

# **Real Earnings Management and Information Asymmetry in the Equity Market**

**David Abad**  
University of Alicante

**M. Fuensanta Cutillas-Gomariz**  
University of Murcia

**Juan Pedro Sánchez-Ballesta**  
University of Murcia

**José Yagüe**  
University of Murcia

Accepted for publication in *European Accounting Review*

*Acknowledgements:* We thank an anonymous reviewer and the editor, Professor Guochang Zhang, for their useful comments and suggestions, which have improved the paper greatly. We also appreciate the comments made by Thomas Gilliam, Francisco Villanueva, José A. Gonzalo Angulo, Begoña Giner, J. Samuel Baixauli and the participants at the European Accounting Association 38th Annual Congress, the XVIII AECA Congress, the XI Workshop on Empirical Research in Financial Accounting, and the XXIV Finance Forum. David Abad acknowledges financial support from the *Ministerio de Ciencia y Competitividad* through grants ECO2013-4409-P and ECO2014-58434-P. José Yagüe acknowledges financial support from *Fundación Caja Murcia*.

*Correspondence address:* José Yagüe, Department of Management and Finance, Faculty of Economics and Business, University of Murcia, Campus de Espinardo s/n, 30100 Murcia, Spain. Email: [ppyague@um.es](mailto:ppyague@um.es)

# **Real Earnings Management and Information Asymmetry in the Equity Market**

## **Abstract**

The literature suggests that real earnings management (REM) activities can increase adverse selection-risk in capital markets. Due to their opacity and the difficulties in understanding their implications, REM strategies may increase the level of information asymmetry among investors. This paper examines the association between earnings management through real activities manipulation and information asymmetry in the equity market. To estimate the level of adverse selection risk we use a comprehensive index of information asymmetry measures proposed by the market microstructure literature. For a sample of Spanish listed firms, we find that firms' strategies of increasing earnings through REM are associated with higher information asymmetry in those firms that meet last year's earnings. Our findings are consistent with the hypothesis that earnings management through real activities manipulation garbles the market, enhances private information production, and exacerbates information asymmetry in the stock market.

**Keywords:** Information Asymmetry, Real Earnings Management, Real Activities Manipulation, Discretionary Accruals.

**JEL Classification:** G14, M41, M48.

## 1. Introduction

Earnings management occurs when managers use their discretion in the financial reporting process and in structuring transactions to misrepresent the true economic performance of the company (Dechow & Skinner, 2000; Healy & Wahlen, 1999). Firms can manage earnings through two types of activities: accrual-based activities and real activities manipulation. While accrual earnings management implies discretionary choices permitted within accounting standards and with no direct cash flow consequences, real earnings management (hereinafter REM) involves deviations from normal operational practices to manipulate earnings numbers, with direct consequences for current and future firm cash flows. In this paper, we are interested in investigating the association between REM and the level of information asymmetry in the stock market.

Although earnings management activities may be informative, most research adopts the opportunistic perspective, and assumes that managers try to mislead stakeholders. According to this view, earnings management reduces earnings quality and garbles the information provided by financial statements. Consequently, as Bhattacharya, Desai, & Venkataraman (2013) hypothesize, based on the model of Kim & Verrecchia (1994), if investors differ in their ability to process earnings related information, then poor earnings quality can lead to differentially informed investors, so exacerbating the information asymmetry in financial markets. Consistent with this hypothesis, empirical evidence shows that accrual based-earnings management is associated with higher information asymmetry and reductions in market liquidity, leading to a higher cost of capital (e.g. Bhattacharya et al., 2013; Jayaraman, 2008; Rajgopal & Venkatachalam, 2011). Research has analyzed the association between earnings management (or earnings quality) and the firm's information environment, with the focus mainly on accrual-based earnings management. Nevertheless, there is little evidence for the effect of REM on the adverse selection problem in financial markets.

Since earnings management through real activities manipulation distorts earnings and cash flows, REM strategies may imply lower earnings quality, as manipulated earnings numbers hinder the evaluation and assessment of the true firm's current performance and the expected level of future cash flows by investors. Hence, a positive association of

REM with information asymmetry could be expected. Moreover, since REM is less subject to external monitoring and scrutiny by board, auditors and regulators than accruals earnings management (Cohen & Zarowin, 2010), and its implications for firm future performance are not clear (e.g. Graham, Harvey, & Rajgopal, 2005; Gunny, 2010), it may be difficult to understand by capital markets (Kothari, Mizik, & Roychowdhury, 2016) and thus, it may contribute to increase the informational asymmetry problem. As Gunny (2010) states, it is complicated to determine whether managers use REM opportunistically to the detriment of shareholders or, on the contrary, they use REM to signal future performance or to attain benefits that will allow the firm to perform better in the future. Hence, REM could increase the uncertainty of investors about the distribution of firm's future cash flows, and, in this case, traders who have better information-processing abilities could take advantage of their superior assessments of firm performance. In addition, since it is difficult to distinguish sub-optimal from optimal business decisions, the opacity of REM activities could lead some investors to engage in acquisition of private information with the aim of exploiting it and obtaining profits from trading on the market. For all these reasons, we expect REM strategies to exacerbate information asymmetry among investors in stock markets.

In order to examine the association between REM and information asymmetry we construct a sample of Spanish non-financial listed firms for the period 2001-2008 and use different measures of REM based on Roychowdhury (2006). However, since the proxies for REM represent abnormal levels of cash flows from operations, production costs and discretionary expenses, they may contain noise that is unrelated to managerial opportunism and that may be capturing situations other than intentional manipulation (e.g., unusual business circumstances). Therefore, the prediction of a positive relation between empirical proxies for REM and information asymmetry may not hold in general and the sign of this relation could depend on the particular underlying factors that determine the values of REM measures in the sample. Hence, we divide our total sample into two subsamples based on the incentives to manage earnings. Specifically, we examine the association of REM measures and information asymmetry in two settings: one where managers are likely to engage in REM activities to meet last year's net income (suspect sample) versus another, delimited by the rest of the sample, where deviations from normal activity may be unrelated to opportunistic earnings management (non-suspect sample).

To capture the extent of information asymmetry among investors we use an adverse selection index based on market microstructure measures estimated from high frequency data: the bid-ask spread, the illiquidity measure developed in Amihud (2002), price impact introduced by Huang & Stoll (1996), the Probability of Informed Trading (*PIN*) of Easley, Nicholas, O'Hara, & Paperman (1996), and the Volume-Synchronized Probability of Informed Trading (*VPIN*) of Easley, López de Prado, & O'Hara (2012). Bid-ask spread is a commonly used proxy for information asymmetry as it compensates liquidity providers for transacting with better-informed traders and increases with the degree of information asymmetry. The measures that capture the price impact of transactions – the illiquidity measure of Amihud (2002) and the price impact of Huang & Stoll (1996) – are important in describing the arrival of new information to market participants. The well-known *PIN* and the novel *VPIN* directly infer the presence of privately informed traders in the market from the computation of order imbalances between buys and sells. Using the index of information asymmetry, we extract the common variation in these information asymmetry proxies, so minimizing the possibility of their being driven by factors other than adverse selection (e.g. inventory costs, transactions costs, monopoly rents, etc.).

Our findings indicate that for firms which just meet last year's earnings, that is, firms with strong incentives to manage earnings, income increasing REM is associated with higher information asymmetry. This is consistent with our prediction that firms that incur in REM strategies distort earnings quality and thus, increase adverse selection among investors, because in this scenario informed investors can take advantage of their private information to assess the implication of REM activities for firm value. On the other hand, for firms which do not have incentives to meet last year's earnings, deviations from normal activity are associated with decreasing information asymmetry in the market. Thus, our findings show that deviations from normal operations affect the level of adverse selection in a contrary manner, depending on the particular underlying factors that determine them.

Our study contributes to the literature in several ways. First, it provides new evidence of the association between REM and information asymmetry on the stock market. The evidence on this topic is scarce, mixed and focuses exclusively on the US market. To the best of our knowledge, ours is the first paper that studies the effect of REM on

information asymmetry outside the US. We examine this association for Spain, a country with clearly different features from the US, not only in terms of the size and liquidity of the stock market, but also of weaker investor protection and lower accounting quality (e.g. La Porta, López de Silanes, Shleifer, & Vishny, 1998; Leuz, Nanda, & Wysocki, 2003). Nevertheless, both countries show similar levels of REM according to the international comparison carried out by Enomoto, Kimura, & Yamaguchi (2015). This study examines the differences in earnings management strategies across 38 countries, finding that Spain and U.S. show similar levels of REM, ranked 17 and 21 out of 38, respectively. Therefore, we think it is interesting to provide new evidence on how REM is perceived by investors in a setting where they have more incentives to acquire private information than in US.

Second, this paper extends the recent literature on the market consequences of REM, which has shown that REM is positively associated with the cost of equity capital (Kim & Sohn, 2013) and the cost of new corporate bonds (Ge & Kim, 2014). Based on the well-documented positive association between information asymmetry and the cost of capital, both findings can be considered as indirect evidence of REM creating information asymmetry in financial markets. Unlike the authors above, we directly test the link between REM and information asymmetry. Third, our findings suggest that private informed investors produce information in those circumstances where firms have incentives to manipulate earnings through REM activities, that is, where earnings quality is lower. However, when such incentives are not clear, private informed investors do not engage in producing private information, since the benefit from producing private information in this context is lower.

Fourth, since REM affects the quality of earnings reported by firms, our paper also extends a large body of research on the economic consequences of earnings quality and disclosure quality (e.g. Bhattacharya et al., 2013; Cormier, Houle, & Ledoux, 2013; Francis, LaFond, Olsson, & Schipper, 2005). Finally, to the best of our knowledge, this is one of the first papers to use a composite index of adverse selection to examine the effect of REM on the levels of information asymmetry in the market. Previous studies have mainly focused on individual proxies and sometimes on indirect measures of information asymmetry, such as the accuracy of financial analysts' forecasts (García Lara, García Osma, & Penalva, 2013) or the cost of capital (Ge & Kim, 2014; Kim &

Sohn, 2013). As Bharath, Pasquariello, & Guojun (2009) argue, the use of an index of information asymmetry based on market microstructure measures is more desirable than using individual proxies proposed by other areas of finance literature (e.g. analyst coverage, dispersion of analysts' forecasts, cost of capital, growth opportunities, tangibility of assets), because these measures are often inconsistent, static, persistent, or have multiple and *ad hoc* interpretations.

The rest of the paper proceeds as follows. Section 2 reviews the related literature and develops our testable hypothesis. Section 3 describes the research design, sample, and data. Section 4 presents the empirical results and the final section concludes.

## **2. Related Literature and Hypothesis Development**

### *2.1. Real Earnings Management*

Earnings management can be achieved through managerial discretion in the application of accounting standards and by changing the timing or structuring of real transactions. Traditionally, the extensive earnings management literature has mainly focused on accrual-based earnings manipulation (Xu, Taylor, & Dugan, 2007). However, there has recently been a growing research interest in the relevance and understanding of how firms manage earnings through real activities manipulation and its consequences.

The survey study conducted by Graham et al. (2005) shows that financial officers of US public firms recognize that most earnings management actions are carried out via real actions, as opposed to accounting manipulations. They also report that approximately 80% of more than 400 U.S. firms' executives surveyed admitted that they would decrease discretionary spending (including R&D, maintenance, and advertising expenses) and 55.3% said that they would delay a project in order to meet an earnings target, both of which are REM decisions. The increased importance of these managerial practices is also borne out by prior empirical research, which indicates that REM activities have increased steadily over the years, in particular substituting accrual-based earnings managements in contexts where managers are more subject to scrutiny and control of auditors and institutions. In this sense, Cohen, Dey, & Lys (2008) find for the US that the level of accrual-based (real) earnings management decreases (increases)

subsequent to the passing of the Sarbanes-Oxley Act (SOX) in 2002. This increase in REM practices could be to avoid auditors' and regulators' scrutiny. In contrast to accrual-based earnings management, where a GAAP framework exists to assess deviations from normal practices, real operations belong to the expertise of managers, and it is more difficult for outsiders (auditors, regulators, external investors, among others) to distinguish suboptimal decisions from optimal ones (Cohen et al., 2008; Cohen & Zarowin, 2010; Kothari et al., 2016).

