Discussion

Discussion of “Attribute differences between U.S. GAAP and IFRS earnings: An exploratory study”

Ole-Kristian Hope

Rotman School of Management, University of Toronto

1. Introduction

Using firms listed on the German Neuer Markt, Van der Meulen, Gaeremynck, and Willekens (henceforth MGW) investigate whether there are significant difference in four attributes of earnings between firms that apply U.S. GAAP and firms that apply IFRS (or IAS). In my opinion, this study provides a timely investigation of an issue of considerable interest to both regulators and academics. Currently about 100 countries either require or allow the use of IFRS for their domestic publicly listed corporations. In the near future, we will likely only have two accounting standards worldwide: IFRS and U.S. GAAP.1 Given the adoption of IFRS around the world and the relative lack of evidence on effects of such adoption, the authors do not have to try hard to motivate their study.

MGW have made several improvements to the paper following their presentation – and my discussion of their paper – at the International Journal of Accounting Conference in Paris. Consequently, my discussion here will be relatively brief and will focus on a few selected issues.

---

1 Whether such a “convergence” to one or two accounting standards is optimal or not is hard to say. It is my feeling that the decision to require IFRS – and thus abolish domestic GAAPs – around the world has been mostly based on “faith” rather than on solid evidence suggesting an overall gain to society (see also Ball, 2006).
2. Relation to literature

In my opinion, MGW complement the extant literature on differences between IFRS and U.S. GAAP. That is, there are several prior studies that examine the quality of financial reporting across countries. Similarly, there are prior studies that focus on the German market but which either use small sample sizes or focus on information–asymmetry outcomes rather than earnings attributes. Since quite a few studies examine IAS versus U.S. GAAP, the main contribution of this study likely lies in its focus on one market, Neuer Markt, as well as its investigation of several (not just one or two) earnings attributes. Focusing on one market is a strength because it holds many institutional factors constant (see, however, my comments below regarding differential enforcement). On the other hand, it also significantly limits the sample size (which, as explained below, turns out to be important in this setting), and limits the generalizability of results.

3. Use of $R^2$ as a test metric and the Cramer test

At the conference, I raised the question of why the authors have chosen to focus on differences in $R^2$'s between two samples of firms. In my opinion, this choice has few (if any) advantages and several important disadvantages. Suffice to say that it is very difficult to compare $R^2$ across different groups of firms (which explains why we do not see such tests often in research papers). The authors do not provide any discussion of why they have made this choice. An alternative would be to instead focus on differences in coefficient estimates. Using such an approach, they could pool all observations and easily test for differences between the two groups using interaction terms.

Similarly, even ignoring the inherent difficulties involved in comparing $R^2$'s across samples, the Cramer test used for testing differences in $R^2$ between samples is an unusually weak test metric. In particular, the Cramer test is extremely sensitive to the number of observations included in the tests. This fact is important for the current paper because MGW want to be able to conclude that there “are no significant differences” between the

---

2 In their literature review, MGW discuss research on 20F reconciliations at some length. Such research is of relevance to regulators (i.e., the SEC), but the limitation of these studies is that they only examine firms already listed, whereas regulators may care even more about foreign firms that are not currently listed on the domestic exchange. In addition, the reconciliation stream of literature tends to focus on incremental value relevance (or information content, depending on the type of study) of additional, required U.S. GAAP disclosures. Such research does not directly address which GAAP system is “best.” Instead, it tests whether there is any additional information conveyed in having extra information disclosed (a subtle but important difference often overlooked in the literature).

3 Readers should note that there are other papers that investigate differences between IAS and U.S. GAAP using German firms. For example, Bartov, Goldberg, and Kim (2005) examine value–relevance differences, Van Tendeloo and Vanstraelen (2005) study earnings–management differences (i.e., “abnormal accruals” differences), and Leuz (2003) focuses on differences in bid–ask spreads and trading volume.

4 In addition, the sample period is marked by declining firm performance. “Value relevance” tests do not perform particularly well during such periods.

5 At the conference I made several specific comments on the four earnings attributes. For example, I do not understand why the authors use the $R^2$ from the Dechow and Dichev model when other studies use the standard deviation of the residuals from this model. Furthermore, I am not quite sure how different the test of “timeliness” is from the test of “value relevance.”
use of IFRS and U.S. GAAP. A quick perusal of the empirical results reveals a number of differences in $R^2$ that seem large in magnitude but which show up as either statistically insignificant or only marginally significant based on the Cramer statistic. For example, in the first regression results reported (Model 1a in Table 4), the $R^2$ for the U.S. GAAP sample is 12 percentage points higher than for the IFRS sample (0.2836 compared with 0.1633). However, the reported results show only very modest (if any) significance, with a reported $z$ statistic of 1.57. To highlight how sensitive this statistic is to the sample size, I recomputed the $z$ statistic after artificially increasing the sample size by 200. The resulting $z$ statistic is 2.36, which is highly significant. As another example, in panel A of Table 6, the $R^2$s for large firms in Model 2 using IFRS and U.S. GAAP are 0.0526 and 0.1212, respectively. Although the latter is 2.3 times larger than the former, the difference is reported as not being significant according to the Cramer test. Similarly, in the last row of panel A of Table 6, the $R^2$ for U.S. GAAP is almost 22 times greater than for IFRS. Still, the difference is deemed insignificant. Such results are indicative of a weak test.

