The prediction of earnings movements using accounting data: An update and extension of Ou and Penman

Received (in revised form): 7th February, 2001

Ron Bird

studied at Monash University prior to taking up a lectureship at Macquarie University in 1970. He moved to the Australian National University in 1973 and was subsequently promoted to Professor and Head of the Commerce Department. After leaving in 1988, the Australian National University awarded him the title of Emeritus Professor. Initially, in his private sector career, he headed up Towers Perrin's asset consulting operations in the Asian region while also being in charge of their global research. He moved to funds management in 1992, first being in charge of research and product development at Westpac Investment Management prior to founding GMO Australia in 1995. He returned to academia in 1999 and took up his current position in the School of Finance and Economics at the University of Technology, Sydney.

Richard Gerlach

is an Associate Lecturer in Statistics in the School of Mathematical and Physical Sciences at the University of Newcastle (Australia). He completed his PhD, majoring in Statistics, at the Australian Graduate School of Management, University of New South Wales in May 2001. From 1999 to 2001, he was a Research Associate in the School of Finance and Economics at the University of Technology, Sydney.

A. D. (Tony) Hall*

is Professor of Finance and Economics at the University of Technology, Sydney, where he is director of the postgraduate research degree programmes in the School of Finance and Economics. His research interest is financial econometrics. He has taught finance and econometrics at the University of California, San Diego, The Australian National University and Bond University. In 2000, he was a Jean Monnet Fellow in the Economics Department at the European University Institute in Florence.

*School of Finance and Economics, University of Technology, Sydney, PO Box 123, Broadway NSW 2007, Australia.
Tel: +61 2 9514 7729; Fax: +61 2 9514 7711; e-mail: tony.hall@uts.edu.au

Abstract The greater interest apparent in the recent academic literature in the impact of corporate earnings information on the valuation of shares has prompted an updating of the seminal work of Ou and Penman (1989) on the role that accounting information can play in predicting future movements in earnings relative to expectations. The Ou and Penman analysis has also been extended by considering two measures of expectations; by covering a more recent time period; by encompassing the UK and Australian markets in addition to the US market; and by applying a new methodology for developing the forecasting models. The study found that a model based on accounting information retains merit as a means of forecasting movements in trend-adjusted earnings, but that these forecasts no longer provide the basis for a profitable investment strategy. The models used to forecast earnings surprise gave mixed results, but nonetheless, some of the resulting investment strategies were consistently profitable.

Keywords: *accounting information; earnings; analysts' forecast; investment strategy*

Introduction

Numerous recent research papers have concentrated on the importance of earnings announcements and/or forecasts in the determination of share prices. The origin of such earnings research dates back to Ball and Brown (1968), who demonstrated the importance of earnings in the minds of equity investors by measuring the market's reaction at, and around, the time of an earnings announcement. Other writers have subsequently demonstrated the existence of a post-announcement earnings drift in returns that lasts for several months, which is consistent with an initial underreaction to the earnings announcement.

The analysis of earnings would seem to have gathered impetus in recent years, as evidenced by the continuing stream of research into earnings forecasts, earnings manipulation and the market's reaction to earnings surprise. The findings of these mainly US studies have included the identification of an optimistic bias in analysts' earning forecasts (although this may have dissipated in recent years; see Richardson *et al.* (1999)) a predictable manipulation of earnings by management which seems to 'fool' the market (eg Sloan, 1996); and a closer specification of the market's reaction to earnings surprise (eg Lopez and Rees, 1999). All of these studies have highlighted that the major market participants — analysts, investors and management — are fixated with earnings. In turn, researchers have attempted to develop various investment strategies designed to exploit the types of behaviour displayed by these market participants (eg Collins and Hribar, 1999).

Ou and Penman (1989) (hereafter O&P) were the first writers to focus on the usefulness of accounting information to predict the direction of the movement of earnings relative to trend-adjusted current earnings. Presumably, anything that enables a better understanding of earnings is of value, given the market's fixation on earnings. The O&P study is important because it evaluates whether accounting information, on which analysts heavily rely in arriving at their forecasts and recommendations, can subsequently be used as the basis for a profitable investment strategy. The authors found that the accounting data could be used to provide superior forecasts of future earnings movements and that these forecasts could be used as the basis for a profitable investment strategy. The evidence from subsequent studies applying the same method as O&P, however, has been mixed, as is highlighted in the next section.

An initial objective of this study was to repeat the original O&P study over a more recent time period. This was felt to be of importance because it would determine whether their results still hold at a time when management (at least, in the US) would appear to have become more involved in engineering a preferred earnings surprise outcome. The study is extended to other markets (in this case, the UK and Australian markets) where differing behaviour by the major market participants might result in different findings. Finally, a superior method of determining the variables to be included in the forecasting models was used to enable better definition of the usefulness of accounting information in predicting future movements in earnings.

The next section of this study provides a brief evaluation of the O&P, and subsequent, studies. The method employed and the data used are outlined in the third section. The next two sections present and discuss the results for the models developed to forecast the future movements in earnings relative to trend-adjusted earnings; first in terms of the accuracy of the forecasting models, and then in terms of using the forecast from these models as the basis for a profitable investment strategy. The study's findings suggest that the market may well have moved beyond using trend-adjusted earnings to using analysts' forecasts as the basis for building expectations of future earnings.

