INCENTIVES AND MANAGERIAL EXPERIENCE IN MULTI-TASK TEAMS: EVIDENCE FROM WITHIN A FIRM

Rachel Griffith
Andrew Neely

THE INSTITUTE FOR FISCAL STUDIES
WP06/22

Incentives and managerial experience in multi-task teams: evidence from within a firm

Rachel Griffith[†] and Andrew Neely[‡]
14 October 2006

Abstract: This paper exploits a quasi-experimental setting to estimate the impact that a multi-dimensional group incentive scheme had on branch performance in a large distribution firm. The scheme, which is based on the Balanced Scorecard, was implemented in all branches in one division, but not in another. Branches from the second division are used as a control group. Our results suggest that the balanced scorecard had some impact, but that it varied with branch characteristics, and in particular, branches with more experienced managers were better able to respond to the new incentives.

Acknowledgements: The authors would like to thank the anonymous firm for use of the data and Laura Abramovsky, Rupert Harrison, Gareth Macartney, Imran Rasul, Helen Simpson and seminar participants at UCL for helpful comments. The analysis contained in this paper was funded by the Economic and Social Research Council (ESRC) and EPSRC under the Advanced Institute of Management Research (AIM). All errors and omissions remain the responsibility of the authors.

JEL classification:

Keywords: incentive design, balanced scorecard, managerial experience

Correspondence: rgriffith@ifs.org.uk; aneely@london.edu

[†] Institute for Fiscal Studies and University College London.
[‡] Cranfield School of Management, London Business School and AIM

1 Introduction

Many organisations provide some form of incentive pay to managers and workers. What form should these incentives take? This is the topic of many papers in economics and management literature. A well known theoretical result (Holmstrom and Milgrom, 1991) suggests that workers should not be jointly responsible for single tasks because sharing responsibility for a task increases the total risk that each worker faces of successfully completing the task without any benefit. Holmstrom and Milgrom also suggest that tasks should be grouped together base on the cost of measuring and rewarding performance in that dimension. Some workers should do easy-to-measure tasks and pay should be contingent on performance, while other workers should focus on hard to measures tasks and received fixed wages. This is because if a worker has both easy and hard to measure tasks they will concentrate on the easy to measure tasks, as the expense of the hard to measure tasks.

However, one of the most popular incentives schemes currently used by firms - the Balanced Scorecard - runs counter to these predictions. The Balanced Scorecard was introduced by Kaplan and Norton (1992). Variants of this have been adopted by a large number of firms and organisations across the globe.[1] The Balanced Scorecard was designed to overcome the dual problems of subjective performance evaluation (which can give rise to various forms of bias and encourage workers to waste effort to curry favour with managers) and focussing on one key performance target (which can lead to dysfunctional behaviour with workers focusing all their energy on tasks that are rewarded, and ignoring those that are not included in the incentive scheme).[2]

---

[1] See, for example, the Balanced Scorecard Hall of Fame at http://www.bscol.com/pdf/BSCHoF-Membersby_Industry_2000-2005.pdf, or see Gates (1999) who surveys 113 "leading" US, European and Asian companies and finds that 81% of respondents use a strategic performance measurement system; Maisel (2001) samples 1990 US management accountants and finds that 47% of respondents use a strategic performance measurement system; Rigby (2001) surveys 214 North American firms and finds that 44% of organisations use the Balanced Scorecard; Speckbacher et al (2003) estimate that 26% of firms in Germany, Switzerland and Austria use the Balanced Scorecard, and Marr et al (2004) find that 35% of North American organisations use it.

[2] See Argyris (1952), Ridgway (1956), Holmstrom and Milgrom (1991), Kerr (1995) and for a recent review of the literature Prendergast (1999).

The idea behind the Balanced Scorecard is that managers can improve performance by monitoring and rewarding a large range (usually 15 or more) of activities that reflect both past actions, that are believed to have lead to good financial performance, and current activities that will lead to future performance. Performance is typically evaluated at the group level (for example, a team or a branch) and incentive payments based on group performance on a large number of indicators.

The use and effectiveness of the Balanced Scorecard has been criticised in the management literature,[3] and seems to run counter to the theoretical literature, which emphasises the potential perverse outcomes when workers are responsible for multiple (substitutable) tasks and employers can not perfectly monitor inputs.[4]

The rapid uptake of the Balanced Scorecard has been largely fuelled by high profile success stories in other firms, yet there is little empirical evidence on how well the scheme works.[5]

In this paper we exploit a quasi-experimental setting within a single firm in order to investigate the effectiveness of the Balanced Scorecard. The firm implementing the Balanced Scorecard in one division to see how well it would work before rolling it out across the firm. This allows us to overcome one of the main problems in evaluating incentive schemes – the fact that who adopts the incentive mechanisms is endogenous.[6] We compare monthly performance across a large number of branches in two different divisions.

The firm initially used profit related pay in all divisions, so the comparison we make is between two different forms of incentive pay - an incentive based only on branch profits compared to one based on a large number of financial and non-financial performance indicators. In particular, we investigate the idea that it is not the

[3] See, inter alia, Cools and van Praag (2003), Fink (2004), Gosling (2003), Jensen (1991, 2001) and Ittner, Larcker and Meyer (1998). There is also an earlier literature, for example, Ridgway (1956) criticises the idea of multidimensional incentives schemes.
[4] Holmstrom and Milgrom (1991)
[5] Hoque and James (2000) survey 66 Austrailian manufacturing firms, Banker, Potter and Srinivsan (2000) look at 18 hotels, Malina and Selto (2001) consider multiple divisions of a large firm, Ittner, Larcker and Meyer (2003) conduct a case study in a large firm, Neely, Martinez and Kennerly (2004) consider 35 branches in one firm, Davis and Albright (2004) look at nine branches of a firm, Burgess et al (2004) evaluate a randomised trial incentive scheme in Job Centres in the UK.

incentive aspect of the Balanced Scorecard that has been successful, but the improved information that is transmitted to managers and workers in the large number of indicators that are monitored.

We find that overall the Balanced Scorecard changed behaviour, but this change did not lead to increased productivity - costs increased by at least as much as sales at the branch level. However, the impact varied across branches - branches with more experienced managers were able to effectively improve performance, while less experienced managers responded, but were not able to effectively improve performance. We use interviews with a large number of individual managers to show that this was due to their greater ability to interpret the large number of indicators and allocate effort within the branch.

The idea is that it is not only incentives that matter, but the ability of managers and workers to respond to them. When a large number of tasks are necessary to perform successfully (for example, in running a retail establishment), it is important that the manager can effectively decide where to put marginal effort. The Balanced Scorecard gives the manager additional information on past performance, but does not tell the manager where additional marginal effort will be most effective. It gives the same incentive on all margins. This has been on of the main features that has been criticised (Jensen, 2001). It requires additional ability (acquired through experience) for the manager to know where the greatest pay off will be.

The structure of this paper is as follows. In the next section we discuss the firm, describe the Balanced Scorecard, how it was implemented and the impact we would expect to find. Section 3 describes the data and our empirical approach. Section 4 presents the results. A final section summarises and concludes.

2 The setting

We start by describing the firm, the incentive scheme, how it was design and implemented, and then discuss what impact we expect to see on performance.

---

[6] Burgess et al (2004) are able to evaluate the introduction of a team based multi-task incentive scheme in Job Centres in the UK, where implementation was randomised.

2.1 The firm

The firm is a multinational distributor of heating and plumbing products. It has thousands of branches in over ten European and North American countries and employs around 50,000 people. In the UK there are four main divisions. We use data on two of these in this study. Both divisions have several sub-divisions, or brands. The organisational structure of the firm is shown in Figure 1.

[Figure 1 around here]

Traditionally branches have dealt primarily with one brand and have acted as relatively small trading units, employing between 2 and 32 staff (with a mean of 10).

We compare performance in Division 1 (where the Balanced Scorecard was introduced) with Division 2. Prior to August 2002 both firms used the same incentive scheme, which was based only on branch profits. Division 2 kept this scheme after August 2002. The two divisions are similar in terms of average sales and profits per branch. They differ in that they sell distinct product (which are both used in building). Another distinction is that the average value of and margin earned on the product sold in Division 2 is higher. However, the branches operate in similar economic conditions, for example, they employ people from the same labour markets, and experience similar demand shocks.

