



Accruals, Cash-Flows and Tobin's q : An Investment Perspective on Firm Accruals

Christian Calmès^a, Denis Cormier^b, François-Éric Racicot^c, Raymond Théoret^{d*}

^a Université du Québec (Outaouais), 101 St. Jean Bosco, Gatineau, Québec, Canada, J8X 3X7. Chaire d'information financière et organisationnelle, École des sciences de la gestion (UQAM). Laboratory for Research in Statistics and Probability (Carleton University).

^b Chaire d'information financière et organisationnelle, École des sciences de la gestion (UQAM). Université du Québec (Montréal), 315 Ste. Catherine est, Montréal, Québec, Canada, H2X 3X2

^c Université du Québec (Outaouais), 101 St. Jean Bosco, Gatineau, Québec, Canada, J8X 3X7. Chaire d'information financière et organisationnelle, École des sciences de la gestion (UQAM). Laboratory for Research in Statistics and Probability (Carleton University).

^d Université du Québec (Montréal), 315 Ste. Catherine est, Montréal, Québec, Canada, H2X 3X2. Professeur associé (UQO). Chaire d'information financière et organisationnelle, École des sciences de la gestion (UQAM).

* Corresponding author. Tel: +1 514 987 3000 4417.

E-mail addresses: christian.calmes@uqo.ca (Christian Calmès), denis.cormier@uqam.ca (Denis Cormier), francoi-eric.racicot@uqo.ca (François-Éric Racicot); raymond.theoret@uqam.ca (Raymond Théoret).

Accruals, Cash-Flows and Tobin's q : An Investment Perspective on Firm Accruals

Abstract

Following Zhang (*Accounting Review* 2007) we cast firm accruals in terms of short-term investment. Since many studies consider accruals as a smoothed measure of cash flows, we first adopt Zhang specification and augment the standard Jones model with a cash-flow variable. Second, if accruals are indeed a form of short-term investment they should also be influenced by firm's performance as measured by Tobin's q . Consequently we propose a new version of the accrual model including a proxy for Tobin's q . Given that accounting data and Tobin's q are generally measured with errors, we also introduce a new estimation method based on a modified version of the Hausman artificial regression, featuring an optimal weighting matrix composed of higher moments instrumental variable estimators. Our results suggest that all the key parameters of the accrual models are indeed systematically biased with measurement errors. More importantly, our findings largely qualify Zhang's conjecture on accruals, as both cash-flows and Tobin's q are found strongly significant regressors of firm accruals. Relatedly we find that the Tobin's q augmented model better isolate discretionary accruals so that the residuals of the equation are particularly well-suited to forecast stock returns.

JEL classification: M41; C12; D92.

Keywords: Discretionary accruals; Earnings management; Investment; Tobin's q ; cash-flows; Measurement errors; Instrumental variable estimators.

We would like to thank the Chaire d'information financière et organisationnelle (ESG-UQAM) for its financial support, and Stephen Young and Mark A. Trombley for their valuable comments.

Accruals, cash-flows et q de Tobin: Les accruals d'une firme dans le cadre de la théorie de l'investissement

Résumé

Dans la lignée des travaux de Zhang (2007), nous considérons les *accruals* d'une firme comme une forme d'investissement à court terme. Puisque plusieurs études voient les *accruals* comme une mesure lissée des cash-flows, nous adoptons d'abord la spécification de Zhang et ajoutons une variable de cash-flow au modèle classique de Jones. En second lieu, si les *accruals* sont effectivement une forme d'investissement à court terme, ils devraient être influencés par la performance de la firme, telle que mesurée par le q de Tobin. Par conséquent, nous proposons une nouvelle version du modèle d'*accruals* qui incorpore une *proxy* pour le q de Tobin. Étant donné que les données comptables et le q de Tobin sont généralement entachés d'erreurs de mesure, nous introduisons également une nouvelle méthode d'estimation basée sur une version modifiée de la régression artificielle d'Hausman qui se fonde sur une matrice de pondération faisant appel à des estimateurs de variables instrumentales construits à partir des moments supérieurs des variables du modèle. Nos résultats suggèrent que les paramètres-clefs des modèles d'*accruals* sont biaisés systématiquement du fait des erreurs de mesure. Plus spécifiquement, nos résultats soutiennent la conjoncture de Zhang puisque les cash-flows et le q de Tobin se révèlent tous deux des variables explicatives importantes des *accruals* d'une firme.

Classification JEL: M41; C12; D92.

Mots-clefs: Accruals discrétionnaires; Manipulation des profits; Investissement; q de Tobin; cash-flows; Erreurs de mesure; Estimateurs de variables instrumentales.

1. Introduction

A fundamental drawback of cash-flows is that they present timing and matching problems that cause them to be a very noisy measure of firm performance. To mitigate these issues, accountants rely on accounting accrual system to intertemporally smooth earnings. Accruals are thus used to separate the timing of cash flows from their accounting recognition. Based on the cash-flows statement, total accruals, the sum of discretionary and non-discretionary accruals, are defined as the difference between firm earnings and cash-flows (Jones 1991, Bartov et al. 2001), and, as such, are influenced by both firm performance and earning management. For this reason, models of non-discretionary accruals, i.e. accrual models, are widely used by financial analysts to assess the level of *discretionary* accruals, an important indicator of earnings management practices, and, more importantly, a significant predictor of stocks returns (Sloan 1996, Dechow and Dichev 2002, Fama and French 2007, Hirshleifer et al. 2009). Actually, one of the main reasons why accrual models have generated such attention in the finance literature is precisely the fact that the residuals of accrual equations carry valuable information about stock returns. This relates to the famous accruals anomaly (Sloan 1996, Dechow and Dichev 2002), which, contrary to the asset growth and profitability anomalies seems to be very robust (Fama and French 2007).

To better isolate the information on earnings management embedded in the accrual equation residuals and account for the nonlinear relationship between accruals and firm performance, researchers often include in their accrual models proxies for economic values and investment potential such as property plant and equipment (*PPE*) and sales. However, as Zhang (2007) notes “Surprisingly, little is known in the literature about the investment perspective of accruals despite the fact that, by definition, accruals measure investment in working capi-

tal” (p.1336). According to Zhang’s view, the standard specification of the accrual models misses some important aspects of accruals, namely the fact that accruals constitute a form of short-term investment, at least in terms of working capital. In particular, he notes that the accruals anomaly might actually relate to the investment information embedded in accruals. In the literature, the standard models used to estimate non-discretionary accruals generally rely on OLS estimation (e.g., Jones 1991, Bartov et al 2001, Xie 2001, Kothari et al. 2005, Wu et al. 2007, among many others). It is also common to use models taking into account various aspects of simultaneity biases and address the problem of measurement errors associated to accruals (e.g. Kang and Sivaramakrishnan 1995, Hansen 1999, Young 1999, Hribar and Collins 2002, Zhang 2007, Ibrahim 2009). To control for firm performance, many studies include a variable such as *ROA*, since accruals and performance are positively related (e.g., Kothari et al. 2005). In this paper, we analyze two commonly used discretionary accrual models, the Jones (1991) model, our benchmark, and the augmented Jones model including cash-flows (Dechow 1994, Dechow et Dichev 2002, Hirshleifer et al. 2004, Zhang 2007, Hirshleifer et al. 2009), and propose a new type of model where accruals are fully specified as short-term investment. In this respect, the first contribution of this paper is to propose a third version of the accrual model where we introduce Tobin’s *q* as a new proxy for firm performance. Tobin’s *q* offers the crucial advantage of being one of the most significant factors explaining firm investment decisions (see, among many others, Fazzari et al. 1988, Gilchrist and Himmelberg 1995, or Brown and Petersen 2009 and Brown et al. 2009 for recent references). In this sense, if accruals are indeed a form of short-term investment, as Zhang (2007) conjectures, they should be influenced by Tobin’s *q*.

However, the inclusion of Tobin's q adds to the endogeneity problem often encountered in accrual models, as this variable is notoriously plagued with measurement errors, a well documented fact in the investment literature (e.g., Hayashi 1982, Erickson and Withed 2000, 2002, among many others). When examining non-discretionary and discretionary accruals – the latter being the error term of the accrual models – we have to bear in mind the fact that this error term is the portion of accruals managed by firms, so that it may sometimes be influenced by various earnings management practices (Dechow et al. 1995, Beneish 1997, Burghstahler and Dichev 1997, DeFond and Park, 1997, Degeorge, Jeter and Shivakumar 1999, Peasnell et al. 2000, Xie 2001, Leuz, Nanda and Wysocky 2003, Marquardt and Wiedman 2004, Hirshleifer et al. 2004, Roychowdhury 2006). This may result in a statistical anomaly worth detecting as discretionary accruals are often used to forecast market returns. Sometimes, the accrual models account for heteroskedasticity with a form of weighted least-squares. However the measurement errors inherent to accounting data are usually assumed to be systematically biased in the same direction even though they can cause a serious bias in the estimation if the orthogonality between the explanatory variables and the equation innovation is not satisfied. Correcting for measurement errors is even more important in our case since we are interested by the significance of Tobin's q , a variable also contaminated by measurement errors. It is thus desirable to resort to a robust estimation method in order to compute the accruals with the greatest possible precision (Ibrahim 2009). The second contribution of this paper is precisely to propose a new method to handle this task. To this end we introduce new instruments based on a weighted optimal matrix of the higher moments and cumulants of the explanatory variables, and apply these optimal instruments to an Hausman artificial regression (named the Haus-C).