Accordingly, the study repeats the analysis using the same accounting information to forecast future earnings surprise. In the final section, a summary of the major findings and recommendation for possible further analysis are presented.

Review of previous research

Ou and Penman study

O&P's foundation paper sought to use 68 accounting variables to model the direction of movements in earnings per share (EPS) one year out. They formed a logistic regression model over the 1968–72 period. This was used in the years from 1973 to 1977 to forecast the probability of a company's EPS in a particular year lying above its value for the previous year, adjusted for its trend over the previous five years. They then used the data from 1973 to 1977 to form a second logistic regression. This was used to forecast the probability of a company's EPS lying above its trend-adjusted EPS in each of the years from 1978 to 1983. Starting with the 68 different accounting variables, they progressively eliminated them, first, by applying a chi-squared test on a univariate basis and, secondly, by utilising a stepwise regression to evaluate the significance of those variables that remained. Using data from 1968 to 1972, the application of this method resulted in a model with 16 explanatory variables. Using data from 1973 to 1977 resulted in 18 explanatory variables remaining in the model developed.

O&P then applied their models each year to obtain an estimate between zero and one, which represented the probability of a company's EPS actually increasing over the next year. When they classified each company with a probability above 0.6 ($p > 0.6$) as one that would realise an increase in EPS (ie EPS$\uparrow$) or each company with a probability below 0.4 ($p < 0.4$) as one that would realise an earnings per share decrease (EPS$\downarrow$), their forecasts proved correct in the case of approximately two out of every three companies. They then extended their study to examine the outcome of an investment strategy where they purchased an equally weighted portfolio of all stocks whose estimated probability was in excess of 0.6 and sold an equally weighted portfolio of all stocks whose probability was below 0.4. This strategy realised a return of 8.3 per cent over a one-year holding period, an incremental 5.7 per cent in the second year and 5.5 per cent in the third year.

Replications of Ou and Penman

There have been a number of replications of the O&P study. Holthausen and Larker (1992) studied US stocks over the period from 1978 to 1988. They found that the O&P method would realise little added value over the period of their study; that is, it performed reasonably well over the period from 1978 to 1983, which was a common period with the O&P study, but performed poorly during the 'new' period from 1984 to 1988. Bernard *et al.* (1997) replicated the O&P study over the period from 1973 to 1992, actually using the exact models developed by O&P during the earlier years. They found that the strategy added on average slightly in excess of 4 per cent in the first year and slightly in excess of 2 per cent in the second year. Of course, the period covered captures the relatively good performance identified by O&P in the period from 1973 to 1984, which suggests a significant drop off in performance in the latter years (consistent with Holthausen and Larker). In contrast, Stober (1992) replicated the O&P study over the O&P data period, finding similar results to O&P. He found that these returns could be further enhanced by using the strategy in combination with analysts' earnings forecasts — in particular, by only going long (short) those stocks for which the models predicted an increase (decrease) in EPS and the analysts' forecasts predicted the opposite. Finally, Setiono and Strong (1998) applied the O&P approach to UK stocks over the period from 1980 to 1988 and found that a long–short portfolio based on the forecasted probabilities realised 9.0 per cent during the first year, 7.7 per cent during the second and 4.9 per cent during the third year.

The present study

The summary of the research to date provides mixed evidence as to the success of using accounting data to predict the direction of future movements in EPS and the use of these predictions as the basis for a profitable investment strategy. The objective of this study is to update the O&P study in a number of dimensions: first, to extend the analysis to a more recent period when management would appear to be more involved in manipulating the earnings outcome; secondly, to review the application of the approach to an expanded set of countries — the US, the UK and Australia; thirdly, to apply a superior Bayesian method for identifying the accounting variables to be included in the model, and comparing this with the O&P method; fourthly, to use the analysts' consensus earnings forecasts rather than trend-adjusted current earnings as the comparison when measuring the trend in reported earnings.

Data and method

Data

The data for the US companies in this study were drawn from the Compustat (North American) database for the years 1983–97 and included data for non-continuing companies. The companies included in the study are all the non-financial companies included in the Compustat database with market capitalisations in excess of an equivalent US$100m as at the end of 1998. This minimum capitalisation figure was applied to ensure that there would be sufficient market liquidity to support any investment strategy that might be developed. The minimum figure was respecified each year, in line with movements in the Russells 3000 Index. The accounting data for both the UK and Australian companies was gathered from the Compustat (Global) database for the period from 1987 to 1998. Again, financial businesses were excluded from the study, but no capitalisation restriction was imposed, as the size of the firms included in the Compustat global universe was already relatively large. The return data used for the UK and Australian stocks was acquired from Grantham, Mayo, Van Otterloo's proprietary database. The number of companies included in the study's US database averaged around 1,700 in each of the years, with the corresponding figures being 450 for the UK and 225 for Australia. Finally, the information on consensus analysts' forecasts was drawn from the I/B/E/S database.[1]

The intention was to include in the

study as many of the 68 O&P accounting variables as possible. For the US, the study was able to capture 63 variables, including all of those that proved significant in the O&P models. Because companies with missing variables were excluded when developing the models, the only variables not included in the study were those whose inclusion would have resulted in the exclusion of a large number of companies. The Compustat global database is much less comprehensive, and only 52 variables were able to be included when developing the UK models and 47 variables when developing the Australian models. For both countries, very few of the variables found to be significant in the O&P study were excluded.