The key elements of profits are the sum of profits across individual branches and volume discounts that head office receives from suppliers. In this study we focus on branch profits. Profits of an individual branch are the revenue earned on the sale of each product minus the costs of sale minus central branch costs. The main elements of cost at the branch level are the cost of goods sold, labour costs, infrastructure (capital costs), distribution and transport costs. Other costs include general and administration costs, information technology, local marketing, advertising and other branch level administrative costs.

Effort of branch staff and branch managers can affect outcomes in the following ways:

- the price paid for a product can vary with each transaction (branch staff have discretion to negotiate individual prices);

- staff deal directly with customers and thus have influence on the quantity, type and range of any products sold;
- branch and regional managers influence quantity sold through setting price levels, marketing and setting special offers;
- staff, branch and regional managers' actions can affect hiring and firing costs, volume discounts and all branch level costs in various other ways.

A typical branch is managed by a branch manager,[7] an administrator who works in the office, two or three sales staff, a driver, and three or four people who work in the warehouse or stockyard. Further details are given in the Data Appendix. The role of the manager, among other things, is to decide on hiring decision in conjunction with head office, allocate staff to tasks, decide on special offers and local advertising.

2.2 The incentive scheme

Prior to August 2002 employees in both divisions received a bonus that was a function of branch level profits. The bonus was allocated to branch staff by the branch manager, at his discretion and in consultation with the regional manager.

Senior management became concerned that the profit-based bonus scheme was leading to dysfunctional behaviours, as is emphasised in the literature. For example, managers had incentives to adjust the timing of capital investment and to compete for business with other local branches.[8] They were also concerned that branch staff were not putting sufficient effort into activities that enhanced long-run profitability, such as maintaining customer loyalty, relationships with suppliers. Senior management decided to change the basis of the performance measurement and incentive scheme in the organisation, and decided to trial a new scheme, based on the Balanced Scorecard, in Division 1.

The new incentives scheme made three big changes (i) it used multiple targets rather than a single target, (ii) it was non-discretionary rather than discretionary, (iii) it

[7] The branch manager reports to a regional manager who reports to a brand operations director, who reports to a brand managing director who reports to a divisional managing director, who reports to the board.

provided more information to managers on a wide range of leading indicators, such as customer behaviour and efficiency, then was previously available.

The specific incentive scheme that the firm implemented works as follows. Incentives were paid to each employee on a six-monthly basis, but effectively earned on a monthly basis. For branch staff and branch managers the payment was based on the number of points the branch earned in the month times the value of a point.[9] Each branch is graded "green", "amber" or "red" on each of the 17 measures contained in the balanced scorecard. The number of points earned by the branch is given by,

$$\text{points} = (3 \times \text{green} + 1.5 \times \text{amber}) \times \text{number of people in branch.}$$

There are 17 measures in total so the maximum number of points per employee a branch can earn in a month is 51. In 2003 a point was worth £1 for branch staff (the value of a point is higher for managers) so the maximum bonus a branch worker could earn in a month was £51, or £612 a year.

The average salary of branch workers is around £12,000, so the maximum bonus represents 5.1% of their salary. For some workers it could represent substantially more, for example, a warehouse worker earns around £8,000, so it would represent 7.7% of their salary.

The total amount the firm allocated for the scheme was £1,836,000. This was set aside in a separate account to signal the firm's commitment to the scheme. The previous scheme, based on profits, cost about the same total amount. The firm estimates the cost of implementation at around £0.5m, including the direct and indirect costs such as management time.

As well as changing the incentive structure, an important aspect of the Balanced Scorecard was the provision of more detailed data on performance. Each branch manager received a detailed report each month on each of the 17 measures.

Why did the firm implement the Balanced Scorecard in Division 1 and not Division 2? This is an important question, as our strategy for identifying the impact of the

---

[8] See Hayes and Abernathy (1980) and Prendergast (1999).

Balanced Scorecard relies on performance being independent of this decision. The pragmatic reason is simply that the Divisions were run relatively independently and the Managing Director of Division 1 felt that the Balanced Scorecard would help his division improve its performance. At the outset the organisation considered whether it would subsequently implement the Balanced Scorecard in other divisions, but in essence decided to delay a decision and instead treat Division 1's implementation as a pilot study. Thus we argue that the introduction can be treated as independent of any expectations about the impact it would have in one Division over the other.

2.3 The design and implementation process

In evaluating the impact of the Balanced Scorecard we need to be careful to distinguish two questions - (i) did the new incentive scheme change behaviour, and (ii) did this behaviour lead to improved performance. The idea behind the Balanced Scorecard is that the firm needs to consider what are the key determinants of future profitability and build these into the incentive scheme.

The board devoted significant time to discussing what were the objectives for the division, what behaviours they wanted to encourage in the branch network, and how these behaviours might be reinforced through the choice of appropriate performance measures. At various stages during the process the directors consulted regional and branch managers. The firm invested considerably in education and training, as well as engagement with workers at all levels throughout this time.[10] In the end the firm adopted 17 measures, shown in Figure 2 and Tables 1.[11]

[Figure 2 around here]

[Table 1 around here]

---

[9] For regional managers it is based on the average of the points earned in branches under their control. For central office staff it is calculated based on a simplified version of the balanced scorecard containing only measures that the central office staff could affect.

[10] The discussions were about what the objectives were and who had control over them. Meetings were held in every branch and discussion was in depth and focussed, so for example, if a warehouse worker said "I can't influence that" others would say "Yes you can if ...".

[11] Personal development was included as a measure, but the data were not adequately collected, so it was not used. The employee satisfaction survey had a low response rate (e.g 32% in December 2002), but the firm used it anyway.

The key dates for introduction of the Balanced Scorecard that we focus on in this paper are:

August 2002	- official implementation; branch staff pay linked to performance; Branch Managers' pay linked to performance in December 2002
August 2003	- Regional Managers' and Directors' pay linked to performance
August 2004	- firm stopped sending out Balanced Scorecard reports to branches

2.4 Anticipated impact on performance

How do we expect the three key changes - (i) multiple measures, (ii) non-discretionary, and (iii) increased information to affect incentives and behaviour? Remember that, as mentioned above, there was no major change in the overall value of incentives, just in their form.

We assume that individuals seek to maximise their income, net of effort. Income consists of a base wage and an incentive payment, which is a function of performance. Performance is a function of the effort of all staff, managers and directors. Effort is costly. Individuals will exert effort up to the point where the marginal cost of effort equals the marginal benefit, in terms of the incentive payment.

Consider the incentives faced by branch staff first. An individual will equate the marginal cost of their effort with the (expected) marginal benefit. For branch staff (and Branch Managers) the incentive payment is a function of branch performance. Performance is a function of the effort made by all staff, managers and directors working within that the branch. This is true under either scheme. What changed is the way performance is measured and rewarded.

The main objectives of the firm were to encourage workers to put more effort into a range of activities that were previously not rewarded, but which the firm believes feed into long-term profits (long term being over several months to a year). An issue highlighted in the literature is the provision of incentives in a setting where workers have to perform multiple tasks. The tasks concerned are substitutes from the individual workers point of view (they each take time and an individual worker can spend time on one task or another), but from the point of view of the firm and value-

maximisation they are complements, in the sense that workers need to spend time on all of the tasks in order to maximise the value of the firm. Where workers undertake tasks that are substitutes for each other, and where the rewards are equal and independent (as is the case here), then workers will devote more time to those tasks they find easiest. If the measurement of some tasks is more precise, or more clearly understood – in the sense that it is easier to identify improvements in performance - then workers will devote more time to those tasks that are measured more precisely.[12] This is essentially the objection that Jensen (2001) and others have made to the Balanced Scorecard - it does not give clear guidance to workers on what is the most important task to perform, but rather allows them to put too much effort into non-profitable tasks (either due to mis-information about the payoff to individual tasks or due to shirking).