Our results suggest that measurement errors have indeed a great influence on the parameters estimation of the basic Jones model. More precisely, our estimation of the accrual models confirms that important measurements errors contaminate the accounting measures of changes in sales and fixed assets, the two explanatory variables of the Jones model. More importantly, our findings indicate that Tobin's q , which has already a very significant positive impact on non-discretionary accruals when using the standard OLS method, displays an increased explanatory power when applying our Haus-C procedure. Furthermore, the Haus-C procedure delivers a coefficient of the error adjustment regressor comparable, in level, to the Tobin's q coefficient itself. We can interpret this result as new evidence that firms expectations are partly incorporated in future cash-flows, a point often mentioned in the empirical literature on investment. Interestingly, if we introduce Tobin's q in the accruals equation, the cash-flow variable has a smaller influence on short-term investment. When measurement errors are properly accounted for, the role of Tobin's q is reinforced, while cash-flows seem to play a minor role, although non-trivial. The evidence we gather also tends to support the idea that market imperfections impend the Modigliani and Miller theorem to hold in the short-run. In the context of our model, these imperfections might be associated to liquidity constraints in earnings management, and the preference to directly self-finance accruals with cash-flows before resorting to external finance.

This paper is organized as follows. In section 2, we present the three accrual models we analyze with pooled data, and some considerations to address the problem of measurement errors in the models. In section 3, we detail the empirical results, and in section 4 we compare the residuals of these accrual equations to assess the performance of the Tobin's q augmented model. Section 5 concludes.

2. Empirical Framework

2.1 The Models

With a balance sheet approach, total accruals would be defined as:

$$TA \equiv (\Delta CA - \Delta CASH) - (\Delta CL - \Delta STD - \Delta TP) - DEP \quad (1)$$

where ΔCA stands for change in current assets; $\Delta CASH$, change in cash or cash equivalents; ΔCL , change in current liabilities; ΔSTD , change in debt included in current liabilities; ΔTP , change in income taxes payable; and DEP represents the depreciation and amortization expenses. Note however that the negative depreciation accruals tend to strongly influence the model's fit as the depreciation to assets ratio is five times higher than the accounts receivable and accounts payable, on average (Barth, Cram and Nelson 2001). One way of dealing with this issue is to look at short-term accruals and omit the long-run accruals of depreciation (Teoh, Welch and Wong 1998a, b). Relatedly, since this study focuses on the short-term investment dimension of firm accruals, i.e. the working capital component of accruals, we also consider an alternative construct, CA , which eliminates the long-term element of equation (1), i.e. total depreciation.

2.1.1 The Jones Model of Accruals

The Jones (1991) model of accruals, our benchmark (model I), can be written as:

$$\frac{TA_{it}}{A_{i,t-1}} = \alpha_s \left(\frac{1}{A_{i,t-1}} \right) + \beta_s \left(\frac{PPE_{it}}{A_{i,t-1}} \right) + \delta_s \left(\frac{\Delta REV_{it}}{A_{i,t-1}} \right) + \varepsilon_{it} \quad (2)$$

where TA is total accruals; A , total assets; PPE , gross property plant and equipment at the end of year t ; and ΔREV represents revenues in year t less revenues in year $(t-1)$. As usually done in the

literature, we scale all variables by $A_{i,t-1}$ to account for heteroskedasticity. This precaution also helps control for size effects. Equation (2) may be decomposed in two parts, the non-discretionary accruals component and the discretionary accruals one. The fitted value of the equation, $\frac{T\hat{A}_{it}}{A_{i,t-1}}$, represents the non-discretionary accruals, while the innovation, ε_{it} , is the discretionary part of accruals. Two control variables of the benchmark model relate, respectively, to the two main components of accruals, working capital, and depreciation. The first control variable, ΔREV_{it} , often replaced by the change in sales in the literature (e.g. Cormier et al. 2000), is associated to working capital, and the second control variable, PPE_{it} , is linked to depreciation. Usually, the ΔREV_{it} coefficient is found positively related to total accruals. Indeed, an increase in ΔREV_{it} should lead to an increase in working capital since accounts receivable are generally more sensitive to changes in sales than accounts payable¹. The coefficient of PPE_{it} should be negative as PPE_{it} determines the depreciation expenses, a negative component of accruals. Note however that there is potentially an endogeneity issue here because the control variable PPE_{it} might be collinear to accruals, the link between depreciation and PPE_{it} being quite strong.

In the short-term version of the benchmark model, we omit DEP on the LHS and PPE on the RHS of equation (2). In this case, equation (3) obtains:

$$\frac{CA_{it}}{A_{i,t-1}} = \alpha_s \left(\frac{1}{A_{i,t-1}} \right) + \delta_s \left(\frac{\Delta REV_{it}}{A_{i,t-1}} \right) + \varepsilon_{it} \quad (3)$$

Note that for the sake of consistency, the same procedure is followed for each model. First, we present the long-term version, with DEP and PPE included in the accrual models, second, we estimate the short-term version without PPE and DEP .

¹ In some cases, the sign of this coefficient may be negative. For more details, see McNichols and Wilson (1988).

2.1.2 An Augmented Version of the Jones Model with Cash-Flows

To account for firm performance it is common to introduce cash-flows (CF) in the accrual models (e.g., Dechow 1994, McNichols 2002, Francis et al. 2005 and Zhang 2007). In its long-term form, this standard accrual model (model II) can be written as:

$$\frac{TA_{it}}{A_{i,t-1}} = \alpha_s \left(\frac{1}{A_{i,t-1}} \right) + \beta_s \left(\frac{PPE_{it}}{A_{i,t-1}} \right) + \delta_s \left(\frac{\Delta REV_{it}}{A_{i,t-1}} \right) + \kappa_s \left(\frac{CF_{it}}{A_{i,t-1}} \right) + \varepsilon_{it} \quad (4)$$

The corresponding short-term version of equation (4) is:

$$\frac{CA_{it}}{A_{i,t-1}} = \alpha_s \left(\frac{1}{A_{i,t-1}} \right) + \delta_s \left(\frac{\Delta REV_{it}}{A_{i,t-1}} \right) + \kappa_s \left(\frac{CF_{it}}{A_{i,t-1}} \right) + \varepsilon_{it} \quad (5)$$

A standard approach often found in the accounting literature is to lag cash-flows to correct for the endogeneity of cash-flows. This procedure is generally adequate to mitigate the error term autocorrelation, but less appropriate to tackle the endogeneity per se (Theil 1953). Hence, in our models, we replace the cash-flows variable by its predicted value to ensure its orthogonality with the error term. An intuitive justification for doing so is that accruals are often value related, particularly for outperforming firms, that is firms characterized by high Tobin's q and strong persistence in sales. Indeed, for these firms, accruals are strongly (positively) autocorrelated and also quite correlated with both firm earnings and firm cash-flows. Consequently, even though accrual persistence can be partly attributable to a cosmetic smoothing, through allocating total assets accruals strategically among few accounting periods, or any earnings management practice, including hiding information on current sales innovation, accrual persistence is also explained by firm performance, and, consequently, expected cash-flows. As a matter of fact, note that in a sense, this properly directly relates to the investment perspective on accruals.

2.1.3. The Tobin's q Augmented Accrual model

To follow Zhang (2007) view about the investment perspective on firm accruals, we propose a new accrual model, model III, where we directly introduce Tobin's q as an explanatory variable in the equation of long-term accruals:

$$\frac{TA_{it}}{A_{i,t-1}} = \beta_1 \left(\frac{1}{A_{i,t-1}} \right) + \beta_2 \frac{PPE_{it}}{A_{i,t-1}} + \beta_3 \frac{\Delta Sales_{it}}{A_{i,t-1}} + \beta_4 \frac{q_{it}}{A_{i,t-1}} + \beta_5 \frac{CF_{it}}{A_{i,t-1}} + \xi_{it} \quad (6)$$

where Tobin's q , the shadow price of capital, is proxied by:

$$\frac{\text{Market value of capital} + \text{Accounting value of debts}}{\text{Accounting value of assets}} \quad (7)$$

Note that there are many empirical measures of Tobin's q in the investment literature. For example, one popular measure used in economic studies defines Tobin's q as the ratio of the market value of assets to their replacement cost. However, since we work in an accounting framework, we rely on an accounting definition of Tobin's q to be more consistent with the literature on accruals. In this literature, Tobin's q is often measured as the market-to-book ratio, i.e., the market value of equity scaled by book value of equity. Given the investment perspective of accruals we investigate, we depart slightly from this common practice to be more in the spirit of Tobin's definition and consider instead the market value of equity plus the book value of debt, divided by assets.

