ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/74718/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Abdallah, Wissam , Goergen, Marc and O'Sullivan, Noel 2015. Endogeneity: how failure to correct for it can cause wrong inferences and some remedies. British Journal of Management 26 (4) , pp. 791-804. 10.1111/1467-8551.12113

Publishers page: http://dx.doi.org/10.1111/1467-8551.12113

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

ENDOGENEITY – HOW FAILURE TO CORRECT FOR IT CAN CAUSE WRONG INFERENCES AND SOME REMEDIES

Wissam Abdallah

Cardiff Business School
Cardiff University, Colum Drive, Cardiff, CF10 3EU, United Kingdom.
Email: abdallahw@cardiff.ac.uk

Marc Goergen*

Cardiff Business School and European Corporate Governance Institute (ECGI)
Cardiff University, Colum Drive, Cardiff, CF10 3EU, United Kingdom.
Email: goergenm@cardiff.ac.uk

Noel O'Sullivan

School of Business and Economics
Loughborough University, Ashby Road, Loughborough LE11 3TU, United Kingdom
Email: c.n.osullivan@lboro.ac.uk

Forthcoming in British Journal of Management

Abstract

Although researchers in business and management are becoming increasingly aware of the importance of endogeneity affecting regression analysis, they frequently do not have the right methodological toolkit to adjust for this issue. This paper discusses such a toolkit. There are also areas in business and management research which to date seem to be mostly oblivious about the endogeneity issue. This paper highlights such an area, which studies the question as to whether firms that are cross-listed on a foreign stock exchange are charged premium fees by their auditors. When the same methodology (pooled ordinary least squares) as in the existing literature is used, the existence of an audit fee premium for cross-listed firms seems to be confirmed. However, once methodologies are used which adjust for the various types of endogeneity (i.e. omitted variable bias, simultaneous and dynamic endogeneity) there is no longer support for the existence of such a generalised premium. Hence, this paper not only illustrates that failure to adjust for endogeneity has severe consequences such as drawing the wrong inferences, but it also reviews various ways to control for the different types of endogeneity.
Keywords: endogeneity; omitted variable bias; panel data

JEL codes: M00, C01, C18, C58

*Corresponding author

Introduction

Endogeneity is a major methodological concern for many areas of business and management research that rely on regression analysis to draw causal inference. Roberts and Whited (2013, p.493) broadly define endogeneity as "a correlation between the explanatory variables and the error term in a regression". Endogeneity may arise due to the omission of explanatory variables in the regression, which would result in the error term being correlated with the explanatory variables, thereby violating a basic assumption behind ordinary least squares (OLS) regression analysis. It may also be caused by the dependent variable being influenced by one or several explanatory variables, which in turn are influenced by the dependent variable. Such endogeneity may be of the simultaneous type whereby *contemporaneous* realisations of both the dependent variable and the explanatory variables in question affect each other. Or it may be of the dynamic type whereby *past* realisations of the dependent variable influence current realisations of one or more of the explanatory variables. For example, the current composition of a football team is likely to be – at least to some extent – the result of its past match performance. In particular, players who failed to play well in the past may no longer be part of the team and may have been replaced by new players who (are expected to) perform better. In other words, this is an example whereby past team or player performance (i.e. past realisations of the dependent variable) has an impact on current team composition (i.e. the current realisation of the explanatory variable(s)). Consequently, a study on the effect on performance of football team composition runs the risk of drawing the wrong inferences if it does not allow for dynamic effects.

Endogeneity arises in empirical research in a broad range of business and management fields, including strategy, international business research, initial public offerings and venture capital, corporate diversification, corporate governance and accounting research. For example, Hamilton and Nickerson (2003) highlight the need for researchers to carefully consider endogeneity when investigating the impact of strategic decisions on subsequent organisational performance. As an example of this, Shaver (1998) examines whether the entry mode (acquisition or greenfield) into a market impacts the survival likelihood of foreign direct investments. The endogeneity concern arises here as performance may itself influence the entry mode. Indeed, in his subsequent empirical analysis, Shaver (1989) illustrates that his findings are very different once endogeneity is appropriately

controlled for. Similarly, Reeb *et al.* (2012) highlight the importance of choosing a research approach that successfully tackles endogeneity concerns in the case of international business research. Consistent with this, in a study of the impact of the presence of foreign venture capitalists (VCs) on the likelihood of an international initial public offering (IPO) by entrepreneurial firms, Humphery-Jenner and Suchard (2013) highlight the need to control for endogeneity arising from both the likelihood of foreign VCs investing in the firm due to the expectation of a subsequent international listing and the likelihood of foreign VCs investing in the firm due to its success and, as a result of this, being more likely to pursue an international IPO. The importance of correcting for endogeneity is also highlighted by Villalonga (2004) who finds that, after using econometric methods that correct for sample selection, corporate diversification does not destroy value, a finding at variance with prior research that has not utilised such controls. Further, in the case of corporate governance research, issues of endogeneity are especially pertinent. For example, Wintoki *et al.* (2012) argue that it is paramount that studies examining the impact of the structure of the board of directors on firm performance adjust for dynamic effects. Indeed, the current board structure is almost certainly a result of past firm performance and in turn is likely to impact on future firm performance. Another related example concerns the effect of managerial stock ownership on firm value. Himmelberg *et al.* (1999) argue that studies investigating the link between the two must allow for the possibility that current managerial stock ownership may have been affected by past firm performance. Finally, accounting research is another area which suffers from the lack of adjusting for endogeneity issues. For example, Gow et al. (2010, p. 484) argue that "much of the empirical accounting literature uses panel data sets, typically repeated observations on the same, or a substantially overlapping, set of firms over time. In these data sets the variables of interest are often both cross-sectionally and serially correlated. As such, the common assumption of independence in regression errors is generally violated. While it is well known that residual dependence can result in misspecified test statistics (Bernard, 1987), prior methodological work in the accounting literature focuses exclusively on cross-sectional dependence and does not examine the issues created by the presence of both cross-sectional and time-series dependence."