The Balance Scorecard scheme is complex and managers were provided with large amounts of information. In order respond effectively the manager needed to be able to assimilate this information and know how to act on it. A key issue is that the data as presented in the BSC are very aggregated, hence managers cannot take action on them. They need to develop local solutions [use the customer retention example to illustrate]. Our thesis is that experienced managers are either more likely to have seen in the past, or are better able now to develop, local solutions that allow them to take action to improve performance according to the BS.

This is where experience comes in. A more experienced manager will both be better able to interpret the large number of indicators, and better able to (e.g. more credible) motivate staff to put effort into the activities that matter for performance; with an inexperience managers workers can easily become overloaded with information and tasks, leading to underperformance.

[12] See Holmstrom and Milgrom (1991) and Burgess et al (2004).

3 Data and econometric method

3.1 Data

The main data is drawn from the monthly Profit and Loss (P&L) accounts of the firm. We have information on sales, gross profits (sales minus cost of goods sold), trading profits, labour costs, infrastructure, transport costs, general and administrative costs, information technology investment, local marketing and advertising expenditure and other costs for each branch. We have these data monthly from August 1999 to July 2005. We also have data on all employees in Division 1 including their job title, length of tenure in the firm and age, at August 2003.

We use the location (postcode) of each branch to match branches from the two divisions. There are a number of factors that affect sales, but are both exogenous (not affected by actions of the firm) and common across both divisions. These include the economic cycle, local economic and labour market conditions and other local factors. There are also factors that will affect demand for the two divisions differently, most notably weather. We use monthly data from twenty-six weather stations in the UK on the minimum temperature (in Celsius) and rain fall (in millimetres).

We also use data on total quarterly construction activity in a range of categories to capture variation in aggregate demand. These data come from the Construction Products Association. We aggregate separate information on aggregate quarterly output for new private sector housing, private and public sector remodelling and commercial and industrial development. The firm has given us the weight of each of these in demand for each of their brands, and we use this to construct a measure of quarterly aggregate demand for each brand.

Tables 2 and 3 show descriptive statistics of the main variables. Tables A.1 and A.2 show descriptive statistics on the number of employees, types of jobs within a branch, and the average experience and age of staff.

[Tables 2, 3, A.1 and A.2 around here]

We match each branch from Division 1 brand A to the geographically nearest branch from Division 2 brand C. The markets for the goods sold by the firm are very local -

the firm tells us that consumers are rarely willing to travel further than 30 kilometres. There are four Division 1 brand A branches which do not have a Division 2 brand C branch within 30 km, and we exclude these from our analysis. The average distance between matched branches is 4.5 kilometres.

3.2 Econometric Method

Our setting allows us to adopt a quasi-experimental design method[13] and use a combination of matching and a difference in difference estimator. As highlighted above, a major problem in the literature attempting to identify the impact of incentive schemes has been the fact that organisations choose whether and which incentive schemes to adopt - the adoption of the incentives scheme is endogenous.

We are able to use the fact that the firm implemented the Balanced Scorecard in one part of the firm, but not the other, to tackle this problem. As discussed above, the two divisions are affected similarly by many economic conditions such as local labour market conditions, but are affected differently by other factors, such as the weather.

We can write the determinants of sales for each type of branch, where we denote branches within Division 1 Brand A with subscript i and within Division 2 Brand C with subscript j

$$
\begin{aligned}
S_{irt} &= \alpha_i + \beta_1 L_{rt} + \beta_2 D_{1t} + \beta_3 \tau_t + \beta_4 W_{rt} + \gamma BSC_t + e_{irt} \\
S_{jrt} &= \alpha_j + \lambda_1 L_{rt} + \lambda_2 D_{2t} + \lambda_3 \tau_t + \lambda_4 W_{rt} + \varepsilon_{jrt}
\end{aligned} \tag{3}
$$

where S: sales, L: local economic factors such as labour markets, supply chain, etc., D: national demand, τ : seasonal (monthly) dummies, W: weather, BSC: indicators of the use of the balanced scorecard, e, ε : idiosyncratic shocks.

Table 4 presents coefficient estimates for the two equations in (3) separately. In column (1) the dependent variable is the monthly level of sales in division 1 brand A branches. Higher aggregate demand and a higher minimum temperature lead to higher sales, while more rain leads to less rain. In column (2) we include branch fixed effects. Columns (4) and (5) repeat the exercise for the branches in division 2 brand C that

were matched to division 1 brand A branches. Here we see that weather has the opposite effect on demand - lower temperatures lead to more demand.

[Table 4 around here]

The problem with looking only at changes over time in the level of sales (or profits) before and after the adoption of the Balanced Scorecard - as in columns (3) and (6) of Table 4 - is that we cannot control for many, in particular unobservable, factors. To control for these factors we match each branch from brand A to the geographically nearest branch from brand C. This is a combined matching and difference-in-difference estimator,[14] and takes the form

$$(4) \qquad \begin{aligned} (S_{irt} - S_{jrt}) &= (\alpha_i - \alpha_j) + \beta_2 (D_{1t} - D_{2t}) + (\beta_3 - \lambda_3)\tau_t \\ &\quad + (\beta_4 - \lambda_4) W_{rt} + \gamma BSC_t + (e_{irt} - \varepsilon_{jrt}) \end{aligned}$$

where we have assumed that $\beta_1 = \lambda_1$, i.e. local market conditions effect the two type of branches in the same way, so that they drop out of the difference equation, and that $\beta_2 = \lambda_2$, so that we can include the difference in aggregate demand for goods sold in the two branches.

We extend this specification to allow the main parameters of interest (on the Balanced Scorecard) to vary with managerial experience,

$\gamma = \gamma_0 + \gamma_1 Experience_i$.

4 Results

4.1 The overall impact of the balanced scorecard

We start in Table 5 by considering the combined matching and difference-in-difference estimator of the coefficients in equation (4) for sales. In the first column we include an indicator just for the two years that the incentive scheme based on the Balanced Scorecard was in place (August 2002 to July 2004) - so we are comparing

---

[13] See, inter alia, Cook and Campbell (1979) and Heckman, Ichimura and Todd (1997). For a recent survey see Blundell and Costa-Dias (2005).
[14] See the recent survey by Blundell and Costa-Dias (2005).

the difference between brand A and C branches in the level of sales during this period with the difference in the level before and after this period. This shows that sales increased by an average of £4,538 per month per branch in brand A branches during the period in which the Balanced Scorecard was implemented. In column (2) we split this period in half and see that the impact was similar across the two years. In column (3) we also include an indicator for the year after the firm put the incentive scheme on hold - so we are now comparing just to the two years before the Balanced Scorecard was implemented. We now see a larger impact around (£8,305) which diminishes slightly over time. In column (4) we consider whether there was an effect of the initial discussion period (i.e. a Hawthorne effect in anticipation of the actual implementation), but find no evidence of this.

[Table 5 around here]

These results suggest that the Balanced Scorecard had an impact on increasing sales, but what about profits? We next implement the combined matching and difference-in-difference estimator using measures that also reflect costs. Specifically, we use sales, gross profits, trading profits, labour costs, infrastructure expenditure, transport costs, general and administrative costs, information technology expenditure, local marketing and advertising expenditure and other costs.

[Table 6 around here]

Table 6 shows these results, where we have assumed a homogeneous impact of the Balanced Scorecard across all branches. All regressions include controls for aggregated demand, the weather, month and branch effects. We consider three time periods - the two years during which the Balanced Scorecard was implemented and the year after. The omitted period is the two years before it was implemented.

Column (1) of Table 6 repeats column (3) of Table 5. In column (2) we see that there was also a corresponding increase in gross profits (the difference between sales and gross profits is the cost of goods sold). In column (3) we see that this did not feed through into an increase in trading profits, and in fact led to a fall in profits in the later period of on average -£1,018. The difference between gross and trading profits is the costs considered in columns (5)-(10). The fall in profits arose largely because transport costs (column 6) and labour costs increased (column 4), and to a lesser

extent because infrastructure (column 5) and general and administrative costs (column 7) increased. Other and IT costs actually fell.