As for the previous models, we also analyse the short-term version of equation (6):

$$\frac{CA_{it}}{A_{i,t-1}} = \beta_1 \left(\frac{1}{A_{i,t-1}} \right) + \beta_3 \frac{\Delta Sales_{it}}{A_{i,t-1}} + \beta_4 \frac{q_{it}}{A_{i,t-1}} + \beta_5 \frac{CF_{it}}{A_{i,t-1}} + \xi_{it} \quad (8)$$

In the investment literature, Tobin's q is, by definition, one of the predominant variables explaining investment. Since we want to follow Zhang's line of reasoning, and to the extent that accruals decisions can be cast in terms of investment strategy, the analysis of Tobin's q explanatory power seems quite natural. The intuition here is that the return on investment should be a key explanatory variable of firm accruals, as economic theory suggests that investment relates to its return. Actually, accrual models often include a return measure such as return on assets, ROA , to control for the non-linear effect of firm performance (Dechow and Dichev 2002, Kothari et al. 2005). However, omitting Tobin's q can potentially lead to a strong colinearity issue as the return on investment and cash-flows are mixed together. Indeed, according to the investment theory, if financial markets are imperfect, the cash-flow variable becomes a potential significant regressor of investment, and financial constraints tend to influence both the capital structure and the performance of the firm. As a matter of fact, this kind of drawback also applies to Tobin's q itself. Theoretically, Tobin's q is a marginal concept, but its observed empirical counterpart is an average one, the marginal measure being unobservable. Therefore, when using the average measure, cash-flows might embed information about Tobin's q , and cash-flows and Tobin's q could be colinear. This matter is dealt with in the following section, where we present the modified Hausman artificial regression used for the estimation.

2.2. Estimation Method

Kang and Sivaramakrishnan (1995) and Kang (2005) argue that OLS accrual estimations can deliver misleading results and advocate instead the use of the IV approach and the GMM to deal with the errors-in-variables, omitted variables and simultaneity problems. In the same vein, since accruals and Tobin's q are generally measured with errors, we apply a tailor-

made specification error correction method² to the three accrual models. To detect specification errors in the accrual models we run two sets of regressions. For example, consider model III (i.e. the Tobin's q augmented model). Following Kothari et al (2005), we first run the two OLS regressions using equations (6) and (8), for long-term and short-term accruals respectively. Second, we run the following Haus-C artificial regressions:

$$\frac{TA_{it}}{A_{i,t-1}} = \beta_1^* \left(\frac{1}{A_{i,t-1}} \right) + \beta_2^* \frac{PPE_{it}}{A_{i,t-1}} + \beta_3^* \frac{\Delta Sales_{it}}{A_{i,t-1}} + \beta_4^* \frac{q_{it}}{A_{i,t-1}} + \beta_5^* \frac{CF_{it}}{A_{i,t-1}} + \sum_{i=1}^5 \varphi_i \hat{w}_{it} + \varepsilon_{it}^* \quad (9)$$

$$\frac{CA_{it}}{A_{i,t-1}} = \beta_1^* \left(\frac{1}{A_{i,t-1}} \right) + \beta_3^* \frac{\Delta Sales_{it}}{A_{i,t-1}} + \beta_4^* \frac{q_{it}}{A_{i,t-1}} + \beta_5^* \frac{CF_{it}}{A_{i,t-1}} + \sum_{i=1}^4 \varphi_i \hat{w}_{it} + \varepsilon_{it}^* \quad (10)$$

where, as explained in the appendix, the \hat{w}_{it} are the residuals obtained by running the regressions of the endogenous variables on the higher moments instrumental variables. These higher moments instruments are robust, and also have the advantage of requiring no extraneous information from the models. Equations (9) and (10) represent the generalized version of the augmented accrual model for long-term and short-term, respectively. Note that the estimated coefficients, φ_i , allow the detection of specification errors, and that their signs indicate whether the corresponding variable is overstated or understated in the OLS regression. The β^* estimated in these equations are equivalent to TSLS estimates, but our method presents the key advantage of providing additional information about the severity of the specification errors. Indeed, the φ_i measure the bias in the sensitivity of accruals to the i^{th} explanatory variable. If the φ_i associated to the i^{th} regressor is significantly positive, then the corresponding β will be lower in the artificial regression (and vice-versa if φ_i is negative). In general, we can expect a high positive correlation between $\hat{\beta}_i - \hat{\beta}_i^*$, the estimated error in the coefficient of variable i , and $\hat{\varphi}_i$, the esti-

² For details on this method, see the appendix.

mated coefficient of the corresponding artificial variable \hat{w}_i . We can sum up the former argument using the following equation:

$$\forall i \quad Spread_i = \pi_0 + \pi_1 \varphi_i + \zeta_i \quad (11)$$

where $Spread_i = \hat{\beta}_i - \hat{\beta}_i^*$. According to equation (11), the φ_i indicate the degree of overstatement or understatement of the OLS estimation, and the goodness of fit of the equation provides information about the severity of the specification errors. This constitutes a straightforward variant of the original Hausman test.

2.3. Data

For the distributional approach, a deviation of the distribution of earnings from the normal one should indicate earnings management. However, to explain the accrual conundrum (McNichols and Wilson 1988, Sloan 1996, Dechow and Dichev 2002) related to the stationarity of revenues and expenses (Yaari et al. 2007) it is often assumed that the ratio of accruals to earnings is actually a random variable. In other words, abnormal accruals might not only reflect earnings management of discretionary accruals but also changes in the underlying economic models and firm performance. Relatedly, if earnings management uses forward-looking information, it is beneficial because it increases the predictability of accruals, but, at the same time, it is precisely high-performance firms which present the most persistence in earnings management. High performance may thus erroneously lead the researcher to classify abnormal accruals as discretionary when the residuals of the accrual models still contain information on firm performance. Hence the challenge here is to arrive at a specification which can better disentangle earnings management from firm performance in the residuals. A main argument in this paper is

that Tobin's q is particularly suited to control for firm performance and isolate earnings management in the discretionary accruals used to forecast stock returns. To show this, we need to analyze data on high performance firms and compare our Tobin's q augmented model to the standard accrual models. Consequently, we use a sample composed of all the non financial firms registered in the *S&P500* index. The observations are retrieved from the U.S. **COMPUSTAT** database. Data are annual and run from December 1989 to December 2006. As previously done by most researchers (e.g. Kothari et al. 2005), we exclude firms displaying missing observations. After having discarded, among the 500 firms constituting the index, the firms with missing information, we have a total of around 10000 pooled observations.

Since the objective of this study is to shed light on the relationship between accruals and key investment factors, instead of analyzing industrial sectors individually, we focus our attention on representative firms. We thus adjust firm data for size, and run our regressions using pooling methods. Each year the sample does not vary much given that, in our data set, most firms are good performers and the sample is quite homogeneous, which, per se, mitigates the issue of composition effect. Consequently we can follow Ye (2006) who advocates pooling to improve the goodness of fit of the accrual models, and instead of slicing the sample by year and industry, we thus chose firm pooling, which also offers the advantage of a more parsimonious approach for testing the presence of measurement errors when estimating accrual models.

Insert table 1 here

Total accruals are computed using a balance sheet approach, i.e. change in non-cash current assets (Compustat #4 – #1) less change in current liabilities (Compustat #5), excluding the current portion of long-term debt (Compustat #44), less depreciation (Compustat #14) and taxes (Compustat #71). The annual cash-flows variable is operating cash-flows computed as

the mean value of the monthly data. $\Delta Sales$ is the difference of revenues in year t and revenues in year $(t-1)$ (Compustat #12). PPE , the acronym for property, plant and equipment, is measured at the end of year t (Compustat #7).

Table 1 provides the correlation between these variables. At first glance, this table shows that some explanatory variables might present a significant degree of correlation with each other, which might cause some multicollinearity issues in the regressions. For instance, the correlation between PPE and CF is about 0.70.

3. Results

3.1 OLS Estimations

In Table 2 we provide the OLS estimation results for the long-term accrual models, correcting for heteroskedasticity and treating size effects. Based on the adjusted R^2 (at 0.74, 0.77 and 0.80 for the three models, respectively) the equations seem to perform quite well. The best model in terms of adjusted R^2 is model III. In the equation, except for $\Delta Sales$, all the coefficients are significant at the 99% confidence level. The Durbin-Watson statistics are quite similar across models, ranging from 1.98 to 2.12 and, in light of the R^2 , there thus seems to be no apparent autocorrelation or non-stationary residuals problems.

Insert Table 2 about here

When comparing the coefficients of models II and III, first note the similarities in terms of values and signs of the coefficients. For instance, the coefficient of PPE is -0.0514 in model II, and -0.0554 in model III. We obtain the same results for the estimated coefficients of CF ,

respectively 0.3336 and 0.3086 (and $\Delta SALES$, 0.0005 versus 0.0003). At first, the positive sign of CF might be somewhat surprising. After all, total accruals (not necessarily short-term) are typically high when cash-flows are low, and vice-versa, and, as a result, accruals are negatively correlated with contemporaneous cash-flows. However, total accruals are also positively correlated with lagged and leaded cash-flows (Dechow and Dichev 2002). Indeed remind that accruals can be viewed as a smoothed measure of CF . Hence, even if the contemporaneous CF are negatively correlated with accruals, since we take the twelve month average of CF it is not so surprising to get an overall positive correlation. In other respects, note that for model I the estimated coefficient of PPE and $\Delta SALES$ are larger, at -0.1686 and 0.0026 respectively. Obviously, this finding is partly attributable to the omission of cash-flows. In fact, Table 1 indicates that the correlation between PPE and cash-flows is equal to 0.70, which suggests that a great proportion of the impact of PPE is transferred to cash-flows when shifting from model I to model II, the coefficient of cash-flows being equal to 0.3336 in model II. Overall these results suggest that accruals are indeed sensitive to cash-flows, a fact which could be consistent with the investment perspective on accruals. The intuition here would be that market imperfections and financial constraints influence earnings management and short-term investment. The introduction of Tobin's q also delivers results in the same vein. According to the investment theory, to the extent that accruals can be viewed as a form of short-term investment, they should be positively influenced by Tobin's q . Our results clearly support this view, as the Tobin's q coefficient is equal to 4.0897 and significant at the 99% confidence level.