Earnings management through REM can be defined as actions taken by managers that deviate from normal business practices to achieve certain earnings targets such as avoiding losses, maintaining or attaining positive growth in earnings, meeting analyst earnings forecasts, and smoothing earnings (Roychowdhury, 2006). Firms are found to manage earnings through manipulation of various operating and investing activities such as (Gunny, 2010; Roychowdhury, 2006): (a) Reducing discretionary expenses, including R&D spending and SGA (selling, general, and administrative) expenses, which boosts earnings and lowers cash outflows in the current period, but could lead to lower future cash flows. (b) Sales manipulation, that is, increasing price discounts (cutting prices) or extending more lenient credit terms to boost sales, increasing reported earnings and lowering current operating cash flow for a given level of sales. (c) Overproduction or increased production in order to report a lower unit cost of goods sold, which leads to increased operating margins and hence increased reported earnings. (d) Timing the income recognition from the sales of fixed-assets.<sup>1</sup> These managerial decisions, which imply changes in the underlying business transactions, have different features of accrual-earnings management: they are undertaken during the fiscal period (Zang, 2012), are hard to detect, since they could be camouflaged as normal activities (Kothari et al., 2016), and, fundamentally, they directly affect the firm's cash flow. Moreover, the deviation from normal business practices may impose a real cost on the firm, although there is a growing debate in the literature on the effects of REM on firm value (Ewert & Wagenhofer, 2005; Roychowdhury, 2006).

The literature is mixed regarding the effects and implications of REM on future performance and the value of the firm. On the one hand, as Roychowdhury (2006) and

---

<sup>1</sup>In their review of REM literature, Xu et al. (2007) consider a wider definition of REM strategies by including financing transactions. Financing activities include stock repurchases, use of stock options in compensation packages, use of financial instruments, and structuring financing transactions.



Gunny (2010) assert, REM may be opportunistic and reduce firm value because actions taken to boost current-period earnings can have a negative effect on cash flow in future periods. In this line, Bhojraj, Hribar, Picconi, & McInnis (2009) provide evidence consistent with managers' undertaking myopic actions to beat benchmarks through earnings management. In particular, they find that firms that beat analysts' forecast by cutting discretionary expenditures underperform in the long-term with respect to firms that increase discretionary expenditures and miss forecasts. Studies have also found that REM around seasoned equity offerings is associated with a subsequent decline in firm operating performance (Cohen & Zarowin, 2010) and with negative returns (Kothari et al., 2016). The perception of REM as opportunistic could lead credit agencies and bondholders to demand a higher risk premium, since the increase of earnings through REM is viewed as a factor that increases credit-risk. In this line, Ge & Kim (2014) find that sales manipulation and overproduction are associated with higher bond yield spreads, and Kim & Sohn (2013) also find a positive association between REM and the cost of capital, providing evidence that suggests that this association stems from managerial opportunism.

The opposite view is that earnings management via real activities is not opportunistic, but informative: managers engage in real earnings management to attain current-period benefits that enable better performance in the future (Bartov, Givoly, & Hayn, 2002; Gunny, 2010). Supporting this argument, Gunny (2010) finds that earnings management through REM is positively associated with firm future performance, and that those firms that engage in real activity manipulation have relatively better subsequent performances than firms that do not. In the same line, Zhao, Chen, Zhang, & Davis (2012) find that abnormal real activities intended just to meet either zero earnings or the prior year's earnings are associated with better future performance.

## *2.2. REM and Information Asymmetry*

According to the microstructure literature, information asymmetry (or adverse selection risk) in the stock market arises when there are traders with superior information who try to obtain profits by trading on the basis of their informational advantage (e.g. Bagehot, 1971; Copeland & Galai, 1983; Easley & O'Hara, 1987; Kyle, 1985). In all these models there are two types of traders in the market, informed and uninformed, trading

an asset of uncertain value. Whereas uninformed traders negotiate in financial markets for liquidity reasons and have no special information, informed traders take a position in the market based on information about the asset's true value. The informational advantages of informed market participants come from two sources. First, the informed investors may have access to private information about firm value that is not accessible to uninformed investors (insider trading). Second, traders who have a greater ability to process and interpret public information become informed traders because they can make superior assessments of the implications of this information for firm performance or value (Kim & Verrecchia, 1994). The information asymmetries among market participants create an adverse selection problem, which is typically manifested in increased trading costs and reduced levels of stock liquidity, because when liquidity providers perceive increases in the adverse selection risk, they protect themselves by widening the bid-ask spread, thereby reducing liquidity, and increasing the cost of capital (e.g. Copeland & Galai, 1983; Easley & O'Hara, 2004; Glosten & Milgrom, 1985; Kyle, 1985).

Based on the above, REM has attributes that can exacerbate the information asymmetry among investors in financial markets. First, financial executives asked in the anonymous survey by Dichev, Graham, Harvey, & Rajgopal (2013) affirmed that REM is difficult to detect and understand for analysts and other market participants outside the firm. This REM opacity could offer sophisticated investors an opportunity to profit from this private information by detecting and analyzing the potential existence of these managerial practices, thereby creating information asymmetry. In this sense, some research shows evidence that specific sophisticated investors may be interested in and concerned about earnings management practices and their implications for the long-term value of the firm. Bushee (1998), for example, shows that certain sophisticated institutional investors can, by monitoring managers, gather, interpret, and value information about managerial investment decisions and R&D spending.<sup>2</sup>

Second, since REM involves management's attempts to alter reported earnings with the aim of misleading some stakeholders, the implications of which on firm value are not

---

<sup>2</sup> Bushee (1998) hypothesizes that the monitor role of institutional investors could affect managerial incentives to manipulate R&D to meet earnings targets. In this study, we do not analyze this aspect, as we only seek to highlight that sophisticated or informed investors, unlike individual investors, are concerned about real activities manipulation and its firm's value implications.

clear, these REM practices may reduce the information content of firm earnings. Therefore, earnings manipulation through real activities could impair the market's ability to infer the firm's future cash flows and could provide the opportunity to obtain benefits to traders with higher abilities to process earnings-related information. As a consequence, REM could contribute to information asymmetry in the stock markets.

Nevertheless, the prediction of a positive relation between empirical proxies for REM and adverse selection may not hold in general. As the theoretical model developed by Zhang (2001) predicts, the level of information asymmetry across firms can be positively or negatively related to the firm's disclosure quality, depending on the factors that cause differences between firms.<sup>3</sup> Since REM measures could be capturing earnings quality or specific business circumstances, the sign of the association between measures of REM and information asymmetry may not be the same for different firms. On the one hand, consistent with our hypothesis, the informed traders have high incentives to produce private information in those settings where managers use REM practices to meet an earnings target.. Consequently, we expect to find a positive association between REM measures and the level of information asymmetry in the market. On the other hand, in those settings where the deviations from normal activities may be just a consequence of business circumstances and not of earnings management, the REM measures may affect information asymmetry in an opposite way. When empirical proxies for REM are less likely to be a proxy for poor earnings quality, the benefit of private information production may be lower and, consequently, we expect that informed investors will not engage in the production of private information. In addition, in this setting, the firm could have higher incentives to publicly disclose more information about the underlying business factors. Thus, the effect of private information production could be dominated by the effect of firm's public disclosure policy, leading to a reduction of the level of information asymmetry among investors.

---

<sup>3</sup> Zhang (2001) theoretically examines incentives behind public disclosure by the firm and trading by informed investors, the interaction between both two forms of information dissemination, and their consequences on the extent of information asymmetry among traders. Assuming that the amount of private information production by informed traders (public disclosure by the firm) increases (reduces) information asymmetry, Zhang's model derives an equilibrium in which the amount of private information production, the level of disclosure, and information asymmetry are all linked to specific characteristics of the firm.

Although prior literature suggests that earnings quality affects the information environment of the firm, most research to date has used accruals-based earnings management as a proxy for earnings quality, finding that poor earnings quality is significantly associated with higher information asymmetry (Bhattacharya et al., 2013; Cormier et al., 2013; Francis et al., 2005). However, to our knowledge, only two papers have analyzed the effect of REM on the firm information environment, and they provide unclear evidence. For a sample of NYSE firms, Ascioglu, Hedge, Krishnan, & McDermott (2012) find mixed results and weak evidence for the association between REM and liquidity. Their results depend on the proxies used: a) in some regressions they find a significant association between abnormal discretionary expenses and liquidity, but with the opposite sign to that expected; b) when they use abnormal cash flow, however, the association with liquidity proxies is, overall, not statistically significant.

Likewise, García Lara et al. (2013) provide mixed evidence for the information consequences of REM for a sample of US firms. Depending on the proxy used for the firm information environment, their findings lead to different conclusions. On the one hand, they find no evidence that REM impacts on analysts' forecast accuracy and dispersion. On the other, they report a positive association between REM and stock return volatility, which indicates that REM garbles the earnings signal and thus increases idiosyncratic volatility. Given these unclear findings regarding the association between REM and information asymmetry among market participants, we consider that it is still an open empirical question and we provide new evidence in a different context to the US market that may shed new light on whether REM is associated with the extent of adverse selection among investors in stock markets. Moreover, unlike our paper, the two previous papers do not consider the alternative interpretations of REM measures and they do not design tests to disentangle the effect of the different underlying factors which may influence the relation between empirical proxies for REM and information asymmetry. Therefore, our paper provides a more refined analysis of the influence of earnings management through real activities on the level of information asymmetry by considering the endogenous character of REM measures.

### 3. Research Design and Data

#### 3.1. Informational Asymmetry Metric

Market microstructure literature has proposed different measures and procedures to capture financial market perception about adverse selection risk, which arises when some traders possess private information not currently reflected in stock prices. In contrast to the measures introduced by corporate finance, market microstructure exploits several sources of information contained in intraday market data to capture the presence of traders with better information (informed traders). Nevertheless, in the literature there has always been a debate about the appropriateness of each proxy in measuring information-based trading. Since information asymmetry is not directly observable, all measures available are imperfect proxies for the financial market's perception of the adverse selection between informed and uninformed traders. Thus, to obtain a more complete information asymmetry measure, prior studies (e.g. Bharath et al., 2009) use principal component analysis to extract the first principal component from individual proxies of information asymmetry. In this paper, we create an adverse selection index (denoted as *ASY* hereafter) from five individual measures of information asymmetry developed by the market microstructure literature: the relative bid-ask spread, illiquidity measure developed in Amihud (2002), price impact, introduced by Huang & Stoll (1996), *PIN*, and *VPIN*.

The first and effortless proxy for asymmetric information is the bid-ask spread, a widely used measure of trading costs (liquidity). Bid-ask spread incorporates a component related to the liquidity providers' protection from being adversely selected. Glosten & Milgrom (1985) and Easley & O'Hara (1992) theoretically show that the mere presence of traders with different levels of information is reason enough for the existence of the bid-ask spread. We compute the relative quoted spread, *RQS*, as the difference between the bid and ask quotes in time  $t$  scaled by the quote mid-point as follows:

$$RQS_t = \frac{(a_t - b_t)}{Q_t} \quad (1)$$

where  $a_t$  and  $b_t$  corresponds to the ask and the bid quotes in  $t$ .  $Q_t = (a_t + b_t)/2$  is the quoted midpoint in  $t$ , commonly used as a proxy for the efficient price. First, we computed  $RQS$  on a daily basis by averaging (time-weighted) all the observations within the day. After that, we obtained an annual  $RQS$  by averaging (equally-weighted) daily values.