4.2 Did the balanced scorecard target the right tasks?

The response of branch managers to the Balanced Scorecard differed markedly. One important question is whether branches that improved on the non-financial measures subsequently experienced improved performance - i.e. did the Balanced Scorecard correctly identify the key drivers of performance? We use the data collected under the Balanced Scorecard to distinguish between branches that successfully put effort into improving on the non-financial measures from those which did not do well on these. For each branch we calculate the proportion of the total non-financial points that were available they earned. This ranges from 16% to 74% and has a median value of 52%.

In Table 7 we split the sample into those who are below and above this median level of points on the non-financial measures. In the top half of Table 7 we consider the change in outcomes of those branches that either did not try, or were not able, to perform well on the non-financial measures, and in the bottom half those branches that did well on the non-financial measures.

[Table 7 around here]

Here we see large differences between the two groups. In the top panel branches that did poorly on non-financial measures did not experience any significant growth in sales, but they did experience some increases in costs, leading to reduced trading profits. In contrast, those branches that did well on the non-financial measures saw an increase in sales, gross profits and trading profits. While costs did increase in these branches, sales increased by more, resulting in higher profits. However, this may simply be showing that good branches (or good branch managers) do well on both financial and non-financial indicators.

To investigate this we make a further comparison. Proponents of the Balanced Scorecard argue that giving branches an incentive to put effort into a broader range of factors that feed into long run performance will lead to better performance in the long run. Opponents of the Balanced Scorecard argue that giving managers and workers so

many different incentive will lead to worse performance as individuals lose focus and put too much effort into the easiest tasks.

We consider branch financial performance during three time periods - prior to the Balance Scorecard (August 1999 - July 2002), during the Balanced Scorecard (August 2002 - July 2004) and after (August 2004 - July 2005).

[Figure 3 and 3b around here]

In Figure 3 we consider just the period during the Balanced Scorecard and compare branches performance on financial and non-financial measures (as measured by the share of potential points earned on the Balanced Scorecard). There is a fairly clear positive correlation between the two measures (correlation coefficient of 0.612, with a p-value of 0.000 so significant at the 1% level). We do not have the Balanced Scorecard measures for the other two time periods, instead we use one of the measures - trading profits over sales - which we do observe over the entire time period. This has a median of 8.3% and varies from positive to negative 87% (we exclude a small number of observations where it is greater than 1). The picture in Figure 3b yields a similar impression (and the correlation coefficient is 0.662 (p-value=0.000)).

In Figure 4 we split by financial performance in the period before the Balanced Scorecard was introduced. The relationship is stronger in those branches which previously performed badly than in those that did well (correlation coefficient (p-value) of 0.727 (0.000) and 0.346 (0.002) respectively).

[Figure 4 around here]

How did performance on non-financial measures affect future financial performance? In Figure 5 we see that those branches that did better on the non-financial measures subsequently did better in financial terms as well. In Figure 6 we show this same picture separately for branches that did poorly in financial terms in the period before the Balanced Scorecard and those that did well, and we see a positive effect in both cases.

[Figure 5 and 6 around here]

These pictures are suggestive, but we haven't controlled for any of the local economic time varying factors we were concerned with above. To do this we return to the

regressions of the form of equation (4) and we compare performance in four groups of branches - (1) those who did badly on financial measures prior to implementation of the Balanced Scorecard, and who subsequently also did badly on non-financial measures during the implementation of the Balanced Scorecard, (2) those who did badly on financial measures prior to implementation and well on non-financial during implementation, (3) those who did well on financial measures prior but badly on non-financial during, and (4) those who did well on financial prior and well on non-financial during.

The top panel of Table 8a shows the estimates for group (1), the bottom panel for group (2), the top panel of Table 8b shows the estimates for group (3) and the bottom panel for group (4).

[Tables 8a and 8b around here]

What we see is a pattern that suggests that branches that do well on the non-financial measures also do well on financial measures. In particular, those branches that previously did badly on financial measures (bottom panel of Table 8a) had substantially higher sales during implementation period and less so after. While costs increased, sales increased by more, so that trading profits also increased.

Consider the branches that did well financially prior to implementation. Those that did badly on non-financial measures (top panel of Table 8b) did badly in financial terms after implementation, while those that did well on non-financial measures (bottom panel of Table 8b) did well in financial terms after implementation.

4.3 The importance of experience

Table 9 repeats the results in Table 6 but allowing the impact of the Balanced Scorecard to vary with the average years of experience of all staff. We see that experience matters. A branch with staff with the average years of experience (6.6 years) will have experienced a £400 higher level of trading profits after the Balanced Scorecard was introduced than a branch with all new staff.

In Table 10 we show that it is the years of experience of senior staff that matters most. In results not shown we show that it is experience, not age, that is important.

One thing that we have not done yet is look at whether inexperienced managers did well on non-financial measures.

5 Summary and conclusions

Our results suggest that the balanced scorecard had some impact, and that this impact varied significantly by branch. Sales increased on average across all branches, but costs increased by at least as much, so that while gross profits did increase, trading profits did not increase. Indeed at the aggregate level trading profits had decreased in the division that implemented the balanced scorecard by the end of the period of study. There is significant variations in the impact. When we separate those branches that perform well on non-financial measures from those that perform poorly on non-financial performance measures we find that the first group – those that perform well on non-financial measures – experience statistically significant increases in sales, gross profits and trading profits. One potential explanation of this finding is that branches that perform well on non-financial measures and financial measures are simply well managed. First note that we are allowing for each branch to have a different average level of performance (we include branch fixed effects). To consider this further we split the sample into branches that perform relatively well and those that perform relatively poorly on financial measures in the pre-implementation period. We then explore whether there are differences in non-financial and financial performance in the implementation period. We find that, regardless of prior performance, branches that perform well on non-financial measures also perform well on financial measures. This finding is particularly important as it suggests that the balanced scorecard, when implemented correctly and adopted by the branches, has a positive impact on branch performance in terms of sales, gross profit and net profit.

We then show that years of experience is an important factor in explaining these differences in performance. More experience managers were able to improve performance. We interpret this as suggesting that the information content of the Balanced Scorecard is what is important, not the incentive per se.

References

Baker, G (1992) “Incentive Contracts and Perforamnce Measurement” *Journal of Political Economy*, 100, 598-614

Baker, George P., Michael C. Jensen and Kevin J. Murphy, 'Compensation and Incentives: Practice vs. Theory,' Journal of Finance (July, 1988)

Banker, R. Potter, G and Srinivsan, D. (2000) “An Empirical Investigation of an Incentive Plan That Includes Nonfinancial Performance Measures”, *Accounting Review*, 75(1), 65-92.

Blundell, R and Costa-Dias, M. (2000) “Evaluation Methods for Non-Experimental Data” *Fiscal Studies* vol. 21, no. 4, pp. 427–468

Blundell and Costa-Dias (2005) “Alternative Approaches to Evaluation in Empirical Microeconomics” UCL mimeo, http://www.econ.ucl.ac.uk/courses/mect2/rbmcd-13-03-2006.pdf

Burgess, S., Propper, C., Ratto, M. and Tominey, E. (2004) “Incentives in the Public Sector: Evidence from a Government Agency” *CMPO Working Paper* No. 04/103

Cook, T.D. and Campbell, D.T. (1979) “Quasi-Experimentation: Design and Analysis Issues for Field Settings”, Houghton Mifflin.

Cools, Kees and Mirjam van Praag (2003), “The Value Relevance of a Single-Valued Corporate Target: An Explorative Empirical Analysis.” *Tinbergen Discussion Paper* TI2003-049/3, http://www.tinbergen.nl/discussionpapers/03049.pdf

Davis, S. and Albright, T. (2004), “An investigation of the effect of Balanced Scorecard implementation on financial performance”, *Management Accounting Research*, 15, 135-153

Fink, R “New Carrots, Old Yardsticks?” CFO Magazine June 2004

Gosling, J. (2003), A Worldly Mindset: Satisfying the needs of stakeholders. Creating an Agenda for European Leadership - Conference Proceedings, Bled 10-11 June 2003 http://www.leadership-studies.com/research/viewabs.asp?id=42

Hayes, R.H. and Abernathy, W.J. “Managing Our Way to Economic Decline”, Harvard Business Review, July-August, 1980, pp. 67-77.

