Comment on ‘Earnings management around UK open offers’

Seth Armitage¹ and John Capstaff²

Pre-publications draft, September 2007

European Journal of Finance 15, 2009, pp. 53-60

¹ School of Management and Economics, University of Edinburgh, UK.

² Department of Accounting and Finance, University of Strathclyde, UK.

Acknowledgements

We are grateful to Marie Lerin (Edinburgh) and Martin Kemmitt (Strathclyde) for research assistance.
The purpose of this note is to document discrepancies between certain of the findings reported in the preceding paper by Iqbal, Espenlaub and Strong, and the findings of the note’s authors using apparently similar data and methods. We were aware of the contents of the paper before its publication because one of us was a referee for the paper.

The results with which we are concerned are those for the long run average buy-and-hold abnormal return (ABHAR) following the open offers, and the two-day average abnormal return (AAR) on their announcement. The discrepancies in relation to our published findings are as follows. First, Ngatuni, Capstaff and Marshall (2007) report statistically significant positive BHARs over five years following open offers during 1991-95, whereas the BHARs in Iqbal et al over four years are negative and significant. Second, Armitage (2002) reports a significant positive announcement AAR of 1.99% for open offers during 1987-96, whereas the AAR in Iqbal et al is –1.10%, though it is not significant. We elaborate on these differences below.

Our puzzlement is with regard to the results derived from share returns. We have no reason to question the findings from accounting data regarding discretionary accruals around open offers. But if the estimates of the long and short run abnormal returns in the paper are not reliable, the results reported for the relationships between these variables and the discretionary current accruals (DCA) variable may likewise not be reliable.

**Long run abnormal returns**

Iqbal et al. refer to Ngatuni et al. (2007)\(^1\) who examined the long term performance of UK firms making SEOs. The main focus of that study was a large sample...
of rights issues made during the period 1986-1995, but they also examined a smaller sample of open offers. They used a BHAR methodology as follows. The BHAR on security \( j \) over the holding period \( \tau \) to \( T \) is the difference between the actual and expected buy-and-hold returns (BHRs):

\[
BHAR_{j,\tau,T} = \prod_{t=\tau}^{T}(1 + R_{j,t}) - \prod_{t=\tau}^{T}(1 + E(R_{j,t}))
\]

BHAR is the monthly BHR on the offer firm minus the expected BHR. The expected BHR is the return on a matched, non-issuing firm. The matched firms were chosen based on (a) size, (b) size and book-to-market ratio, and (c) size and industry.

The average BHAR for \( N \) firms over the holding period was calculated as:

\[
ABHAR_{\tau,T} = \frac{1}{N} \sum_{j=1}^{N} BHAR_{j,\tau,T}
\]

Ngatuni et al.’s (2007) sample covered open offers made during the period 1991-95, which is the same period examined by Iqbal et al. Ngatuni et al. reported a significant positive ABHAR in the 5-year period following open offer announcements. In contrast Iqbal et al. report significant negative abnormal returns in the 4-year post-offer period using a Fama-French 4-factor model, and a significant negative ABHAR over the 3-year post-offer period using a market portfolio benchmark. To validate their findings Iqbal et al. replicated the tests in Ngatuni et al. (2007) and found negative abnormal post-offer returns. They suggested that the contrasting results are most likely due to a survivorship bias in Ngatuni et al.’s (2007) analysis, where the offer firms were required to survive the full 5-year post-offer period but the matched firms were not.\(^2\)

\(^2\) See footnote 10 in Iqbal et al.
Iqbal et al. are correct in pointing out this inconsistency in Ngatuni et al.’s (2007) framework of analysis, and it does have an impact on the results. The results are different when a requirement to survive the 5-year post-offer period is also imposed on the matched firms, but there is only limited evidence of negative performance. The revised results are summarised in Table 1. There are notable changes in the magnitude and significance of the findings and, to some extent, they are benchmark sensitive. But, in general, the abnormal returns tend to be positive rather than negative, particularly over the 5-year post-announcement period.

