ABSTRACT

Commentators suggest that between 30 and 60% of large US firms have adopted the Balanced Scorecard, first described by Bob Kaplan and David Norton in their seminal Harvard Business Review paper of 1992 (Kaplan and Norton, 1992; Neely and Marr, 2003). Empirical evidence that explores the performance impact of the balanced scorecard, however, is extremely rare and much that is available is anecdotal at best. This paper reports a study that set out to explore the performance impact of the balanced scorecard by employing a quasi-experimental design. Up to three years worth of financial data were collected from two sister divisions of an electrical wholesale chain based in the UK, one of which had implemented the balanced scorecard and one of which had not. The relative performance improvements of matched pairs of branches were compared to establish what, if any, performance differentials existed between the branches that had implemented the balanced scorecard and those that had not. The key findings of the study include: (i) when analyzing just the data from Electrical – the business that implemented the balanced scorecard - it appears that implementation of the balanced scorecard might have had a positive impact on sales, gross profit and net profit; but (ii) when comparing Electrical’s performance with its sister company these findings can be questioned. Clearly further work on this important topic is required in similar settings where natural experiments occur.

Keywords: performance measurements, impact and balanced scorecard.

BACKGROUND AND CONTEXT

Commentators suggest that between 30 and 60% of large US firms have adopted the Balanced Scorecard, first described by Bob Kaplan and David Norton in their seminal paper of 1992 (Frigo and Krumwiede, 1999; Kaplan and Norton, 1992; Neely and Marr, 2003). Despite this impressive take up, however, there is a paucity of empirical evidence that explores the performance impact of the balanced scorecard and indeed of performance measurement systems more generally (Franco and Bourne, 2003). In fact the extant literature has tended to focus on the problems with traditional measurement systems and how these can be overcome with alternative measurement methods and frameworks, such as the Balanced Scorecard and the Performance Prism (Kaplan and Norton, 1992; Neely et al, 2002). As a result much work has been carried out on the design and deployment of measurement systems, but relatively little on their impact (Bourne et al, 2000; Neely et al, 2000).

The aim of this paper is to explore the performance impact of balanced scorecards. The paper draws on data gathered over three-year period from a major wholesaler of electrical components in the UK, referred to as Electrical. The board of Electrical decided to implement a balanced scorecard in late 1999 and began working...
on the design of their balanced scorecard in early 2000. They spent six months completing the design phase of the process and a further six months rolling the balanced scorecard out across their 140 UK branches. On 1st January 2001 the balanced scorecard was formally launched and from that day the business stopped releasing information on branch profitability, which previously had been the main method of branch measurement. They also changed the firm’s incentive scheme, moving away from a bonus based on branch profitability and to a bonus based on performance against the balanced scorecard.

Importantly, mid-way during the rollout phase of the project (in the 3rd quarter of 2000), Electrical was acquired by another UK wholesaler of electrical components. The acquiring company – referred to in the paper as Sister – continued to use traditional methods of performance reporting at the branch level, namely profit and loss accounts, throughout 2001. This situation presented the researchers with a unique opportunity. In essence the research team was able to construct a sample of 35 pairs of matched branches (based on location), drawn from the samples provided by Electrical and Sister. One branch in each matched sample had adopted the balanced scorecard while the other had not. This matched sampling technique, known as quasi-experimental design, is a powerful methodology for assessing the impact of organizational changes (Cook and Campbell, 1979).

The paper consists of four main sections. In the first the relevant literature is explored, as this sets the scene for the study. In the second the methodology used in this study is explained and justified. The third section presents a quantitative analysis of the impact of the balanced scorecard, using branch level performance data. These data are compared to performance data from Electrical’s sister company, which as already discussed had not implemented the balanced scorecard at that time. The fourth and final section of the paper summarises the implications of this research for both the practitioner and academic audiences.

THE IMPACT OF PERFORMANCE MEASURES – RELEVANT LITERATURE
The shortcomings and dysfunctional consequences of performance measurement systems have been discussed in the academic literature for at least fifty years (Ridgway, 1956), but recently there has been a flurry of activity. Throughout the 1980s vocal and influential authors criticised the measurement systems used by many firms (Johnson and Kaplan, 1988; Hayes and Abernathy, 1980). By the 1990s the noise made by these voices had grown to a crescendo (Neely et al., 1995; Marr and Schiuma, 2002) and increasing numbers of firms appeared to be "re-engineering" their measurement systems, with data suggesting that between 1995 and 2000, 30 to 60% of companies transformed their performance measurement systems (Frigo and Krumwiede, 1999). Evidence suggests, for example, that by 2001 the balanced scorecard had been adopted by 44% of organisations worldwide (57% in the UK, 46% in the US and 26% in Germany and Austria). And more recent data suggests that 85% of organisations will have performance measurement system initiatives underway by the end of 2004 (Rigby, 2001; Silk, 1998; Williams, 2001; Speckbacher et al, 2003, Marr et al, 2004). However, cautionary evidence from three Austrian academics reported that 8% of 174 companies from German speaking countries decided not to implement a performance measurement system (and a balanced scorecard in particular) because they could not see advantages or ‘positive impact’, especially given the implementation effort required (Speckbacher et al, 2003).