While the above examples we use to illustrate the consequences from the failure to adjust for endogeneity relate to fairly narrow research questions, this does not imply that this type of failure is limited to a few isolated areas in business and management research. Indeed, a 2012 editorial in the Journal of International Business Studies (Reeb *et al.*, 2012) highlights the prevalence of endogeneity issues in international business (IB) research, but also points out that many researchers fail to address these issues in their research. The editors write that "in IB studies that do not take into account the [endogeneity] problem, we routinely observe that JIBS reviewers recommend rejection" (Reeb *et al.*, 2012, p.212). The fact that this editorial has been written as recently as 2012, by the editors of a top business and management journal, suggests that researchers are still in need of increased awareness of endogeneity and guidance on how to address this important methodological issue. Moreover, in relation to leadership studies Antonakis *et al.* (2014, p.93) stress that "… most researchers in leadership studies (and management in general) ignore one key design and estimation problem rendering parameter estimates uninterpretable: endogeneity". Furthermore, Antonakis *et al.* (2010) review the methodological rigour with which causal claims are being made in a sample of 110 articles on leadership published in top tier journals. They find that "… researchers fail to address at least 66% and up to 90% of design and estimation conditions that make causal claims invalid" (p.1086). In addition, Semadeni *et al.* (2013, p.1071) state that their review of articles in the Strategic Management Journal between 2005 and 2010 "reveals alarming inconsistencies regarding how strategy researchers approach and remedy endogeneity". Similarly, Wintoki *et al.* (2012,p.581) write that: "[e]mpirical corporate finance research, which attempts to explain the causes and effects of financial decisions, often has serious issues with endogeneity […] [T]he implications for the empirical work's usefulness if it does not properly deal with endogeneity can be substantial." This broad methodological issue gives rise to a need for a road map providing guidelines to researchers about how to correct for endogeneity once the issue has been identified. We provide such guidance throughout the paper, including a road map towards the end of the paper.

In what follows, we apply an estimation technique to a particular research question which adjusts for the three types of endogeneity (i.e. omitted variable bias, simultaneous endogeneity, and

dynamic endogeneity). It is important to note that simultaneous endogeneity can be addressed by estimating a system of equations, i.e. 2-stage (2SLS) or 3-stage least squares (3SLS) regression. However, such a system requires as many external exogenous instruments as there are equations. A practical difficulty that arises is that it may be very difficult to find such instruments. The estimation techniques we review and illustrate in this paper do not suffer from that difficulty. With the help of a concrete example, we show how failure to correct for endogeneity may have led to the wrong causal inferences. The example we use is that of UK firms that are cross-listed in the USA. A cross-listed firm is a firm, which in addition to its domestic stock exchange listing, is listed on a foreign exchange. The existing literature suggests that UK firms that are cross-listed in the USA are charged a premium on their audit fees, possibly reflecting the greater risk of litigation in the USA. However, when the econometric estimation procedure allows for endogeneity the results obtained from the existing literature are no longer upheld. More precisely, we find that, when allowing for endogeneity, our results fail to corroborate those from the existing literature as to the existence of a cross-listing premium.

Our study makes three important contributions to the broader business and management literature. First, it highlights the severity of the biases that may result when one fails to adjust for endogeneity. As our example suggests, failure to adjust for endogeneity can have major consequences and result in the researcher drawing the wrong inferences. Second, it applies a well-known estimator to a new empirical problem and shows that the results may differ depending on the assumptions made (the ones required to define the correct estimator). This estimator corrects for the three types of endogeneity, i.e. omitted variable bias and dynamic and simultaneous endogeneity. Contrary to e.g. 2SLS and 3SLS, the estimator we apply does not rely on additional, external sets of instruments that may be difficult to come by. Finally and as mentioned above, we also provide a road map for researchers on how to deal with these methodological issues.

Literature Review – Audit Pricing and Cross-listing

An important factor in audit pricing models is perceived client risk, with greater client risk implying higher audit fees (Cobbin, 2002; Hay *et al.*, 2006). The rationale for this is that riskier audits

increase the likelihood that the auditors will be subject to liability claims which may result in both financial and reputational costs to auditors. Auditors are expected to react to greater risk by increasing the effort they put into the audit and/or by adding an insurance premium to the audit fee (Simunic and Stein, 1996). However, as most audit pricing studies are based on a single country and the risk of litigation is therefore virtually the same across firms, assessing the impact of liability risk on audit fees is difficult. One way of overcoming this difficulty is to investigate companies that are subject to more than one legal jurisdiction. This is the case for companies that are cross-listed on a foreign stock exchange. In particular, as the USA is thought to have a more litigious environment than the UK, a cross-listing on a US stock exchange is likely to increase liability exposure for auditors (Seetharaman *et al.*, 2002). Hence, companies cross-listed in the USA are likely to be charged a premium on their audit fees.

The level of the audit fee is also likely to be related to the type of cross-listing. Unless it is Canadian or Israeli, a foreign company can only cross-list indirectly, i.e. via American Depositary Receipts (ADRs), on a US exchange. These ADRs are issued by an American bank and each represents a fraction or a multiple of one unit of the foreign stock (Ayyagari, 2004). A US cross-listing comes in three different types: a listing on a US exchange, i.e. NYSE, Nasdaq or Amex, a listing on the over-the-counter market (OTC) and a listing on PORTAL. An exchange listed firm may consist of a Level III ADR or a Level II ADR. The former requires full compliance with US GAAP and enables the firm to issue capital on a US exchange. The latter only requires partial compliance with US disclosure and accounting rules. Both Level II and Level III ADRs require the company to file with the Securities and Exchange Commission (SEC) the last three years of its financial accounts, audited according to US audit standards (Abdallah, 2008). If there is a cross-listing premium on audits, exchange-listed firms are likely to be those subject to such a premium. In contrast, a Level I ADR, which is listed and traded on OTC, does not require the company to comply with US GAAP and all of the SEC disclosure requirements. Finally, a Rule 144a private offering consists of a listing on PORTAL, an electronic market, which is not regulated by the SEC and only open to so called qualified institutional buyers (QIBs) for trading.

Although studies that investigate the evidence of a US cross-listing premium for audits are as yet sparse, they tend to agree that such a premium exists (Seetharaman *et al.*, 2002; Choi *et al.*, 2009).

More recently, research has been focusing on the impact of the Sarbanes-Oxley Act (SOX) on audit pricing in the USA. Several studies (see e.g., Ghosh and Pawlewicz, 2007; Griffin and Lont, 2007; Eldridge and Kealey, 2005; Charles *et al.*, 2010; Iliev, 2010) report significant increases in audit fees after Sarbanes-Oxley. This suggests that Sarbanes-Oxley may also have increased audit fees for companies with a US cross-listing.