Methods

As one approach to developing the model, the O&P method was replicated, using the same two-stage approach to selecting variables to be included in the model. Initially, a chi-squared test was used on a variable-by-variable basis with the objective of excluding those variables that did not have a strong relationship with the dependent variable (ie the directional movement in trend-adjusted EPS). A stepwise regression was then used to determine the variables to be included in the final model. This involved a cycle of including all the remaining variables in a single regression, and then in this multifactor setting, progressively removing those that did not prove significant at the 10 per cent level. A different model was developed for each of the years for which forecasts were made, using the previous five years of observations — the forecast period being ten years for the US (1988–1997) and six years for both the UK and Australia (1993–1998). This approach contrasts with that in O&P, who also used a five-year period to develop each model, but then used it to arrive at a probability of the directional movement in EPS for the subsequent five years.

The study applied a Bayesian method as an alternative to the O&P two-stage method for variable selection, in the belief that it would overcome the statistical drawbacks of using stepwise regression. In particular, the variables selected by stepwise regression are not chosen to maximise the likelihood function of the model nor any other reasonable objective criterion (see Weisburg (1985) for further discussion). Bayesian methods estimate models by maximising the likelihood function of the model multiplied by a prior, with the choice of flat or uninformative priors leading to maximum likelihood estimates.

The mechanics of the Bayesian approach is that an initial (prior) probability is proposed for each of the variables; this expresses the likely importance of that variable being an important explanatory of the dependent variable. The prior probability applied in each instance was 0.5, which is equivalent to being neutral as to whether a particular variable should be included in the model. This prior probability attached to each variable is updated to provide an estimate of the posterior probability that the variable should be included in the model, based on the five years of data used to develop each model. The decision to include a particular variable in the final model was based upon the value of its posterior probability. Again, a five-year moving window was used to develop each model to forecast the probable outcomes in each of the ten years studied. The major advantage of the Bayesian approach is that it should result in superior in-sample forecasts relative to the O&P two-stage approach. This provides a superior starting point for arriving at the out-of-sample forecasts.[2]

Table 1 Details of the forecasting models developed using the Bayesian and Ou and Penman methods

Method	Bayesian	O&P
	US models	
Average number (range) of explanatory variables	18 (14–21)	12 (4–32)
Most common explanatory variables (times included)	— return on total assets (10) — change in sales to inventory (10) — change in total assets (10) — % change in inventory (9) — change in cap. exp. to total assets (9) — change in cap. exp. to total assets lagged one year (9) — operating income to sales (8) — pre-tax income to sales (8)	— change in cap. exp. to total assets lagged one year (7) — return on equity (6) — change in cap. exp. to total assets (5) — % change in quick assets (5) — change in debt to equity (4) — change in sales to total assets (4) — operating income to sales (4)
	UK models	
Average number (range) of explanatory variables	8 (4–14)	7 (4–11)
Most common explanatory variables (times included)	— return on total assets (6) — % change in depreciation (6) — % change in sales (5) — operating income to total assets (5) — pre-tax income to sales (4) — % change in total assets (4)	— % change in depreciation (6) — return on total assets (5) — % change in equity to fixed assets (5) — net profit margin (3) — % change in total assets (3)
	Australian models	
Average number (range) of explanatory variables	6 (4–9)	6 (3–7)
Most common explanatory variables (times included)	— return on total assets (6) — pre-tax income to sales (5) — net profit margin (3) — cash flow to debt (3)	— return on total assets (6) — % change in long term-debt to equity (5) — return to opening equity (3) — pre-tax income to sales (3)

Models

A summary of the more important details of the models developed using both methods is contained in Table 1. In the case of the US models, the average number of variables included is 18 using the Bayesian method and 12 using the O&P method, both of which are similar to the number of explanatory variables in the O&P models. In the case of the UK and Australian models, the numbers of explanatory variables included were much lower, typically lying in the range from six to eight. The overlap between the variables chosen using the two methods is approximately 20 per cent, meaning that one in every five variables chosen to be included in a model for a particular year by one of the Bayesian or the O&P methods is also chosen by the other method.

Perhaps the matter of most interest is the extent to which variables appear across a large number of the models. There are 22 models developed by each method — ten for the US, used to forecast each year from 1988 to 1997, and six each for the UK and Australia, used to forecast each year from 1993 to 1998. Using the Bayesian method, one variable (return to total assets) appears in all 22 models, while pre-tax earnings to sales appears in 17 out of 22 models. In some cases, a variable would only seem to be important in a single country, such as the '% change in depreciation' variable, which is included in all six UK models using both methods for variable

Table 2 Forecasting accuracy based upon Bayesian (B) and Ou and Penman (OP) techniques

