations are related to long-term capital expenditure accruals rather than working capital accruals (Allen et al., 2010). We can add that not including depreciations, the models that consider only the working capital accruals as dependent variable can detect earnings management in the extent to which changes of working capital accruals map changes of cash flow included between the explicative variables. Methodologically, McNichols (2002) compared the Dechow and Dichev (2002) model, the Jones (1991) and her model, which combines the former with the latter. By using a sample of non-financial firms over the period 1988-1998, the scholar found that the model with cash flow between the independent variables is the most correctly specified and that both the Dechow and Dichev (2002) model and the Jones (1991) model are not correctly specified, since their residuals – assumed by Dechow and Dichev (2002) to be a measure of earnings quality – are associated with lagged, current and future cash flow and also with the change in sales. In addition to being less misspecified, her model is the one that better estimates discretionary accruals. McNichols (2002) achieved this conclusion by comparing the adjusted  $R^2$  and the F-statistic of the three different accrual estimation models. McNichols (2002) found that her model compared to the other ones has the highest explanatory power (adjusted  $R^2$ ) and F-statistic. In addition, she argued that the current operating cash flow ( $CFO_{it}$ ) has a regression coefficient and t-statistic higher than both the lagged and the future cash flow ( $CFO_{it-1}$ ;  $CFO_{it+1}$ ); for this reason, the current operating cash flow should be considered among the explicative variables of an accrual estimation model, to reduce problems due to omitted variables.

Even though the new specification of the Jones (1991) model contributed to the literature, McNichols (2002) knew that several other improvements still needed to be made to the various models. This is probably the reason why, in the conclusion of her paper, we can read that a possible direction of future research is to focus on specific accruals, rather than aggregate accruals or the total accruals as a whole. Notwithstanding this suggestion, other scholars, instead of studying specific accruals, continued to extensively investigate aggregate accruals. Unlike McNichols (2002), they added cash flow in the modified Jones (1991) model, and not in the prior Jones (1991) model. Among them are Larcker and Richardson (2004). These scholars add, between regressors, the current operating cash flow from operations that controls for the current operating performance. They also include the book-to-market ratio as a proxy for the expected growth in the firm’s operations. The reason why they do not also add the lagged and the future cash flow is due to the results found by McNichols (2002) that the current operating cash flow has the highest regression coefficients and t-statistics. The specification has been used also by Callao and Jarne (2010). In their study, published in *Accounting in Europe*, Callao and Jarne (2010) explained that the model was estimated for the period 1997–2002, assuming that non-discretionary accruals (NDA) are a function of year-to-year change in sales, property, plant and equipment, book-to-market ratio and current operating cash flows. Having estimated the parameters for 1997–2002, they predicted discretionary accruals in the event period (2003-2006), once non-discretionary accruals are subtracted from total accruals (Callao and Jarne, 2010, p. 168). Callao and Jarne (2010) used the prediction error to interpret the discretionary part of accruals. In their model, discretionary accruals are assumed to be a measure of abnormal accruals, which is their proxy of earnings management. From a statistical point of view, they are the residuals of a regression model. Once discretionary accruals are calculated, analysis of the increase or decrease of misrepresentation due to earnings management requires the use of a test statistic (such as the t-test or the non-parametric Wilcoxon (1945) signed-rank test) that compares discretionary accruals

estimated at each reporting date and check whether the difference is statistically significant.

Other scholars introduced new models that try to overcome concerns of the “first-generation” models.

Kothari *et al.* (2005) have been among those who adopted a methodology to detect earnings management that overcomes most of their limitation. In 2005, they published a paper in the *Journal of Accounting and Economics* titled “Performance matched discretionary accrual measures.” In this paper, they examined the specification and power of tests based on performance-matched discretionary accruals and compared them with tests using traditional discretionary accruals measures (e.g., Jones (1991) and modified Jones (1991) models) (Kothari *et al.*, 2005, p. 163).

Kothari *et al.* (2005) used ROA to control for the performance of the entities included in the sample analysed. Moreover, they abandoned the firm-specific approach to assess discretionary accruals and adopted the industry-specific procedure by using the so-called performance-matched methodology coherently with Barber and Lyon (1997).