Insert Table 3 about here

Table 3 reports the corresponding results for the short-term versions of the three models. In spite of the omission of the PPE variable, and the associated removal of the depreciation

component of accruals, the results remain very comparable to those of the long-term versions, both in terms of sign and magnitude of the coefficients. In particular they confirm that the variables traditionally used as regressors in investment equations are also significant explanatory variables of firm accruals.

However, there are some differences between the results obtained from the estimation of the short-term versus the long-term versions of our models. First, the impact of $\Delta SALES$ is more important in the short-term versions. For instance, in model II, the coefficient of $\Delta SALES$ is respectively 0.0005 (not significant) and 0.0086 (significant) in the long-term and short-term versions. Second, the influence of the Tobin's q coefficient is also larger in the short-term version (5.2361) compared to the long-term one (4.0897). The greater value of the Tobin's q coefficient is partly attributable to the cash-flows variable, whose coefficient decreases from 0.3086 to 0.2234 when shifting from the long-term to the short-term model.

Finally note that the DW statistics are rarely reported in the accruals studies. Yet, the residuals of the estimated accrual models – i.e., the discretionary accruals – should not be autocorrelated, because if they were, the returns forecast on which they are based could be biased. Looking at the data we find that accruals are indeed autoregressive. However, the influence of earnings management cannot last indefinitely and the residuals ought to converge to zero eventually. Dechow and Dichev (2002) regress working capital on lead and lag of cash-flows, which, as noted previously, might constitute an indirect way to control for accruals autoregressivity. In our case, we follow Beneish (1997) and Dechow et al. (2003) and add autoregressive terms in the regressions going backwards up to five periods to control for reversals. Using this method to control for the accruals autocorrelation improves the fit of the models, and the DW statistics suggests no remaining autocorrelation in the residuals.

3.2 Haus-C Estimations

Tables 4 and 5 present the results of the corresponding Haus-C estimations for the three models corrected for heteroskedasticity, for the long-run and short term accruals respectively. For the three models the levels of the Durbin-Watson do not seem to indicate any significant autocorrelation. Although most variables are significant at the 95% confidence level, Table 4 indicates that the Haus-C regressions systematically yield lower R^2 , 0.28, 0.48 and 0.39 respectively. This confirms that measurement errors in the explanatory variables indeed cause significant biases in the OLS regressions. More importantly, looking at the significance levels, model III seems clearly to outperform the other models. For instance, as reported in Table 4, the coefficient of $\Delta SALES$ is significant in model III and not significant in model II, and, once again, the estimated impact of PPE on accruals seems overstated in model II relative to model III. Actually, when shifting from model II to model III, the decrease in the PPE coefficient from -0.1605 to -0.0786 coincides with an increase of the cash-flows coefficient from 0.1435 to 0.2866. This suggests that the greater influence of PPE in model II is likely due to misspecification.

Insert Tables 4 and 5 about here

To analyze further the issue of measurement errors let focus on model III. As expected, the $\hat{\phi}_i$ indicate the presence of substantial measurement errors for all the explanatory variables. First, Table 4 reveals that the most commonly used explanatory variables of accruals, $\Delta SALES$, PPE and $1/A_{i,t-1}$, seem to be measured with significant error which translates into misspecification. One common explanation for this is that these accounting variables are used in accrual models as proxies for economic values. The error on $1/A_{i,t-1}$ is particularly severe, which could

explain the great instability of this coefficient and its changing sign when moving from one specification to another. Second, the coefficient of $\Delta SALES$ changes substantially from one model to another, and it also seems to be quite contaminated. More precisely, for model III, the $\hat{\phi}_i$ coefficient of $\Delta SALES$ is equal to -0.0612, significant at the 99% confidence level, whereas in the OLS estimation the coefficient of this variable is almost 0. The Haus-C result thus suggests a severe understatement of this coefficient in the OLS estimation. There is also a significant measurement error for the PPE variable, its $\hat{\phi}_i$ being equal to 0.0840 in model III, significant at the 99% confidence level. In this case, there is an overstatement of the coefficient in the OLS regression.

More importantly, the cash-flow coefficient doubles when Tobin's q is introduced. In model III, the cash-flow coefficient is equal to 0.2866, significant at the 99% confidence level, with a coefficient of understatement of -0.3182, significant at the 99% level, whereas in model II, from which Tobin's q is absent, the coefficient is lower, at 0.1435, with a coefficient of understatement of -0.1313. Since the correlation between cash-flows and Tobin's q is very low (cf. Table 1), this result cannot be attributed to collinearity. Actually, the introduction of Tobin's q generally increases the sensitivity of accruals to the other explanatory variables as well, and improves the fit of the model once errors-in-variables are accounted for. Furthermore, the Haus-C results confirm the expected positive relationship between accruals and Tobin's q , the influence of this regressor being significant at the 99% confidence level. Consistent with the conventional view that proxies of marginal Tobin's q are usually badly measured, the coefficient of Tobin's q estimated by OLS is about 4.0897 in model III, and much higher, at 6.3103, significant at 95%, when estimated with the Haus-C procedure. The error adjustment variable, \hat{w}_{5t} , at 9.6702, confirms the substantial measurement error of this variable.

In other respects, the short-term models perform poorly in the Haus-C estimations. For example, as reported in Table 5, the adjusted R^2 is almost halved when *PPE* is excluded. In the short-term version of the models, the levels of all the coefficients are also lower. However, the specification remains qualitatively robust and the explanatory variables are still significant and of the right sign. Besides note that, consistent with the OLS results, when excluding *PPE* from model III the influence of Tobin's q is strengthened, while the impact of cash-flows is divided by two, and its coefficient becomes only significant at the 90% confidence level. Not only is the influence of Tobin's q relative to cash-flow higher in the short-term model III, but it is even larger if we account for measurement errors. Going from OLS to Haus-C, the coefficient increases from 5.2361 to 8.6652. Hence, when correcting for measurement errors, the explanatory power of cash-flows seems to weaken in the short-term accrual models. This result could be paralleled to the one reported by Erickson and Whited (2000) for (long-term) investment. Quite counter-intuitively however, everything works as if the classical theory of investment applied more at a shorter horizon, firm performance influence on short-term investment and earnings management being reinforced whilst cash-flows influence would be dampened further. To understand this finding, we have to bear in mind the fact that in the short-term accrual models *PPE* is excluded. Since this variable, as often documented in previous studies, is highly correlated with cash-flows, it should not be too surprising to find cash-flows more significant in the long-term version of the models, including *PPE*, and this, regardless of the way errors-in-variables are treated. In other words, the lack of significance of cash-flows in the short-term accrual models is partly an artefact of the *PPE* cash-flows correlation.

4. Accruals Residuals Analysis

The problem with the residuals of the accrual models used to forecast stock returns is that discretionary accruals do not necessarily reflect earnings management only since accruals are also related to firm performance, so that differences in estimated discretionary accruals can be due to performance characteristics rather than incentives to manage earnings – particularly so if the relationship between accruals and performance is nonlinear. In this paper, we rely on OLS but also Haus-C to estimate Tobin’s q augmented model of accruals and argue that this model provides a better specification of firm accruals. If this is the case, our model should deliver residuals which better isolate the earnings management of discretionary accruals. To check this conjecture and gauge the respective merits of each model specification, it is thus very instructive to study the residuals of our regressions – i.e., the discretionary part of accruals. From an econometric perspective, the mean of the residuals of a regression ought to be equal to 0 in order to avoid any bias in the estimation. In this respect, TDA , total discretionary accruals, also ought to be 0 in the long-run since no earnings management practice can influence financial results indefinitely (Ronen and Yaari 2008). Figure 1 and Figure 2 provide the distributions of total discretionary accruals, TDA , and current discretionary accruals, CDA , respectively, both expressed in terms of total assets for models I and III. Once again, compared to model I, model III seems to better perform along this dimension. Indeed in Figure 1 note that the mean of TDA is equal to 0.055³ when estimated with model I, whereas it is practically 0 when estimated with model III. Therefore, having a mean of zero, the TDA associated to model III seem more appropriate to forecast returns. Relatedly, regardless of the model considered,

³ Remind that accruals are defined in terms of assets.

the *TDA* distribution seems positively skewed. For instance, for model III, the skewness coefficient is equal to 7.28, which suggests that discretionary accruals are likely influenced by various earnings management practices. As a matter of fact, it is remarkable to see that while the mean of model III residuals is lower, the skewness of the residuals is higher. One obvious explanation for this is that, by effectively controlling for firm performance, Tobin's q is indeed able to better isolate the earnings management information contained in discretionary accruals.