Somewhat surprisingly (especially given all of this activity) there has been relatively little research into whether the balanced scorecard actually works. In fact this criticism can be leveled at the field of performance
measurement more generally, which has seen much prescription, but relatively little empirical research (Franco and Bourne, 2003). Kaplan and Norton have made some efforts to demonstrate the impact of the balanced scorecard, but their approach has been to use largely anecdotal cases (Kaplan and Norton, 2000). An important and notable effort is the work of Chris Ittner and David Larcker, which reports that only 23% of organizations that they surveyed consistently built and tested causal models to underpin their measurement systems, but that these 23% achieved 2.95% higher return on assets and 5.14% higher return on equity (Ittner and Larcker, 2003).

Others have sought to undertake similar studies, but have tended to execute them less robustly. Survey data collected by consultancy and commercial research companies suggests that organisations managed through ‘balanced’ performance measurement systems perform better than those that are not (Lingle and Schiemann, 1996; Gates, 1999). Lingle and Schiemann (1996) report evidence that organisations making more extensive use of financial and non-financial measures and linking strategic measures to operational measures have higher stock market returns. While Lawson et al’s (2003) study shows that the use of a performance measurement system as a management control tool reduces overhead costs by 25% and increases sales and profits. Other authors such as de Waal (2003) and Sandt et al (2001) have found less tangible benefits from the use of performance measurement systems. Dumond (1994) and Sandt et al (2001) suggest that the using balanced performance measurement systems improves the decision-making performance of managers and employees. Lawson et al (2003) and Dumond (1994) found that using performance measurement systems and linking scorecards to compensation significantly increased employee satisfaction, although Ittner et al. (2003) present evidence to the contrary.

Ketelhohn (1999) and Vasconcellos (1988) found that the identification and selection of appropriate measures and key performance indicators enhance the implementation and acceptance of business strategy, at the same time as enhancing employee understanding of the business. Furthermore, Forza and Salvador’s research (2000, 2001) supports the suggestion that employee communication that focuses on feedback from measures increases collaboration and facilitates buy-in.

**RESEARCH METHODOLOGY AND QUESTIONS**

As stated already, the aim of the research reported in this paper was to address the question - what is the performance impact of the balanced scorecard? To address this question the authors decided to adopt the quasi-experimental design methodology advocated by Cook and Campbell (1979). Core to the experimental design was one of the authors’ involvement (over an extended period) in the design and deployment of a balanced scorecard in a multi-branch electrical wholesale business based in the UK. This author worked with the board of the business and other members of the firm's senior management team over a two-year period, facilitating the design and deployment of their balanced scorecard. The length and extent of this involvement delivered significant benefits to the research team in several ways. Firstly, the author's involvement in the entire design and deployment process gave him detailed insight and valuable contextual information, which was used in the subsequent data analysis. Second, the research team was able to obtain unparalleled access to the organisation and particularly highly sensitive and confidential performance data.

In essence the research team were able to access sales and profitability data from two sister organizations. The first, Electrical, was the one which implemented the balanced scorecard in January 2001, following a one year
design and deployment process. Electrical provided the research team with data for 122 branches. The second company, Sister, continued to use traditional methods of performance reporting throughout the period of the study and provided the research team with data from 190 branches. These two sets of data were compared and branches based in the same location were matched. This matching by location enabled the research team to compare changes in organizational performance over the duration of the study, while controlling for local economic conditions, product range and customer base.

In the next section of the paper two phases of analysis will be presented. In the first the performance of all of Electrical’s branches will be considered, for the time period 2000-2002. Electrical made available to the research team monthly data on sales, gross profit and net profit for the entire period 2000-2002. The business moved from reporting profit and loss at the branch level to using a balanced scorecard on 1st January 2001, but reverted back to reporting profit and loss at the branch level of 1st January 2002, following an internal reorganization. This enabled the research team to explore the performance impact of implementing, operating and then removing a balanced scorecard and associated incentive scheme over a three-year period across a large sample of branches. In addition to the firm's performance data the research team also had access to extensive qualitative data to supplement this analysis. These data, which were gathered via participant observation during the design and deployment phase of the balanced scorecard project and more formally, through a series of semi-structured interviews conducted six to nine months after the balanced scorecard had been implemented.