As to the inclusion of ROA, Kothari *et al.* (2005) explained that by including ROA, scholars avoid misspecification due to the correlation between accruals and past and contemporaneous firms’ performance if we assume ROA a measure of the performance of the entity (Prencipe, 2006). The misspecification is due to the omission in the “first-generation” models of variables that regard the performance of the entity. In this regard, Kothari *et al.* (2005, p. 167) argued that the correlation between accruals and the performance of the entity is suggested by economic intuition, extant models of accruals, earnings and cash flow whose empirical evidence has been provided by several scholars such as Guay *et al.* (1996), Healy (1996), Dechow *et al.* (1998), Barth *et al.* (2001). Being that ROA is considered a measure of the performance of the entity, its inclusion among regressors allowed scholars to control for firms with extreme financial performance. Dechow *et al.* (1998) recognized that this variable is very useful to control for the effect of performance on measured discretionary accruals, despite the presence of a large variety of variables that can be used for the same purposes.

As to the use of the performance-matched methodology, it avoids the use of the event study methodology which is not always accurate in detecting earnings management. Avoiding event studies, Kothari *et al.* (2005) calculated discretionary accruals by estimating the residuals directly from their annual cross-sectional industry regression model. In doing so, Kothari *et al.* (2005) opened a new strand of literature without the use of event studies that distinguish between an estimation and an event period. In this strand, the concept of discretionary accruals perfectly overlaps with that of residuals of the accrual estimation models, which continue to include, between regressors, variables that are assumed to be associated with non-discretionary accruals. Kothari *et al.* (2005, p. 174) did not have a “pre-event” period, helping them to overcome concerns of the event studies that involve the strong hypothesis that the non-discretionary accruals are assumed constant over time. Thus, subjectivity is required to identify both the estimation period and the event period and the assumption of linearity of the relation between the magnitude of performance and accruals is very strong.

If the use of the performance-matched methodology and the inclusion of ROA among the independent variables of an accruals estimation model avoid several limitations of the “first-generation” models, other (minor) aspects of the Kothari *et al.* (2005) model overcome other concerns regarding the “first-generation” models. For instance, the methodology avoids the concern that all of the models generate well-specified test statistics when applied to random samples. In this regard, to randomise their sample, Kothari *et al.* (2005) used a simulation analysis that allowed them to identify 250 different simulated samples of 100 firms each. The scholars drew samples without replacement from the full sample or from stratified subsets. The subsets are the lowest and highest quartiles of firms ranked on book-to-market, past sales growth, earnings-to-price, size (market value of equity, referred to as large and small firms) and operating cash flow. To construct the subsets, each year they ranked all firm-year observations based on each partitioning characteristic (e.g., book-to-market or size, measured at the beginning of the year). Each year, they only retained the upper and lower quartiles of the sample. For each partitioning variable, they then pooled observations across all years to form two sub-samples: one based on pooling all data from the annual upper quartiles and another based on pooling all data from the annual lower quartiles (Kothari *et al.*, 2005, p. 172).

In their empirical analysis, they collected data by using COMPUSTAT for a sample of 123,000 firm-year observations over the period 1962-1999. Kothari *et al.* (2005) required a large amount of data not only to guarantee the robustness of their findings but also because they used randomised samples to detect earnings management with both the Jones (1991) and the modified Jones (1991) models and also with the performance matching approach, discovering that all of them are misspecified without any controls for performance. Since Kothari *et al.* (2005) did not use the event study approach, we recall that, in comparison to previous studies, abnormal

(discretionary) accruals are the residuals of their regression model. Another distinctive feature of the “second-generation” models is the presence of a constant term between the regression parameters. They justified this methodological choice by arguing that the constant term, together with deflating variables with the lagged total assets, reduces heteroskedasticity. Moreover, by having the work of Brown *et al.* (1999) as a reference, they explain that the constant term allows controlling for omitted size variables. On the contrary, a model without the constant term produces biases in the test statistics used to compare the different models.

We can conclude the presentation of Kothari *et al.* (2005) by arguing that their model is not the panacea to assess earnings management. Kothari *et al.* (2005) conceded that their methodology cannot and does not solve all the problems arising from bad discretionary accruals models or from a researcher’s failure to recognize the accrual management incentives that are unique to the research question being addressed (Kothari *et al.*, 2005, p. 165).