In other respects, as it is the case for *TDA*, the *CDA* mean goes down to 0 when the financial variables (cash flows and Tobin's q) are introduced (Figure 2). As shown in Figure 2, the *CDA* mean for model I, at 0.088, is higher than the corresponding mean of *TDA*, and the *CDA* skewness coefficient, at 10.47, is much higher than its *TDA* counterpart. This confirms that discretionary accruals are indeed manipulated especially at short horizons. A look at the skewness coefficient confirms this view as the larger coefficient observed for model III (10.47 versus 8.54 for model I) suggests that the distribution of the residuals can no longer be attributed to firm performance.

Overall, our results suggest that the residuals of model III are better suited for financial analysis as they are better purged from the normal link between stock returns and investment.

5. Conclusion

Some econometric challenges are related to the estimation of our augmented accrual model. A well-known issue relates to the difficulty in properly estimating Tobin's q given that it is not directly observable. As usually done in the investment literature, we rely on an average measure to proxy Tobin's q . There is thus an inherent measurement error related to the computation of this ratio. Ignoring the appropriate correction entails an empirical interaction between

cash-flows and Tobin's q which biases the estimated coefficients of both variables. We thus resort to a specific estimation procedure to tackle this measurement error. Based on this methodology we are also able to detect serious measurement errors in the basic accrual model and its augmented versions. The estimation of the augmented accrual model we propose reveals that the differences between the coefficients obtained from the IV method and those resulting from OLS are substantial, which suggests the presence of significant measurement errors in all variables.

More importantly, we find Tobin's q to be a significant explanatory variable of firm accruals. We can also confirm the influence of financial constraints on accruals management, but also that the impact of cash-flows is actually reduced when simultaneously introducing Tobin's q and accounting for measurement errors in the short-term version of the model.

Obviously, many questions remain open to investigation. Accrual models aim at analyzing the fundamental factors which normally influence CF smoothing in order to identify earnings management patterns with the residuals – i.e., with the discretionary accruals. With the introduction of Tobin's q we find more appropriate estimated residuals in the sense that they are closer to zero on average. This is due to the fact that our approach removes more effectively from accruals residuals the information related to firms performance. This might prove particularly useful for portfolio managers and financial analysts since discretionary accruals provide an important information to forecast stocks returns. However, we do not know *per se* whether the residuals of the Tobin's q augmented model are better suited to forecast stock returns. Given the nonlinear relationship between accruals and firm performance, autoregressive behavior of accruals and the skewness coefficients we still obtain, it would be interesting to study this question with a GARCH approach. This is left for future work.

References

- Bartov, E., Gul, F.A., Tsui, J.S.L. (2001). "Discretionary-accrual models and audit qualifications", *Journal of Accounting and Economics*, 30, 421-452.
- Beneish, M.D. (1997). "Detecting GAAP violation: Implications for assessing earnings management among firms with extreme financial performance", *Journal of Accounting and Public Policy*, 16, 271-309.
- Brown, J.R., Fazzari, S.M., Petersen, B.C. (2009). "Financing innovation and growth: cash-flow, external equity, and the 1990s R&D boom", *Journal of Finance*, 64, 151-185.
- Brown, J.R., Petersen, B.C. (2009). "Why has the investment-cash flow sensitivity declined so sharply? Rising R&D and equity market developments", *Journal of Banking & Finance*, 33, 971-984.
- Burgstahler, D., Dichev, I. (1997). "Earnings management to avoid earnings decreases and Losses", *Journal of Accounting & Economics*, 24, 99-126.
- Coën A, Racicot F.E. (2007). "Capital asset pricing models revisited: Evidence from errors-in-variables", *Economics Letters*, 95, 443-450.
- Cormier, D., Magnan, M., Morard, B. (2000). "The Contractual and Value Relevance of Reported Earnings in a Dividend-Focused Environment s", *European Accounting Review*, 9 (3), 387-417.
- Dagenais, M.G., Dagenais, D.L. (1997). "Higher moment estimators for linear regression models with errors in the variables", *Journal of Econometrics*, 76, 193-221.
- Dechow, P. (1994). "Accounting earnings and cash-flows as measures of firm performance", *Journal of Accounting and Economics*, 18, 3-42.
- Dechow, P.M., Dichev, I.D. (2002). "The quality of accruals and earnings: The role of accruals estimation errors", *The Accounting Review*, 77, 35-59.
- Dechow, P.M., Richardson, S.A., Tuna, I.A. (2003) "Why are earnings kinky? An examination of the earnings management explanation.", *Review of Accounting Studies*, 8, 355-384.
- Dechow, P., Sloan, R., Sweeny, A. (1995). "Detecting earnings management", *The Accounting Review*, 70, 193-225.
- DeFond, M., Park, C. (1997). "Smoothing income in anticipation of future earnings", *Journal of Accounting and Economics*, 23, 115-139.

- Degeorge, F., Patel, J., Zeckhauser, R. (1999). "Earnings Management to Exceed Thresholds", *Journal of Business*, 72, 1-33.
- Durbin, J. (1954). "Errors in variables", *International Statistical Review*, 22, 23-32.
- Erickson, T., Whited, T.M. (2002). "Two-step GMM estimation of the errors-in-variables model using high-order moments", *Econometric Theory*, 18, 776-799.
- Erickson, T., Whited, T.M. (2000). "Measurement errors and the relationship between investments and Q", *Journal of Political Economy*, 108, 1027-1057.
- Fama, E.F., French, K.R. (2007). "Dissecting anomalies", Working Paper, University of Chicago.
- Fazzari, S.M., Hubbard, R.G., Petersen, B.C. (1988). "Financing constraints and corporate investment", *Brookings Papers on Economic Activity*, 1, 141-195.
- Francis, J., LaFond, R., Olsson, P., Schipper, K. (2005). "The market pricing of accruals quality", *Journal of Accounting and Economics*, 39, 295-327.
- Fuller WA. (1987). *Measurement error models*. John Wiley & Sons: New York.
- Geary R.C. (1942). "Inherent relations between random variables." *Proceedings of the Royal Irish Academy*, 47, 23-76.
- Gilchrist, S., Himmelberg, C.P. (1995). "Evidence on the role of cash-flow for investment", *Journal of Monetary Economics*, 36, 541-572.
- Hansen, G. (1999). "Bias and measurement error in discretionary accrual models", Working paper, Pennsylvania State University.
- Hausman, J. A. (1978). "Specification tests in econometrics." *Econometrica*, 46(6), 1251-1271.
- Hayashi, F. (1982). "Tobin's marginal q and average q : a neoclassical interpretation", *Econometrica*, 50, 215-224.
- Hirshleifer, D., Hou, K., Teoh, S.H. (2009). "Accruals, cash flows, and aggregate stock returns", *Journal of Financial Economics*, 91, 389-406.
- Hirshleifer, D., Hou, K., Teoh, S.H., Zhang, Y. (2004). "Do investors overvalue firms with bloated balance sheets?", *Journal of Accounting and Economics*, 38, 297-331.
- Hribar, P., Collins, D. (2002). "Errors in estimating accruals: implications for empirical research", *Journal of Accounting Research*, 40, 105-134.

- Ibrahim, S. (2009). "The usefulness of measures of consistency of discretionary components of accruals in the detection of earnings management", *Journal of Business Finance & Accounting*, 36, 1087-1116.
- Jeter, D.C., Shivakumar, L. (1999). "Cross-sectional estimation of abnormal accruals using quarterly and annual data: Effectiveness in detecting event-specific earnings management", *Accounting & Business Research*, 29, 299-319.
- Jones, J.J. (1991). "Earnings management during import relief investigations", *Journal of Accounting Research*, 29, 193-228.
- Kang, S.H. (2005). "A conceptual and empirical evaluation of accrual prediction models", Working paper, SSRN147259.
- Kang, S.H., Sivaramakrishnan, K. (1995). "Issues in testing earnings management and an instrumental variable approach", *Journal of Accounting Research*, 33 (2), 353-367.
- Kendall, M. G, Stuart, A. (1963). *The Advanced Theory of Statistics*. Vol.1. London: Charles Griffin.
- Kothari, S.P., Leone, A.J, Wasley, C.E. (2005). "Performance matched discretionary accrual measures", *Journal of Accounting and Economics*, 39, 163-197.
- Leuz, C., Nanda, D., Wysocky, P. (2003). "Earnings management and investor protection: An international comparison", *Journal of Financial Economics*, 69, 506-527.
- Lewbel A. (1997). "Constructing instruments for regressions with measurement error when no additional data are available, with an application to patents and R&D." *Econometrica* , 65, 1201-1213.
- MacKinnon, J.G. (1992). "Model specification tests and artificial regressions", *Journal of Economic Literature* ", 30, 102-146.
- Malevergne, Y., Sornette, D. (2005). "Higher moments portfolio theory, capitalizing on behavioral anomalies of stock markets." *Journal of Portfolio Management*, 31(4), 49-55.
- Marquardt, C., Wiedman, C. (2004). "How are earnings managed? An examination of specific accruals", *Contemporary Accounting Research*, 21, 461-489.
- McNichols M.F.(2002). "Discussion of the quality of accruals and earnings; The role of accrual estimation error", *The Accounting Review*, 77, 61-69.
- McNichols, M., Wilson, G.P. (1988). "Evidence of earnings management from the provision for bad debts", *Journal of Accounting Research*, 26, 1-31.