Table 1 about here

In their paper Ngatuni et al. (2007) only reported the post-announcement ABHARs for open offers based on the size and book-to-market benchmark, so we concentrate on this revised result to begin with. Imposing the survival criterion on the size and book-to-market matched firms reduced the original sample in Ngatuni et al from 132 to 112, as 20 of the matched firms did not survive the full 5-year period. The sample sizes in Iqbal et al. are not stated but they start at 176 in year 0 and decline with the holding period due to delistings. The non-surviving matches were not replaced to permit a more effective test of whether the positive findings in Ngatuni et al. (2007) were entirely driven by the 20 non-surviving matched firms. The revised results are all positive but no longer statistically significant, although still quite large for the 5-year period at 39.78%. The result for the 1-year period is now small and positive at 4.51% as opposed

---

3 The unreported results using a size benchmark, and using a size and industry benchmark, were qualitatively similar but of lesser magnitude and significance.
to small and negative in the original sample (−6.70%). The 3-year post-announcement returns are reduced from 15.60% to 8.80% (not statistically significant).4

As predicted by Iqbal et al. the survivorship bias in Ngatuni et al. (2007) led to a large overstatement of the positive returns in the 5-year post-offer period. Although the results for the size and book-to-market benchmark remain positive, they are now less emphatic. An examination of the detail of the results highlighted the potential for outliers to dominate the results. In a relatively small sample both survivorship bias and outliers can be distorting. A survivorship bias is likely to generate outliers but they can also occur from genuine observations.

The presence of outliers, even after removing the non-surviving matched firms, could account for the difference between the findings in Iqbal et al. and those in Ngatuni et al. We therefore checked the revised results for the impact of outliers. Figure 1 is a plot of the BHARs for the 112 offer firms over the 5-year period, using the size and book-to-market benchmark.

Figure 1 about here

One notable outlier is +1391%, whilst there are three positives and two negatives around 1000%. The outliers here more or less cancel each other out but it is easy to envisage a situation in which the negatives or the positives dominate, and hence drive the average results. One method of checking for the undue influence of outliers is by trimming. For example, setting boundaries of + and −1000% gives an ABHAR of

4 The results reported in Ngatuni et al. (2007) for the 1-year and 3-year periods were also not statistically significant.
+38.17% (N = 107) that is close to being statistically significant (p = 0.115). So the revised results for size and book-to-market do not appear to be driven by outliers. Iqbal et al. may wish to check whether there is any outlier dominance in their results.

Another possible source of the conflicting results lies in the choice of matched firms. Finding suitable matching firms can be problematical when combined with the requirement that matches did not raise new capital from share issues in the previous five years. The latter criterion rules out a substantial number of possible matches, and requires consideration of a wider range of firms. It may be that a matched firm strictly fits the criteria but has drifted sufficiently to cast doubt on its efficacy. This is more likely when a benchmark contains two criteria. For example, when matching by size and industry the criterion is that the matched firm is in the same industry with the closest, but higher market value. The latter requirement occasionally means that the closest match has a much larger MV, which may undermine the validity of the match. There are other reasons why some researchers do not use an industry based criterion. Iqbal et al. avoided matching problems by using market return as a buy-and-hold benchmark for the results in their Table 5. However, they may wish to examine more closely the choice of the matched firms for all benchmarks in their replication of Ngatuni et al. (2007).

The adjusted Ngatuni et al. (2007) results for the size and industry benchmark are also shown in Table 1 above. The imposition of survival on matched firms reduced the sample to 112. The 5-year post-offer ABHAR is 24.75% but it is not statistically significant. There is a small positive ABHAR for the 3-year period, and a small negative ABHAR for the 1-year holding period, both statistically insignificant. Hence, there is no

---

5 This is the criterion used by Speiss and Affleck-Graves (1995). They choose matched firms larger than the sample firms mainly because the size of the issuing firm size is expected to increase post-issue.

6 See Loughran and Ritter (1995) who argue against using an industry based criterion (pp. 27-28).
evidence of long term negative performance using the size and industry benchmark, although there are small negative median ABHARs for the two shorter holding periods.