Cheng *et al.* (2012) demonstrate that both the McNichols (2002) specification and the Kothari *et al.* (2005) model are useful to detect earnings management. In detail, the best-performing model is the one in which operating cash flows are used as the controlling variable in firm-specific regressions, followed by the one in which ROA is used as the control variable in industry-specific regressions. However, despite the effectiveness of both the specifications of the “second-generation models”, none of them controlled for reversals of accruals from one period to another. This is the purpose of Dechow *et al.* (2012) titled “Detecting earnings management: A new approach” which was published in the *Journal of Accounting Research*. Dechow *et al.* (2012) criticize the performance-matching procedure adopted by Kothari *et al.* (2005) to mitigate misspecification and introduce a new approach to detect earnings management, using a methodology more similar to that of McNichols and Wilson (1988) direct to overcome its limitations though the analysis of a sample of 209,530 firm-year observations between 1950 and 2009. In their paper, they compared three different specifications that correspond to three different scenarios that involve the presence of reversals.

In the first scenario, Dechow *et al.* (2012) hypothesized that there are no reversals in their model, and so the model specification that they adopted is the same as McNichols and Wilson (1988). This specification does not consider the dummy variable  $PART_{it}$ , and is representative of the different specifications of all the “first-generation” models (e.g., sales and property, plant and equipment for the Jones (1991) model and “second-generation” models that do not control for reversals, being  $X_{k,i,t}$  the variables that scholars used to control for non-discretionary accruals.

In the second scenario, the scholars hypothesized that all the accrual-accounting earnings management reverses in the following period. In this model, if the difference between the coefficients of  $PART_{it}$  and  $PARTR_{it}$  is positive, it suggests an upward earnings management with respect to its reversal over the period analysed; on the contrary, if negative, it suggests a downward earnings management with respect to its reversal over the period analysed. More concretely, in the first case ( $a_2 - a_3 > 0$ ), manipulations of accruals done during the period investigated will reverse in future periods not investigated by the researcher. In the second case ( $a_2 - a_3 < 0$ ), reversals that refer to manipulations done during previous periods not investigated by the researcher reverse during the period analysed. If researchers could detect earnings management by analysing the entire firm’s life cycle, they would find the equivalence between  $a_2$  and  $a_3$  because all the earnings management carried out in prior periods has to reverse in subsequent periods.

In the third scenario, Dechow *et al.* (2012) hypothesized that the accrual-accounting earnings management reverses both in the first and in the second year following the period during which the researcher assumes the presence of manipulations.

Other than controlling for reversal of earnings management, Dechow *et al.* (2012) consider that all of the models generate well-specified test statistics when applied to random samples. In this regard, they used an iterative methodology, based on simulations, to evaluate the specifications of the Healy (1985), the Jones (1991), the modified Jones (1991), the Dechow and Dichev (2002) and the McNichols (2002) models. The methodology used by Dechow *et al.* (2012) reduced misspecification due to omitted variables and increased the power of the test statistics by around 40% (Dechow *et al.*, 2012, p. 275).

#### 4. Conclusions

By using an evolutionary approach, this paper has systematized the models used in literature to assess earnings management. The paper has discussed the different model specifications useful to detect opportunistic behaviours in annual reports proving a critical review of the existing literature. Academic scholars have detected earnings management by using specifications that this paper has conveniently classified in “first-generation” and “second-generation” models.

The paper contributes to the literature being the first paper that has provided a review of the literature on the