	Upward movement in EPS (EPS↑)							Downward movement in EPS (EPS↓)						
	$p > 0.8$		$p > 0.6$		$p > 0.5$		Actual	$p < 0.5$		$p < 0.4$		$p < 0.2$		Actual
	B	OP	B	OP	B	OP		B	OP	B	OP	B	OP	
US models														
In-sample accuracy (%)	83	79	74	73	62	62	**52**	58	58	66	64	80	78	**48**
Out-of-sample accuracy (%)	79	75	70	69	61	59	**51**	57	56	63	58	64	60	**49**
UK models														
In-sample accuracy (%)	87	88	73	75	66	66	**54**	64	61	71	64	80	78	**46**
Out-of-sample accuracy (%)	85	86	74	72	68	65	**61**	45	50	47	49	55	45	**39**
Australian models														
In-sample accuracy (%)	87	84	79	82	68	70	**53**	65	65	75	78	97	92	**47**
Out-of-sample accuracy (%)	78	79	71	73	59	61	**50**	59	63	64	60	60	54	**50**

selection. This variable was never included in the US or Australian models, however. The general lack of stability of the variables is demonstrated by the fact that 41 of the evaluated variables feature across the aggregate of 243 variables included in the 22 Bayesian models, while 51 variables feature across the aggregate of 194 variables included in the 22 O&P models. On balance, the Bayesian method selects a more stable set of variables than the O&P method.

The model forecasts

The logistic models are derived as described in the previous section and then used to provide a forecast of the probability for each company of its EPS for the next year being above its current EPS adjusted by a five-year trend. Based on these forecasted probabilities, it is then possible to classify a particular stock as one for which the EPS are assumed to either increase or decrease. For example, one might classify each company with a probability of greater than 0.6 as one where EPS are expected to increase (EPS↑) and each company with a probability of less than 0.4 as one where EPS are expected to fall (EPS↓). The accuracy of the forecasts are then judged on the basis of the percentage of companies classified as EPS↑ that actually experience an increase in EPS and those classified as EPS↓ that actually experience a decrease in EPS.

Accuracy of the forecasts

The accuracy of the forecasts from the models developed under the Bayesian (B) and O&P (OP) methods for each of the countries are reported in Table 2. As an illustration, it can be seen from the first number on the left-hand side in Table 2, that where the Bayesian models are used to forecast in-sample (ie the forecasts are made over the same period used to develop the model), 83 per cent of all firms with a forecasted probability greater than 0.8 actually experience an increase in EPS. This 83 per cent can be compared with the 52 per cent, which is the actual percentage of companies that actually experience an increase in EPS. First considering the US models, it can be seen for both the in-sample and

Table 3 Summary of forecasting accuracy of Bayesian (B) and Ou and Penman (OP) techniques (%)

	$p > 0/6, p < 0.4$		$p > 0.5, p < 0.5$	
	B	OP	B	OP
US	66.5	63.5	62.5	59.5
UK	60.5	60.5	56.5	57.5
Australia	67.5	66.5	59.0	62.0

out-of-sample results that the accuracy of the forecasts for both EPS↑ and EPS↓ are much higher than would be expected from random allocation of companies to each classification. O&P used $p > 0.6$ ($p > 0.5$) and $p < 0.4$ ($p < 0.5$) as the break points for allocating stocks to EPS↑ and EPS↓ and found that this resulted in an overall accuracy of 66.5 per cent (61 per cent) for out-of-sample classifications. The equivalent figures in this study when the O&P method is used to develop the US models are 63.5 per cent (59.5 per cent),[3] but these are improved to 66.5 per cent (62.5 per cent) with the use of the Bayesian method. Therefore, the Bayesian superior in-sample accuracy carries over to the out-of-sample (forecasting) accuracy. Further, it is only the Bayesian results that achieved the same level of accuracy as those achieved in the O&P 1989 study.

A summarised version of the information contained in Table 2 is set out in Table 3. It demonstrates that the level of accuracy achieved by the US model is matched by the Australian models, while the UK models achieved a lower level of out-of-sample accuracy. In contrast to the US results, the study found that the O&P method provided a very similar level of accuracy to the Bayesian method for forecasting the direction of future movements in EPS for both UK and Australian companies. This in part reflects that the true advantages of the Bayesian method only become apparent when one has available relatively large amounts of data.

Investment strategy

'Perfect foresight' performance

Having identified models that have demonstrated some predictive power in forecasting the movement in a company's EPS over the next year, the next question was whether this information would be sufficient to identify mispriced stocks. Returns were calculated from a perfect foresight (PF) strategy composed of a long position in all stocks whose EPS for the next financial year is above trend (EPS↑) and a short position in all stock whose EPS is below trend (EPS↓). Over the ten-year measurement period for the US market, this long–short portfolio yielded an annual return of 14.2 per cent if the stocks in the portfolios were equally weighted and 6.5 per cent if the stocks were market weighted.[4] This indicates a small cap bias in terms of the value of information about the directional movement of a company's EPS over the next year. Overall, the 'perfect foresight' performance was much lower than the 22 per cent realised by the PF long/short portfolios in the O&P study. Indeed, the PF long/short returns in each of the three markets, as reported in Table 4, do not approach those of O&P. The reason is that an advance knowledge of the directional movement in EPS relative to trend was not nearly as valuable in the period from 1988 to 1998 as it was from 1973 to 1983. Nonetheless, the PF returns remain significantly positive, which suggests the possibility that the forecasts from the

Table 4 The performance of alternative investment strategies: 1988–1997 (% per annum)