Finding suitable matching firms is less of a problem when a single criterion is used, and this was the case with matching by size only, but the results may be less compelling as fewer of the offer firms’ characteristics are captured. Size-matched ABHARs show evidence of negative abnormal returns over the 3-year period but not over the 5-year period. The sample in Ngatuni et al. (2007) is reduced to 119 when the survivorship criterion is imposed on size matched firms. The previous result for 3 years was positive but small at just under 10%, and it contained some large positive outliers linked with the non-surviving matched firms. In that particular case, therefore, the survivorship bias appears to have moved Ngatuni et al.’s (2007) results from positive to negative. The revised mean of $-24.40\%$ is statistically significant at the 10% level. The median of $-33.70\%$ is statistically significant at the 1% level, and therefore offers support to a finding of negative performance over a 3-year holding period. The 5-year positive finding is not driven by outliers. Trimming to within $+\text{ or } -1000\%$ gives a 5-year ABHAR of $19.22\%$, with only two observations removed.

In summary, the application of the same survival criterion to both offer firms and matched firms in the Ngatuni et al. (2007) sample changes the scale and significance of the findings. However, following the adjustment, there is only limited evidence of negative post-announcement abnormal returns. The ABHARs are all positive for the 5-year holding period, and positive for two out of three benchmarks for the 3-year holding period, although they are not statistically significant.\footnote{There remains a clear distinction between performance following rights issues and performance following open offers as emphasised in Ngatuni et al. (2007).} Taking account of Iqbal et al.’s
important point on survivorship bias has moved Ngatuni et al.’s findings closer to the findings of Iqbal et al., but a discrepancy remains. It may simply be due to difference in the samples. The re-examination of Ngatuni et al. (2007) also led to deliberation of the possible problems posed by outliers, and the possible distorting effects of unsuitable matched firms. These concerns need to be explicitly recognised in research using relatively small samples, and should feature in the discussion of the associated results.

Event study

Iqbal et al. report an AAR for the day before the announcement plus the announcement day of –1.10%, which is not statistically significant (presumably at the 10% level or better). Their starting sample is all 286 open offers by UK listed industrial firms during 1991-95 recorded in the *FT Extel Record of Takeovers, Offers and New Issues*. From this sample they subtract 28 offers of less than £1m, 32 repeat offers and 94 offers for which there was insufficient accounting data in Datastream, leaving a final sample of 132 open offers. Their method of calculating the abnormal return for share \( j \) on day \( t \), \( AR_{jt} \), is the index model, ie

\[
AR_{jt} = R_{jt} - R_{M,t}
\]  

(3)

where \( R_{jt} \) is the return for share \( j \) and \( R_{M,t} \) is the return on a market index, in this case the FTSE All-Share Index. To enable daily returns to be summed, each return will be calculated as

\[
R_{jt} = \ln(P_{jt}/P_{jt-1})
\]  

(4)

where \( P_{jt} \) is the share price at the close of day \( t \). Their source for both accounting and share price data is Datastream.
We use the abnormal returns calculated for Armitage (2002) to check the AAR of –1.10%. His source for both prospectuses and share data was Primark Extel, which no longer exists. His database contains 318 open offers with share price data for the years 1991-95. He calculated abnormal returns using two methods. One is the familiar market model:

$$AR_{j,t} = R_{j,t} - (\alpha_j + \beta_j R_{M,t})$$

(5)

where the alpha and beta coefficients were estimated via OLS regression using daily returns and an estimation period of 180 days, and the index used was the FTSE All-Share, as in Iqbal et al. The second method was introduced by Eckbo and Masulis (1992); the results reported in Armitage (2002) are from this method. For each offer a regression was run using daily data and dummy variables to distinguish sub-periods of interest:

$$R_{j,t} = \alpha_j + \beta_j R_{M,t} + \gamma_{1,j} D_{1,t} + \gamma_{2,j} D_{2,t} + \gamma_{3,j} D_{3,t} + \gamma_{4,j} D_{4,t} + \epsilon_{j,t}$$

(6)

where $$D_{1,t} =$$ one for event days –1 to 0, and zero otherwise, day 0 being the announcement day. The other dummy variables picked out subsequent sub-periods, with which we are not concerned here. The combined estimation and event period was from 85 days before the announcement to 100 days after the offer close. The coefficient $$\gamma_{1,j}$$ is a measure of the abnormal return for each day of the announcement sub-period, so the two-day announcement abnormal return is $$2\gamma_{1,j}$$.