Correcting for Endogeneity

The existing literature, such as Seetharaman *et al.* (2002), typically analyses a cross-section repeated over a number of years. However, instead of treating the data as a panel of data, the sample is pooled across the years and then used to estimate ordinary least squares (OLS) regressions. This pooling of the data is not as innocuous as it seems at first sight. Indeed, it is based on the assumptions that 1) there are no omitted variables and 2) there is no dynamic and simultaneous endogeneity. Both assumptions seem highly unrealistic. The first assumption is unrealistic as one of the key factors determining audit fees is the risk of the audit as perceived by the auditors. As this risk is perceived and cannot be directly measured, researchers have used a variety of proxies. Hence, it is possible that the proxies are of poor quality and/or exclude important aspects of risk. In other words, given the way audits are priced it is highly likely that any model of audit pricing suffers from omitted variables. The second assumption is also unrealistic for at least three reasons. First, it is likely that the decision to remain cross-listed depends on the current and past costs of the cross-listing, including the (potential) cross-listing premium for the audit. SOX may have increased this premium. Indeed, following Sarbanes-Oxley there were cases of UK firms that cancelled their US cross-listing given the increase in compliance costs.[1] Second, it could be the case that some firms may have historically been charged higher audit fees, maybe because of their greater complexity and/or lower bargaining power vis-à-vis the auditors. While audit pricing models include some measures of complexity and bargaining power, clearly such measures are proxies as complexity and bargaining power cannot be directly observed. Moreover, auditors are highly likely to use the past year's audit fees as a reference point when negotiating the audit fees for the coming year. This would suggest that there might be dynamic effects, calling for the inclusion of the lagged dependent variable on the right hand side. Third, to

confirm the assumption of stationarity, we perform a unit root test specifically designed for unbalanced panel data (Maddala and Wu, 1999). The null that there is a unit root is rejected at the 1% confidence level (p-value= 0.0001).[2] Hence, there is no evidence of a unit root or non-stationarity in audit fees. In other words, there is evidence that audit fees are stationary or persistent over time. This suggests that there may be dynamic effects. Given the latter, it makes sense to include the lagged level of the dependent variable on the right hand side of the regression:

$$y_{it} = ay_{i,t-1} + bx_{it} + f_i + v_{it} \text{ where } |a| < 1; i = 1,...,N; t = 2,3,...,T \qquad (1)$$

y_{it} is the audit fees paid by firm *i* for period *t*; $y_{i,t-1}$ is audit fees for the preceding year; x_{it} is a set of observable firm, audit and auditor characteristics as suggested by audit pricing models; f_i is an unobserved time-invariant characteristic of that firm and v_{it} is an error term,[3] which is not serially correlated. The number of firms *N* is assumed to be large and the number of periods *T* is fairly small. In the empirical analysis that follows, we check the statistical significance of the coefficient *a*. If the coefficient proves to be statistically significant, then dynamic effects will be important.

If f_i is correlated with x_{it}, then the OLS estimator will not be a consistent estimate of *b* (i.e. the bias is not reduced with a larger sample size). This is the OLS omitted variable inconsistency or OLS omitted variable bias (Wooldridge, 2002; Verbeek, 2012). In addition, the OLS estimator of a will be upward biased and inconsistent as $y_{i,t-1}$ is correlated with f_i (see e.g. Bond, 2002). This is the dynamic endogeneity issue. The Within Groups or fixed-effects estimation technique deals with the former issue by taking for each variable the deviation of each observation from the variable's time mean. As f_i is time-invariant, it is hence eliminated. Still, as the number of periods *T* is small, the Within Groups estimator of a is also biased, but in the opposite direction. This bias stems from the correlation between the transformed lagged dependent variable and the transformed error term. Similar to OLS, the Within Groups estimator is inconsistent as this bias does not disappear as *N* increases.

Blundell and Bond (1998) suggest the Generalized Method of Moments as a system (System GMM) which corrects for all three types of endogeneity. However in contrast to 2SLS and 3SLS, it

does not rely on external exogenous instruments, which in practice may be difficult to identify (Wintoki *et al.*, 2012). It consists of a system of two sets of equations, each with its own internal instruments. The first set of equations is in levels and is therefore still subject to omitted variable bias; just like an OLS regression. However, by assuming that these omitted variables are time invariant, which is likely a reasonable assumption for a short period of time (i.e. *T* is small), the lagged differences of the dependent variable and the independent variables can then be used as instruments for the first set. The second set consists of the equations in first differences with the lagged levels of the dependent variable and the independent variables as instruments. The validity of the instruments can be confirmed via a Sargan (1958) test and Hansen (1982) J test. The Sargan test checks for over-identifying restrictions and it is asymptotically distributed as a $\chi^2(n)$ with n degrees of freedom under the null hypothesis of valid instruments. It is not robust to heterogeneity and autocorrelation. While the Hansen test is robust to these, in turn it tends to perform poorly in the presence of a large number of instruments. Finally, we use the Arellano-Bond test for autocorrelation to test for first-order correlation (m_1) and second-order correlation (m_2) in the residuals. While first-order correlation is not an issue (given that the equations are in first differences, second-order serial correlation would be.

Data Sources, Sample Selection and Methodology

Our period of study is 1999 to 2008, which includes the passing of SOX in 2002. Our sample consists of the 137 cross-listed UK firms with complete data that we could identify via stock exchange websites and direct correspondence with the Bank of New York, the depository bank with the largest market share (i.e. 64%) in the ADR market.[4] Eighty-three of these firms are cross-listed on a US exchange (NYSE, Nasdaq, Amex), 53 on the OTC and one on PORTAL. We then build a control sample of non-cross-listed firms. These are matched with the cross-listed firms by industry (Datastream Industrial Classification) and with the closest size (as measured by total assets in 2000). We match the cross-listed firms not only by industry, but also by size as cross-listed firms tend to be the largest firms in their industry. We are able to perform this matching for 128 of the sample firms.

Table A in the Appendix reports the definitions and sources for the variables we use. We include the following variables, which prior research has identified as being influential in the

determination of audit fees: firm size as measured by the natural logarithm of total assets (Hay *et al.*, 2006), audit complexity and risk as measured by the total number of subsidiaries, the number of non-UK subsidiaries, the number of US subsidiaries (O'Sullivan, 2000; McMeeking *et al.*, 2006, 2007), account receivables and also inventories, both expressed as a proportion of total assets, as well as return on assets and return on equity to proxy firm performance (Hay *et al*., 2006). We also include debt as a proportion of total assets to measure leverage (Choi *et al.*, 2008; Zaman *et al.*, 2011). Finally, we also control for the characteristics of the auditors by including binary variables indicating whether the audit was undertaken by one of the large audit firms[5] (McMeeking *et al.*, 2006, 2007; Zaman *et al.*, 2011) and whether the auditors are based in London (O'Sullivan, 2000).