Heckman, J, Ichimura, H. and Todd, P. (1997), ‘Matching as an econometric evaluation estimator’, *Review of Economic Studies*, vol. 64, pp. 605–54.

Hoque, Z. and James, W. (2000), “Linking Balanced Scorecard Measures to Size and Market Factors: Impact on Organisational Performance”, *Management Accounting Research*, 12, 1-18.

Holmstrom, B. (1982) “Moral Hazard in Teams” *The Bell Journal of Economics*, Vol. 13, No. 2 (Autumn 1982), 324-340

Holmstrom, B. and Milgrom, P. (1991) " Multitask Principal Agent Analyses: Incentive Contracts, Asset Ownership and Job Design" *Journal of Law, Economcs and Organisation*, 7, 24-52

Ittner, Christopher D., David F. Larcker, and Marshall W. Meyer (2003) "Subjectivity and the Weighting of Performance Measures: Evidence From a Balanced Scorecard", *Accounting Review*, 78 (3), 725-758

Jensen, M. "Value Maximization, Stakeholder Theory, and the Corporate Objective Function" *Journal of Applied Corporate Finance, Fall 2001* http://papers.ssrn.com/abstract=220671

Kahneman, D. & Tversky, A. (1979) Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47, 263-291.

Kandel, E and Lazear, E (1994) "Peer Pressure and Partnerships" *Journal of Political Economy*, 100 (4), 801-817

Kaplan, R and D. Norton (1992) "The Balanced Scorecard - Measures that Drive Performance" *Harvard Business Review* January-February, 71-79.

Kerr, S. (1995) "On the Folly of Rewarding A While Hoping for B", Academy of Management Executive, 9 (1), 7-14.

Knez, M and Simester, D (2001) "Firm-wide incentives and mutual monitoring at Continental Airlines" *Journal of Labour Economics*, 19 (4), 743-772

Lazear, E (2001) "Performance pay and productivity", *American Economic Review*, 90: 5, 1346-1361

Maisel, L. (2001) "Performance measurement practices survey results", AICPA, USA.

Malina, M.A and Selto, F.H. (2001) "Communicating and Controlling Strategy: An Empirical Study of the Effectiveness of the Balanced Scorecard", *Journal of Management Accounting Research*, 13, 47-90.

Marr, B., Neely, A., Franco, M., Wilcox, M., Adams, C. and Manson, S. (2004), "Business Performance Measurement - What Is the State of the Art?" Proceedings of the Fourth International Performance Measurement Association Conference, Edinburgh, UK.

Neely, A.D.; Kennerley, M. and Martinez, V. "Does the Balanced Scorecard Work: An Emprical Investigation", Proceedings of the Fourth International Performance Measurement Association Conference, Edinburgh, UK.

Prendergast, C (1999) "The Provision of Incentives in Firms" *Journal of Economic Literature*, Vol. XXXVII (March 1999), 7-63

Ridgway, V.F. (1956), "Dysfunctional Consequences of Performance Measurements", *Administrative Science Quarterly*, 1(2), pp. 240-247.

Rigby, D. (2001), "Management Tools and Techniques: A Survey", California Management Review 43(2), 139-160.

Speckbacher, G., Bischof, J. and Pfeiffer, T. (2003), "A Descriptive Analysis on the Implementation of Balanced Scorecards in German-speaking Countries", *Management Accounting Research*, 14, pp. 361-387.

Tversky, A. & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference Dependent Model. *Quarterly Journal of Economics* 106, 1039-1061.

Table 1: Balanced Scorecard measures

Financial measures	
Return on Capital Employed	*[PBIT/(Debtors + Stock + Fixed Assets)] x 100*
Growth in Profit	*[(Contribution This Year To Date-Contribution Last Year To Date) / Contribution Last Year To Date] x 100*
PBIT as a % of Sales	*(Contribution YTD/ Sales YTD) x 100*
Positive Cash Flow	*[[Contribution – (+/- Stock Movement £'s) + (+/- Debtors Movement £'s) = Basic Cash Flow] / Total Sales] x 100*
Sales Growth	*[(Sales PWD This Year To Date – Sales PWD Last Year To Date) / Sales PWD Last Year To Date] x 100*
Customer measures	
Customer Satisfaction[b]	*Score achieved via an external survey*
Customer Retention	*[(No. of Customers retained in rolling 12 months to current month – No. of Customer retained in rolling 12 months to last month) / No. of Customers retained in rolling 12 months to last month] x 100*
Sales Mix	*[(Sales of Selected SPGs This Year to Date – Sales of Selected SPGs Last Year to Date) / Sales of Selected LLSPGs Last Year to Date] x 100*
Availability of Stock Range	*(Sum of Number of Days where Stock Ins for your MBR are equal to or greater than 90% / Number of Trading Days) x 100*
Internal measures	
Operational Efficiency	*Stock/Debtors/Labour/Transport – Yes/No against individual targets: Stock 40 days, Debtors 0.5% against Sales, Labour 10% against Ex-Stock Sales, Transport 8% against Delivered Sales, where 25% is awarded per point*
Operational Standards	*(Score from Operational Standards Check List / Total possible score from Operational Standards) x 100*
Inter-company Co-operation	*[(Number of Customers trading with foreign Branches This YTD – Number of Customers trading with foreign Branches Last YTD) / Number of Customers trading with foreign Branches LYTD] x 100*
People measures	
Staff retention	*(Number of voluntary leavers on a rolling 12 month basis / Average head count in rolling 12 months) x 100*
Employee satisfaction	*(The number of people who indicate they are satisfied at work / average number of employees over the period) x 100*
Communication	*(Number of people who feel they have been made aware of businesses activities / Average number of employees over the period) x 100 (By Region)*
Supplier measures	
Spend with Approved Suppliers	*(Purchases from preferred Suppliers This Year To Date / Total purchases from Suppliers This Year To Date) x 100*

Table 2: Mean (in £,000) and standard deviation for Division 1 branches and matched Division 2 branches

	Division 1 Brand A	Division 2 Brand C (matched only)
Branches	*156*	*121*
Observations	*11076*	*11076*
Sales	152.5	166.4
	(87.8)	(103.4)
Gross profits	45.7	39.9
	(28.3)	(19.1)
Trading profits	15.3	22.8
	(19.4)	(15.6)
Labour costs	14.9	8.8
	(7.7)	(4.2)
Infrastructure	7.6	4.5
	(4.7)	(2.4)
Transport costs	4.2	1.1
	(2.6)	(1.1)
General and administration	1.4	0.8
	(1.3)	(0.6)
Other	0.9	0.7
	(0.7)	(0.5)
IT	0.48	0.49
	(0.22)	(0.13)
Marketing and advertising	0.15	0.02
	(0.34)	(0.15)

Notes: Values are monthly in nominal £,000 over the period August 1999 to July 2005.

Table 3: Descriptive statistics, weather and demand

	Mean (standard deviation)
Minimum temperature (in Celsius), measured at 26 points throughout the UK	7.28 (4.17)
Rain fall (in mm)	65.7 (41.0)
National quarterly demand for activities using Brand A products (in £m)	2628 (200)
National quarterly demand for activities using Brand products (in £m)	2588 (168)

Notes: Data on weather if from http://www.met-office.gov.uk/climate/uk/stationdata/

Table 4: Descriptive Statistics, level of sales

	(1)	(2)	(3)	(4)	(5)	(6)
Dep Var: sales	Division 1 Brand A			Division 2 Brand C (matched only)		
Aug 2002 - Jul 2003			4958			8276
			(1571)***			(1705)***
Aug 2003 - Jul 2004			6986			14239
			(2535)***			(3483)***
Aug 2004 - Jul 2005			10970			20639
			(2645)***			(4258)***
Demand - Div 1	38	40	20			
	(4)***	(2)***	(6)***			
Demand - Div 2				35	31	-14
				(7)***	(3)***	(10)
Min temp in C	3295	778	738	-6903	-604	-601
	(543)***	(294)***	(294)**	(766)***	(435)	(437)
Rain fall in mm	-107	-93	-89	-18	-11	-0
	(22)***	(8)***	(8)***	(30)	(12)	(12)
Observations	11,076	11,076	11,076	8,591	8,591	8,591
Branches	156	156	156	121	121	121
R-squared	0.06	0.34	0.35	0.03	0.15	0.15
Branch fixed effects		yes	yes		yes	yes