	Market weighted portfolios							Equally weighted portfolio						
	PF	S1			S2			PF	S1			S2		
	L/S	L	S	L/S	L	S	L/S	L/S	L	S	L/S	L	S	L/S
US models														
B	6.5	20.7	28.7	−6.2	21.7	24.6	−2.3	14.2	24.3	21.4	2.4	24.6	19.3	4.4
OP	6.5	20.5	22.6	−1.7	21.0	22.3	−1.1	14.2	20.9	21.5	−0.5	21.3	24.1	−2.3
UK models														
B	8.9	20.5	17.7	2.4	20.4	21.1	−0.6	6.6	15.1	14.1	0.9	15.5	14.3	1.1
OP	8.9	16.9	14.3	2.3	18.7	22.4	−3.0	6.6	14.4	10.7	3.3	14.7	14.1	0.5
Australian models														
B	8.6	13.3	8.9	4.0	15.7	9.5	5.7	7.4	12.1	9.3	2.6	17.2	10.0	6.5
OP	8.6	19.0	19.7	−0.7	19.8	20.3	−0.4	7.4	9.1	16.3	−7.2	12.0	19.9	−7.6

models may still give rise to a profitable investment strategy.

Developing profitable investment strategies

The study evaluated two strategies across the three markets to determine whether a profitable strategy could be developed. The first strategy (S1) involved taking a long position in all stocks for which the model produces a probability estimate of EPS↑ greater than 0.6 and a short position in all stocks where this estimate is less than 0.4. The second strategy (S2) involved taking a long position in the top quartile of stocks ranked by their probability estimate of EPS↑ and a short position in the bottom quartile of stocks ranked by this same probability estimate. The stocks to be included in each portfolio were determined based on both the Bayesian (B) and O&P (OP) methods, with the allocations to each stock being either equally weighted or determined by market weights. This resulted in a total of eight different strategies being evaluated in each market. The performances of each of these strategies, along with the perfect foresight performance, are reported in Table 4.

For the US market, only two of the eight strategies yielded a positive annualised return over the sample period, the strategies being the Bayesian method and equal portfolios weights. The best performing strategy was S2, using the Bayesian method and equal weights. That returned 4.4 per cent, which represents about one-third of the perfect foresight returns. The returns from six of the eight strategies in the UK market were positive, with the only negative returns coming from S2 implemented using market weights. The best performing strategy (SI using O&P and equal weights), however, yielded a low 3.3 per cent per annum. In the Australian market, all four strategies based upon the Bayesian method yielded a positive return, with all four based on the O&P method yielding a negative return. The best outcome came from S2 using equal weights, which realised 6.5 per cent per annum. The same strategy with market weights, however, yielded only a slightly lower 5.7 per cent per annum.

In summary, returns of over 4 per cent in the US market, and less in the UK, did not seem sufficient to form the basis for a profitable investment strategy, especially after taking account of the transaction costs involved even on a one-year buy-and-hold strategy. The returns of around 6 per cent in the Australian market may well suggest a profitable strategy, but it should be noted

Table 5 Perfect foresight returns for the US, UK and Australian markets of earnings surprise relative to analysts' forecasts (trend-adjusted earnings)

Country	USA	UK	Australia
Market weighted returns (% pa)	12.2 (6.5)	4.1 (8.9)	11.6 (8.6)
Equally weighted returns (% pa)	17.9 (14.2)	8.1 (6.6)	12.0 (7.4)

that this figure is based on six years of fairly volatile experience. Comparing the use of the Bayesian and O&P methods as the basis for an investment strategy, there would seem to be little to choose between them based upon the evidence in the US and UK markets. In the Australian market, however, the case clearly points in favour of the Bayesian approach, as it outperformed the O&P method by amounts varying between 5 per cent and 14 per cent across the various strategies.

Analysts' forecasts

Up to this point of the study, the authors, like O&P, had used trend-adjusted EPS as the basis against which to judge one-year earnings movements. DeGeorge *et al.* (1999) established that management regards it as important to report an earnings figures in excess of the previous year's earnings. They also found, however, that another important reference point for management is the market's expectations for earnings as indicated by the consensus earnings forecast of financial analysts. Indeed, the weight of evidence would seem to suggest that the degree of earnings surprise associated with an earnings announcement is better measured with reference to the consensus analysts' forecasts rather than trend-adjusted earnings. This may at least partially explain why the returns associated with perfect foresight as to earnings movements have declined so much since the completion of the O&P study.

The increased importance of analysts' earnings forecasts suggest the use of these as the basis against which to judge reported earnings. Several studies have documented the impact on a firm's market value of reporting an EPS either above or below the consensus number.