The results for 1991-95 using Armitage’s data are reported in Table 2. The AAR for the full sample of 318 open offers is 1.55% (from the market model; the results from the Eckbo-Masulis method are very similar). In an attempt to replicate the sample of Iqbal et al., the following were excluded: offers of less than £1m; offers by issuers that made more than one open offer during the sample period; and offers by issuers that could
not be found in Datastream as at July 2007 or did not have accounting data in Datastream for the year in which the offer took place. This resulted in a sample of 129, which is close in size and, we trust, in composition to the sample of 132 studied by Iqbal et al. The AAR for our replicating sample is 2.91%. Thus, the exclusions to try to replicate the sample in Iqbal et al. result in an increase in the AAR using Armitage’s data and an increase to 4.04 percentage points in the difference between the AAR of Iqbal et al. and the AAR reported here. The distribution of the abnormal returns is somewhat skewed to the right, as the median is 1.52%. But the AAR of 2.91% is not due to outliers; it is 2.83% excluding the two highest and two lowest abnormal returns, and 2.53% excluding the four abnormal returns with an absolute value in excess of 20%.

Some of the observations in the forgoing samples should, arguably, be excluded because of problems with the data, as discussed in Armitage (2002, p. 1258). In particular, (i) some shares make offers during periods in which their shares have been suspended from trading, and Extel continues to show a (constant) price for the share; (ii) Extel does not always make an adjustment for the share’s going ex-rights; and (iii) sometimes it was not possible to check for (i) and (ii). Removing the observations with problems (i), (ii) or (iii) results in small increases in the AARs for the full sample and the replicating sample.

In their discussion of previous event studies on UK seasoned equity offers, Iqbal et al. infer that the reaction to private placings is likely to be more positive than the reaction to open offers. Our sample of 238 open offers with unproblematic price data contains 36 offers that were accompanied by a pure private placing (the shares in an open offer are ‘placed with clawback’, ie placed and also offered pro rata to existing
shareholders, whereas there is no clawback in a pure private placing). But excluding the open offers accompanied by a placing increases the sample AAR, so the presence of open offers accompanied by a pure placing does not account for the positive AARs in our samples.

The method of calculating the test statistic for significance makes a big difference to its numerical value, though in this case not to our conclusions. Many event studies, including Armitage (2002), report a test statistic based on standardised abnormal returns (SARs). For the market model, the abnormal return for each share is standardised by dividing it by the standard error of the market model regression for the relevant share, and then a test statistic is calculated from the average of the SARs:

\[ t_{1MM} = \sqrt{N[\text{av}(SAR_j)]}/\sqrt{T} \]  

where \( N \) is the number of offers in the sample and \( T \) is the number of days over which the SAR is accumulated. The standardisation has the effect that shares with more volatile returns are given less weight in arriving at the test statistic. For the Eckbo-Masulis method,

\[ t_{EM} = \sqrt{N[\text{av}(\gamma_j/s_{\gamma_j})]} \]  

where \( s_{\gamma_j} \) is the standard error of the \( \gamma_j \) coefficient for share \( j \) (Eckbo & Masulis, 1992, p. 319).

An alternative for the market and index models is to calculate a conventional \( t \)-statistic for the mean of a sample. In our case the sample consists of the two-day abnormal returns, so

\[ t_{2MM} = \sqrt{N(AAR)/\text{stdev}(AR_j)} \]
where \( \text{stdev}(AR_j) \) is the standard deviation of the sample of two-day abnormal returns. It is striking that the test statistics based on SARs are several times larger than the conventional \( t \)-statistics. SARs can not be calculated using the index model, so (9) is probably the test statistic used by Iqbal et al. The AAR for the sample of 129 offers closest to that of Iqbal et al is significantly greater than 0 at the 1% level, according to the statistics in (7), (8) and (9). Since the AAR in Iqbal et al is negative, the difference in the AARs between the two studies is almost certainly highly statistically significant.