4. GAAP differences

Given that MGW focus on properties of accounting earnings, the paper would benefit from a richer discussion of the actual accounting differences between IFRS and U.S. GAAP during the sample period. Such a discussion would be interesting for several reasons. First of all, some readers may not know if there were any material GAAP differences and if so, which they were. Second, if the GAAP differences were small during the sample period, one would expect differences between U.S. GAAP and IFRS to be (even) smaller today than they were back then. Finally, knowing more about specific accounting differences would help readers interpret the reported results. In particular, are the results due to accounting differences or to differences in firm characteristics? For example, MGW find that U.S. GAAP firms have more predictable earnings. A reader would like to know which accounting differences are likely to cause such a result—or whether the difference in predictability is unrelated to the use of different accounting standards (i.e., explained by firm characteristics instead).

5. The potential for differential enforcement

One of the strengths of the study is that by focusing on one market, Neuer Markt, the authors are able to hold many institutional factors constant. However, it is not clear that all such factors are constant between the two sets of firms. One such important institutional factor is the enforcement of accounting standards. In practice, enforcement can be as important in explaining accounting outcomes as the standards themselves (e.g., Hope, 2003). Applied to the MGW setting, I am not entirely convinced that the auditing of IAS was as stringent as the auditing of U.S. GAAP financial statements.

---

6 In my conference discussion I provided several such examples, which presumably explains why the authors chose to present results of pooled regressions (rather than year-by-year regressions) in the published version.

7 A general issue in the literature is whether predictability is good or bad. That is, smoothing is considered “bad” but smooth earnings are considered “good.”
Generally speaking, prior research has found considerable evidence of noncompliance with IAS (e.g., Cairns, 1999). To me this is especially interesting given that IAS allowed wide latitude in accounting measurement and recognition, and IAS as applied in practice was often referred to as “IAS light.” More specifically, during the sample period there was no strong market regulator for the German market, and none of the authorities were responsible for checking the accuracy and completeness of the financial statements (d’Arcy, 2001; Glaum & Street, 2003). In fact, Glaum and Street (2003, 92) state the following regarding the collapse of the Neuer Markt: “ironically, the sharp downfall has been attributed in part to the low quality of accounting and disclosure.” Glaum and Street (2003) document that the level of GAAP compliance is significantly lower for IAS companies than for U.S. GAAP companies listed on the Neuer Markt, suggesting that IAS was applied less rigorously than U.S. GAAP.

I appreciate the fact that the authors have attempted the address the issue of differential enforcement I raised at the conference. Specifically, MGW report multivariate results excluding firms that are not audited by Big 5 audit firms. This is a nice test. However, Glaum and Street (2003) show that GAAP disclosure compliance is significantly higher for Neuer Markt U.S. GAAP firms than for Neuer Markt IAS firms, even after controlling for auditor type.

6. Differences in firm characteristics (other than auditor)

One of the main concerns with the study is that readers may not be convinced that reported results can be attributed to accounting differences (or lack of accounting differences) between IFRS and U.S. GAAP and not to the fact that we are dealing with two different sets of firms (i.e., omitted firm variables).

In the post conference version, MGW show results separately for small and large firms and for less and more levered firms. This is a useful addition to the paper. Nevertheless, it also highlights the weakness of the research design employed (which is a result of focusing on differences in \( R^2 \)). Since the sample size likely is not large enough to use a matched sample design, a complementary (and efficient) test would be to add intercept controls for important firm characteristics (e.g., as in Van Tendeloo & Vanstraelen, 2005).

7. Conclusion

This is an interesting study that examines an issue of great importance to both practitioners and academics. It is an unresolved issue—whether there exists material differences in accounting outcomes between the two (main) remaining GAAP systems worldwide, IFRS and U.S. GAAP, and whether such differences (if they exist) have economic consequences. The article by Van der Meulen, Gaeremynck and Willekens will

---

8 Note that the \( R^2 \) for Model 3a (earnings predictability—the only test for which significant differences are found) in Table 5 is only 0.0363 as compared with 0.1938 in Table 4. This huge difference in explanatory power suggests that either “something else is going on” or that there is very significant noise in the estimation of these models.
provide motivation for future research on the economic consequences of IFRS implementation around the world.

References