Note: Robust standard errors in parentheses. Period covered is August 1999 to July 2005. Constant and month dummies included in all regressions.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5: Matched/Diff-in-Diff results on sales, different time periods

Dep var: difference in Sales	(1)	(2)	(3)	(4)
Nov 2001 - Jul 2002				404
				(1727)
Aug 2002 - Jul 2004	4538			
	(1433)***			
Aug 2002 - Jul 2003		3752	8305	8671
		(1670)**	(1815)***	(2520)***
Aug 2003 - Jul 2004		4992	8903	9216
		(1557)***	(1694)***	(2299)***
Aug 2004 - Jul 2005			6916	7124
			(1525)***	(1843)***
Demand	-25	-23	-50	-52
	(12)**	(12)*	(13)***	(17)***
Min temp in C	1686	1694	1354	1348
	(446)***	(446)***	(451)***	(453)***
Rain fall in mm	-90	-89	-80	-79
	(12)***	(12)***	(12)***	(12)***
R-squared	0.15	0.15	0.15	0.15

Note: Robust standard errors in parentheses. 11,076 observations on 156 branches over the period August 1999 to July 2005. Constant, month and branch dummies included in all regressions. Dependent variable is the difference in sales in a division 1 branch and the geographically nearest division 2 branch. Demand is the difference in national demand for division 1 products and demand for division 2 products.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Matched/Diff-in-Diff results on all variables

Dep var: difference in	(1) Sales	(2) Gross profits	(3) Trading profits	(4) Labour costs	(5) Infrastructure	(6) Transport costs	(7) General and administration	(8) Other	(9) IT	(10) Marketing and advertising
Aug 2002 - Jul 2003	8305 (1815)***	1512 (548)***	771 (664)	192 (114)*	91 (120)	704 (73)***	28 (54)	-12 (32)	-59 (8)***	12 (14)
Aug 2003 - Jul 2004	8903 (1694)***	2565 (560)***	995 (619)	675 (104)***	183 (122)	1014 (58)***	115 (50)**	-48 (30)	-78 (8)***	19 (22)
Aug 2004 - Jul 2005	6916 (1525)***	1835 (450)***	-1018 (517)**	933 (93)***	667 (126)***	1154 (50)***	88 (52)*	-72 (25)***	-91 (8)***	-4 (15)
Demand	-49 (12)***	-4.08 (3.58)	-7.61 (4.17)*	-2.161 (0.785)***	1.732 (0.809)**	-0.594 (0.471)	0.200 (0.382)	-0.577 (0.218)***	-0.333 (0.067)***	-0.163 (0.108)
Min temp in C	1354 (451)***	527 (123)***	624 (148)***	-6.60 (27.95)	-88.9 (33.2)***	-40.6 (19.9)**	34.9 (11.1)***	-4.593 (8.180)	-5.457 (2.301)**	0.372 (2.821)
Rain fall in mm	-79 (12)***	-19.0 (3.3)***	-22.8 (3.9)***	0.144 (0.774)	-0.957 (0.807)	0.167 (0.484)	-0.724 (0.333)**	0.840 (0.255)***	0.017 (0.066)	0.100 (0.095)
R-squared	0.15	0.20	0.13	0.05	0.01	0.08	0.01	0.01	0.05	0.00

Note: Robust standard errors in parentheses. 11,076 observations on 156 branches over the period August 1999 to July 2005. Constant, month and branch dummies included in all regressions. Dependent variable is the difference in variable indicated in a division 1 brand A branch and the geographically nearest division 2 brand C branch. Demand is the difference in national demand for division 1 products and demand for division 2 products.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Matched/Diff-in-Diff results on all variables - comparison of branches on non-financial balanced scorecard measures

Dep var: difference in	(1) Sales	(2) Gross profits	(3) Trading profits	(4) Labour costs	(5) Infrastructure	(6) Transport costs	(7) General and administration	(8) Other	(9) IT	(10) Marketing and advertising
5,538 observations on 78 branches scoring below median on non-financial balanced scorecard measures										
Aug 2002 - Jul 2003	1947 (2706)	-450 (772)	-838 (955)	-91 (169)	-64 (177)	667 (102)***	18 (77)	-2 (47)	-54 (11)***	17 (20)
Aug 2003 - Jul 2004	3535 (2508)	-40 (790)	-979 (871)	528 (160)***	-212 (173)	970 (94)***	98 (68)	-5 (44)	-79 (11)***	40 (27)
Aug 2004 - Jul 2005	3 (2328)	-353 (690)	-2725 (785)***	681 (147)***	490 (216)**	1121 (80)***	147 (89)*	-62 (37)*	-83 (13)***	-13 (20)
5,538 observations on 78 branches scoring above median on non-financial balanced scorecard measures										
Aug 2002 - Jul 2003	14676 (2405)***	3500 (767)***	2385 (911)***	475 (152)***	251 (162)	740 (103)***	37 (77)	-21 (45)	-65 (13)***	8 (21)
Aug 2003 - Jul 2004	14266 (2263)***	5188 (771)***	2980 (860)***	822 (133)***	580 (171)***	1057 (67)***	130 (72)*	-91 (39)**	-78 (12)***	-0 (36)
Aug 2004 - Jul 2005	13583 (1960)***	3985 (570)***	680 (663)	1182 (113)***	838 (126)***	1186 (61)***	31 (53)	-84 (33)**	-99 (8)***	6 (22)

Note: Robust standard errors in parentheses. Period August 1999 to July 2005. Constant, month and branch dummies, demand, min temperature and monthly rainfall included in all regressions. Dependent variable is the difference in variable indicated in a division 1 brand A branch and the geographically nearest division 2 brand C branch. Demand is the difference in national demand for division 1 products and demand for division 2 products.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 8a: Matched/Diff-in-Diff results on all variables - branches scoring below median on financial measures in pre-Aug2002 period

Dep var: difference in	(1) Sales	(2) Gross profits	(3) Trading profits	(4) Labour costs	(5) Infrastructure	(6) Transport costs	(7) General and administration	(8) Other	(9) IT	(10) Marketing and advertising
3408 observations on 48 branches below median on non-financial balanced scorecard measures										
Aug 2002 - Jul 2003	1665 (3533)	-564 (965)	-418 (1227)	-385 (213)*	-304 (211)	552 (127)***	20 (87)	-25 (57)	-49 (14)***	14 (27)
Aug 2003 - Jul 2004	678 (3301)	-231 (956)	-147 (1027)	-18 (207)	-550 (206)***	828 (124)***	53 (80)	-35 (57)	-77 (14)***	47 (38)
Aug 2004 - Jul 2005	-7080 (3131)**	-1369 (929)	-2269 (1023)**	-151 (195)	105 (177)	940 (96)***	96 (134)	-164 (47)***	-73 (20)***	-8 (26)
2,130 observations on 30 branches above median on non-financial balanced scorecard measures										
Aug 2002 - Jul 2003	19833 (4035)***	4001 (972)***	2605 (1352)*	490 (266)*	183 (296)	730 (164)***	19 (108)	118 (73)	-48 (17)***	4 (23)
Aug 2003 - Jul 2004	20045 (3951)***	5155 (984)***	2626 (1225)**	981 (249)***	296 (213)	1077 (110)***	75 (123)	25 (64)	-47 (19)**	-22 (26)
Aug 2004 - Jul 2005	16018 (3365)***	3238 (708)***	-546 (986)	1065 (195)***	837 (230)***	1308 (90)***	-61 (92)	-69 (53)	-85 (13)***	-14 (14)

Note: Robust standard errors in parentheses. Period August 1999 to July 2005. Constant, month and branch dummies, demand, min temperature and monthly rainfall included in all regressions. Dependent variable is the difference in variable indicated in a division 1 brand A branch and the geographically nearest division 2 brand C branch. Demand is the difference in national demand for division 1 products and demand for division 2 products.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 8b: Matched/Diff-in-Diff results on all variables - branches scoring above median on financial measures in pre-Aug2002 period