The substantial difference in the AARs is perplexing. The samples should not differ much. The studies use different methods of calculating abnormal returns, but this is extremely unlikely to account for a difference in the AAR of 4.04 percentage points for a two-day event window. Brown and Warner (1985), for example, find via simulation that the market and index models produce very similar results for event windows of one day and eleven days. The sources of the price data differ. One possibility is that Datastream’s ‘adjusted prices’ (code P) are not adjusted for the impact of the discount in open offers before the shares go ex-entitlement, which is usually on the announcement day in open offers. Most of the prices in Extel were adjusted in this way. If Datastream’s prices are not adjusted, a lower announcement AAR would be expected from Datastream data than from Extel data. Alternatively, or in addition, there may be differences in the data due to errors in either database. Ince and Porter (2006) compare Centre for Research in Securities Prices and Datastream US share price data and report, *inter alia*, that ‘the closing prices used by each source often do not agree’ (p. 472). Screens for errors alter results substantially, for example for the mean monthly returns of size-ranked portfolios.
Conclusion

Our checks have confirmed the initial impression of materially different results across the studies for both long and short run average abnormal returns, though the difference has narrowed in the case of the long run returns. We have discussed some possible explanations but remain uncertain as to why the results differ.

References


Table 1
Average buy-and-hold abnormal returns (ABHARs) following open offers in the UK during 1991-95: adjusted results from Ngatuni et al. (2007)

The table reports ABHARs calculated as the average of the buy-and-hold returns of offer firms minus the buy-and-hold returns of matched firms. Firms are matched by size, by size and book-to-market ratio, and by size and industry. ABHARs are calculated for the 1-year, 3-year and 5-year post-offer holding periods. The sample is from Ngatuni et al. (2007) but has been reduced following the removal of matched firms that do not survive the full 5-year period.

<table>
<thead>
<tr>
<th>Matching by</th>
<th>Mean</th>
<th>[median] ABHAR %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-year</td>
<td>3-year</td>
</tr>
<tr>
<td>Size (n=119)</td>
<td>-1.73</td>
<td>-24.40*</td>
</tr>
<tr>
<td></td>
<td>[4.05]</td>
<td>-[33.70]***</td>
</tr>
<tr>
<td>Size and book-to-market (n=112)</td>
<td>4.51</td>
<td>8.80</td>
</tr>
<tr>
<td></td>
<td>[3.31]</td>
<td>[0.46]</td>
</tr>
<tr>
<td>Size and industry (n=112)</td>
<td>-1.38</td>
<td>2.60</td>
</tr>
<tr>
<td></td>
<td>[-1.65]</td>
<td>[-7.80]</td>
</tr>
</tbody>
</table>

***statistically significant at the 1% level (t-test)
*statistically significant at the 10% level (Wilcoxon signed rank sum test)
The table reports the average abnormal return (AAR) for the day before the announcement plus the announcement day, using abnormal returns calculated for Armitage (2002). The market model AAR is given by equation (5) in the text and the two test statistics, $t_{1MM}$ and $t_{2MM}$, are given by (5) and (7). The Eckbo-Masulis AAR and its test statistic are given by equations (7) and (8) respectively.

<table>
<thead>
<tr>
<th>Market model</th>
<th>$t_{1MM}$</th>
<th>$t_{2M}$</th>
<th>Eckbo-Masulis</th>
<th>$t_{EM}$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAR</td>
<td></td>
<td></td>
<td>AAR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full sample</td>
<td>1.55%</td>
<td>19.18</td>
<td>1.89</td>
<td>1.65%</td>
<td>17.34</td>
</tr>
<tr>
<td>Excluding problem data</td>
<td>1.77%</td>
<td>19.11</td>
<td>2.51</td>
<td>1.95%</td>
<td>16.32</td>
</tr>
<tr>
<td>Excluding problem data and offers with placings</td>
<td>2.30%</td>
<td>19.91</td>
<td>3.46</td>
<td>2.50%</td>
<td>16.86</td>
</tr>
<tr>
<td>Replication of Iqbal et al</td>
<td>2.91%</td>
<td>17.64</td>
<td>4.66</td>
<td>2.83%</td>
<td>13.39</td>
</tr>
<tr>
<td>Replication excluding problem data</td>
<td>3.14%</td>
<td>17.98</td>
<td>4.49</td>
<td>3.13%</td>
<td>13.59</td>
</tr>
</tbody>
</table>