Strategy using analysts' forecasts

The study evaluated the returns that could be earned from a strategy of going long (short) all stocks whose actual earnings were above (below) the consensus analysts' forecast made 12 months in advance of the end of the reporting period. These perfect foresight returns over a 12-month holding period are reported in Table 5 for both the Australian and UK markets over the period from 1993 to 1998, and for the US market over the period from 1988 to 1997. The numbers in parentheses are returns previously calculated where the perfect foresight returns are based on trend-adjusted earnings (see Table 4). The findings for the US and Australian markets substantiate the claim that superior investment performance might be achieved by being able to forecast earnings surprise accurately one year in advance rather than forecasting whether the actual earnings will be above or below trend-adjusted earnings. In the case of the UK market, the finding is less clear, with any improvement from attempting to forecast earnings surprise

Table 6 Performance of market weighted (equally weighted) portfolios formed on the basis of model predictions and analysts' forecasts

	USA returns (% pa)	
25.6 (20.5)	**Model > trend**	**Model < trend**
IBES > trend	24.5 (19.1)	25.2 (21.3)
IBES < trend	27.3 (22.2)	27.0 (23.3)
	UK returns (% pa)	
18.4 (14.3)	**Model > trend**	**Model < trend**
IBES > trend	20.7 (11.0)	11.8 (15.1)
IBES < trend	16.9 (15.7)	22.3 (15.1)
	Australian returns (% pa)	
	Model > trend	**Model < trend**
IBES > trend	13.8 (15.2)	18.5 (19.7)
IBES < trend	22.9 (16.3)	30.6 (25.3)

only occurring for those portfolios formed on an equally weighted basis.

The findings largely confirmed that the markets gain more information by accurately forecasting future earnings relative to the consensus analysts' forecast number rather than relative to trend-adjusted earnings. This suggested it would be interesting to rework the previous analysis using the consensus earnings forecasts as the point of comparison.

Strategy integrating forecasts of the model and analysts

The first stage in integrating both forecasts into the study was to evaluate the performance of portfolios formed on the basis of the forecasts from both the existing models and from the analysts' forecasts. The findings are reported in Table 6. Both the I/B/E/S consensus forecasts and those derived from the models developed using the Bayesian method were used. For each country, the return in the top left-hand cell represents the return from investing in those stocks where forecasts from both the analysts and the model are above the trend-adjusted earnings, while the bottom right-hand cell represents the return from investing in those stocks where the forecasts from both the analysts and the model are below the trend-adjusted earnings. The other two cells represent investing in stocks where there is disagreement between the two sets of forecasts as to whether reported earnings will be above or below the trend-adjusted earnings.

Stober (1992) found that the best returns were obtained where the models' forecasts differed from those of the analysts, suggesting that the models provide better information than could be obtained from the analysts' forecasts. This study found similar, but less strong, results in each of the three markets examined, using a portfolio composed of stocks whose model forecast was above trend-adjusted earnings, but whose analysts' forecast was below trend-adjusted earnings. This outperforms a portfolio formed on the reverse basis. Somewhat surprisingly, the best performing portfolio was that made up of stocks for which both the model and analysts' forecasts were below the trend-adjusted earnings. Consistent with some previous findings (Bird *et al.*, 2000), this suggests that the trend-adjusted earnings may be on

Table 7 Details of the forecasting models developed using the Bayesian method

	US model	UK model
Average number (range) of explanatory variables	8 (5–11)	5 (3–9)
Most common explanatory variables (times included)	— inventory to total assets (8) — return on total assets (7) — operating profit to sales (5) — % change in current ratio (4) — inventory turnover (4) — sales to total assets (4) — % change in working capital to total assets (4) — pre-tax income to sales (4)	— return on total assets (6) — net profit margin (4) — inventory to total assets (3) — current ratio (3) — operating profit to sales (3) — pre-tax income to sales (3)

average more accurate than either of the forecasts. As already noted, however, the accurate prediction of reported earnings relative to earnings forecasts is likely to give rise to an exploitable investment strategy, and this is addressed in the next section.

Forecasting earnings surprise

The final step in the study was to replicate the analysis described in the previous three sections using 12-month ahead consensus analysts' forecasts rather than trend-adjusted earnings as the reference point when judging movements in future earnings. In other words, the study developed models based on the same accounting data to forecast the probability that the earnings realised over the next 12 months will be either above or below the current consensus forecasts of those earnings. It was then evaluated whether these models could be used to derive forecasts which would form the basis for a profitable investment strategy.

The study's analysis was somewhat constrained by the limited number of companies for which both I/B/E/S consensus forecast data and all of the other data requirements for estimating the models were available. Typically, there was available information each year on approximately 60–70 Australian firms, 150 UK firms and 800 US firms. Effectively, this data constraint prevented the undertaking of any analysis of the Australian market, so the results reported in this section only relate to the US and UK markets. Further, given the previous finding of the superiority of the Bayesian technique for developing forecasting models, the analysis was restricted to the use of this approach and so only considered the one set of models to forecast earnings surprise.[5]

The models

Details of the models developed for both the US and the UK markets are summarised in Table 7. The most obvious characteristic of these models is that they contain approximately half the explanatory variables previously included when estimating future earnings relative to trend-adjusted earnings (see Table 1). Further, there is less consistency in the variables included in the various models than was found previously. The overall appearance of slightly inferior models to those developed previously to forecast movements relative to trend-adjusted earnings is likely at least partially to reflect the almost halving of the sample size in each of the countries.