- Pal, M. (1980). "Consistent moment estimators of regression coefficients in the presence of errors in variables", *Journal of Econometrics*, 14, 349-364.
- Peasnell, K., Pope, P., Young, S. (2000). "Detecting earnings management using cross-sectional abnormal accrual models", *Accounting and Business research*, 30, 313-326.
- Pindyck, R.S., Rubinfeld, D.L. (1998). *Econometric Models and Economic Forecasts*, 4th edition, McGraw-Hill.
- Racicot, F.E., Théoret, R. (2009). "Integrating volatility factors in the analysis of the hedge funds alpha puzzle", *Journal of Asset Management*, 10 (1), 37-62.
- Ronen, J., Yaari, V. (2008). "Earnings Management: Emerging Insights in Theory, Practice and Research". Springer.
- Roychowdhury, S. (2006). "Earnings management through real activities manipulation", *Journal of Accounting and Economics*, 42, 335-370.
- Sloan, R.G. (1996). "Do stock prices fully reflect information in accruals and cash flows about future earnings?", *The Accounting Review*, 71, 289-315.
- Spencer, D.E., Berk, K.N. (1981). "A limited information specification test." *Econometrica*, 49(4), 1079-1085.
- Teoh, S.H., Welch, I., Wong, T.J. (1998a). "Earnings management and the underperformance of seasoned equity offerings." *Journal of Financial Economics*, 50, 63-99.
- Teoh, S.H., Welch, I., Wong, T.J. (1998b). "Earnings management and the long-run market performance of initial public offerings". *Journal of Finance*, 53, 1935-1974.
- Theil, H. (1953). "Repeated least-squares applied to complete equations systems". The Hague: Central Planning Bureau (mimeo).
- Wu, D.M. (1973). "Alternative tests of independence between stochastic regressors and disturbances." *Econometrica*, 41(4), 733-750.
- Wu, J.G., Zhang, L., Zhang, X.F. (2007). "The accrual anomaly: Exploring the optimal investment hypothesis", NBER working paper series.
- Xie, H. (2001). "The mispricing of abnormal accruals", *The Accounting Review*, 76, 357-373.
- Yaari, H., DaDalt, P., Ronen, J., Yaari, V. (2007) "On accrual conundrum in earnings management research", mimeo.
- Ye, J. (2006). "Accounting accruals and tests of earnings management", Working paper, Baruch College.

Young, S. (1999). "Systematic measurement error in the estimation of discretionary accruals: An evaluation of alternative modelling procedures", *Journal of Business Finance & Accounting*, 26, 833-862.

Zhang, X.F. (2007). "Accruals, investment and the accrual anomaly", *The Accounting Review*, 82, 1333-1363.

Appendix

A.1 The Choice of Instruments⁴

Since accrual models usually present specification errors, the condition of orthogonality is generally violated and the estimators of the coefficients of the models are not unbiased and consistent. To reduce the estimation biases in these coefficients, we thus regress, in a first pass, the endogenous explanatory variables on instrumental ones. The delicate part is to judiciously choose the instruments. To deal with specification errors, Geary (1942), Durbin (1954), Kendall and Stewart (1963), Pal (1980), Fuller (1987), and more recently Dagenais and Dagenais (1997) and Lewbel (1997) have proposed instruments based on higher moments and cumulants. Racicot and Théoret (2009) generalize these instruments and apply them to financial models of returns, testing and correcting specification errors in a GMM framework. As a matter of fact, the literature on financial risk relies increasingly on the cumulants as more reliable measures of risk (Malevergne and Sornette 2005).

The set of new instruments we propose to build an estimator accounting for specification errors (and more specifically measurement errors) is based on an *optimal* combination of the estimators of Durbin (1954) and Pal (1980). Let us first assume the following general form $y = \alpha + \mathbf{X}\beta$, where y is the vector ($n \times 1$) representing the dependent variable, here accruals, and \mathbf{X} is the matrix ($n \times k$) of the explanatory variables⁵. β is the ($k \times 1$) vector of parameters to estimate. Assume further the existence of specification errors in the explanatory variables which might create inconsistency in the estimation of the vector β . To tackle this issue, Durbin (1954) proposes to use as instruments the following product: x^*x , where x is the \mathbf{X} matrix of

⁴ For this section, see: Racicot and Théoret 2009.

⁵ In model III of short-term accruals, the matrix \mathbf{X} is equal to $[1/A_i \Delta Sales_i/A_i CF_i q_i/A_i]$.

the explanatory variables expressed in deviation from the mean, and where the symbol * stands for the Hadamard element by element matrix multiplication operator. In the same vein, Pal (1980) introduces as instruments cumulants based on the third power of x instead of the squares, as Durbin. Combining these instruments, we obtain a new matrix of instruments \mathbf{Z} based on the cumulants and co-cumulants of \mathbf{x} and \mathbf{y} , these being the matrix \mathbf{X} and the vector \mathbf{y} expressed in deviation from the mean. This \mathbf{Z} matrix may be partitioned into k vectors or series, i.e. $\mathbf{Z} = [z_1 \ z_2 \ \dots \ z_k]$. The vector z_l , built with the first explanatory variable, is the instrument of the first explanatory variable, and so on. We regress the explanatory variables on this vector \mathbf{Z} to obtain $\hat{\mathbf{x}}$:

$$\hat{\mathbf{x}} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}'\mathbf{x} \quad (12)$$

Then the new optimal instruments $\hat{\mathbf{w}}^c$ based on cumulants of the explanatory variables are defined as:

$$\hat{\mathbf{w}}^c = \mathbf{x} - \hat{\mathbf{x}} = [\hat{w}_1^c \ \hat{w}_2^c \ \dots \ \hat{w}_k^c] \quad (13)$$

In their study, Racicot and Théoret (2009) find that these instruments are orthogonal to the estimated residuals. The correlation between \hat{w}_i^c and the corresponding explanatory variable x_i is around 90%, and the correlation is close to 0 with the other explanatory variables. In this sense, these instruments can be considered optimal. To improve the existing instrumental methods used to tackle the endogeneity issue in accrual models, we thus adopt the \hat{w}_i^c instruments they developed with a modified version of the Hausman (1978) artificial regression.

A.2. The Augmented Hausman Artificial Regression⁶

To detect specification errors in our sample of firms, we could use the original Hausman h test⁷ with the following classical linear regression model: $y = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where \mathbf{y} is a $(n \times 1)$ vector representing the dependent variable; \mathbf{X} , a $(n \times k)$ matrix of the explanatory variables; $\boldsymbol{\beta}$, a $(k \times 1)$ parameters vector, and $\boldsymbol{\varepsilon} \sim iid(0, \sigma^2)$. The Hausman test compares two estimates of the parameters vector, $\boldsymbol{\beta}_{OLS}$, the least-squares estimator (OLS), and $\boldsymbol{\beta}_A$, an alternative estimator taking a variety of specifications, the instrumental variables estimator $\boldsymbol{\beta}_{IV}$ in our case. The hypothesis H_0 is the absence of specification errors, and H_1 , their presence. First note that the vector of estimates $\boldsymbol{\beta}_{IV}$ is consistent under both H_0 and H_1 , whereas $\boldsymbol{\beta}_{OLS}$ is only consistent under H_0 and not consistent under H_1 . Consequently, under H_0 , $\boldsymbol{\beta}_{IV}$ is less efficient than $\boldsymbol{\beta}_{OLS}$. Second, the Hausman test aims at verifying if “the endogeneity” of some variables – in our case the variables measured with errors – has any significant effect on the estimation of the parameters vector. Therefore, the Hausman test is an *orthogonality* test, that is, helping verify if $plim(1/T) \mathbf{X}'\boldsymbol{\varepsilon} = 0$ in large samples. To implement the test, we define the following vector of contrasts or distances: $\boldsymbol{\beta}_{IV} - \boldsymbol{\beta}_{OLS}$. The resulting h test statistic reads:

$$h = (\hat{\boldsymbol{\beta}}_{IV} - \hat{\boldsymbol{\beta}}_{OLS})' \left[\mathbf{Var}(\hat{\boldsymbol{\beta}}_{IV}) - \mathbf{Var}(\hat{\boldsymbol{\beta}}_{OLS}) \right]^{-1} (\hat{\boldsymbol{\beta}}_{IV} - \hat{\boldsymbol{\beta}}_{OLS}) \sim \chi^2(g), \text{ with } \mathbf{Var}(\hat{\boldsymbol{\beta}}_{IV}) \text{ and } \mathbf{Var}(\hat{\boldsymbol{\beta}}_{OLS})$$

the respective estimates of the covariance matrices of $\hat{\boldsymbol{\beta}}_{IV}$ and $\hat{\boldsymbol{\beta}}_{OLS}$, and g the number of potentially endogenous regressors. H_0 is rejected if the p -value of this test is less than α , the critical threshold of the test (e.g. 5%). Third, and more importantly, note that, according to MacKinnon

⁶ For previous applications of this method see Coën and Racicot (2007) and Racicot and Théoret (2009).