Data Analysis

Table 1 reports the descriptive statistics for the cross-listed and non-cross-listed firms. The table pools firm-year observations across the period of study of 1999 to 2008. While cross-listed firms pay higher audit fees, they are also larger and have greater leverage. In contrast, there are no differences in terms of the importance of inventories and receivables between cross-listed firms and non-cross-listed firms. Cross-listed firms also have much less volatile profitability as evidenced by the lower standard deviations for the return on equity (ROE) and return on assets (ROA). Although the means for ROE and ROA give a mixed message, the medians suggest that the cross-listed firms are more profitable. Cross-listed firms also tend to be more complex organisations as reflected by the greater number of subsidiaries. Finally, cross-listed firms are more likely to have an auditor from the Big Four or Big Five auditing firms.[6]

INSERT TABLE 1 ABOUT HERE

Table 2 shows the results from the pooled OLS regressions. The regressions are similar to those estimated by Seetharaman *et al.* (2002). Each regression in the table contains the same control variables, except for the performance measure (which is ROA in regressions (1), (3) and (5) and ROE in the remaining ones) and the cross-listing indicator variables. In terms of the latter, the first two

regressions include an indicator variable which equals one if the firm is cross-listed, and zero otherwise. This indicator variable tests for the presence of an audit premium affecting all cross-listed firms. The next two regressions include two indicator variables: the first one indicates whether the firm is exchange-listed whereas the second one indicates whether the firm is OTC-listed. These two indicator variables allow for a differential audit premium across the two types of host markets for the cross-listing. The final two regressions include an indicator variable accounting for a possible post-SOX effect for the exchange-listed firms and a post-SOX effect for the OTC-listed firms, respectively.

INSERT TABLE 2 ABOUT HERE

Except for ROE, the results for the control variables reflect those from other studies. There is strong evidence (at the 1% significance level) that UK firms with a US cross-listing pay premium audit fees (see regressions (1) and (2)) and that the premium is greater for exchange-listed firms compared to OTC-listed firms. Hence, our pooled OLS regressions seem to confirm the existence of a cross-listing premium and the premium is greater the greater the exposure of the cross-listed firm to US legislation and hence litigation. Nevertheless, these results need to be interpreted with caution as the pooled OLS regressions do not adjust for possible omitted variable bias and simultaneous or dynamic endogeneity.

Table 3 shows how the various estimation techniques – OLS in levels (regressions (1) to (4)), Within Groups (regressions (5) to (8)) and System GMM (regressions (9) to (12)) – perform in terms of addressing these two issues. All of the regressions in Table 3 are identical to those of Table 2, except for the fact that the data are no longer pooled and that the lagged dependent variable is now included on the right hand side. In other words, the models are dynamic panel regressions rather than static regressions. The first OLS in levels regression confirms the existence of a cross-listing premium. However, contrary to Table 2, this premium is only observed for exchange-listed firms and no longer for OTC-listed firms (see regression (2)). Also, this cross-listing premium can only be observed for exchange-listed firms after SOX was passed (see regressions (3) and (4)). This result is

already in stark contrast with the existing literature which finds evidence of a pre-SOX cross-listing premium.

When the Within Groups estimator is used, there is no longer evidence of a cross-listing premium whatever the type of cross-listing (see regressions (5) and (6)). However, similar to OLS in levels, there is evidence of a cross-listing premium for exchange-listed firms post-SOX. As expected, the value of the coefficient on the lagged dependent variable is much smaller (0.49-0.52) for the Within Groups regressions than for the OLS in levels regressions (0.74). The large difference between the two values reflects the likely upward bias of the OLS in levels estimator and the likely downward bias of the Within Groups estimator. This indicates that there is endogeneity. Further, the System GMM regressions confirm the results from the Within Groups regressions as to the absence of a generalised cross-listing premium (see regressions (9) and (10)). However, although significance levels have now dropped from 1% to 10%, there is still evidence of a post-SOX cross-listing premium for the exchange-listed firms (see regressions (11) and (12)). The significant coefficient on the lagged dependent variable in the GMM regressions confirms the presence of dynamic effects, and neglecting such effects would cause dynamic endogeneity.

INSERT TABLE 3 ABOUT HERE

This case study suggests that failure to adjust for omitted variable bias and simultaneous or dynamic endogeneity can lead to the wrong inferences. We have applied an estimation technique which adjusts for both. In what follows we provide a road map for the use of the econometric techniques that addresses endogeneity in panel data arising from omitted variable bias, simultaneous and dynamic endogeneity.

Guidance for empirical researchers

If the researcher suspects endogeneity, the first step would be to determine the source of it. If it is likely to arise because of omitted variables then the Within Groups estimator can mitigate the ensuing bias. However, the Within Groups estimator will only produce consistent parameters if the

independent variables are strongly exogenous, i.e. past realisations of the dependent variable are uncorrelated with current values of the independent variable(s) (Wintoki *et al.*, 2012). If the assumption of exogeneity of the independent variables is violated, then Within Groups is not an adequate estimation technique to address the endogeneity. In addition, Flannery and Hankins (2013) find that Within Groups performs poorly for short panels and is sensitive to high persistence in the data (i.e. coefficient *a* in equation (1) would be close to 1).

If simultaneity is the likely cause of endogeneity (i.e. the current observations of the dependent variable affect the current observations of one or more of the independent variables and vice versa), then the OLS and Within Groups estimators will be biased. A comparison of the OLS estimator of the coefficient on the lagged dependent variable with the equivalent Within Groups estimator will act as a diagnostic test of endogeneity. If both estimators are very different, then endogeneity is likely an issue. One solution to the problem is to use a system of equations (2SLS/3SLS). However, using this estimation technique requires the use of external exogenous variables which in many cases are difficult to find.

Therefore, given the strict exogeneity assumptions of Within Groups and the need for finding external instruments in an instrumental variable system, the researcher may have to rely on a dynamic system GMM regression which produces consistent and efficient parameters that are robust to the biases and inconsistencies caused by omitted variables, simultaneity and dynamic endogeneity. To test the assumption of exogeneity of the lagged independent variables the researcher may use the test of second-order serial correlation. Two other tests are the Sargan and the Hansen tests which both test for over-identification of the instruments. While for most settings, one chooses a significance level of 10%, for the Sargan and Hansen tests the cut-offs used under System GMM are typically 5%. In other words, with certain data sets if one were to choose a significance level of 10% it would be virtually impossible to find sets of valid instruments. A way forward would be to use e.g. instruments dated *t-4* instead of e.g. instruments dated *t-2* or *t-3*. However, this may not be ideal as such instruments may be too weakly correlated with the dependent and independent variables measured in *t*.

The researcher should also determine how many lags of the dependent variable should be used. Wintoki *et al.* (2012) argue that this is important to ensure dynamic completeness. If full

information from the past that affects the present is not incorporated then we still have endogeneity. Also, as the earlier lags are exogenous to current residuals, they can be used as instruments. One way to check how many lags of the dependent variable to include is to regress the dependent variable on several of its lags. The cut off for the lags is determined when a particular lag of the dependent variable becomes insignificant. For instance, if the researcher includes four lags of the dependent variable in the regression model, and after the estimation of the model the first two lags are significant whereas the last two are not, then the first two lags of the dependent variable are appropriate in this case.

By using System GMM one controls for omitted variable bias. System GMM should be used over a short period of time as the equations in levels still suffer from omitted variable bias and to deal with this one assumes that the correlation between the independent variables and the omitted variable(s) is constant over time which is a reasonable assumption over a short period of time (Wintoki *et al.*, 2012).