Since adverse selection is an important determinant of stock liquidity, we estimate the index of illiquidity introduced by Amihud (2002), which is a volume-based liquidity indicator and is defined as

$$AMH_t = \frac{1}{D_t} \sum_{d=1}^{D_t} \frac{|R_{dt}|}{V_{dt}} \quad (2)$$

where  $R_{d,t}$  is the return on day  $d$  of year  $t$ ,  $V_{dt}$  is the volume in euros on day  $d$  of year  $t$ , and  $D_t$  is the number of days for which data are available in year  $t$ . Like Amihud (2002), we multiply  $AMH$  by  $10^6$ . Amihud's illiquidity measure gives the average of the daily price impact of the order flow or absolute percentage price change associated with a unit of trading volume. When a stock is liquid, large trading volumes provoke small price changes. Therefore, higher values of  $AMH$  indicate higher price moves in response to trading volume, and thus higher stock illiquidity. It is expected that the greater the information asymmetry, the worse the stock liquidity, and the higher the  $AMH$  value.

Both bid-ask spread and illiquidity ratio are noisy proxies for asymmetric information given that they commonly include other components that are not related to information (inventory costs, order processing cost, monopoly rents, etc.), but that they also influence stock liquidity. Moreover, the illiquidity index of Amihud (2002) provides a rough measure of the price impact. Trades initiated by noise traders lead to transitory changes in transaction prices, while information-based trades provoke permanent price changes. Thus, Huang & Stoll (1996) introduce the realized spread (or price reversal) and the price impact by considering the quote adjustment that takes place a period of time after a trade to extract the presence of new information. Price impact ( $PI$ ) is the permanent price change (or information content) of a trade and is defined as

$$PI_{t+\tau} = (Q_{t+\tau} - Q_t)X_t \quad (3)$$

where  $Q_t$  is the quote midpoint defined previously,  $X_t$  is a trade indicator variable taking the value -1 if the trade in  $t$  is initiated in the sell side and 1 if it is initiated in the buy side. Finally,  $\tau$  is the period of time for prices to fully reflect the information content in trade  $t$ . Like Huang & Stoll (1996), we set  $\tau$  equal to 30 minutes. A daily  $PI$  is computed in trade-time by averaging (volume-weighted) all the trades within the day. Then, we obtain an annual value by averaging (equally weighted) all the trading days in the year. A large and positive  $PI$  indicates a high frequency of information-based trades.

The fourth measure of information asymmetry considered to compute our index is the probability of information-based trading ( $PIN$ ), a measure that can be included in the group of the asymmetric information measures based on the computation of order imbalances between buys and sells to extract the information content of the trading process. The  $PIN$  is a measure based on the theoretical work of Easley & O'Hara (1987, 1992), with the original  $PIN$  model introduced by Easley et al. (1996). The  $PIN$  is the unconditional probability that a randomly selected trade originates from an informed trader. The  $PIN$  is not directly observable but as a function of the theoretical parameters of a microstructure model that have to be estimated by numerical maximization of a likelihood function. Once the parameters of interest are estimated, the  $PIN$  is calculated as the ratio of orders from informed traders to the total number of orders. For reasons of space, the description of the model and the estimation process of this well-known methodology are presented in Appendix A.

As an update of the  $PIN$  model, Easley et al. (2012) have developed a new measure for adverse selection risk called Volume-Synchronized Probability of Informed Trading or  $VPIN$ . The  $VPIN$  approach has some practical advantages over the  $PIN$  methodology that make it particularly attractive for both practitioners and researchers. The main advantage is that  $VPIN$  does not require the estimation of non-observable parameters using optimization or numerical methods, thereby avoiding all the associated computational problems and biases. In particular,  $VPIN$  measures order flow toxicity, which can be considered as a broader concept for adverse selection applied to the particular world of liquidity providers in a high frequency trading (HFT) environment. However,  $VPIN$  can be considered as a more flexible measure of asymmetric information that can be applied in a wide range of frameworks by choosing the appropriate values of the variables involved in the estimation process (Abad & Yagüe,

2012). There are three relevant variables in the *VPIN* approach: time bar, volume bucket and sample length. At bar level, trade flow is split between buys and sells. At bucket level, order imbalances are computed. Finally, order imbalances are smoothed in the sample length by computing a moving average and the *VPIN* series is obtained. An annual *VPIN* is computed by averaging all the values of the result *VPIN* series into the year. A brief description of this procedure can be found in Appendix B.

### 3.2. REM Measures

Roychowdhury (2006) develops three measures of real activities manipulation (abnormal cash flows, abnormal production costs, and abnormal discretionary expenses) to focus on three methods of manipulating real activities in order to manage earnings upwards: (1) sales manipulation through increased price discounts or more lenient credit terms, to temporarily boosts sales revenues, which will have the effect of unusually low cash flow levels from operations; (2) overproduction, to report a lower cost of goods sold; and (3) reduction of discretionary expenses. Following previous research on REM (Cohen et al., 2008; Cohen & Zarowin, 2010; Ge & Kim, 2014; Kim & Sohn, 2013; Roychowdhury, 2006), we employ the three models proposed by Roychowdhury (2006) to construct REM measures. We use model (4) to estimate the normal level of cash flow from operations:

$$\frac{CFO_t}{Assets_{t-1}} = \alpha_0 + \alpha_1 \left( \frac{1}{Assets_{t-1}} \right) + \beta_1 \left( \frac{Sales_t}{Assets_{t-1}} \right) + \beta_2 \left( \frac{\Delta Sales_t}{Assets_{t-1}} \right) + \varepsilon_t \quad (4)$$

where *CFO* is cash flow from operations estimated as operating income less total accruals, *Sales* and  $\Delta Sales$  represents sales and change in sales, respectively. All variables, including the intercept, are scaled by lagged total assets (*Assets*). We also include an unscaled intercept (Roychowdhury, 2006).

We use model (5) to estimate the normal level of production costs:

$$\frac{PROD_t}{Assets_{t-1}} = \alpha_0 + \alpha_1 \left( \frac{1}{Assets_{t-1}} \right) + \beta_1 \left( \frac{Sales_t}{Assets_{t-1}} \right) + \beta_2 \left( \frac{\Delta Sales_t}{Assets_{t-1}} \right) + \beta_3 \left( \frac{\Delta Sales_{t-1}}{Assets_{t-1}} \right) + \varepsilon_t \quad (5)$$



where *PROD* is production costs defined as the sum of costs of goods sold, which we estimate from the profits and losses account, plus the change in inventory in the year. The other variables have been defined previously.

We estimate the normal level of discretionary expenses with model (6):

$$\frac{DISPEX_t}{Assets_{t-1}} = \alpha_0 + \alpha_1 \left( \frac{1}{Assets_{t-1}} \right) + \beta_1 \left( \frac{Sales_{t-1}}{Assets_{t-1}} \right) + \varepsilon_t \quad (6)$$

where *DISPEX* is discretionary expenses and the other variables are calculated as defined previously. Since in Spain firms do not usually report advertising or general and administrative expenses specifically, we measure *DISPEX* with the item *other operating expenses* in the profits and losses statement, which includes R&D, advertising, and other general expenses.

We estimate models (4), (5) and (6) cross-sectionally for each year and industry group using all the data available on Spanish listed firms in the period. Based on the industry classification of the Madrid Stock Exchange, we classify firms into three big industries in order to have a minimum of 15 observations for each regression. For every firm-year, the residuals of the regressions represent, respectively, the abnormal cash flow from operations (*ACFO*), the abnormal productions costs (*APROD*), and the abnormal discretionary expenses (*ADISEXP*). Firms that manage earnings upwards will show abnormally low cash flows from operations, and/or abnormally high productions costs, and/or abnormally low discretionary expenses (Cohen & Zarowin, 2010). Accordingly, for abnormal cash flows and abnormal discretionary expenses, we multiply the residuals of models (4) and (6) by (-1), so that higher values of these variables represent greater increases of earnings.<sup>4</sup> We separately analyze each measure and also define three combined measures of REM. Following Cohen & Zarowin (2010), we define *REMI* as *APROD+ADISEXP*; and *REM2* as *ACFO+ADISEXP*. Thus, higher values of *REMI* and *REM2* indicate higher probability of real decisions to increase earnings, in particular,

---

<sup>4</sup> Income increasing real earnings management does not always affect cash flows and earnings in the same direction (Roychowdhury, 2006) because, whereas price discount and overproduction have a negative effect on cash flows, cutting discretionary expenses has a positive effect. Although this has led some studies to disregard abnormal cash flows in REM measures, and thus focus only on abnormal production costs and abnormal discretionary expenses, other authors include abnormal cash flows in order to take into account the possibility of sales manipulation.

that the firm is engaged in higher production costs and cutting discretionary expenses (*REM1*), and in sales manipulation and cutting discretionary expenses (*REM2*). Finally, we construct *REM3* as an overall measure of *REM* as  $ACFO + APROD + ADISEXP$  (Kim, Park, & Wier, 2012; Ge & Kim, 2014).<sup>5</sup>

### 3.3. Incentives to Engage in Earnings Management

REM measures, as defined in the previous section, are the residuals of the models developed by Roychowdhury (2006) to estimate normal levels of cash flow, production costs and discretionary expenses. However, the abnormal levels of these measures, apart from earnings management activities, could be capturing abnormal situations caused by incompetent managers or unusual changes in the business conditions. Prior research addresses this concern by analysing REM activities and their effects in settings in which earnings management is likely to occur, such as firms that use REM to meet earnings benchmarks (Gunny, 2010; Kim & Sohn, 2013; Roychowdhury, 2006; Zang, 2012). To avoid the association between proxies for information asymmetry and REM being explained by factors unrelated to managerial opportunism, we implement our analyses for two different settings: (a) firms with strong incentives to opportunistically manage earnings (suspect firms); and (b) the rest of the sample (non-suspect firms). In particular, we consider as a sample of suspect firm-year observations those that just meet last year's earnings. For each firm-year, we compute net income on total assets and suspect firm-years are those whose change in net income divided by total assets is between 0 and 0.01. Thus, in the first scenario we assume that deviations from normal operations represent REM decisions, whereas in the second setting these deviations are more likely to represent unusual business conditions.

### 3.4. Regression Model

We test the association between REM and the level of information asymmetry between traders in the stock market with the following model:

---

<sup>5</sup>Note that *ACFO* and *ADISEXP* are the residuals of models (4) and (6) multiplied by (-1), so these are the values we add to *APROD* in *REM1*, *REM2* and *REM3*.

$$ASY = \beta_0 + \beta_1 REM + \beta_2 DiscAcc + \beta_3 Size + \beta_4 ROA + \beta_5 Turnover + \beta_6 Volat + \beta_7 Analysts + \beta_8 Own + \sum_t \beta_t Year + \sum_j \beta_j Ind + \varepsilon \quad (7)$$

where *ASY* is our index of information asymmetry and *REM* corresponds to each of the different REM measures described in the previous section. We include variables in the regression to control for factors that, according to the previous literature, affect the information environment of a firm and that are likely to be associated with information asymmetry among investors in the capital markets. These control variables are: discretionary accruals (*DiscAcc*), firm size (*Size*), return on assets (*ROA*), trading volume (*Turnover*), stock volatility (*Volat*), financial analyst following (*Analyst*) and ownership concentration (*Own*).