⁷ For details on the Hausman test, see : Hausman (1978), Wu (1973), MacKinnon (1992) and Pindyck and Rubinfeld (1998). A very good presentation of the version of the Hausman test using an artificial regression in the context of correction of errors-in-variables may be found in Pindyck and Rubinfeld (1998). They present the case of one explanatory variable, whereas we apply it to the case of multiple explanatory variables.

(1992), the h test might also run into difficulties if the matrix $[\mathbf{Var}(\hat{\boldsymbol{\beta}}_{IV}) - \mathbf{Var}(\hat{\boldsymbol{\beta}}_{OLS})]$, which weights the vector of contrasts, is not positive definite. Since this is the case with most of the accrual models we study, we rely instead on an alternative method to run our Hausman test. For example, assume a five-variables linear regression model (e.g., the long-term version of the accrual model incorporating Tobin's q and cash-flows model III):

$$y_t = \beta_0 + \sum_{i=1}^5 \beta_i x_{it}^* + \varepsilon_t \quad (14)$$

with $\varepsilon \sim iid(0, \sigma^2)$.

and that the variables x_{it}^* ⁸ are measured with errors, that is:

$$x_{it} = x_{it}^* + \nu_{it} \quad (15)$$

with x_{it} the corresponding observed variables measured with errors. By substituting equation (15) in equation (14), we have:

$$y_t = \beta_0 + \sum_{i=1}^5 \beta_i x_{it} + \varepsilon_t^* \quad (16)$$

with $\varepsilon_t^* = \varepsilon_t - \sum_{i=1}^5 \beta_i \nu_{it}$. As explained before, estimating the coefficients of equation (16) by the

OLS method leads to biased and inconsistent coefficients because the explanatory variables are correlated with the innovation. Consistent estimators can be found if we can identify an instrument vector \mathbf{z}_t which is correlated with every explanatory variable but not with the innovation of equation (16). Then we regress the five explanatory variables on \mathbf{z}_t . We have:

$$x_{it} = \hat{x}_{it} + \hat{w}_{it} = \hat{\gamma}_i \mathbf{z}_t + \hat{w}_{it} \quad (17)$$

⁸ We use the asterisks to designate the unobserved variables.

where \hat{x}_{it} is the value of x_{it} estimated with the vector of instruments, and \hat{w}_{it} the residuals of the regression of x_{it} on \hat{x}_{it} . Substituting equation (17) into equation (16), the following artificial regression obtains:

$$y_t = \beta_0 + \sum_{i=1}^5 \beta_i \hat{x}_{it} + \sum_{i=1}^5 \beta_i \hat{w}_{it} + \varepsilon_t^* \quad (18)$$

The explanatory variables of this equation are, on the one hand, the estimated values of x_{it} , obtained by regressing the five variables on the vector of instruments \mathbf{z}_t , and on the other hand, the respective residuals of these regressions. Therefore equation (18) is an augmented version of equation (16).

We can show that:

$$p \lim \left[\frac{\sum \hat{w}_{it} \varepsilon_t^*}{N} \right] = p \lim \left[\frac{-\beta_1 \sum x_{it} v_{it}}{N} \right] = -\beta_1 \sigma_{v_1}^2 \quad (19)$$

If there is no specification error, $\sigma_{v_1}^2 = 0$, the OLS estimation results in a consistent estimator for β_1 , the parameter of \hat{w}_{it} in equation (18), and the coefficient is then equal to the one of the corresponding explanatory variable. In the case of specification errors, $\sigma_{v_1}^2 \neq 0$ and, therefore, the estimator is not consistent. For detecting the presence of specification errors, as we do not know a priori if there are such errors, we first have to replace the coefficients of the \hat{w}_{it} in equation (17) by θ_i . We thus have:

$$y_t = \beta_0 + \sum_{i=1}^5 \beta_i \hat{x}_{it} + \sum_{i=1}^5 \theta_i \hat{w}_{it} + \varepsilon_t^* \quad (20)$$

Since according to equation (17), $\hat{x}_{it} = x_{it} - \hat{w}_{it}$, we can then rewrite equation (20) as:

$$y_t = \beta_0 + \sum_{i=1}^5 \beta_i x_{it} + \sum_{i=1}^5 (\theta_i - \beta_i) \hat{w}_{it} + \varepsilon_t^* \quad (21)$$

If there is no specification error for x_{it} , $\theta_i = \beta_i$. In the opposite case, $\theta_i \neq \beta_i$, and the coefficients of the residuals terms \hat{w}_{it} are significantly different from 0. A significantly positive estimate of $(\theta_i - \beta_i)$ indicates that the estimated coefficient of the corresponding explanatory variable, x_{it} , is overstated in the OLS regression. In this case, the estimated coefficient for this variable in equation (21) is lower compared to the OLS one. On the other hand, if the estimated coefficient $(\theta_i - \beta_i)$ is significantly negative, it suggests that the estimated coefficient of the corresponding explanatory variable x_{it} is understated by OLS, and consequently, the estimated coefficient for this variable is higher in equation (21). In other respects, the estimated coefficients β_i are identical to those produced by a TSLS procedure with the same instruments (Spencer and Berk 1981), except that, compared to a strict TSLS, equation (21) also provides additional information which proves quite helpful when estimating accruals. In the procedure we propose to test for specification errors, we first regress the observed explanatory variables x_{it} on the instruments vector to obtain the residuals \hat{w}_{it} . Then, we regress y_t on the observed explanatory variables x_{it} and on these residuals \hat{w}_{it} . This is the auxiliary (or artificial) regression we just described. If the coefficient of the residuals associated to an explanatory variable is significantly different from 0, we can directly infer the presence of a specification error. In this case, a t test is used to assess the severity of the specification error. To our knowledge, such a test has never been used in this context. Usually, a Wald test (F test) is performed to check whether the whole set of $(\theta_i - \beta_i)$ coefficients is significantly different from zero, but this ignores the case of specification errors associated to a *specific* subset of explanatory variables. We can generalize the former procedure to the case of k explanatory variables with our modified Hausman regression. Let \mathbf{X} be a $(n \times k)$ matrix of explanatory variables not orthogonal

to the innovation, and let \mathbf{Z} be a $(n \times s)$ matrix of instruments ($s > k$). We regress \mathbf{X} on \mathbf{Z} to obtain $\hat{\mathbf{X}}$:

$$\hat{\mathbf{X}} = \mathbf{Z}\hat{\boldsymbol{\theta}} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X} = \mathbf{P}_z\mathbf{X} \quad (22)$$

where \mathbf{P}_z is the “predicted value maker”. Having run this regression, we can compute the matrix of residuals $\hat{\mathbf{w}}$:

$$\hat{\mathbf{w}} = \mathbf{X} - \hat{\mathbf{X}} = \mathbf{X} - \mathbf{P}_z\mathbf{X} = (\mathbf{I} - \mathbf{P}_z)\mathbf{X} \quad (23)$$

and perform the following artificial regression:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \hat{\mathbf{w}}\boldsymbol{\lambda} \quad (24)$$

A F test on the $\boldsymbol{\lambda}$ coefficients indicates whether the $\hat{\mathbf{w}}$ are significant as a group. But we also introduce a t test on each individual coefficient to check whether the corresponding $\boldsymbol{\beta}$ is understated or overstated. The vector of $\boldsymbol{\beta}$ estimated in equation (24) is identical to the TSLS estimates, that is:

$$\boldsymbol{\beta} = \boldsymbol{\beta}_{iv} = (\mathbf{X}'\mathbf{P}_z\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}_z\mathbf{y} \quad (25)$$

TABLES

Table 1 Variables Correlation Matrix

	<i>TA</i>	<i>CF</i>	$\Delta SALES$	<i>ROA</i>	<i>PPE</i>	<i>q</i>
<i>TA</i>	1.00	0.30	-0.04	0.02	-0.02	0.14
<i>CF</i>	0.30	1.00	0.04	-0.01	0.70	-0.01
$\Delta SALES$	-0.04	0.04	1.00	-0.01	0.07	-0.01
<i>ROA</i>	0.02	-0.01	-0.01	1.00	-0.02	0.04
<i>PPE</i>	-0.02	0.70	0.07	-0.02	1.00	-0.10
<i>q</i>	0.14	-0.01	-0.01	0.04	-0.10	1.00

Note. *TA* represents total assets; *CF*, cash-flows; $\Delta SALES$, the change in sales; *ROA*, the return on assets; *PPE*, property, plant and equipment and *q*, Tobin's *q*.

Table 2 OLS estimation of the three models (long-term versions)

	Model I	Model II	Model III
$I/A_{i,t-1}$	-0.0873	2.7080***	0.3892***
$PPE_{it}/A_{i,t-1}$	-0.1686***	-0.0514***	-0.0554***
$\Delta SALES_{it}/A_{i,t-1}$	0.0026	0.0005	0.0003
$CF_{it}/A_{i,t-1}$		0.3336***	0.3086***
$q_{it}/A_{i,t-1}$			4.0897***
<i>Adjusted R</i> ²	0.74	0.77	0.80
<i>DW</i>	2.01	2.12	1.98

Note. The long-term versions of models I, II and III are described respectively in equations 2, 4 and 6. The definition of the variables is provided in Table 1. The explanatory variables are scaled by lagged assets to account for heteroskedasticity. Asterisks indicate the significance levels: * stands for 10%, ** stands for 5% and *** stands for 1%.