Table 8 Forecasting accuracy based upon Bayesian techniques

	Positive earnings surprise				Negative earnings surprise			
	$p > 0.8$	$p > 0.6$	$p > 0.5$	Actual	$p < 0.5$	$p < 0.4$	$p < 0.2$	Actual
				US models				
In-sample accuracy (%)	68	55	52	**30**	70	77	77	**70**
Out-of-sample accuracy (%)	32	32	29	**29**	71	71	68	**71**
				UK models				
In-sample accuracy (%)	88	78	77	**75**	58	64	67	**25**
Out-of-sample accuracy (%)	83	80	80	**79**	44	47	67	**21**

Table 9 Summary of forecasting accuracy of Bayesian techniques (%)

	$p > 0.6, p < 0.4$	$p > 0.5, p < 0.5$
US	51.5 (66.5)	51.5 (62.5)
UK	63.5 (60.5)	62.0 (56.5)

Accuracy of the forecasts

The accuracy of the forecasts from the models developed under the Bayesian technique for the US and the UK are reported in Table 8. As an illustration, it can be seen from the first number on the left-hand side in Table 8, that firms with a *p* value greater than 0.8 actually enjoy a positive earnings surprise on 68 per cent of all occasions. The accuracy of this in-sample forecast compares with an 83 per cent success rate (see Table 1) that was realised when developing models to forecast earnings relative to trend-adjusted earnings.

Overall, the accuracy of the in-sample forecasts for the US market were not nearly as high as those obtained previously for the models developed to forecast earnings relative to trend-adjusted earnings for US companies. The same was not true for the UK market, where the accuracy of the in-sample forecasts was almost identical to those obtained previously when forecasting earnings relative to trend-adjusted earnings. Table 9 provides a summary of the accuracy of the out-of-sample forecasts of the models and, for comparison purposes, provides in parentheses the percentages obtained previously for the accuracy of the models developed to forecast earnings relative to trend-adjusted earnings.

The information in this table confirms that the models developed previously for the US market to forecast earnings relative to trend-adjusted earnings are much superior to those developed to forecast earnings surprise. Indeed, the US models would appear to have no out-of-sample forecasting power. An examination of the results for the UK market shows that the out-of-sample forecasts from the models were actually superior to those obtained previously when forecasting earnings movements relative to trend-adjusted earnings. In summary, the findings are mixed, with a slight improvement in forecasting accuracy for the UK market, but an appreciable decrease in forecasting

Table 10 The performance of alternative investment strategies based on the Bayesian models (% per annum)

	Market weighted portfolios							Equally weighted portfolio						
	PF	S1			S2			PF	S1			S2		
	L/S	L	S	L/S	L	S	L/S	L/S	L	S	L/S	L	S	L/S
							US models							
B	12.2	13.6	25.6	−9.6	27.8	17.7	8.6	17.9	11.8	19.5	−6.5	23.0	14.1	7.8
							UK models							
B	4.1	18.1	5.0	12.5	23.6	17.0	5.6	8.1	15.3	1.5	12.6	18.6	10.9	6.9

accuracy for the US market to the point where it totally disappears.

Investment strategy based on earnings surprise forecasts

The final piece of the study's analysis was to examine whether the model forecasts could be used as the basis for a profitable investment strategy, as was found to be the case in the O&P study. The study had previously identified for the US market, but not the UK market, that using a model which correctly forecasts earnings surprise is likely to realise greater added value than one that correctly forecasts earnings relative to trend-adjusted earnings. This results in the apparent situation that the US models did not forecast earnings surprise nearly as well as they did movements in trend-adjusted earnings, but the rewards for accurate forecasts were higher for the former than the latter. In contrast, the UK situation would appear to be that the models for forecasting earnings surprise are superior, but that the rewards are at best no better.

Strategy performance

Table 10 reports the performance of the two strategies previously evaluated in the section on 'Investment strategy': S1 where the strategy is to form a long portfolio of all stocks with a $p > 0.6$, and a short portfolio of all stocks with a $p < 0.4$; S2 where the strategy involves forming a long portfolio of the quartile of stocks with the highest p values and a short portfolio of the quartile of stocks with the lowest p values. The findings for the US market in relation to S1 are somewhat in line with expectations, as the strategy performs even worse than it did previously when the forecasts were based on trend-adjusted earnings. In contrast, the performance of S2 is not only superior to that obtained previously, but also extremely consistent in that it is positive in nine of the ten years evaluated. One would have to be sceptical of what would appear to be a potentially profitable strategy coming from a model with no apparent forecasting ability. The answer might be, however, that the probabilities from the model provide more information in terms of their implied ranking of stocks than they do in terms of the classification of stocks on the basis of whether they are forecast to have a positive or negative earnings surprise.

The performances of the two UK strategies were far superior to those determined previously using the models developed to forecast the direction of the movement in earnings relative to trend-adjusted earnings. As reported in Table 10, S1 captures added value well in excess of that which could be obtained from perfect foresight forecasting, while S2 captures the majority of the perfect foresight added value. Although the authors had good reason to expect that the performance of the UK models would improve, the results obtained are again questionable, as they suggest added value far greater than one would expect to realise based on the improved forecasting power of the models.

Conclusion

The focus of this study has been on replicating the Ou and Penman (1989) paper, which developed models using accounting data to forecast the direction of movements in EPS. The use of a superior Bayesian method to choose the explanatory variables improved the accuracy of the forecasting models for the US market relative to that obtained using the O&P method. The study found that a level of accuracy similar to the US market could be achieved in the Australian market, but that the Bayesian method was less successful at forecasting in the UK market.