The second phase of analysis involved comparing the performance of the 35 matched pairs of branches. This quasi-experimental design enabled the research team to explore the relative differences in performance achieved by Electrical and Sister, while attempting to control for local economic conditions, product range and customer base that might have otherwise affected the results of the study. Unfortunately, for reasons of commercial sensitivity, Sister was only willing to make monthly branch level sales and gross profit data available to the research team for 2000 and 2001. However this still enabled the research team to explore the relative changes in performance between Electrical and Sister during the period when the balanced scorecard was introduced.

THE IMPACT OF THE BALANCED SCORECARD IN ELECTRICAL
As mentioned in the previous section, Electrical provided the research team with monthly branch level figures for sales, gross profit and net profit for 122 branches for the entire period 2000-2002. 77 of these were complete records and hence the analysis presented in this section is limited to these. Between them these 77 branches account for £150 million annual sales, out of a total for Electrical of £180 million. Figure 1 shows how Electrical’s sales and net profit changed over the period under study.
Table 1 summarises these data and compares sales, gross profit and net profit in 2000 with 2001 and 2002. These data suggest that while sales and net profit were higher in 2001 (during the period when the balanced scorecard was operating), both sales and net profit dropped back in 2002 (once the company had reverted to traditional profit and loss reporting).

Table 1: Summary Performance Data for Electrical

<table>
<thead>
<tr>
<th></th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Sales</td>
<td>100%</td>
<td>105%</td>
<td>98%</td>
</tr>
<tr>
<td>Gross Profit</td>
<td>100%</td>
<td>99%</td>
<td>98%</td>
</tr>
<tr>
<td>Net Profit</td>
<td>100%</td>
<td>115%</td>
<td>51%</td>
</tr>
</tbody>
</table>

To explore these findings in more detail the research team conducted paired sample t-tests comparing the sample means. The data in Figure 1 suggest some seasonality, certainly in terms of sales, so the decision was taken to compare sales in January 2000 with sales in January 2001 and to compare separately sales in February 2000 with sales in February 2001, etc. The process resulted in 72 paired samples being tested, 24 for each of sales, gross profit and net profit. Table 2 summarises the results of these tests, highlighting only the statistically significant findings.
As Table 2 shows of the 72 tests, 41 provided a statistically significant result. Nine of these related to sales, nine related to gross profit and the remaining 23 related to net profit. Of the nine significant results relating to sales, three relate to the comparisons between 2000 and 2001 and all of these show that sales significantly increased after the balanced scorecard had been introduced. The other six, all of which relate to the comparisons between 2001 and 2002 show that sales significantly decreased after the balanced scorecard had been removed. The nine statistically significant results relating to gross profit show a less clear picture. For the comparisons between 2000 and 2001, one result shows a statistically significant increase in growth profit after
the introduction of the balanced scorecard, while another shows a statistically significant decrease in gross profit after the balanced scorecard had been introduced. Similarly for the comparisons involving the time periods 2001 and 2002. For this period, four results show a statistically significant decrease in gross profit after the withdrawal of the balanced scorecard, while three show a statistically significant increase in gross profit after the withdrawal of the balanced scorecard. Like sales, the results relating to net profit are more straightforward to interpret intuitively. Of the eleven statistically significant results for the time periods 2000 and 2001 nine show an increase in net profit, while two show a decrease in net profit after the balanced scorecard had been introduced. For the time periods 2001 and 2002 there are eleven that show a statistically significant decrease in net profit after the removal of the balanced scorecard, while only one shows a statistically significant increase in net profit after removal of the balanced scorecard.

On the surface, these data appear to suggest that introduction of the balanced scorecard has had a positive impact in terms of both sales and net profit and its removal has had a negative impact on both of these variables. Clearly there are questions of what else was happening in the business at the same time, especially given the fact that Electrical was taken over during the third quarter of 2000, during the deployment of the balanced scorecard. It could therefore be argued that many members of the organization would have been distracted and concerned about the implications of the takeover and therefore it would not be surprising to see performance deteriorate during the last quarter of 2000 and the early part of 2001. Indeed it was clear to the author of this paper who was working with the company at the time that many people were distracted, but even so the performance impact appears to have been minimal with performance improving during 2001. Of course the takeover was not the only thing that was happening in the organization. Changes were being made constantly at a local level, in terms of stocking policies, branch openings and closings, etc. All of these will have impact the branch managers and hence their local performance, but it is important to understand the magnitude of the change that the organization felt it was undertaking by implementing the balanced scorecard. The Chief Executive frequently described this as the biggest change that the business had made in a decade and it had clearly captured the attention of many of the business’ most senior managers. Indeed, in the 3rd quarter of 2000 the board of Electrical declared that they were willing collectively to resign if their new parent company blocked the implementation of the balanced scorecard.\footnote{Discussion between the board of Electrical and the Chairman and Finance Director of the company that acquired them in Sept., 2000} The point is that while other events were clearly occurring in the organization at the time of the study one of the most significant was widely perceived to be the implementation of the balanced scorecard and associated incentive scheme.