As commented on above, previous studies document that accrual-based earnings management and REM can be used as substitutes to manipulate earnings (Cohen et al., 2008; Cohen & Zarowin, 2010; Zang, 2012) and that earnings quality is associated with information asymmetry proxies. Empirical studies, such as Francis et al. (2005) and Bhattacharya et al. (2013), use discretionary accruals as a proxy for earnings quality and suggest that poor earnings quality enhance information asymmetry among investors. This supports the opportunistic view of accrual-based earnings management, which assumes that the objective of these accounting practices is to garble the market, resulting in an increase in the adverse selection risk. However, some studies support an informational view of discretionary accounting choices. According to these studies, if investors detect accrual-based earnings management, discretionary accruals might not be a noisy signal but could, in contrast, be informative about firm future cash flows. This would improve the informativeness of earnings (e.g. Subramanyam, 1996) and, as a consequence, more informative financial reporting could minimize the informational advantages of informed traders. Therefore, since we control for discretionary accruals, *DiscAcc*,  $\beta_1$  represents the incremental effect on information asymmetry of REM once accrual-based earnings management is taken into account. *DiscAcc* is calculated as the value of discretionary accruals estimated by the Jones (1991) model, modified by Dechow, Sloan, & Sweeney (1995).

Market microstructure literature provides extensive empirical evidence of firms' characteristics that are related to the probability of informed trading and, consequently,

stock liquidity. In particular, stocks of larger and more profitable firms and stocks with larger trading volumes and lower return volatility suffer lower adverse selection problems and are more liquid (e.g. Easley et al., 1996; Stoll, 2000; Goh, Lee, Ng, & Yong, 2016). This is consistent with the widely known argument that larger and more profitable firms, and firms whose stocks are more frequently traded, have richer information environment as a consequence of their higher levels of information production and publicly available information. Additionally, the positive relation between information asymmetry and stock volatility suggests a higher presence of informed traders due to the greater profit opportunities in stocks that have higher information uncertainty (Bhattacharya et al., 2013). Hence, we include *Size*, the company's size measured as the natural logarithm of total assets; *ROA*, defined as operating income divided by total assets; *Turnover*, the logarithm of the average daily trading volume in euro scaled by the market value of the firm's equity at the end of the year; and *Volat*, a proxy for stock return volatility calculated as the standard deviation of daily returns.

Disclosure literature also predicts that the information environment of a firm is affected by the activities of producing and disseminating information performed by financial analysts following the firm. However, neither theoretical nor empirical studies are totally conclusive about the sign of the relation between analyst following and information asymmetry. For example, Easley, O'Hara, & Paperman (1998) state that the number of analysts following the firm can be either positively or negatively associated with the level of disclosure and with the probability of informed trading depending on whether financial analysts create new private information or disseminate public information among investors. Although it is possible to find some empirical studies that provide findings suggesting that the number of analyst following a stock is positively correlated with information asymmetry (e.g. Chung, McInish, Wood, & Wyhowski, 1995), the great majority report that analyst coverage is negatively related to information asymmetry (e.g. Easley et al., 1998; Roulstone, 2003). This inverse relation supports the argument that more analyst following increases publicly available information on the firm, which results in a reduction in the risk of information-based trading and an improvement in stock liquidity. Therefore, we include in our model the variable *Analyst*, which represents the natural log of the total number of analysts following a firm.

In addition, the distribution of private information among investors can be affected by the predominance of large shareholders in the firm ownership. For this reason, we include ownership concentration, *Own*, as a control variable measured by the percentage of common shares held by the largest five shareholders of the company. A more highly concentrated ownership is expected to be positively associated with information asymmetry because the larger shareholders are likely to control the firm and therefore to have access to, or generate, private information about the firm, so exacerbating adverse selection problems in the market (e.g. Heflin & Shaw, 2000).

Finally, we also include year and industry dummy variables to control for temporal and industry effects.

### 3.5. Sample and Data

Our sample is made up of stocks traded on the electronic trading platform of the Spanish Stock Exchange, known as the SIBE (*Sistema de Interconexión Bursátil Español*). The SIBE is an order-driven market where liquidity is provided by an open limit order book. Trading is continuous from 9:00 a.m. to 5:30 p.m. There are two regular call auctions each day: the first determines the opening price (8:30-9:00 a.m.), while the second sets the official closing price (5:30-5:35 p.m.). Three basic types of orders are allowed: limit orders, market orders, and market-to-limit orders. In the continuous session, a trade occurs whenever an incoming order matches one or more orders on the opposite side of the limit order book. Orders submitted that are not instantaneously executed are stored in the book, waiting for a counterparty, according to a strict price-time priority rule. Unexecuted orders can always be cancelled and modified. Continuous trading can be temporally interrupted, since a system of stock-specific intraday price limits and short-lived call auctions is implemented to handle unusual volatility levels. In all auctions (open, close and volatility) orders can be submitted, modified or cancelled, but no trades occur.

Trade and quote data for this study come from SM data files provided by the *Sociedad de Bolsas, S.A.* SM files contain detailed time-stamped information about the first level of the limit order book for each stock listed on the SIBE. Any trade, order submission and cancellation affecting best prices in the book generates a new entry in the file. The

distinction between buyer-initiated and seller-initiated trades is straightforward and no classification algorithm is needed. Firms' financial statement data were taken from the SABI database, made by Bureau Van Dijk, and from the annual reports at the Spanish Securities Market Commission (*Comisión Nacional del Mercado de Valores, CNMV*). Ownership concentration and analysts' data were collected from Thomson Reuters Eikon Datastream.

Our sample consists of non-financial firms listed on the main segment of the SIBE in the period 2001-2008, with full data available for all the period. After applying the usual filters to detect and eliminate errors in the preparation of the intraday trading data and combining the different databases, we obtain 468 firm-year observations, for which we have been able to collect the information asymmetry measures, the complete financial-accounting information and data on analyst following and ownership concentration.

## 4. Results

### 4.1. Descriptive Statistics

Table 1 reports the estimations of normal levels of cash flow from operations, production costs and discretionary expenses –models (4), (5), and (6). We estimate these models using all the available information for Spanish listed firms during the period 2001-2008. The regressions are estimated for industry-year groups with at least 15 observations. The table reports the mean coefficients across all industry-years and  $t$ -statistics calculated using the standard error of the mean across industry-years, as well as the mean  $R^2$  across industry-years. We can see that the models explain the real operations quite well, and our results are similar to those reported by Roychowdhury (2006) for US firms.

[Table 1 near here]

Table 2 presents descriptive statistics for the information asymmetry measures (Panel A), REM measures (Panel B), and control variables (Panel C). The mean, median, standard deviation, 10th percentile and 90th percentile are reported for each. With

regard to proxies for information asymmetry, the mean (median) of *RQS* is 0.6% (0.4%). The mean value of *AMH* in our sample (0.3) is similar to that reported by Amihud (2002). The average (median) of price impact measure (*PI*) is 0.37% (0.30%). According to Abad & Yagüe (2012), the *PIN* and the *VPIN* show similar mean values of around 19% and 20%, respectively. *PIN* values are also consistent with those reported in prior studies that use this information asymmetry proxy (e.g., Brown & Hillegeist, 2007; Easley, Hvidkjaer, & O'Hara, 2002). The statistical distributions of the above measures show that there are clear differences in the degree of asymmetric information among firms included in our sample. As discussed in Section 3, in order to isolate the common adverse selection component underlying the former proxies, we constructed an index of information asymmetry (*ASY*) by employing principal components analysis (PCA) for each firm and year of our sample. The mean of *ASY* is zero (by construction) and its median is -0.40. The first (and only) factor with an eigenvalue greater than one explains 64.3% of the variance and each component of *ASY* enters with a positive sign and loadings as follows:

$$ASY = 0.511RQS + 0.257AMH + 0.477PI + 0.459PIN + 0.484VPIN \quad (8)$$

Therefore, each proxy for information asymmetry plays its role in the index.<sup>6</sup> A higher value of the index means a higher level of adverse selection. As seen in Table 3, which provides the Pearson correlation matrix between the variables used in the study, all the information asymmetry proxies are positively correlated with each other, which indicates that these measures are likely to be driven by adverse selection, but each contains unique information. Moreover, the index is positive and significantly correlated with each information asymmetry variable, varying from a correlation of 92% between *ASY* and *RQS* to a correlation of 46% between *ASY* and *AMH*. Additionally, correlations between all five proxies for information asymmetry and the index are

---

<sup>6</sup> A potential concern about the use of *ASY* as proxy for information asymmetry for our sample is that the PCA is sensitive to sample size. To check the robustness of the index, we evaluate the performance of the PCA by applying computer-based resampling (bootstrap) techniques. Thus, we draw a large number of samples (1,000, 5,000, and 10,000) of different sizes —smaller than (234 observations), equal to (468), and larger than (1,000) our sample size. We perform PCA analysis to all the samples and compute confidence intervals (basic percentile) at the 1% level. We observe that our full-sample estimations for all relevant parameters (the eigenvalues and the component weights for the first factor) are always included in the bootstrap intervals.

generally higher than between them, which suggests that the index is a parsimonious way of measuring information asymmetry.

[Table 2 near here]

Mean values of *ACFO*, *APROD*, *ADISPEXP* and *DiscAcc* are very close to zero, as expected. Their deviation from zero is due to these variables having been estimated with all the available information for listed firms in the period, which is higher than the size of our sample. With regard to the control variables (*Size*, *ROA*, *Turnover*, *Volat*, *Analysts*, and *Own*), these show a significant level of dispersion in their values, reflecting the heterogeneity of our firm-year sample. The correlations between REM variables are positive and significant, which means that firms simultaneously use different strategies of real activities manipulation to achieve their earnings objectives. Firms also simultaneously apply sales manipulation and discretionary accruals strategies to manipulate earnings, since *ACFO* and the REM variables that include *ACFO* are highly and positively correlated with *DiscAcc*. The positive correlations between *Size* and most of REM variables show that bigger firms are more likely to engage in REM activities to increase earnings, and the negative correlations between *ROA* and REM measures suggest that firms with better performance are less prone to managing earnings through real activities manipulation.

[Table 3 near here]

#### 4.2. Analysis for the Whole Sample

Table 4 reports the results of model (7) for the different measures of REM in the whole sample. Columns (1) to (3) present the results for individual proxies while columns (4) to (6) do so for the aggregate measures. Since Table 3 shows high correlations between *DiscAcc* and *ACFO* and *DiscAcc* and *REM2*, models in columns (1) and (5) are estimated using orthogonalized variables with respect to *DiscAcc*, that is, they incorporate the residuals of the regression of *ACFO* on *DiscAcc* and the residuals of the regression of *REM2* on *DiscAcc*, respectively. We report OLS coefficients and *t*-statistics (in brackets) based on robust standard errors, which are clustered by firm.



-Insert Table 4-

The results do not show significant associations between REM proxies and information asymmetry. These insignificant effects could be explained by the heterogeneity of the sample, composed by firms with different incentives to engage in REM activities. Thus, the findings for the overall sample could be showing an offset between a positive and a negative effect in different scenarios, depending on whether or not there are incentives to engage in REM. Neither is accrual-based earnings management significantly associated with information asymmetry in the Spanish market.<sup>7</sup> This finding, which is not consistent with most of the empirical evidence in the literature, along with the findings reported below for the analysis of two subsamples, suggests that, like the REM effect on information asymmetry, the effect of accrual-based earnings management may depend on the incentives to produce private information in relation to accrual-based manipulation of earnings.

Regarding the other control variables, the signs of their coefficients are as expected according to the literature. We find that the stocks of larger and more profitable firms, with higher trading volume, and those being followed by more analysts show less information asymmetry, whereas firms with more volatile stock returns are associated with higher information asymmetry. All these variables are significant at the 1% level in all models estimated. The coefficient on *Own*, as expected, always presents a positive sign, but it is not significant.