As in the O&P study, this study evaluated whether the forecasts from the models could be used as the basis for developing a profitable investment strategy. It was found that using the probabilities generated by the models as the basis for separating stocks into those to be included in a long and in a short portfolio provided the basis for a potentially profitable one-year buy-and-hold strategy for the Australian market. Such strategies, however, achieved no better than half the returns realised by O&P for the US market. Only very small returns were realised for the UK market, which is consistent with the study's findings as to the predictive accuracy of these models in that market.

A potential reason why the investment strategies did not match the performance of those in the O&P study is that the rewards for having perfect foresight as to the direction of earnings movements may have significantly declined in the intervening period. The authors hypothesised that the investment community has moved beyond making a simple trend-adjustment to current earnings when forming expectations of future earnings, and are now much more guided in these expectations by the forecasts of the financial analysts. When an evaluation was made of the information content of earnings diverging from analysts' forecasts (ie earnings surprise), it was found that in both the US and Australian markets it had a much greater association with price performance than did divergences from trend-adjusted earnings.

The previous analysis was repeated using the Bayesian technique to develop models to forecast earnings surprise rather than divergences from trend-adjusted earnings. The findings for the US and UK markets were mixed, while it was not possible to complete the analysis for Australia owing to insufficient data. In the case of US stocks, the accuracy of the forecasts was significantly lower than for those previously obtained, but one of the resulting investment strategies consistently realised added value. For the UK market, the models achieved a higher level of accuracy than previously and gave rise to an investment strategy which realised a higher added value.

The overall conclusion drawn is that the accounting information evaluated in this study would seem to still be useful, as suggested by O&P, for forecasting the directional movement of future earnings. The findings are not as strong as those of O&P, however, nor do they provide the basis for the extent of profitability of their investment strategies. This comes as no surprise to the authors of this paper, who have always doubted the productivity of a fishing exercise such as that embarked on by O&P. A much more useful approach to utilising accounting information is to use it to supplement other types of information. A very interesting example of how that might be done is to be found in Piotroski (2000), who used accounting information to develop a fundamental signal as a supplement to a value-based investment strategy.

Acknowledgments

The financial support of Grantham Mayo Van Otterloo & Co LLC is acknowledged.

Notes

1 The authors would like to thank I/B/E/S for providing this data.
2 A technical appendix, which provides a detailed explanation of the Bayesian approach, is available from the authors.
3 These figures are calculated by taking the average of the EPS↑ and EPS↓ percentages. For example, using $p > 0.5$ and $p < 0.5$, we have 61 per cent and 64 per cent which averages as 62.5 per cent.
4 It is assumed that the investments were made three months after the end of the financial year in order to take account of the delay in releasing accounting information.
5 See Bird *et al.* (1999).

References

Ball, R. and Brown, P. (1968) 'An Empirical Evaluation of Accounting Income Numbers', *Journal of Accounting Research*, 6(2), 159–78.

Bernard, V., Thomas, J. and Wahlen, J. (1997) 'Accounting-based Stock Price Anomalies: Separating Market Inefficiencies from Risk', *Contemporary Accounting Research*, 14(2), 89–136.

Bird, R., Gerlach, R. and Hall, A. (1999) 'A Bayesian Approach to Variable Selection in Logistic Regression with the Application to Predicting Earnings Direction from Accounting Data', Working Paper, University of Technology, Sydney.

Bird, R., McElwee, B. and McKinnon, J. (2000) 'A global perspective of analysts' earnings forecasts', *Journal of Investing*, 9(1), 76–82.

Colling, D. W. and Hriber, P. (1999) 'Earnings-based and Accrual-based Market Anomalies: One effect or Two?', Working Paper, College of Business Administration, University of Iowa.

DeGeorge, F., Patel, J. and Zeckhauser, R. (1999) 'Earnings management to exceed thresholds', *Journal of Business*, 72(1), 1–33.

Holthausen, R. W. and Larcker, D. (1992) 'The prediction of stock returns using financial statement information', *Journal of Accounting and Economics*, 15(2/3), 373–411.

Lopez, T. J. and Rees, L. (1999) 'The Effect of Meeting Analysts' Forecasts and Systematic Positive Forecast Errors on the Information Content of Unexpected Earnings', Working Paper, Lowry Mays College & Graduate School of Business, Texas A&M University.

Ou, J. A. and Penman, S. H. (1989) 'Financial Statement Analysis and the Prediction of Stock Returns', *Journal of Accounting and Economics*, 11(4), 295–329.

Piotroski, J. J. (2000) 'Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers', Working Paper, University of Chicago.

Richardson, S., Teoh, S. H. and Wysocki, P. (1999) 'Tracking Analysts' Forecasts over the Annual Earnings Horizon: Are Analysts' Forecasts Optimistic or Pessimistic?', Working Paper, University of Michigan Business School.

Setiono, B. and Strong, N. (1998) 'Predicting Stock Returns using Financial Statement Information', *Journal of Business Finance & Accounting*, 25(5/6), 631–57.

Sloan, R. G. (1996) 'Do Stock Prices fully Reflect Information in Accurals and Cash Flows about Future Earnings?', *Accounting Review*, 71(3), 289–315.

Stober, T. L. (1992) 'Summary Financial Statement Measures and Analysts' Forecasts of Earnings', *Journal of Accounting and Economics*, 15(2/3), 347–72.

Weisberg, S. (1985) *Applied Linear Regression*, 2nd edn, John Wiley, New York.