Conclusion

Although endogeneity is not a new issue in business and management research, a recent editorial in the Journal of International Business Studies (Reeb *et al.*, 2012) suggests that many researchers are either unaware of the issue or do not know how to address it. With the help of a concrete example, this paper has illustrated that failure to adjust for omitted variable bias and dynamic endogeneity may have severe consequences as it may lead to the wrong inferences being made. A caveat applies. We do not claim here that System GMM is the only valid methodology to adjust for endogeneity and, certainly, it only applies to the context of panel data. In fact, the results we obtained from the Within Groups estimations were fairly similar to those obtained from the System GMM estimations. Both sets of results rejected the presence of a generalised audit premium as suggested by the OLS regressions. What is important to realise is that, in the example we used, the omitted variable bias was clearly the most severe of all three types of endogeneity. As the Within Groups estimator eliminates the unobserved firm characteristic (which is assumed not to vary across time) by taking the deviation of each observation from its time mean, it does not suffer from omitted variable bias as the OLS estimator does. However,

this radical way of dealing with omitted variable bias in turn may cause problems. Indeed, the estimate obtained from Within Groups for the coefficient on the lagged dependent variable is much smaller (it ranges from 0.493 to 0.524) than that obtained from OLS (it ranges from 0.739 to 0.744) and that obtained from System GMM (which ranges from 0.659 to 0.694). In other words, the downward bias the Within Groups estimator suffers from is much more pronounced than the upward bias the OLS estimator suffers from, assuming the System GMM estimator is closest to the true value of the coefficient. If obtaining a precise estimate of the coefficient on the lagged dependent variable were important (as this would be the case for e.g. studies on the persistence of profitability) System GMM would nevertheless have been the preferred option.

In turn, System GMM relies on instruments that are in the form of lagged levels of the dependent and the independent variables (for the equations in first differences) and lagged differences (for the equations in levels). This way of dealing with omitted variable bias and simultaneous or dynamic endogeneity may be less radical than the way adopted by Within Groups. Nevertheless, it relies on a sufficiently long time series. Also, one of the advantages of System GMM, compared to Within Groups, is that it allows for the inclusion of time invariant binary variables whereas these are effectively differenced away under the latter. Another issue that the researcher intent on using System GMM faces is to determine the number of lags used in the instrument set. In practice, the Sargan and Hansen tests can be very helpful in terms of determining the set of instruments. However, the researcher may encounter cases where there is a need to balance the requirement for a set of valid instruments with a model that still has explanatory power.

In summary, the issues discussed in this paper should help inform future business and management research in two respects. First, by highlighting and illustrating the importance of endogeneity business and management researchers should be more aware of both the need to consider it in their research design, analysis and in the interpretation of their empirical findings. Second, the case study presented here provides some insights into how endogeneity may be controlled for in empirical analysis.

[1] For example, the following extract from the 2005 annual report of Provalis plc captures the likely impact of compliance costs on a UK company's cross-listing decision:
During the year [2005] it was decided to terminate the Group's ADR programme in the USA and to de-list from NASDAQ thus reducing compliance costs significantly, freeing management time and avoiding unnecessary and onerous commitments.
Source: 2005 annual report of Provalis plc, p.2.

[2] The null is also rejected when we allow for a trend or a drift and when the Phillips-Perron test rather than the Augmented Dickey Fuller test is used.

[3] We assume that the error terms are independently and identically distributed across firms.

[4] See http://www.adrbnymellon.com/dr_why.jsp?paramUserType=issuers, consulted on 4 April 2013

[5] Between 1999 and 2002 the large audit firms were Arthur Andersen, Deloitte & Touche, Ernst & Young, KPMG and PricewaterhouseCoopers. After the demise of Arthur Andersen in 2002 most of its UK audit business merged with Deloitte & Touche.

[6] For 2002-2008, Big Four is set to one if the auditor is one of the Big Four auditors and zero otherwise. Similarly, Big Five is set to one during the 1999-2001 for firm-year observations with a Big Five auditor, and zero otherwise. See also Table A.

References

Abdallah, A. (2008). 'Do regulations matter? The effects of cross-listing on analysts' coverage and forecast errors: a comparative analysis', Review of Accounting and Finance, 7, pp. 285-307.

Antonakis, J., Bendahan, S., Jacquart, P. and R. Lalive (2010). 'On making causal claims: A review and recommendations', Leadership Quarterly, 21, pp. 1086-1120.

Antonakis, J., Bendahan, S., Jacquart, P. and R. Lalive (2014). Causality and endogeneity: Problems and solutions. In D.V. Day (ed), The Oxford Handbook of Leadership and Organizations, pp. 93-117. Oxford: Oxford University Press.

Ayyagari, M. (2004). 'Does cross-listing lead to functional convergence? Empirical evidence', Policy Research Working Paper Series 3264, The World Bank.

Bernard, V.L. (1987). 'Cross-sectional dependence and problems in inference in market-based accounting research', Journal of Accounting Research, 25, pp. 1-48.

Blundell, R.W. and Bond, S.R. (1998). 'Initial conditions and moment restrictions in dynamic panel data models', Journal of Econometrics, 87, pp. 115-143.

Bond, S.R. (2002). 'Dynamic panel data models: a guide to micro data methods and practice', Portuguese Economic Journal, 1, pp. 141-162.

Charles, S.L., Glover, S.M. and N.Y. Sharp (2010). 'The association between financial reporting risk and audit fees before and after the historic events surrounding SOX', Auditing: A Journal of Practice and Theory, 29, pp. 15-39.

Choi, J-H, Kim, J-B., Liu, X. and D.A. Simunic (2008). 'Audit pricing, legal liability regimes, and big 4 premiums: theory and cross-country evidence', Contemporary Accounting Research, 25, pp. 55-99.

Choi, J-H, Kim, J-B., Liu, X. and D.A. Simunic (2009). 'Cross-listing audit fee premiums: Theory and evidence', The Accounting Review, 84, pp. 1429-1463.

Cobbin, P.E. (2002). 'International dimensions of the audit fee determinants literature', International Journal of Auditing, 6, pp. 53-77.

Eldridge, S.W. and B.T. Kealey (2005). 'SOX costs: auditor attestation under section 404', Available at SSRN: http://ssrn.com/abstract=743285

Ghosh, A. and R. Pawlewicz (2007). 'The impact of regulation on audit fees: evidence from the Sarbanes-Oxley Act', available at SSRN: http://ssrn.com/abstract=1032642.

Gow, I.D, Ormazabal, G., and D.J. Taylor (2010). 'Correcting for cross-sectional and time-series dependence in accounting research', Accounting Review, 85, pp. 483-512.