**CONTRASTING ELECTRICAL WITH SISTER**

As already discussed, in addition to the sales, gross profit and net profit data made available by Electrical the company that acquired Electrical owned another wholesaler of Electrical components in the UK – called Sister for the purposes of this paper. Sister provided monthly data on sales and gross profit for some 190 branches for 2000 and 2001. This data set was combined with the data provided by Electrical and 35 matched pairs of branches (branches based in the same town/city) were identified. This process allowed the research team to control for local economic conditions, product range and customer base, as Electrical and Sister branches tended to stock a similar range of products and deal with a similar range of customers. The data was normalized with the values for January 2000 in each case being set to 100. Figure 2 summarises the average growth in sales and gross profit performance for these matched sets of branches during 2000 and 2001 using these data.
To explore these findings in more detail the research team conducted paired sample t-tests comparing the sample means. Again, to take account of seasonality the research team decided to compare growth in both sales and gross profit on a year by year basis. Sales and gross profit growth were calculated for both Electrical and Sister separately and then the rates of growth compared. Interestingly these tests revealed no significant differences in terms of either sales growth or gross profit growth between the matched pairs of branches.

![Figure 2: The Matched Pairs Sample](image)

These findings are important because they suggest that perhaps the observed growth in sales and gross profit in the Electrical data set are perhaps related to growth in the industry rather than the implementation of the balanced scorecard. Unfortunately the research team have not yet been able to access data on net profit for Sister. Negotiations with Sister are continuing as access to these data would enable the research to explore whether the observed differences in net profit growth in Electrical are also due to industry related factors rather than the balanced scorecard specifically.

**DISCUSSION AND IMPLICATIONS**

These findings are interesting because at first sight the data suggest that implementation of the balanced scorecard has had a positive impact on Electrical’s performance. Further investigation, however, reveals that the observed changes in Electrical’s performance are not significantly different from observed changes in Sister’s performance in terms of sales growth and gross profit growth. Clearly further analysis needs to be carried out on these data and similar studies need to be replicated in other settings. One particular issue the research team would like to address is whether the changes observed in Electrical’s net profit performance are also consistent with changes observed in Sister’s net profit performance. Testing this requires the research team to gain access to Sister’s net profit performance data and clearly this is very sensitive data to the business. But testing it is important as it could be argued that the performance impact of the balanced scorecard is most likely to be observed, in the short term, in terms of net profit performance as this would encapsulate both sales and margin growth, as well as better control of an organisation’s cost base. It will take some time for
customers to recognize improved service and hence become more loyal to the business, but branch managers can, relatively quickly, take control over issues such as operational efficiency, improvements in which would show up in the net profit figures, but not the sales or gross profit data. The data presented on Electrical’s performance appear to support this claim as it is in terms of net profit growth that the most noticeable changes can be observed. However, the study reported in this paper also draws attention to the fact that one has to be careful in making such generalizations without appropriate “control groups” that can be used to take account of sectoral and local economic changes.

In terms of implications for practitioners and researchers, this study highlights the fact that further research is required into the performance impact of balanced scorecards and the timescale over which this performance impact can be observed. Certainly the data presented in this study suggest that the balanced scorecard implemented in Electrical had no significant impact in terms of sales growth or gross profit growth over a twelve month period. It may be that the balanced scorecard would have had a performance impact had it been retained for a longer period and the research team is currently in negotiation with another organization to access data that will allow them to test this claim. Similarly, if the findings of this study are replicated in other similar naturally occurring experiments, then work is required to understand why balanced scorecards do not have the impact one would expect. Intuitively people accept that organizations need to keep track of their performance so that they can identify how well they are doing and what they need to improve. Well-designed balanced scorecards provide access to such measures and so if correctly implemented one would assume that they should enable performance improvement. Yet in the case of Electrical performance improvement does not appear to have accrued. Why is this the case? Is it due to the fact that the organization did not give the balanced scorecard long enough to work? Is it that the organization did not supplement the balanced scorecard with an appropriate improvement methodology and/or programme? Is it that the balanced scorecard suffers from the same criticisms that can be made of many measurement systems – too much data arriving too late for managers to act on them? These issues need to be explored and understood much more fully so that we can advise practitioners not simply on how better to measure, but on how better to perform.

ACKNOWLEDGEMENT
The authors would like to acknowledge the support of the EPSRC under grant number [GR/S28846], which partially supported this research.

REFERENCES