#### *4.3. Analysis for the Suspect and Non-Suspect Samples*

In the previous analysis, we have examined the association between REM measures and information asymmetry in the whole sample. However, since empirical proxies for REM represent abnormal levels of real transactions, they could be capturing the consequences of opportunistic managerial practices, but also specific business circumstances unrelated to earnings management, such as changes in business or unique business models. Consequently, the sign of the association between measures of REM

---

<sup>7</sup>The results do not change (the coefficient on discretionary accruals is not significant) if we include the absolute value of discretionary accruals, as in Kim et al. (2012).

and the level of information asymmetry may depend on these sources of variation among firms in a particular sample.

Hence, in order to extend the understanding of the association between REM practices and information asymmetry, we analyse this association in two subsamples: (a) *suspect sample*, that is, firm-years observations with strong incentives to manage earnings in order to just meet zero earnings growth (last year's earnings), and (b) *non-suspect sample*, that is, firm-years observations without incentives to meet this target. In settings where managers have strong incentives to manage earnings and the reason for earnings management is well-understood, such as the aim to meet an earnings target (Graham et al., 2005), we would expect investors to enhance the private information production to increase their trading profits. This would have the effect of raising the level of information asymmetry in the market. However, in settings without incentives to engage in earnings management, deviations from normal operations can be attributed to other circumstances, rather than the firm's disclosure quality. Since in this case there are fewer incentives to produce private information, we could expect that these deviations from normal operations do not create information asymmetry in the market.

As a preliminary analysis, we compare REM in the suspect sample versus the non-suspect sample. Following Roychowdhury (2006), we run the following regression:

$$REM = \beta_0 + \beta_1 Size + \beta_2 BTM + \beta_3 ROA + \beta_4 Suspect + \sum_t \beta_t Year + \sum_j \beta_j Ind + \varepsilon \quad (9)$$

where *REM* corresponds to each of the different REM measures as described previously, *BTM* is the book to market ratio, *ROA* is operating income divided by total assets, and *Suspect* is a dummy variable that takes the value of 1 if the change in net income divided by total assets is between 0 and 0.01, and 0 otherwise. Table 5 displays the estimate of the model for the six REM measures. The results in columns 2, 3, 4, and 6, show that the coefficients on *Suspect* are positive and significant (at the 10% level, in the *REM3* model; at the 5% level in the abnormal production costs model; and at the 1% level in the abnormal discretionary expenses and *REMI* models). This suggests that suspect firm-years have higher abnormal production costs and lower discretionary expenses than the rest of firms, which is consistent with their engagement in REM

activities to meet last year's earnings. The non-significant coefficients for the abnormal cash flow model and for *REM2* could be explained by the opposite effect on cash flows of cutting discretionary expenses in relation to sales manipulation and increasing production costs.

[Table 5 near here]

Table 6 reports the results of model (7) for the two subsamples. Panel A corresponds to the suspect sample and Panel B to the non-suspect sample. We report OLS coefficients and *t*-statistics (in brackets) based on robust standard errors that are clustered by firm.

[Table 6 near here]

In the sample with strong incentives to manage earnings (suspect sample), we find that, with the exception of the cash flow model, all coefficients on REM measures are positive and significant (at the 5% level, in the abnormal discretionary expenses and *REM2* models, and at the 1% level in abnormal production costs, *REM1* and *REM3* models). These findings are consistent with the hypothesis that managerial opportunism to increase earnings through REM creates information asymmetry in the market in those contexts where managers have incentives to engage in REM activities, and consequently, informed investors also have incentives to produce private information. However, as in the whole sample, the coefficient on *DiscAcc* is not statistically significant. Therefore, the strong and clear effect of REM and the non-significant effect of accrual-based earnings management on information asymmetry could be explained by the different implications of both types of earnings management. The larger opacity and real effects of REM for firm value in comparison to accrual-based management may imply a higher marginal benefit for the production of private information. As a consequence, informed traders may have more incentives to produce information about REM than in relation to accrual-based management. The signs and significance of the other control variables are quite similar to those of the analysis with the whole sample, but in the subsample of suspect firms we also find that *ROA* is not significant, which can be explained by its low variability or because earnings numbers are noisier or less credible due to the high likelihood of REM in this setting. In contrast, we find a significant positive association between ownership concentration and information

asymmetry, which is consistent with lower disclosure levels in firms with a predominance of large shareholders.

In the rest of the sample (non-suspect sample), we find that, with the exception of the cash flow model, all coefficients on REM measures are significantly negative (at the 10% level in the abnormal production costs and at the 1% level in abnormal discretionary expenses, *REM1*, *REM2* and *REM3* models). Thus, in this setting there is a negative association between information asymmetry and REM measures after taking into account the effect of accrual-based earnings management. This finding suggests that when REM measures are not reflecting low earnings quality but change in business, the informed traders have fewer incentives to produce private information. Regarding the control variables, the coefficient on discretionary accruals (*DisAcc*) is positive and significant at the 1% level in the *REM3* model or quite close to being significant at conventional levels in the rest of models (with the exception of the cash flow model) at two-tail tests. Even if we consider one-tail tests the coefficients on *DisAcc* are significant in all estimations with the exception of the cash flow model. This is consistent with previous research, which has found that accruals earnings management may create information asymmetry in the market. Finally, the coefficients of the rest of control variables show similar signs and significance levels to those presented for the whole sample.

## 5. Conclusion

This study examines the consequences of real activities manipulation on information asymmetry in Spain. Previous studies have examined this association basically for US markets, providing inconclusive evidence. We consider that the analysis of the Spanish market may shed new light because it exhibits a weaker investor protection, lower accounting quality and stock market liquidity, and higher incentives for investors to search for private information than US. We use 468 firm-year observations from 2001 to 2008 and an information asymmetry index built on microstructure measures such as the bid-ask spread, illiquidity measure developed in Amihud (2002), price impact introduced by Huang & Stoll (1996), *PIN*, and *VPIN*.

In line with previous literature, we find that firms with high strong incentives to engage in earnings management to just meet last year's earnings, show higher levels of income increasing REM. Overall, our evidence on the association between REM and information asymmetry is consistent with the prediction that firms' strategies of REM garble the market and create information asymmetry among traders. Thus, in a setting where REM measures are highly likely to indicate low earnings quality (suspect sample) we find a significant and positive association between proxies for earnings management through real activities manipulation and information asymmetry among investors. In contrast, in a setting where the empirical proxies for REM could be capturing situations rather related with business circumstances than with earnings manipulation, we find that deviations from normal activity are significantly and negatively associated with the level of information asymmetry. Thus, we show that the private information production and its influence on the level of information asymmetry in the market depend on firm's circumstances.

Our results have implications for managers, regulators and researchers. Our evidence confirms that managers will possibly manipulate earnings with real activities to meet earnings benchmarks. We add to the literature that these practices may distort the market by creating information asymmetry between traders in those contexts where managers have incentives to engage in REM, since it raises the production of private information by sophisticated investors. We extend previous research on the economic consequences of earnings management in general, and REM in particular, to the study of the adverse selection problem in financial markets, where prior literature is scarce and focused exclusively on the US market. By examining a direct link between REM and information asymmetry, we show that previous findings that associate the engagement in REM activities with a higher cost of capital can be explained by the increase in information asymmetry produced by REM. Finally, from a methodological point of view, this is one of the first papers to use a composite index of adverse selection to examine this issue.

## References

- Abad, D., & Yagüe, J. (2012). From PIN to VPIN: An introduction to order flow toxicity. *The Spanish Review of Financial Economics*, 10 (2), 74-83.
- Akay, O., Cyree, H.B., Griffiths, M.D., & Winters, D.B. (2012). What does PIN identify? Evidence from the T-bill Market. *Journal of Financial Markets*, 15 (1), 29-46.
- Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5 (1), 31-56.
- Ascioglu, A., Hedge, S.P., Krishnan, G.V., & McDermott, J.B. (2012). Earnings management and market liquidity. *Review of Quantitative Finance and Accounting*, 38 (2), 257-274.
- Aslan, H., Easley, D., Hvidkjaer, S., & O'Hara, M. (2011). The characteristics of informed trading: Implications for asset pricing. *Journal of Empirical Finance*, 18 (5), 782-801.
- Bagehot, W. (1987). The only game in town. *Financial Analysts Journal*, 27 (2), 12-14.
- Bartov, E., Givoly, D., & Hayn, C. (2002). The rewards for meeting-or-beating earnings expectations. *Journal of Accounting and Economics*, 33 (2), 173-204.
- Bharath, S.T., Pasquariello, P., & Guojun, W. (2009). Does asymmetric information drive capital structure decisions?. *The Review of Financial Studies*, 22 (8), 3211-3243.
- Bhattacharya, N., Desai, H., & Venkataraman, K. (2013). Does earnings quality affect information asymmetry? Evidence from trading costs. *Contemporary Accounting Research*, 30 (2), 482-516.
- Bhojraj, S., Hribar, P., Picconi, M., & McNinnis, J. (2009). Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts. *The Journal of Finance*, 64 (5), 2359-2386.
- Brown, S., & Hillegeist, S. (2007). How disclosure quality affects the level of information asymmetry. *Review of Accounting Studies*, 12 (2-3), 443-477.
- Bushee, B.J. (1998). The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review*, 73 (3), 305-333.
- Chung, K.H., McInish, T.H., Wood, R.A., & Wyhowski, D.J. (1995). Production of information, information asymmetry, and the bid-ask spread: Empirical evidence from analysts' forecast. *Journal of Banking and Finance*, 19 (6), 1025-1046.
- Cohen, D.A., Dey, A., & Lys, T.Z. (2008). Real and accrual based earnings management in the Pre and Post Sarbanes Oxley periods. *The Accounting Review*, 83 (1), 757-787.
- Cohen, D.A., & Zarowin, P. (2010). Accrual-based and real earnings management activities around seasoned equity offerings. *Journal of Accounting and Economics*, 50 (1), 2-19.
- Copeland, T.E., & Galai, D. (1983). Information effects on the bid-ask spread. *The Journal of Finance*, 38 (5), 1457-1469.

- Cormier, D., Houle, S., & Ledoux, M.J. (2013). The incidence of earnings management on information asymmetry in an uncertain environment: Some Canadian evidence. *Journal of International Accounting, Auditing & Taxation*, 22 (1), 26-38.
- Dechow, P.M., & Skinner, D.J. (2000). Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons*, 14 (2), 235-250.
- Dechow, P.M., Sloan, R.G., & Sweeney, A.P. (1995). Detecting earnings management. *The Accounting Review*, 70 (2), 193-225.
- Dichev, I.D., Graham, J.R., Harvey, C.R., & Rajgopal, S. (2013). Earnings quality: Evidence from the field. *Journal of Accounting and Economics*, 56 (2), 1-33.
- 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., Kiefer, N., & O'Hara, M. (1997). One day in the life of a very common stock. *Review of Financial Studies*, 10 (3), 805-835.
- Easley, D., López de Prado, M.M., & O'Hara, M. (2012). Flow toxicity and liquidity in a high-frequency world. *The Review of Financial Studies*, 25 (5), 1457-1493.
- Easley, D., Nicholas, M.K., O'Hara, M., & Paperman, J. (1996). Liquidity, information, and infrequently traded stocks. *The Journal of Finance*, 51 (4), 1405-1436.
- Easley, D., & O'Hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics*, 19 (1), 69-90.
- Easley, D., & O'Hara, M. (1992). Time and the process of security price adjustment. *The Journal of Finance*, 47 (2), 577-605.
- Easley, D., & O'Hara, M. (2004). Information and the cost of capital. *The Journal of Finance*, 59 (4), 1553-1583.
- Easley, D., O'Hara, M., & Paperman, J. (1998). Financial analysts and information-based trade. *Journal of Financial Markets*, 1 (2), 175-201.
- Enomoto, M., Kimura, F., & Yamaguchi, T. (2015). Accrual-based and real earnings management: An international comparison for investor protection. *Journal of Contemporary Accounting and Economics*, 11 (3), 183-198.
- Ewert, R., & Wagenhofer, A. (2005). Economic effects of tightening accounting standards to restrict earnings management. *The Accounting Review*, 80 (4), 1101-1124.
- Francis, J., LaFond, R., Olsson, P., & Schipper, K. (2005). The market pricing of accruals quality. *Journal of Accounting and Economics*, 39 (2), 295-327.
- García Lara, J.M., García Osma, B., & Penalva, F. (2013). Differential information consequences of real versus accruals earnings management. Working paper, 36th Annual Congress of the European Accounting Association, Paris.
- Ge, W., & Kim, J.B. (2014). Real earnings management and the cost of new corporate bonds. *Journal of Business Research*, 67 (4), 641-647.
- Glosten, L.R., & Milgrom, P.R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14 (1), 71-100.