Dep var: difference in	(1) Sales	(2) Gross profits	(3) Trading profits	(4) Labour costs	(5) Infrastructure	(6) Transport costs	(7) General and administration	(8) Other	(9) IT	(10) Marketing and advertising
2,130 observations on 30 branches below median on non-financial balanced scorecard measures										
Aug 2002 - Jul 2003	2275 (4124)	-216 (1197)	-1428 (1446)	388 (270)	287 (311)	853 (168)***	17 (145)	34 (80)	-61 (19)***	23 (29)
Aug 2003 - Jul 2004	7908 (3779)**	283 (1298)	-2291 (1506)	1417 (244)***	316 (307)	1196 (143)***	171 (123)	47 (70)	-83 (19)***	29 (34)
Aug 2004 - Jul 2005	10617 (3337)***	1148 (982)	-3601 (1198)***	2016 (211)***	1128 (482)**	1406 (140)***	231 (90)**	111 (62)*	-99 (13)***	-18 (28)
3,408 observations on 48 branches above median on non-financial balanced scorecard measures										
Aug 2002 - Jul 2003	11454 (2929)***	3195 (1079)***	2265 (1212)*	466 (181)**	291 (189)	746 (133)***	48 (105)	-107 (57)*	-75 (17)***	11 (31)
Aug 2003 - Jul 2004	10631 (2678)***	5227 (1078)***	3227 (1158)***	722 (149)***	756 (245)***	1044 (85)***	166 (89)*	-165 (50)***	-97 (16)***	14 (56)
Aug 2004 - Jul 2005	11981 (2371)***	4442 (810)***	1444 (880)	1254 (137)***	834 (147)***	1110 (82)***	89 (64)	-94 (42)**	-109 (11)***	18 (35)

Note: Robust standard errors in parentheses. 3,672 observations on 51 branches over the period August 1999 to July 2005. Constant, month and branch dummies included in all regressions. Dependent variable is the difference in variable indicated in a division 1 brand A branch and the geographically nearest division 2 brand C branch. Demand is the difference in national demand for division 1 products and demand for division 2 products.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 9: Years of service

	(1) Difference in sales with nearest plumb	(2) Difference in gross profit with nearest plumb	(3) Difference in trading profit with nearest plumb	(4) Difference in labour cost with nearest plumb	(5) Difference in infrastructure with nearest plumb	(6) Difference in transport costs with nearest plumb
[interaction is evaluated at the mean level of service = 6.6 years]						
service x Aug02Jul03	1724	735	401	183	80	52
	(363)***	(107)***	(119)***	(21)***	(31)**	(14)***
service x Aug03Jul04	3181	1250	659	230	168	76
	(353)***	(117)***	(122)***	(25)***	(26)***	(12)***
service x Aug04Jul05	3857	1023	459	248	122	58
	(420)***	(121)***	(130)***	(29)***	(31)***	(14)***
Aug 2002 - Jul 2003	8320.609	1514.562	771.632	192.991	91.801	703.843
	(1816.261)***	(548.967)***	(664.487)	(113.290)*	(119.922)	(72.523)***
Aug 2003 - Jul 2004	8939.784	2573.438	998.047	677.274	184.193	1014.065
	(1682.480)***	(552.524)***	(616.013)	(103.234)***	(121.489)	(57.878)***
Aug 2004 - Jul 2005	6938.202	1838.946	-1016.917	934.714	666.945	1154.239
	(1512.552)***	(448.213)***	(516.483)**	(91.864)***	(125.289)***	(50.343)***

Note: Robust standard errors in parentheses. 11,076 observations on 156 branches over the period August 1999 to July 2005. Constant, month and branch dummies, demand, min temperature and monthly rainfall included in all regressions. Dependent variable is the difference in variable indicated in a division 1 brand A branch and the geographically nearest division 2 brand C branch. Demand is the difference in national demand for division 1 products and demand for division 2 products.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: Years of service of senior staff

	(1) Difference in sales with nearest plumb	(2) Difference in gross profit with nearest plumb	(3) Difference in trading profit with nearest plumb	(4) Difference in labour cost with nearest plumb	(5) Difference in infrastructure with nearest plumb	(6) Difference in transport costs with nearest plumb
[interaction is evaluated at the mean level of service = 6.6 years]						
service x Aug02Jul03	798	377	110	94	105	33
	(433)*	(124)***	(143)	(24)***	(36)***	(16)**
service x Aug03Jul04	2458	724	192	127	184	49
	(411)***	(129)***	(140)	(28)***	(32)***	(14)***
service x Aug04Jul05	3026	535	48	111	148	32
	(473)***	(140)***	(157)	(30)***	(41)***	(16)**
[interaction is evaluated at the mean level of service = 11.3 years]						
senior service x Aug02Jul03	589.236	228.051	185.144	56.479	-16.040	11.958
	(146.320)***	(44.486)***	(50.396)***	(8.919)***	(8.016)**	(6.797)*
senior service x Aug03Jul04	460.188	334.536	297.151	65.685	-10.119	17.220
	(148.061)***	(52.536)***	(57.476)***	(8.965)***	(8.848)	(5.151)***
senior service x Aug04Jul05	528.470	310.504	261.701	87.282	-16.621	16.782
	(168.709)***	(52.602)***	(57.636)***	(10.062)***	(10.726)	(5.127)***
Aug 2002 - Jul 2003	8317.042	1513.671	770.852	193.107	91.837	703.808
	(1818.743)***	(549.294)***	(664.568)	(113.358)*	(119.905)	(72.506)***
Aug 2003 - Jul 2004	8932.893	2569.558	994.651	676.582	184.339	1013.869
	(1684.153)***	(549.788)***	(613.455)	(103.234)***	(121.502)	(57.857)***
Aug 2004 - Jul 2005	6932.842	1837.945	-1017.793	935.032	666.987	1154.207
	(1512.138)***	(447.028)***	(515.718)**	(91.249)***	(125.417)***	(50.314)***

Note: Robust standard errors in parentheses. 11,076 observations on 156 branches over the period August 1999 to July 2005. Constant, month and branch dummies, demand, min temperature and monthly rainfall included in all regressions. Dependent variable is the difference in variable indicated in a division 1 brand A branch and the geographically nearest division 2 brand C branch. Demand is the difference in national demand for division 1 products and demand for division 2 products.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A.1

Job	Mean number per branch	Std. Dev.	Min	Max	Mean experience	Mean age
Manager	0.883	0.359	0	2	12.971	42.088
					(9.876)	(8.570)
Office	1.090	0.808	0	3	8.821	41.240
					(9.062)	(11.825)
Sales_job	2.548	1.667	0	9	7.088	38.532
					(8.673)	(12.804)
Driver	1.651	1.209	0	6	6.520	46.443
					(7.747)	(10.576)
Warehouse and stockyard	3.670	2.107	0	14	5.771	38.970
					(7.666)	(13.330)
Other_job	0.354	0.543	0	2	7.579	47.572
					(7.476)	(13.564)
Total	*10.2*	*4.631*	*3*	*30*	*7.234*	*40.883*
					(8.520)	*(12.637)*

Table A.2

Job	Specific job titles included
Branch Manager	Branch Manager, Branch Manager (Designate), Heavyside Manager, Senior Branch Manager
Office	Administration Assistant, Administration Supervisor, Administrator, Assistant Branch Manager, Assistant Depot Manager, Branch Supervisor , Deputy Manager, Estimator, General Clerk, Typist, Office Manager, Pt General Clerk, Secretary, Temp General Clerk
Sales	Contracts/Sales Administrator, Credit Controller, Goods Inwards Assistant, Inside Sales, Pt Sales Clerk, Pt Sales Counter Supervisor, Pt Showroom Sales Asst, Sales Clerk, Sales Counter Assistant, Sales Counter Supervisor, Sales Negotiator, Sales Representative, Sales Supervisor, Showroom Sales Assistant, Showroom Sales Supervisor, Showroom Supervisor, Stock Controller
Driver	Driver, Driver LGV, Glass Cutter/Driver, Warehouse Assistant/Driver
Warehouse/Yard	Depot Manager, Depot Manager Designate, Drainage Supervisor, Foreman, Heavy Supervisor, Hire Assistant, Ironmongery Supervisor, Lightside Supervisor, Logistics Manager, Maintenance Supervisor, Office Supervisor, Operations Manager, Plumbing Supervisor, Product Supervisor, Pt Stores/Warehouse Assistant, Saturday Assistant, Shop/Yard Assistant, Stores/Warehouse Assistant, Temp Yard Assistant, Timber Supervisor, Transport Supervisor, Warehouse Supervisor, Yard Assistant, Yard Assistant/Driver, Yard Manager, Yard Supervisor, Yard/Warehouse Assistant, Yard/Warehouse Supervisor
Other	Fixer, Machine Operator, Mill Operative, Pt Cleaner, Timber Machinist, Trainee B2