Griffin, P.A. and D.H. Lont (2007). An analysis of audit fees following the passage of Sarbanes-Oxley', Asian-Pacific Journal of Accounting and Economics, 14, pp. 161-192.

Hamilton, B.H. and J.A. Nickerson. (2003). 'Correcting for endogeneity in strategic management research', Strategic Organization, 1, pp. 51-78.

Hansen, L.P. (1982). 'Large sample properties of generalise method of moment estimators', Econometrica, 50, pp. 1029-1054.

Hay, D.C., Knechel, W.R. and N. Wong (2006). 'Audit fees: a meta-analysis of the effects of supply and demand attributes', Contemporary Accounting Research, 23, pp. 141-191.

Heckman, J. (1979). 'Sample selection bias as a specification error', Econometrica, 47, pp. 153-161.

Himmelberg, C.P., Hubbard, R.G. and D. Palia (1999). 'Understanding the determinants of managerial ownership and the link between ownership and performance', Journal of Financial Economics, 53, pp. 353-384.

Humphery-Jenner, M. and J.A. Suchard (2013). 'Foreign venture capitalists and the internationalization of entrepreneurial companies: evidence from China', Journal of International Business Studies, 44, pp. 607–621

Iliev, P. (2010). 'The effect of SOX section 404: costs, earnings quality and stock prices', Journal of Finance, 65(3), pp. 1163-1196.

Maddala, G.S., & Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. Oxford Bulleting of Economics and Statistics, (special issue), 631-652.

McMeeking, K., Peasnell, K.V. and P.F. Pope (2006). 'The determinants of the UK big firms premium', Accounting and Business Research, 36, pp. 207-231.

McMeeking, K., Peasnell, K.V. and P.F. Pope (2007). 'The effect of large firm mergers on audit pricing in the UK', Accounting and Business Research, 37, pp. 301-319.

O'Sullivan, N. (2000). 'The impact of board composition and ownership on audit quality: evidence from large UK companies', British Accounting Review, 32, pp. 397-414.

Reeb, D., and Sakakibara, M. and I. Mahmood. (2012). From the editors: endogeneity in international business research', Journal of International Business Studies, 43, pp.211–218.

Roberts, M.R. and T.M. Whited (2013), 'Endogeneity in empirical corporate finance'. In Handbook of the Economics of Finance, Volume 2, Part A, pp. 493-572.

Sargan, J.D. (1958). 'The estimation of economic relationships using instrumental variables', Econometrica, 26, pp. 329-338.

Seetharaman, A., Gul, F.A. and S.G. Lynn (2002). 'Litigation risk and audit fees: evidence from UK firms cross-listed on US markets', Journal of Accounting and Economics, 33, pp. 91-115.

Semadeni, M., Withers, M., and S. T. Certo (2014). 'The perils of endogeneity and instrumental variables in strategy research: understanding through simulations', Strategic Management Journal, 35, pp.1070-1079.

Shaver, J.M. (1998). 'Accounting for endogeneity when assessing strategy performance: does entry mode choice affect FDI survival?', Management Science, 44, pp. 571-585.

Simunic, D.A., and M.T. Stein (1996). 'The impact of litigation risk on audit pricing: a review of the economics and the evidence', Auditing: A Journal of Practice and Theory, 15, pp. 119-134.

Verbeek, M. (2012). *A guide to modern econometrics*. 4th edition. UK: John Wiley & Son Ltd.

Villalonga, B. (2004). 'Does diversification cause the "diversification discount"?', Financial Management, 33, pp. 5-27.

Wintoki, M.B., Linck, J.S., and J.M. Netter (2012). 'Endogeneity and the dynamics of internal corporate governance', Journal of Financial Economics, 105, pp. 581-606.

Zaman, M., Hudaib, M., and R. Haniffa (2011). 'Corporate governance quality, audit fees and non-audit services fees', Journal of Business Finance and Accounting, 38(1). pp. 165-197.

Table 1: Descriptive statistics for non-cross-listed firms and cross-listed firms

The data reported are for the 137 cross-listed firms in the US and the 128 industry-and-size matched non-cross-listed firms. The data are pooled across the years 1999 to 2008. Not all firms have data for all the years. N is the number of firm-year observations.

	Cross-listed firms			Non-cross-listed firms		
Variable	Mean	Median	Std Dev (N)	Mean	Median	Std Dev (N)
Audit fee (£'000)	2,593	1,100	4,588 (1,139)	648	300	1,432 (1,153)
Non-audit fee paid to auditor (£'000)	3,252	1,400	5,405 (1,138)	748	253	1,703 (1,150)
Non-audit fee paid to others (£'000)	916	0	11,321 (1,119)	66	0	564 (1,129)
Ln(Total assets) (£'000)	14.56	15.04	2.81 (1,158)	13.04	13.11	2.34 (1,218)
Total inventories / Total assets	0.07	0.04	0.09 (1,051)	0.07	0.03	0.09 (1,112)
Receivables / Total assets	0.14	0.11	0.11 (1,051)	0.16	0.13	0.14 (1,111)
Debt / Total assets	0.28	0.24	0.35 (1,157)	0.19	0.16	0.18 (1,218)
Return on assets (%)	-3.77	5.66	54.93 (1,150)	-2.58	4.10	69.76 (1,209)
Return on equity (%)	-7.90	11.64	227.96 (1,042)	-19.13	8.11	299.63 (1,156)
Number of subsidiaries	28.02	18	31.91 (1,136)	19.79	11	21.4 (1,153)
Number of non-UK subsidiaries	18.44	8	27.51 (1,136)	10.34	5	15.94 (1,153)
Number of US subsidiaries	3.5	2	4.83 (1,136)	2.29	0	4.6 (1,153)
Big Five	0.32	0	0.47 (1,152)	0.24	0	0.43 (1,149)
Big Four	0.61	1	0.49 (1,152)	0.57	1	0.5 (1,153)
London Based Auditor	0.69	1	0.46 (1,298)	0.51	1	0.5 (1,260)