- Goh, B.W., Lee, J., Ng, J., & Yong, K.O. (2016). The effect of board independence on information asymmetry. *European Accounting Review*, 25 (1), 155-182.
- Graham, J.R., Harvey, C.R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, 40 (1-3), 3-73.
- Gunny, K.A. (2010). The relation between earnings management using real activities manipulation and future performance: Evidence from meeting earnings benchmarks. *Contemporary Accounting Research*, 27 (3), 855-888.
- Healy, P., & Wahlen, J. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting Horizons*, 13 (4), 365-383.
- Heflin, F., & Shaw, K.W. (2000). Blockholder ownership and market liquidity. *The Journal of Financial and Quantitative Analysis*, 35 (4), 621-633.
- Huang, R., & Stoll, H. (1996). Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and NYSE. *Journal of Financial Economics*, 41 (3), 313-357.
- Jayaraman, S. (2008). Earnings volatility, cash flow volatility and informed trading. *Journal of Accounting Research*, 46 (4), 809-851.
- Jones, J.J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research*, 29 (2), 193-228.
- Kim, J.B., & Sohn, B.C. (2013). Real earnings management and cost of capital. *Journal of Accounting and Public Policy*, 32 (6), 519-543.
- Kim, O., & Verrecchia, R.E. (1994). Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics*, 17 (1-2), 41-67.
- Kim, Y., Park, M.S., & Wier, B. (2012). Is earnings quality associated with corporate social responsibility?. *The Accounting Review*, 87 (3), 761-796.
- Kothari, S.P., Mizik, N., & Roychowdhury, S. (2016). Managing for the moment: The role of real activity versus accruals earnings management in SEO valuation. *The Accounting Review*, 91 (2), 559-586.
- Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6), 1315-1335.
- La Porta, R., López de Silanes, F., Shleifer, A., & Vishny, R. (1998). Law and finance. *Journal of Political Economy*, 106 (16), 1113-1155.
- Leuz, C., Nanda, D., & Wysocki, D. (2003). Earnings management and investor protection: An international comparison. *Journal of Financial Economics*, 69 (3), 505-527.
- Rajgopal, S., & Venkatachalam, M. (2011). Financial reporting quality and idiosyncratic return volatility. *Journal of Accounting and Economics*, 51 (1-2), 1-20.
- Roulstone, D.T. (2003). Analyst following and market liquidity. *Contemporary Accounting Research*, 20 (3), 551-578.
- Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of Accounting and Economics*, 42 (3), 335-370.
- Stoll, H.R. (2000). Friction. *The Journal of Finance*, 55 (4), 1479-1514.



- Subramanyam, K. (1996). The pricing of discretionary accruals. *Journal of Accounting and Economics*, 22 (1-3), 249-281.
- Xu, R.Z., Taylor, G.K., & Dugan, M.T. (2007). Review of real earnings management literature. *Journal of Accounting Literature*, 26 (1), 195-228.
- Yan, Y., & Zhang, S. (2012). An improved estimation method and empirical properties of the probability of informed trading. *Journal of Banking and Finance*, 36 (2), 454-467.
- Zang, A.Y. (2012). Evidence on the trade-off between real activities manipulation and accrual-based earnings management. *The Accounting Review*, 87 (2), 675-703.
- Zhang, G. (2001). Private information production, public disclosure, and the cost of capital: theory and implications. *Contemporary Accounting Research*, 18 (2), 363-384.
- Zhao, Y., Chen, K.H., Zhang, Y., & Davis, M. (2012). Takeover protection and managerial myopia: Evidence from real earnings management. *Journal of Accounting Public Policy*, 31(1), 109-135.

## Appendix A: PIN Model and Estimation

The *PIN* model views trading as a game between liquidity providers and traders (position takers) that is repeated over trading days. Trades can come from informed or uninformed traders. For any given trading day the arrival of buy and sell orders from uninformed traders, who are not aware of the new information, is modeled as two independent Poisson processes with daily arrival rates  $\varepsilon_b$  and  $\varepsilon_s$ , respectively. The model assumes that information events occur between trading days with probability  $\alpha$ . Informed traders only trade on days with information events, buying if they have seen good news (with probability  $1-\delta$ ) and selling if they have seen bad news (with probability  $\delta$ ). The orders from the informed traders follow a Poisson process with a daily arrival rate  $\mu$ .

Under this model, the likelihood of observing  $B$  buys and  $S$  sells on a single trading day is

$$L((B,S) | \theta) = (1-\alpha) e^{-\varepsilon_b} \frac{(\varepsilon_b)^B}{B!} e^{-\varepsilon_s} \frac{(\varepsilon_s)^S}{S!} + \alpha \delta e^{-\varepsilon_b} \frac{(\varepsilon_b)^B}{B!} e^{-(\varepsilon_s+\mu)} \frac{(\varepsilon_s+\mu)^S}{S!} + \alpha(1-\delta) e^{-(\varepsilon_b+\mu)} \frac{(\varepsilon_b+\mu)^B}{B!} e^{-\varepsilon_s} \frac{(\varepsilon_s)^S}{S!} \quad (\text{A.1})$$

where  $B$  and  $S$  represent total buy trades and sell trades for the day respectively, and  $\theta = (\alpha, \delta, \mu, \varepsilon_b, \varepsilon_s)$  is the parameter vector. This likelihood function is a mixture of three Poisson probabilities, weighted by the probability of having a “good news day”  $\alpha(1-\delta)$ , a “bad news day”  $\alpha\delta$ , and “no-news day”  $(1-\alpha)$ . Assuming cross-trading day independence, the likelihood function across  $J$  days is just the product of the daily likelihood functions:

$$L(M | \theta) = \prod_{j=1}^J L(\theta | B_j, S_j) \quad (\text{A.2})$$

where  $B_j$  and  $S_j$  are the numbers of buy and sell trades for day  $j=1, \dots, J$ , and  $M = [(B_1, S_1), \dots, (B_J, S_J)]$  is the data set. Maximization of (2) over  $\theta$  given the data  $M$  yields maximum likelihood estimates for the underlying structural parameters of the model  $(\alpha, \delta, \mu, \varepsilon_b, \varepsilon_s)$ . Once the parameters of interest have been estimated, the Probability of Informed Trading, *PIN*, is calculated as

$$PIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_b + \varepsilon_s} \quad (\text{A.3})$$

where  $\alpha\mu + \varepsilon_b + \varepsilon_s$  is the arrival rate of all orders,  $\alpha\mu$  is the arrival rate of informed orders. The *PIN* is thus the ratio of orders from informed traders to the total number of orders.

An attractive feature of the *PIN* methodology is its apparently modest data requirement. All that is necessary to estimate the model is the number of buy- and sell-initiated trades for each stock and each trading day. However, one shortcoming of the methodology is that, although the estimation procedure is straightforward, it often encounters numerical problems when performing the estimation in practice, especially with stocks with a huge number of trades when the optimization program may clash with computational overflow or underflow (floating-point exception) and, as a consequence, it may not be

able to obtain an optimal solution. These difficulties in estimating *PIN* have been exacerbated in recent years due to the steady increase in the number of trades which are a consequence, among other reasons, of the growth in automated trading and structural changes in the market, which have greatly reduced market depth (Aslan, Easley, Hvidkjaer, & O'Hara, 2011). We estimate first the *PIN* model via maximum likelihood for each stock and month in each year. The use of one-month transaction data should be a wide enough period to produce reliable estimates and allows us to maximize the number of estimations (Easley, Kiefer, & O'Hara, 1997) indicate that a 30 trading-day window allows sufficient trade observations for the *PIN* estimation procedure, and Akay, Cyree, Griffiths, & Winters (2012) use 20 trading days to estimate *PIN* finding numerical solutions for all their estimations). Finally, we calculate an annual *PIN* by averaging monthly values. We use the optimization algorithm of the Matlab software. We run the maximum likelihood function 100 times for each stock in our sample, except for several large stocks, for which we increase the iterations to 1000 to ensure that a maximum is reached. We follow Yan & Zhang's (2012) proposal to set initial values for the five parameters in the likelihood function.

## **Appendix B: VPIN Estimation Procedure**

In this appendix we briefly review the three levels in which *VPIN* calculation takes place (for a more accurate description of the procedure, see the original paper of Easley et al., 2012).

### *(1) Time bars*

The original procedure begins with trade aggregation in *timebars*. *Bar size* is the first key variable of the *VPIN* computation process. Easley et al. (2012) initially use 1-minute time bars. In each time bar, trades are aggregated by adding the volume of all the trades in the bar (if any) and by computing the price change for this period of time. Afterwards, and in order to take into account trade size, the sample is "expanded" by repeating each bar price change a number of times equal to the number of shares traded in the bar. Thus, the original raw sample became a sample of one-unit trades, each of them associated with the price change of the corresponding bar.

### *(2) Volume buckets, bulk classification and order imbalance*

Volume bucket is the second essential variable in *VPIN* metric. Volume buckets represent pieces of homogeneous information content that are used to compute order imbalances. In Easley et al. (2012) volume bucket size (VBS) is calculated by dividing the average daily volume (in shares) by 50, which is the number of buckets they initially consider. Therefore, if we depart from the average daily volume, it is the number of buckets which fully determines VBS. Consequently, we consider the number of buckets as our second key variable.

Buckets are filled by adding the volume in consecutive time bars until completing the VBS. If the volume of the last time bar needed to complete a bucket is for a size greater than required, the excess size is given to the next bucket. In general, a volume bucket needs a certain number of time bars to be completed, although it is also possible that the volume in one time bar could be enough to fill one (or more) volume buckets.

At the same time of bucket completion, time bar volume is classified as buyer- or seller-initiated in probabilistic terms by employing the Normal distribution. Thus, we label as "buy" the volume that results from multiplying the volume bar by the value of the normal distribution evaluated in the standardized price change  $Z$  ( $\Delta P/\sigma_{\Delta P}$ ). To standardize, we divide the corresponding price change by the standard deviation of all price changes for the whole sample. Analogously, we categorize as "sell" the volume that results from multiplying the volume bar by the complementary of the normal distribution for the buy side,  $1-Z$  ( $\Delta P/\sigma_{\Delta P}$ ).

Order imbalance ( $OI$ ) is then computed for each bucket by simply obtaining the absolute value of the difference between buy volume and sell volume in the assigned time bars.

### (3) $VPIN$ and sample length

Finally, in the last step we obtain  $VPIN$  values. Here, it is necessary to define a new variable: *sample length* ( $n$ ). This variable establishes the number of the buckets with which  $VPIN$  is computed. Following the link established in Easley et al. (2012),

$$VPIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_b + \varepsilon_s} \approx \frac{E[V_\tau^{Sell} - V_\tau^{Buy}]}{E[V_\tau^{Sell} + V_\tau^{Buy}]} = \frac{\sum_{\tau=1}^n OI_\tau}{n * VBS} \quad (A.4)$$

where  $VPIN$  is simply the average of order imbalances in the sample length, that is, the result of dividing the sum of order imbalances for all the buckets in the sample length (proxy of the expected trade imbalance) by the product of volume bucket size ( $VBS$ ) multiplied by the sample length ( $n$ ) (proxy for the expected total number of trades).  $VPIN$  metric is updated after each volume bucket in a rolling-window process. For example, if the sample length is 50, when bucket #51 is filled, we drop bucket #1 and we calculate the new  $VPIN$  based on buckets #2 to #51. Easley et al. (2012) first consider sample length equal to the number of buckets (50), but throughout the paper the authors change this variable to 350 or 250, depending on what they want to analyze. A sample length of 50 buckets when the number of buckets is also 50 is equivalent to obtaining a daily  $VPIN$ . A sample length of 250 (350) when the number of buckets is 50 is equivalent to obtaining a five-day (seven-day)  $VPIN$ .