methodologies used over time to detect such a specific form of misrepresentation. The paper has also implications for the practice, especially for investors. The review of the literature has provided evidence of the high interest of scholars involved in the misrepresentation field to find the most appropriate specification useful to detect accrual accounting earnings management in the annual reports. Despite earnings management can be considered a mature field of study, there are several motivations that might justify such interest. First, the fact that opportunistic behaviours of managers change “cross section” and “over time” is the first plausible motivation that explains the high (actual) interest of academics and practitioners in investigating the accrual-accounting earnings management phenomenon. Second, scholars involved in topics that belong to the positive accounting theory (e.g., value relevance, earnings management) use empirical analyses and models to test their research hypotheses and to answer their research questions. This further reason that justifies the high interest of scholars in earnings management is based on the limitations of the models used to detect accrual-accounting earnings management. Such limitations are leading academics to improve the existing models and to introduce new and more sophisticated statistical models, useful to detect misrepresentation of financial information. The willingness to improve the existing models has emerged, we think, also from our evolutionary approach that presented, the different accrual specification models adopted by scholars in the last 40 years to detect this phenomenon. The essence of our evolutionary approach is that such models have been improved under several aspects over time. However, the most recent ones are still far from perfect and need of further adjustments that allow them to model the “real world”, that is by far more complicated than the “modelled world”. In this regard, for the future research directions, academics and practitioners might improve the existing models furthermore. In addition, future research should adapt such models to the firms’ business model. For instance, the most appropriate accrual accounting earnings management specifications should be identified for financial firms and distinguished from those used, for instance, by the public sector entities. To protect the interests of investors and other stakeholders, standard setters might reduce accounting discretion in order to mitigate the possible earnings management effects in annual reports. Regarding the models’ improvements, probably, scholars will never find a model that will fill the gap between the “real” and the “modelled world”, because earnings management is something that involves also the motivations of insiders that, we think, models cannot catch. Meanwhile, scholars will continue to introduce new improvements to the earnings management models’ specification, being aware that such improvements will only minimize the gap between the “real” and the “modelled world”.

#### **Informed consent**

Obtained.

#### **Ethics approval**

The Publication Ethics Committee of the Canadian Center of Science and Education.

The journal and publisher adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

#### **Provenance and peer review**

Not commissioned; externally double-blind peer reviewed.

#### **Data availability statement**

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

#### **Data sharing statement**

No additional data are available.

#### **Open access**

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

#### **Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

#### **References**

- Allen, E. J., Larson, C. R., & Sloan, R. G. (2010). Accrual Reserrals, Earnings and Stock Returns. *Journal of Accounting and Economics*, 56(1), 113-129. <https://doi.org/10.1016/j.jacceco.2013.05.002>
- Barber, B., & Lyon, J. (1997). Detecting long-run abnormal stock returns: the empirical power and specification of test statistics, *Journal of Financial Economics*, 43(3), 341-372. <https://doi.org/10.1016/S0304->