Table 2: Pooled OLS regressions

The dependent variable in all models is the natural logarithm of audit fees. The models are estimated using OLS. All models include industry dummies. All models include time dummies, except for models (5) and (6). The values in parentheses are the p-values. The standard errors are asymptotically robust to heteroskedasticity and serial correlation.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Total assets)	0.508 (0.000)	0.499 (0.000)	0.506 (0.000)	0.496 (0.000)	0.505 (0.000)	0.495 (0.000)
Inventories/Total assets	1.141 (0.002)	1.335 (0.000)	1.125 (0.002)	1.320 (0.001)	1.077 (0.002)	1.231 (0.001)
Receivables/Total assets	1.318 (0.000)	1.145 (0.000)	1.321 (0.000)	1.148 (0.000)	1.314 (0.000)	1.131 (0.000)
Proportion of non-UK subs.	0.500 (0.000)	0.506 (0.000)	0.498 (0.000)	0.501 (0.000)	0.512 (0.000)	0.507 (0.000)
Proportion of US subs.	0.892 (0.000)	0.838 (0.000)	0.888 (0.000)	0.827 (0.000)	0.988 (0.000)	0.934 (0.000)
Debt/Total assets	0.189 (0.001)	0.204 (0.029)	0.188 (0.001)	0.199 (0.039)	0.228 (0.000)	0.236 (0.012)
ROA	-0.002 (0.000)	–	-0.002 (0.000)	–	-0.002 (0.000)	–
ROE	–	0.000 (0.009)	–	0.000 (0.009)	–	0.000 (0.002)
Big Four	0.163 (0.083)	0.168 (0.080)	0.164 (0.081)	0.170 (0.076)	0.196 (0.036)	0.203 (0.030)
Big Five	-0.03 (0.776)	-0.036 (0.725)	-0.03 (0.779)	-0.034 (0.742)	0.007 (0.950)	0.000 (0.999)
London	0.325 (0.000)	0.301 (0.000)	0.325 (0.000)	0.299 (0.000)	0.360 (0.000)	0.334 (0.000)
Ln(Non-audit fee to auditor)	0.131 (0.000)	0.132 (0.000)	0.13 (0.000)	0.130 (0.000)	0.134 (0.000)	0.136 (0.000)
Ln(Non-audit fee to other)	0.024 (0.129)	0.026 (0.126)	0.024 (0.138)	0.025 (0.137)	0.025 (0.134)	0.027 (0.130)
Cross-listed	0.215 (0.000)	0.207 (0.000)	–	–	–	–
Exchange-listed	–	–	0.247 (0.001)	0.258 (0.001)	–	–
OTC-listed	–	–	0.185 (0.007)	0.162 (0.021)	–	–
Post-SOX effect for exchange-listed firms	–	–	–	–	0.320 (0.000)	0.340 (0.000)
Post-SOX effect for OTC-listed firms	–	–	–	–	0.024 (0.728)	0.028 (0.693)
Constant	-2.383 (0.000)	-2.164 (0.000)	-2.375 (0.000)	-2.119 (0.000)	-2.383 (0.000)	-2.008 (0.000)
R-squared	0.910	0.909	0.910	0.909	0.910	0.909
No. of observations	1901	1768	1901	1768	1896	1763

Table 3: Dynamic panel data models with dummies for US cross-listing

The dependent variable in all models is the natural logarithm of audit fees. The regressions are run on the whole sample of firm-year observations. The values in parentheses are the p-values. m1 and m2 are tests for the absence of first-order and second-order serial correlation in the residuals, asymptotically distributed as N(0,1) under the null of no serial correlation. The Sargan statistic is a test of the over-identifying restrictions, asymptotically distributed as $\chi^2(k)$ with k degrees of freedom under the null of valid instruments. While the Sargan statistic is not robust to heteroskedasticity and autocorrelation, it is not weakened by many instruments. Hence, the Hansen test which is robust, but weakened by many instruments, is also reported. Regressions (9) to (12) use instruments in levels dated t-4 for the equations in first differences and first-differenced instruments dated t-3 for the equations in levels. These regressions are based on orthogonal deviations. All regressions contain industry dummies. Apart from regressions (3), (4), (7), (8), (11) and (12), all regressions include time dummies.

	OLS in levels				Within Groups				System GMM			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Ln(Audit fee$_{t-1}$)	0.744 (0.000)	0.744 (0.000)	0.739 (0.000)	0.741 (0.000)	0.493 (0.000)	0.493 (0.000)	0.522 (0.000)	0.524 (0.000)	0.694 (0.000)	0.690 (0.000)	0.659 (0.000)	0.659 (0.000)
Ln(Total assets)	0.142 (0.000)	0.140 (0.000)	0.142 (0.000)	0.142 (0.000)	0.261 (0.000)	0.262 (0.000)	0.270 (0.000)	0.269 (0.000)	0.175 (0.000)	0.173 (0.000)	0.191 (0.000)	0.191 (0.000)
Inventories/Total assets	0.158 (0.079)	0.141 (0.120)	0.152 (0.093)	0.156 (0.082)	0.333 (0.214)	0.330 (0.218)	0.296 (0.253)	0.300 (0.245)	0.374 (0.240)	0.364 (0.251)	0.155 (0.616)	0.155 (0.617)
Receivables/Total assets	0.113 (0.232)	0.120 (0.203)	0.133 (0.163)	0.125 (0.183)	0.163 (0.506)	0.160 (0.515)	0.193 (0.471)	0.197 (0.460)	-0.034 (0.894)	-0.016 (0.950)	0.043 (0.874)	0.043 (0.876)
Proportion of non-UK subs.	0.174 (0.000)	0.172 (0.000)	0.174 (0.000)	0.177 (0.000)	0.068 (0.337)	0.068 (0.343)	0.104 (0.146)	0.107 (0.134)	0.249 (0.007)	0.246 (0.007)	0.242 (0.014)	0.242 (0.011)
Proportion of US subs.	0.318 (0.000)	0.316 (0.000)	0.314 (0.000)	0.328 (0.000)	0.300 (0.018)	0.299 (0.018)	0.289 (0.013)	0.283 (0.017)	0.466 (0.014)	0.496 (0.009)	0.614 (0.001)	0.615 (0.001)
Debt/Total assets	-0.006 (0.797)	-0.006 (0.783)	-0.006 (0.805)	-0.001 (0.968)	0.032 (0.290)	0.031 (0.310)	0.017 (0.586)	0.022 (0.471)	0.061 (0.221)	0.065 (0.204)	0.070 (0.116)	0.070 (0.107)
ROA	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.488)	-0.001 (0.556)	-0.001 (0.142)	-0.001 (0.132)
BiG Four	0.012 (0.703)	0.011 (0.719)	0.006 (0.857)	0.008 (0.800)	-0.022 (0.845)	-0.019 (0.866)	0.011 (0.920)	-0.001 (0.995)	0.104 (0.295)	0.092 (0.359)	0.058 (0.519)	0.057 (0.514)
Big Five	-0.126 (0.004)	-0.127 (0.004)	-0.108 (0.002)	-0.101 (0.003)	-0.145 (0.251)	-0.141 (0.265)	-0.174 (0.103)	-0.186 (0.074)	-0.138 (0.163)	-0.148 (0.136)	-0.094 (0.309)	-0.095 (0.300)
London	0.072 (0.003)	0.072 (0.003)	0.070 (0.004)	0.073 (0.002)	0.037 (0.381)	0.037 (0.382)	0.036 (0.379)	0.036 (0.379)	0.049 (0.481)	0.049 (0.482)	0.065 (0.368)	0.065 (0.358)
Ln(Non-audit fee to auditor)	0.040 (0.000)	0.039 (0.000)	0.040 (0.000)	0.041 (0.000)	0.038 (0.000)	0.038 (0.000)	0.034 (0.000)	0.034 (0.000)	0.027 (0.078)	0.027 (0.081)	0.044 (0.008)	0.044 (0.006)
Ln(Non-audit fee to other)	0.005 (0.374)	0.004 (0.456)	0.004 (0.467)	0.004 (0.419)	0.000 (0.998)	0.000 (0.983)	-0.001 (0.838)	-0.001 (0.885)	0.003 (0.815)	0.004 (0.722)	-0.007 (0.549)	-0.007 (0.554)
Cross-listed	0.031 (0.058)	–	–	–	-0.027 (0.488)	–	–	–	0.079 (0.235)	–	–	–
Exchange-listed	–	0.060 (0.008)	0.040 (0.121)	–	–	-0.049 (0.473)	-0.100 (0.169)	–	–	0.124 (0.148)	-0.001 (0.994)	–
OTC-listed	–	0.008 (0.659)	0.018 (0.417)	–	–	-0.015 (0.739)	0.032 (0.486)	–	–	0.033 (0.711)	0.002 (0.978)	–
Post-SOX effect for exchange-listed firms	–	–	0.057 (0.057)	0.081 (0.003)	–	–	0.100 (0.000)	0.089 (0.001)	–	–	0.102 (0.099)	0.102 (0.068)
Post-SOX effect for OTC-listed firms	–	–	-0.013 (0.679)	-0.005 (0.846)	–	–	-0.012 (0.754)	-0.004 (0.914)	–	–	0.021 (0.804)	0.024 (0.669)