**Table 1.** Estimation of the normal levels of cash flow, production costs and discretionary expenses

	$CFO_t/A_{t-1}$	$PROD_t/A_{t-1}$	$DISEXP_t/A_{t-1}$
$I/A_{t-1}$	-2.2742** (-2.01)	1.9773 (1.12)	1.1752 (1.44)
$S_t/A_{t-1}$	0.1078*** (8.04)	0.7743*** (48.44)	
$S_{t-1}/A_{t-1}$			0.1426*** (15.77)
$\Delta S_t/A_{t-1}$	0.0941 (1.46)	0.2469*** (2.86)	
$\Delta S_{t-1}/A_{t-1}$		-0.0308 (-0.46)	
<i>Intercept</i>	0.0208 (1.58)	-0.0508*** (-4.34)	0.0444*** (11.09)
<i>Ad. R<sup>2</sup></i>	0.132	0.911	0.366

Notes: This table reports OLS coefficients of the regressions (4), (5) and (6). The regressions are estimated for industry-year groups with at least 15 observations.

$$CFO_t / A_{t-1} = \alpha_0 + \alpha_1(1 / A_{t-1}) + \beta_2(S_t / A_{t-1}) + \beta_3(\Delta S_t / A_{t-1}) + \varepsilon_t$$

$$PROD_t / A_{t-1} = \alpha_0 + \alpha_1(1 / A_{t-1}) + \beta_2(S_t / A_{t-1}) + \beta_3(\Delta S_t / A_{t-1}) + \beta_4(\Delta S_{t-1} / A_{t-1}) + \varepsilon_t$$

$$DISEXP_t / A_{t-1} = \alpha_0 + \alpha_1(1 / A_{t-1}) + \beta_2(S_{t-1} / A_{t-1}) + \varepsilon_t$$

$CFO_t$  is cash flow from operations estimated as operating income less total accruals;  $PROD_t$  is the production costs;  $DISEXP_t$  is the discretionary expenses.  $S$  and  $\Delta S$  represents sales and change in sales, respectively. All variables, including the intercept, are scaled by lagged total assets ( $A_{t-1}$ ). \*\*\*, \*\*, \* denote significance at the 1%, 5%, and 10% (two-tailed) level, respectively.

**Table 2.** Descriptive statistics

Panel A: Information asymmetry measures						
	#obs.	Mean	SD	10 <sup>th</sup> perc.	Median	90 <sup>th</sup> perc.
<i>RQS</i>	468	0.0064	0.0061	0.0013	0.0042	0.0143
<i>AMH</i>	468	0.3020	2.0676	0.0004	0.0090	0.3517
<i>PI</i>	468	0.0037	0.0027	0.0010	0.0030	0.0073
<i>PIN</i>	468	0.1890	0.0603	0.1224	0.1787	0.2648
<i>VPIN</i>	468	0.2048	0.1235	0.0728	0.1740	0.3650
<i>ASY</i>	468	0.0000	1.7928	-1.9042	-0.3964	2.4342
Panel B: REM measures						
<i>ACFO</i>	468	0.0022	0.1097	-0.1187	0.0003	0.1321
<i>APROD</i>	468	-0.0033	0.1013	-0.1260	0.0042	0.0982
<i>ADISEXP</i>	468	-0.0011	0.0725	-0.0888	0.0061	0.0669
<i>REM1</i>	468	-0.0045	0.1618	-0.2038	0.0191	0.1486
<i>REM2</i>	468	0.0011	0.1374	-0.1641	-0.0033	0.1469
<i>REM3</i>	468	-0.0022	0.2202	-0.2653	0.0058	0.2174
Panel C: Control variables						
<i>DiscAcc</i>	468	-0.0081	0.0993	-0.1233	-0.0044	0.1026
<i>Size</i>	468	14.2279	1.7352	12.0895	14.1441	16.6284
<i>ROA</i>	468	0.0703	0.0626	0.0160	0.0657	0.1323
<i>Turnover</i>	468	-6.2675	1.0471	-7.5481	-6.2789	-5.1232
<i>Volat</i>	468	1.8301	0.7641	1.0864	1.6512	2.7705
<i>Analysts</i>	468	1.9567	0.9570	0.0000	2.1972	3.0910
<i>Own</i>	468	0.4901	0.2348	0.1549	0.4900	0.7879

Notes: This table reports descriptive statistics of the variables employed in the present study. *RQS* is the relative quote bid-ask spread; *AMH* is the illiquidity measure of (2002). *PI* is the price impact measure proposed by Huang and Stoll (1996). *PIN* is Probability of Informed Trading based on the Easley et al. (1996) model. *VPIN* is Volume-Synchronized Probability of Informed Trading developed in Easley et al. (2012). *ASY* is the composite index of information asymmetry based on the before market microstructure measures: *RQS*, *AMH*, *PI*, *PIN*, and *VPIN*. *ACFO* is the abnormal level of cash flows according to model (4) multiplied by (-1); *APROD* is abnormal production costs according to model (5); *ADISEXP* is abnormal discretionary expenses according to model (6) multiplied by (-1); *REM1*, *REM2*, and *REM3* are aggregate measures of real earnings management defined as *APROD*+*ADISEXP*, *ACFO*+*ADISEXP*, and *ACFO*+*APROD*+*ADISEXP*, respectively. *DiscAcc* is the value of discretionary accruals estimated by the Jones (1991) model modified by Dechow et al. (1995). *Size* is the natural logarithm of total assets. *ROA* is operating income divided by total assets. *Turnover* is the natural logarithm of the average daily trading volume in euro scaled by market value of the firm's equity at the end of the year. *Volat* is the standard deviation of daily returns. *Analysts* is the natural logarithm of the total number of analysts following a firm. *Own* is the proportion of common shares held by the largest five shareholders.

**Table 3. Correlation matrices**

## Panel A: Measures of information asymmetry

	<i>RQS</i>	<i>AMH</i>	<i>PI</i>	<i>PIN</i>	<i>VPIN</i>	<i>ASY</i>
<i>RQS</i>	1					
<i>AMH</i>	0.306***	1				
<i>PI</i>	0.920***	0.218***	1			
<i>PIN</i>	0.599***	0.347***	0.526***	1		
<i>VPIN</i>	0.700***	0.310***	0.594***	0.766***	1	
<i>ASY</i>	0.916***	0.461***	0.855***	0.824***	0.868***	1

## Panel B: Explanatory variables

	<i>ACFO</i>	<i>APROD</i>	<i>ADISEXP</i>	<i>REM1</i>	<i>REM2</i>	<i>REM3</i>	<i>DiscAcc</i>	<i>Size</i>	<i>ROA</i>	<i>Turnover</i>	<i>Volat</i>	<i>Analysts</i>	<i>Own</i>
<i>ACFO</i>	1												
<i>APROD</i>	0.391***	1											
<i>ADISEXP</i>	0.100**	0.728***	1										
<i>REM1</i>	0.289***	0.952***	0.903***	1									
<i>REM2</i>	0.851***	0.696***	0.608***	0.707***	1								
<i>REM3</i>	0.711***	0.894***	0.714***	0.879***	0.944***	1							
<i>DiscAcc</i>	0.708***	0.008	-0.025	-0.006	0.552***	0.348***	1						
<i>Size</i>	0.021	0.168***	0.146***	0.170***	0.094**	0.136***	0.005	1					
<i>ROA</i>	-0.347***	-0.515***	-0.113***	-0.373***	-0.336***	-0.447***	0.119***	-0.009	1				
<i>Turnover</i>	0.132***	0.032	-0.047	-0.001	0.081*	0.065	0.067	0.178***	-0.119***	1			
<i>Volat</i>	0.047	-0.011	-0.108**	-0.055	-0.019	-0.017	-0.091**	-0.145***	-0.234***	0.416***	1		
<i>Analysts</i>	-0.069	-0.059	0.032	-0.023	-0.038	-0.051	0.003	0.635***	0.265***	0.275***	-0.147***	1	
<i>Own</i>	-0.054	-0.035	-0.020	-0.031	-0.054	-0.050	-0.078	0.125***	0.134***	-0.407***	0.008	0.060	1

Notes: This table reports the pairwise correlation coefficients between the measures used in the study. *RQS* is the relative quote bid-ask spread; *AMH* is the illiquidity measure of Amihud (2002). *PI* is the price impact measure proposed by Huang and Stoll (1996). *PIN* is Probability of Informed Trading based on the Easley et al. (1996) model. *VPIN* is Volume-Synchronized Probability of Informed Trading developed in Easley et al. (2012). *ASY* is the composite index of information asymmetry based on the before measures: *RQS*, *AMH*, *PI*, *PIN*, and *VPIN*. *ACFO* is the abnormal level of cash flows according to model (5) multiplied by (-1); *APROD* is abnormal production costs according to model (6); *ADISEXP* is abnormal discretionary expenses according to model (7) multiplied by (-1); *REM1*, *REM2*, and *REM3* are aggregate measures of real earnings management defined as *APROD*+*ADISEXP*, *ACFO*+*ADISEXP*, and *ACFO*+*APROD*+*ADISEXP*, respectively. *DiscAcc* is the value of discretionary accruals estimated by the Jones (1991) model modified by Dechow et al. (1995). *Size* is the natural logarithm of total assets. *ROA* is operating income divided by total assets. *Turnover* is the natural logarithm of the average daily trading volume in euro scaled by market value of the firm's equity at the end of the year. *Volat* is the standard deviation of daily returns. *Analysts* is the natural logarithm of the total number of analysts following a firm. *Own* is the proportion of common shares held by the largest five shareholders. \*\*\*, \*\*, \* denote significance at the 1%, 5%, and 10% (two-tailed) level, respectively.