Table 3 OLS estimation of the three models (short-term versions)

	Model I	Model II	Model III
$I/A_{i,t-1}$	-0.1868***	-0.2032***	-0.1338***
$\Delta SALES_{it}/A_{i,t-1}$	0.0040***	0.0086***	0.0047
$CF_{it}/A_{i,t-1}$		0.3595***	0.2234***
$q_{it}/A_{i,t-1}$			5.2361***
<i>Adjusted R</i> ²	0.71	0.73	0.74
<i>DW</i>	2.04	1.93	2.23

Note. The short-term versions of models I, II and III are associated respectively to equations 3, 5 and 8. The definition of the variables is provided in Table 1. The explanatory variables are scaled by lagged assets to account for heteroskedasticity. Asterisks indicate the significance levels: * stands for 10%, ** stands for 5% and *** stands for 1%.

Table 4 Haus-C estimations of the three models (long-term versions)

	Model I	Model II	Model III
$1/A_{i,t-1}$	1.3504***	0.2161***	0.6714***
$PPE_{it}/A_{i,t-1}$	-0.0142	-0.1605***	-0.0786***
$\Delta SALES_{it}/A_{i,t-1}$	0.0175***	0.0021	0.0165***
$CF_{it}/A_{i,t-1}$		0.1435***	0.2866***
$q_{it}/A_{i,t-1}$			6.3103***
\hat{w}_{1t}	7.9290***	-0.8610*	7.2810***
\hat{w}_{2t}	0.0150	0.1455***	0.0840***
\hat{w}_{3t}	-0.0778***	0.0104***	-0.0612***
\hat{w}_{4t}		-0.1313***	-0.3182***
\hat{w}_{5t}			9.6702***
Adjusted R²	0.28	0.48	0.39
DW	2.10	1.20	2.11

Note. The long-term versions of models I, II and III are described respectively in equations 2, 4 and 6. The definition of the variables is provided in Table 1. The explanatory variables are scaled by lagged assets to account for heteroskedasticity. Asterisks indicate the significance levels: * stands for 10%, ** stands for 5% and *** stands for 1%. The Haus-C procedure is explained in the appendix of this article. There is one artificial Hausman variable, designated by \hat{w}_i , for each explanatory variables of the models (e.g., $i = 1, \dots, 5$ for model III). For instance, \hat{w}_5 is the Hausman artificial variable associated to Tobin's q . It is the residuals of the OLS regression of Tobin's q on the chosen instruments provided in the appendix. The variable \hat{w}_5 gauges the measurement error of this variable. A positive sign indicates that the impact of the variable is overstated in the OLS regression, while a negative sign indicates the opposite.

Table 5 Haus-C estimations of the three models (short-term versions)

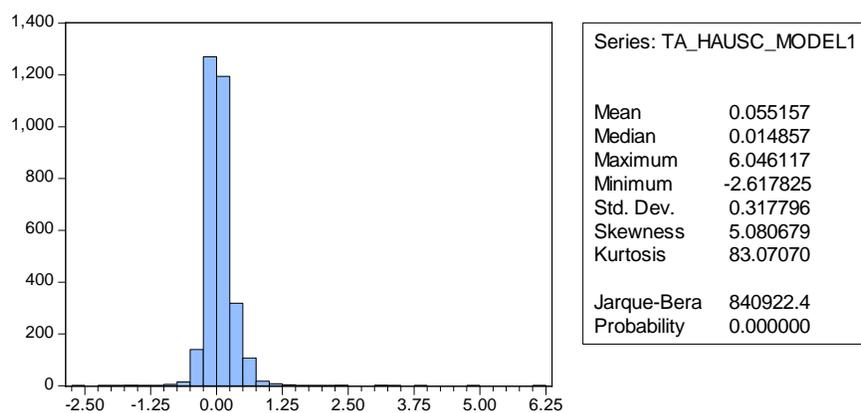
	Model I	Model II	Model III
$1/A_{i,t-1}$	1.6067***	1.4031***	0.4142**
$\Delta SALES_{it}/A_{i,t-1}$	0.0787***	0.0647***	0.0622**
$CF_{it}/A_{i,t-1}$		0.3835***	0.1290*
$q_{it}/A_{i,t-1}$			8.6652***
\hat{w}_{1t}	3.1363***	3.2952***	5.6220***
\hat{w}_{2t}	-0.0749***	-0.0589***	-0.0613***
\hat{w}_{3t}		-0.4926	-0.2492***
\hat{w}_{4t}			-11.1660***
Adjusted R²	0.13	0.18	0.21
DW	2.06	2.02	2.04

Note. The short-term versions of models I, II and III are given respectively by equations 3, 5 and 8. The definition of the variables is provided by Table 1. The explanatory variables are scaled by lagged assets to account for heteroskedasticity. Asterisks indicate the significance levels: * stands for 10%, ** stands for 5% and *** stands for 1%. The Haus-C procedure is explained in the appendix of this article. There is one artificial Hausman variable, designated by \hat{w}_i , for each explanatory variables of the models (e.g., $i = 1, \dots, 5$ for model III). For instance, \hat{w}_5 is the Hausman artificial variable associated to Tobin's q . It is the residuals of the OLS regression of Tobin's q on the chosen instruments provided in the appendix. The variable \hat{w}_5 gauges the measurement error of this variable. A positive sign indicates that the impact of the variable is overstated in the OLS regression, while a negative sign indicates the opposite.

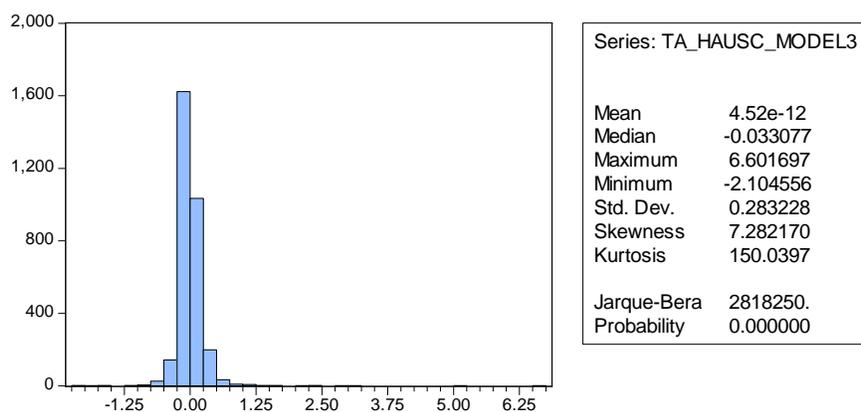
FIGURES

Figure 1

Total discretionary accruals estimated with the Haus-C method, model I



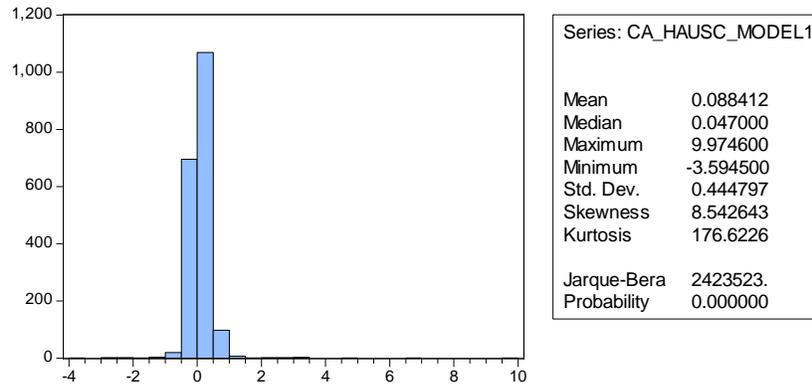
Total discretionary accruals estimated with the Haus-C method, model III



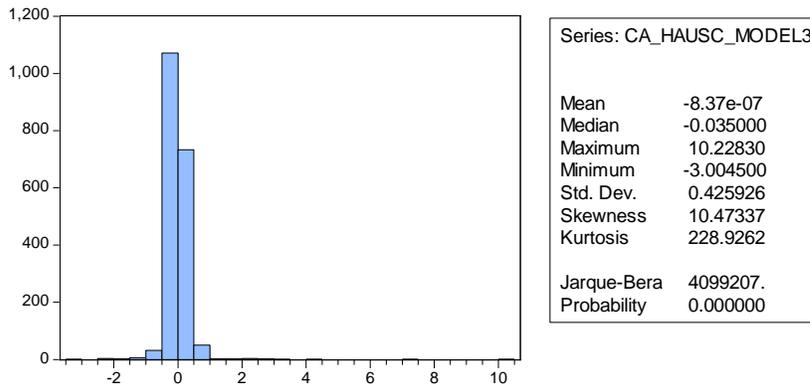
Note. These histograms are built using the residuals of the accrual (long-term) models I and III whose estimation appears in Table 4, discretionary accruals being the residuals of the estimated accrual models.

Figure 2

Current (short-term) discretionary accruals estimated with the Haus-C method, model I



Current (short-term) discretionary accruals estimated with the Haus-C method, model III



Note. These histograms are built using the residuals of the accrual (short-term) models I and III whose estimation appears in Table 4, discretionary accruals being the residuals of the estimated accrual models.