Table 3 Cont'd

	OLS in levels				Within Groups				System GMM			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	-0.632 (0.000)	-0.613 (0.000)	-0.648 (0.000)	-0.650 (0.000)	-0.751 (0.053)	-0.757 (0.051)	-0.975 (0.010)	-0.973 (0.010)	-1.250 (0.000)	0.434 (0.028)	-1.365 (0.000)	-1.363 (0.000)
m1	0.57 (0.568)	0.52 (0.602)	0.63 (0.530)	0.62 (0.534)	0.47 (0.000)	0.43 (0.000)	0.44 (0.000)	0.43 (0.000)	-7.47 (0.000)	-7.46 (0.000)	-7.33 (0.000)	-7.32 (0.000)
m2	0.84 (0.400)	0.79 (0.431)	0.68 (0.497)	0.73 (0.464)	0.45 (0.000)	0.40 (0.000)	0.41 (0.000)	0.40 (0.000)	-0.63 (0.527)	-0.57 (0.568)	-0.59 (0.556)	-0.58 (0.562)
Sargan test (p-value)	–	–	–	–	–	–	–	–	353.02 (0.000)	352.37 (0.000)	392.08 (0.000)	392.02 (0.000)
Hansen test (p-value)	–	–	–	–	–	–	–	–	183.67 (0.152)	186.12 (0.114)	191.73 (0.144)	196.21 (0.119)
No. of observations	1674	1674	1670	1670	1717	1717	1713	1713	1674	1674	1670	1670

Appendix

Table A: Definition of variables	
Variable	**Definition**
Panel A: Financial variables	
Total assets	The sum of total current assets, long term receivables, investment in unconsolidated subsidiaries, other investments, net property plant and equipment and other assets (source: Datastream).
Inventories/ Total assets	Tangible items or merchandise net of advances and obsolescence acquired for either (1) resale directly or (2) included in the production of finished goods manufactured for sale in the normal course of operation over total assets (source: Datastream).
Receivables/ Total assets	The amounts due to the company resulting from the sale of goods and services on credit to customers (after applicable reserves) over total assets. These amounts should reasonably be expected to be collected within a year or within the normal operations (source: Datastream).
Debt/ Total assets	The sum of long and short term debt over total assets (source: Datastream).
Return on assets (ROA)	The net income the company uses to calculate its earnings per share (before extraordinary items) plus interest expense on debt (the service charge for the use of capital before the reduction for interest capitalized; if interest expense is reported net of interest income, and interest income cannot be found the net figure is shown; it includes but is not restricted to interest expense on short term debt, interest expense on long term debt and capitalized lease obligations, amortization expense associated with the issuance of debt) over total assets times 100 (source: Datastream).
Return on equity (ROE)	The net income the company uses to calculate its earnings per share (before extraordinary items) over common equity times 100 (source: Datastream).
Number of subsidiaries	The total number of subsidiaries listed in the annual report (source: annual report).
Number of non-UK subsidiaries	The total number of subsidiaries incorporated and operating overseas (source: annual report).
Number of US subsidiaries	The total number of subsidiaries incorporated and operating in the US (source: annual report).

Table A Cont'd	
Variable	**Definition**
Panel B: Audit variables	
Audit fee	The total fee paid for the statutory audit of the company and its subsidiaries. Where more than one auditor is involved the combined audit fee is used (source: annual report).
Non-audit fee to auditor	The total amount paid to the company's auditor for non-audit services. Where figures for non-audit services in the UK and overseas are separately mentioned the figures are combined in calculating this variable. In instances where companies mention the payment of additional non-audit fees to auditors, which are capitalised, we also include these in calculating this variable (source: annual report).
Non-audit fee to other	The total amount paid to accounting firms other than the auditor for the provision of non-audit services (source: annual report).
Big Four	Time-varying indicator variable equal to one if during 2002 to 2008 the audit is performed by one of the following accounting firms: Deloitte and Touche, Ernst and Young, KPMG and PricewaterhouseCoopers, and zero otherwise (source: annual report).
Big Five	Time-varying indicator variable equal to one if during 1999 to 2001 the audit is performed by one of the following accounting firms: Arthur Andersen, Deloitte and Touche, Ernst and Young, KPMG and PricewaterhouseCoopers, and zero otherwise (source: annual report).
London	Time-varying indicator variable equal to one if the auditor is located in London and zero otherwise (source: annual report).
Panel C: Cross-listing dummies	
Cross-listed	Time-varying indicator variable equal to one if the firm is cross-listed in the US during a specific year and zero otherwise (source: stock exchanges and Bank of New York).
Exchange-listed	Time-varying indicator variable equal to one if the firm is cross-listed on a US exchange during a specific year and zero otherwise (source: stock exchanges and Bank of New York).
OTC-listed	Time-varying indicator variable equal to one if the firm is cross-listed on the US over-the-counter market (OTC) during a specific year and zero otherwise (source: stock exchanges and Bank of New York).
Post-SOX effect for exchange-listed firms	Time-varying indicator variable equal to one during post-SOX period for exchange-listed firms
Post-SOX effect for OTC-listed firms	Time-varying indicator variable equal to one during post-SOX period for OTC-listed firms