**Table 4.** Information asymmetry and REM in the whole sample

	(1)	(2)	(3)	(4)	(5)	(6)
<i>ACFO</i>	0.046 (0.14)					
<i>APROD</i>		-0.076 (-0.29)				
<i>ADISEXP</i>			-0.389 (-0.94)			
<i>REM1</i>				-0.115 (-0.63)		
<i>REM2</i>					-0.172 (-0.76)	
<i>REM3</i>						-0.074 (-0.57)
<i>DisAcc</i>	0.063 (0.66)	0.069 (0.73)	0.057 (0.62)	0.068 (0.72)	0.078 (0.82)	0.132 (0.86)
<i>Size</i>	-0.202*** (-10.63)	-0.202*** (-10.68)	-0.202*** (-10.87)	-0.201*** (-10.69)	-0.203*** (-11.10)	-0.202*** (-10.90)
<i>ROA</i>	-1.404*** (-3.39)	-1.511*** (-3.43)	-1.510*** (-4.21)	-1.567*** (-3.80)	-1.614*** (-3.88)	-1.584*** (-3.60)
<i>Turnover</i>	-0.223*** (-9.09)	-0.223*** (-8.93)	-0.225*** (-9.25)	-0.224*** (-9.07)	-0.223*** (-8.95)	-0.223*** (-8.94)
<i>Volat</i>	0.121*** (4.17)	0.120*** (4.04)	0.115*** (3.94)	0.118*** (3.95)	0.118*** (3.97)	0.118*** (3.98)
<i>Analysts</i>	-0.102*** (-3.00)	-0.101*** (-3.07)	-0.102*** (-3.02)	-0.102*** (-3.05)	-0.101*** (-3.06)	-0.101*** (-3.07)
<i>Own</i>	0.109 (0.98)	0.111 (1.04)	0.111 (1.05)	0.112 (1.05)	0.118 (1.10)	0.114 (1.07)
<i>Intercept</i>	2.435*** (7.29)	2.444*** (7.19)	2.448*** (7.43)	2.446*** (7.28)	2.472*** (7.39)	2.457*** (7.28)
<i>Year</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Ind</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Adj. R<sup>2</sup></i>	0.861	0.861	0.863	0.862	0.862	0.862
<i>#obs.</i>	468	468	468	468	468	468

Notes: This table reports OLS coefficients of our information asymmetry index on real activities manipulation and control variables following the regression model:

$$ASY = \beta_0 + \beta_1 REM + \beta_2 DiscAcc + \beta_3 Size + \beta_4 ROA + \beta_5 Turnover + \beta_6 Volat + \beta_7 Analysts + \beta_8 Own + \sum_i \beta_i Year + \sum_j \beta_j Ind + \varepsilon$$

*ASY* is the log of 3 plus the composite index of information asymmetry based on the following market microstructure measures: *RQS*, *AMH*, *PI*, *PIN*, and *VPIN*. *REM* refers to each of our six proxies of real earnings management: *ACFO* is the abnormal level of cash flows according to model (4) multiplied by (-1); *APROD* is abnormal production costs according to model (5); *ADISEXP* is abnormal discretionary expenses according to model (6) multiplied by (-1); *REM1*, *REM2*, and *REM3* are aggregate measures of real earnings management defined as *APROD*+*ADISEXP*, *ACFO*+*ADISEXP*, and *ACFO*+*APROD*+*ADISEXP*, respectively. *DiscAcc* is the value of discretionary accruals estimated by the Jones (1991) model modified by Dechow et al. (1995). *Size* is the natural logarithm of total assets. *ROA* is operating income divided by total assets. *Turnover* is the natural logarithm of the average daily trading volume in euro scaled by market value of the firm's equity at the end of the year. *Volat* is the standard deviation of daily returns. *Analysts* is the natural logarithm of the total number of analysts following a firm. *Own* is the proportion of common shares held by the largest five shareholders. *Year* and *Ind* represent year and industry dummies, respectively. *ACFO* and *REM2* are orthogonalized respect to *DiscAcc* in models (1) and (5). Robust *t*-statistics clustered at the firm level in parentheses. \*\*\*, \*\*, \* denote significance at the 1%, 5%, and 10% (two-tailed) level, respectively.



**Table 5.** Comparison of suspect-firm years with the rest of the sample in relation to REM activities

	<i>ACFO</i>	<i>APROD</i>	<i>ADISEXP</i>	<i>REM1</i>	<i>REM2</i>	<i>REM3</i>
<i>Size</i>	-0.001 (-0.29)	0.005 (1.11)	0.001 (0.13)	0.005 (0.64)	-0.001 (-0.13)	0.004 (0.39)
<i>BTM</i>	0.001 (0.04)	-0.017 (-1.21)	-0.015 (-1.48)	-0.032 (-1.40)	-0.014 (-0.85)	-0.031 (-1.05)
<i>ROA</i>	-0.618*** (-8.24)	-0.859*** (-7.52)	-0.133 (-1.33)	-0.992*** (-5.11)	-0.751*** (-5.52)	-1.610*** (-6.99)
<i>Suspect</i>	-0.004 (-0.32)	0.037** (2.36)	0.038*** (2.85)	0.075*** (2.64)	0.034 (1.57)	0.070* (1.97)
<i>Intercept</i>	0.067 (0.89)	0.022 (0.31)	0.038 (0.57)	0.060 (0.46)	0.105 (1.02)	0.127 (0.77)
<i>Year</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Ind</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Adj. R<sup>2</sup></i>	0.135	0.402	0.214	0.314	0.190	0.314
<i>#obs.</i>	468	468	468	468	468	468

Notes: This table reports OLS coefficients of our REM measures and control variables following the regression model:

$$REM = \beta_0 + \beta_1 Size + \beta_2 BTM + \beta_3 ROA + \beta_4 Suspect + \sum_i \beta_i Year + \sum_j \beta_j Ind + \varepsilon$$

*REM* refers to each of our six proxies of real earnings management: *ACFO* is the abnormal level of cash flows according to model (4) multiplied by (-1); *APROD* is abnormal production costs according to model (5); *ADISEXP* is abnormal discretionary expenses according to model (6) multiplied by (-1); *REM1*, *REM2*, and *REM3* are aggregate measures of real earnings management defined as *APROD+ADISEXP*, *ACFO+ADISEXP*, and *ACFO+APROD+ADISEXP*, respectively. *Size* is the natural logarithm of total assets. *BTM* is the book to market ratio. *ROA* is operating income divided by total assets. *Suspect* is a dummy variable that takes the value of 1 if the change in net income divided by total assets is between 0 and 0.01, and 0 otherwise. *Year* and *Ind* represent year and industry dummies, respectively. Robust *t*-statistics clustered at the firm level in parentheses. \*\*\*, \*\*, \* denote significance at the 1%, 5%, and 10% (two-tailed) level, respectively.

**Table 6.** Information asymmetry and REM in suspect and non-suspect samples

Panel A: Suspect sample						
	(1)	(2)	(3)	(4)	(5)	(6)
<i>ACFO</i>	0.135 (0.48)					
<i>APROD</i>		0.611*** (3.41)				
<i>ADISEXP</i>			0.531** (2.23)			
<i>REM1</i>				0.322*** (2.94)		
<i>REM2</i>					0.239** (2.22)	
<i>REM3</i>						0.213*** (2.86)
<i>DisAcc</i>	-0.147 (-0.80)	-0.081 (-0.45)	-0.082 (-0.43)	-0.073 (-0.40)	-0.123 (-0.64)	-0.267 (-1.35)
<i>Size</i>	-0.226*** (-9.11)	-0.235*** (-9.67)	-0.229*** (-9.38)	-0.233*** (-9.57)	-0.228*** (-9.25)	-0.231*** (-9.49)
<i>ROA</i>	-1.049 (-0.99)	-0.516 (-0.51)	-0.992 (-0.96)	-0.720 (-0.70)	-0.897 (-0.85)	-0.703 (-0.68)
<i>Turnover</i>	-0.182*** (-5.12)	-0.175*** (-5.08)	-0.178*** (-5.00)	-0.176*** (-5.03)	-0.181*** (-5.11)	-0.179*** (-5.10)
<i>Volat</i>	0.106*** (3.10)	0.110*** (3.58)	0.120*** (3.53)	0.116*** (3.67)	0.111*** (3.40)	0.112*** (3.57)
<i>Analysts</i>	-0.093** (-2.32)	-0.094*** (-2.57)	-0.091** (-2.44)	-0.092** (-2.53)	-0.092** (-2.36)	-0.093** (-2.44)
<i>Own</i>	0.300** (2.01)	0.268* (1.88)	0.291** (2.00)	0.276* (1.92)	0.284* (1.95)	0.272* (1.89)
<i>Intercept</i>	2.926*** (7.29)	3.070*** (8.01)	2.942*** (7.49)	3.013*** (7.79)	2.946*** (7.51)	2.995*** (7.79)
<i>Year</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Ind</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Adj. R<sup>2</sup></i>	0.911	0.917	0.914	0.916	0.912	0.914
<i>#obs.</i>	148	148	148	148	148	148

**Table 6. Continued**

Panel B: Non-suspect sample						
	(1)	(2)	(3)	(4)	(5)	(6)
<i>ACFO</i>	-0.038 (-0.09)					
<i>APROD</i>		-0.550* (-1.69)				
<i>ADISEXP</i>			-0.982** (-2.20)			
<i>REM1</i>				-0.428** (-1.99)		
<i>REM2</i>					-0.552** (-2.12)	
<i>REM3</i>						-0.322** (-2.00)
<i>DisAcc</i>	0.133 (0.95)	0.191 (1.39)	0.165 (1.34)	0.193 (1.47)	0.215 (1.58)	0.464** (2.16)
<i>Size</i>	-0.198*** (-9.23)	-0.197*** (-9.78)	-0.199*** (-10.55)	-0.198*** (-10.12)	-0.204*** (-10.48)	-0.201*** (-10.21)
<i>ROA</i>	-1.468*** (-3.05)	-1.899*** (-3.91)	-1.526*** (-4.14)	-1.835*** (-4.24)	-1.946*** (-4.22)	-2.005*** (-4.10)
<i>Turnover</i>	-0.230*** (-8.42)	-0.227*** (-8.13)	-0.229*** (-8.49)	-0.227*** (-8.32)	-0.225*** (-8.39)	-0.225*** (-8.26)
<i>Volat</i>	0.126*** (3.17)	0.115*** (2.89)	0.114*** (2.97)	0.112*** (2.86)	0.114*** (2.82)	0.113*** (2.80)
<i>Analysts</i>	-0.106*** (-2.74)	-0.109*** (-2.88)	-0.109*** (-2.87)	-0.109*** (-2.87)	-0.104*** (-2.77)	-0.106*** (-2.84)
<i>Own</i>	0.049 (0.38)	0.047 (0.38)	0.043 (0.35)	0.045 (0.37)	0.061 (0.50)	0.055 (0.45)
<i>Intercept</i>	2.373*** (6.11)	2.469*** (6.29)	2.473*** (6.78)	2.493*** (6.59)	2.579*** (6.90)	2.553*** (6.63)
<i>Year</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Ind</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Adj. R<sup>2</sup></i>	0.845	0.849	0.855	0.852	0.850	0.851
<i>#obs.</i>	320	320	320	320	320	320

Notes: This table reports OLS coefficients of our information asymmetry index on real activities manipulation and control variables following the regression model for the suspect sample (Panel A) and non-suspect sample (Panel B):

$$ASY = \beta_0 + \beta_1 REM + \beta_2 DiscAcc + \beta_3 Size + \beta_4 ROA + \beta_5 Turnover + \beta_6 Volat + \beta_7 Analysts + \beta_8 Own + \sum_i \beta_i Year + \sum_j \beta_j Ind + \varepsilon$$

*ASY* is the log of 3 plus the composite index of information asymmetry based on the following market microstructure measures: *RQS*, *AMH*, *PI*, *PIN*, and *VPIN*. *REM* refers to each of our six proxies of real earnings management: *ACFO* is the abnormal level of cash flows according to model (4) multiplied by (-1); *APROD* is abnormal production costs according to model (5); *ADISEXP* is abnormal discretionary expenses according to model (6) multiplied by (-1); *REM1*, *REM2*, and *REM3* are aggregate measures of real earnings management defined as *APROD+ADISEXP*, *ACFO+ADISEXP*, and *ACFO+APROD+ADISEXP*, respectively. *DiscAcc* is the value of discretionary accruals estimated by the Jones (1991) model modified by Dechow et al. (1995). *Size* is the natural logarithm of total assets. *ROA* is operating income divided by total assets. *Turnover* is the natural logarithm of the average daily trading volume in euro scaled by market value of the firm's equity at the end of the year. *Volat* is the standard deviation of daily returns. *Analysts* is the natural logarithm of the total number of analysts following a firm. *Own* is the proportion of common shares held by the largest five shareholders. *Year* and *Ind* represent year and industry dummies, respectively. *ACFO* and *REM2* are orthogonalized respect to *DiscAcc* in models (1) and (5). Robust *t*-statistics clustered at the firm level in parentheses. \*\*\*, \*\*, \* denote significance at the 1%, 5%, and 10% (two-tailed) level, respectively.