405X(96)00890-2

- Barth, M., Cram, D., & Nelson, K. (2001). Accruals and the prediction of future cash flows, *The Accounting Review*, 76(1), 27-58. <https://doi.org/10.2308/accr.2001.76.1.27>
- Binder, J.J. (1998). The Event Study Methodology Since 1969, *Review of Quantitative Finance and Accounting*, 11(2), 111–137. <https://doi.org/10.1023/A:1008295500105>
- Brown, S.J., LO, K., & Lys, T. (1999). Use of R<sup>2</sup> in accounting research: measuring changes in value relevance over the last four decades? *Journal of Accounting and Economics*, 28(2), 83-115. [https://doi.org/10.1016/S0165-4101\(99\)00023-3](https://doi.org/10.1016/S0165-4101(99)00023-3)
- Callao, S., & Jarne, J. I. (2010). Have IFRS Affected Earnings Management in the European Union? *Accounting in Europe*, 7(2), 159-189. <https://doi.org/10.1080/17449480.2010.511896>
- Cheng, A., Liu, Z., & Wayne, C. T. (2012). Abnormal Accrual Estimates and Evidence of Mispricing, *Journal of Business Finance & Accounting*, 39(1-2), 1-34. <https://doi.org/10.1111/j.1468-5957.2011.02274.x>
- Cimini, R. (2015). The selective misrepresentation of financial information due to earnings management: theoretical background, models and empirical evidence from the European Union. Giappichelli (IT).
- DeAngelo, L. (1986). Accounting numbers as market valuation substitutes: a study of management buyouts of public stockholders, *The Accounting Review*, 61(3), 400-420.
- Dechow, P., & Dichev, D. (2002). The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors, *The Accounting Review*, 77(s-1), 35-59. <https://doi.org/10.2308/accr.2002.77.s-1.35>
- Dechow, P., Hutton, A. P., Kim, J. H., & Sloan, R. (2012). Detecting Earnings Management: A New Approach, *Journal of Accounting Research*, 50(2), 275-334. <https://doi-org.unitus.idm.oclc.org/10.1111/j.1475-679X.2012.00449.x>
- Dechow, P., Kothari, S. P., & Watts, R. (1998). The relation between earnings and cash flow, *Journal of Accounting and Economics*, 25(2), 133-168. [https://doi.org/10.1016/S0165-4101\(98\)00020-2](https://doi.org/10.1016/S0165-4101(98)00020-2)
- Dechow, P., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management, *The Accounting Review*, 70(2), 193-225.
- Guay, W., Kothari, S., & Watts, R. (1996). A market-based evaluation of discretionary accrual models, *Journal of Accounting Research*, 34(Supplement), 83-105. <https://doi.org/10.2307/2491427>
- Healy, P. M. (1985). The effect of bonus schemes on accounting-based decisions, *Journal of Accounting and Economics*, 7(1-3), 85-107. [https://doi.org/10.1016/0165-4101\(85\)90029-1](https://doi.org/10.1016/0165-4101(85)90029-1)
- Healy, P. M. (1996). Discussion of a market-based evaluation of discretionary accrual models, *Journal of Accounting research*, 34(Supplement), 107-115. <https://doi.org/10.2307/2491428>
- IASB. (2006). IAS 38 “Intangible Assets”.
- IASB. (2010). Conceptual Framework for Financial Reporting, 2010.
- Izan, H. Y. (1978). An Empirical Analysis of the Economic Effects of Mandatory Government Audit Requirements, *Ph.D. dissertation*, University of Chicago.
- Jones, J. J. (1991). Earnings management during import relief investigations, *Journal of Accounting Research*, 29(2), 193-228. <https://doi.org/10.2307/2491047>
- Kaplan, R. S. (1985). Comments on Paul Healy: Evidence on the Effect of Bonus Schemes on Accounting Procedure and Accrual Decisions, *Journal of Accounting and Economics*, 7(1-3), 109-113. [https://doi.org/10.1016/0165-4101\(85\)90030-8](https://doi.org/10.1016/0165-4101(85)90030-8)
- Kothari, S. P., Leone, A. J., & Wasley, C. E. (2005). Performance matched discretionary accruals measures, *Journal of accounting and economics*, 39(1), 163-197. <https://doi.org/10.1016/j.jacceco.2004.11.002>
- Larcker, D., & Richardson, S. (2004). Fees paid to audit firms, accrual choices, and corporate governance, *Journal of Accounting Research*, 42(3), 625-658. <https://doi.org/10.1111/j.1475-679X.2004.t01-1-00143.x>
- Leuz, C., Nanda, D., & Wysocki, P. (2003). Earnings management and investor protection: an international comparison, *Journal of Financial Economics*, 69(3), 505-527. [https://doi.org/10.1016/S0304-405X\(03\)00121-1](https://doi.org/10.1016/S0304-405X(03)00121-1)
- McNichols, M. F. (2000). Research design issues in earnings management studies, *Journal of Accounting & Public Policy*, 19(4-5), 313-345. [https://doi.org/10.1016/S0278-4254\(00\)00018-1](https://doi.org/10.1016/S0278-4254(00)00018-1)

- McNichols, M. F. (2002). Discussion on the quality of accruals and earnings: the role of accruals estimation errors, *The Accounting Review*, 77(Supplement), 61-69. <https://doi.org/10.2308/accr.2002.77.s-1.61>
- McNichols, M. F., & Wilson, P. (1988). Evidence of earnings management from the provision for bad debts, *Journal of Accounting Research*, 26(Supplement), 1-31. <https://doi.org/10.2307/2491176>
- Prencipe, A. (2006). *Earnings quality. Analisi della qualità degli earnings in una prospettiva internazionale*, Pearson Editore, Milano.
- Roychowdhury, S. (2006). Earnings management through real activities manipulation, *Journal of Accounting and Economics*, 42(3), 335-370. <https://doi.org/10.1016/j.jacceco.2006.01.002>
- Wilcoxon, F. (1945). Individual comparisons by ranking methods, *Biometrics*, 1(6), 80-83. <https://doi.org/10.2307/3001968>

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).