Figure 1: Firm Structure

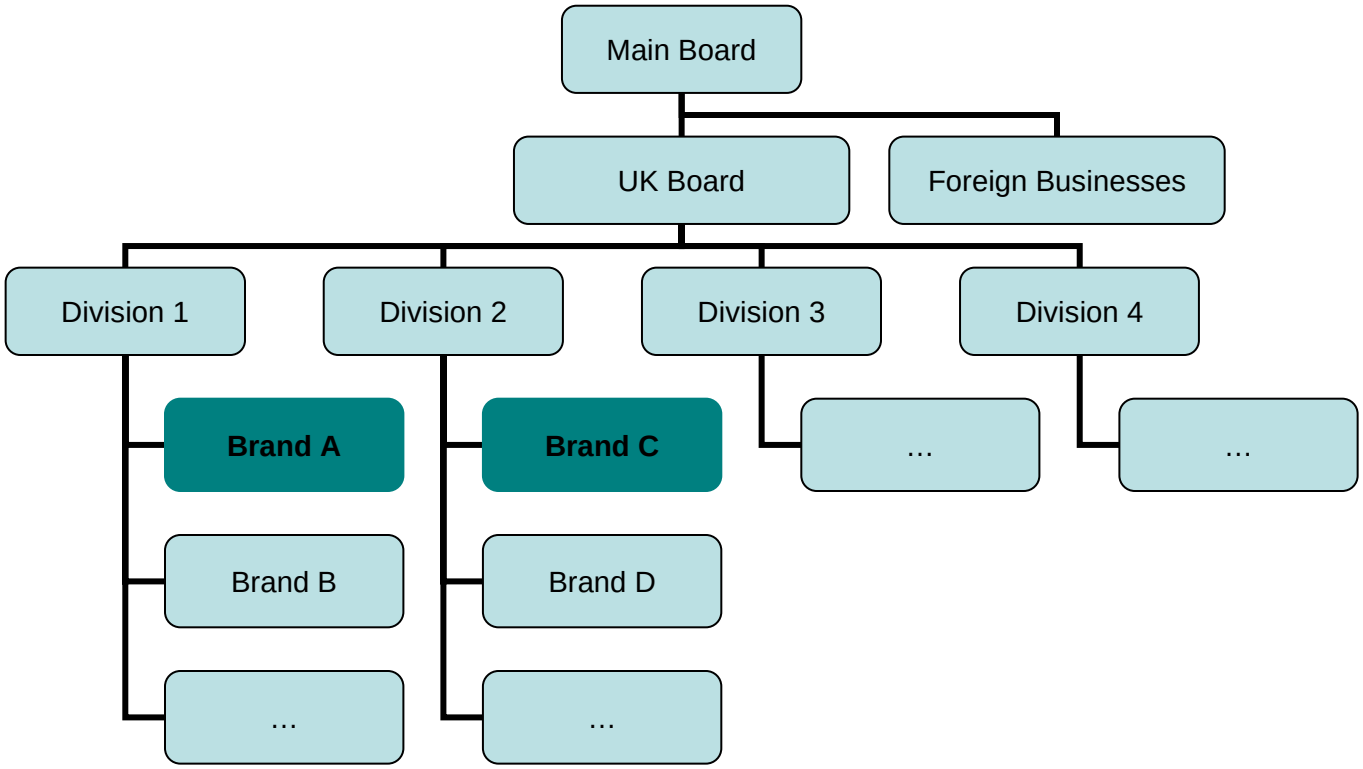


Figure 2: The Scorecard measures

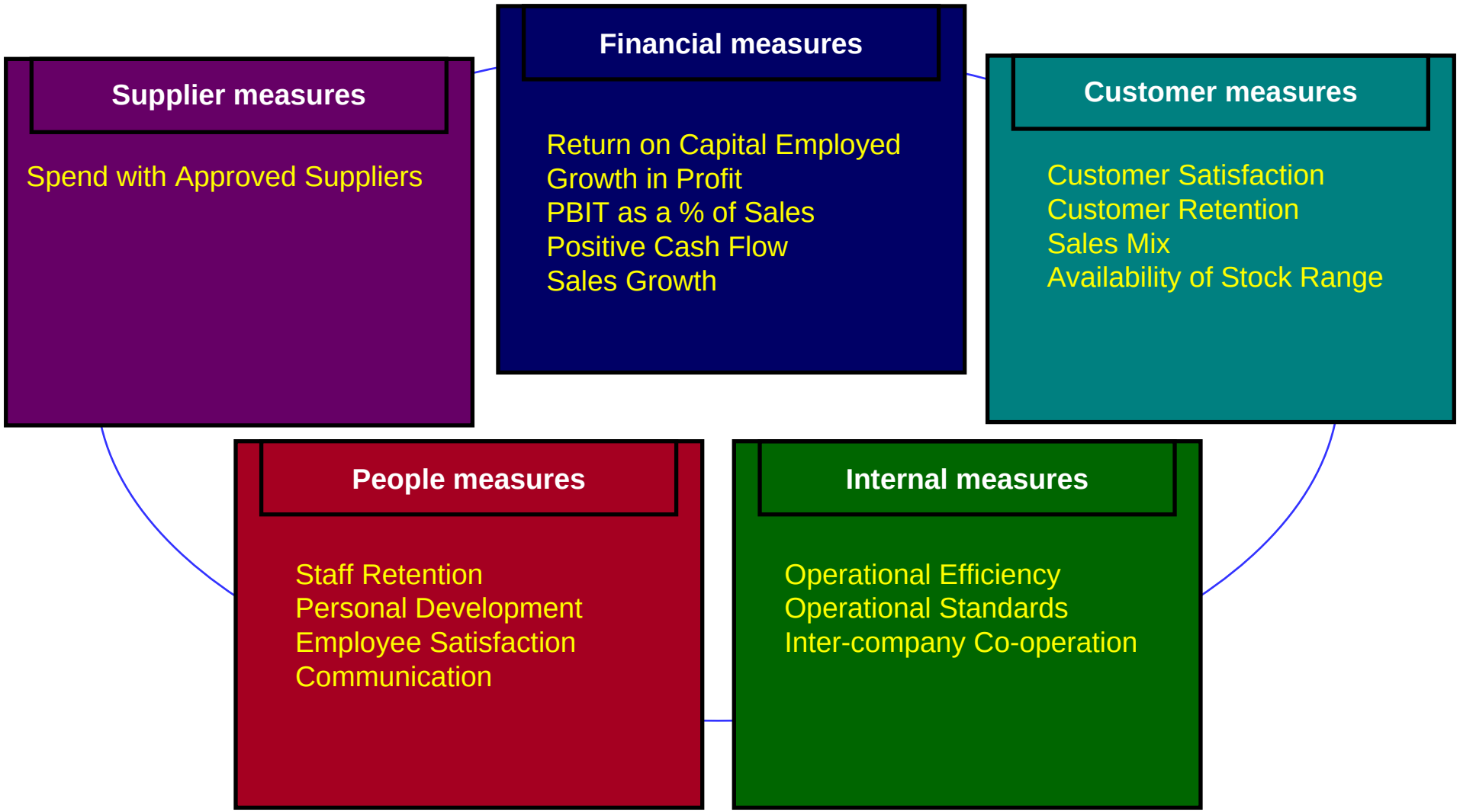
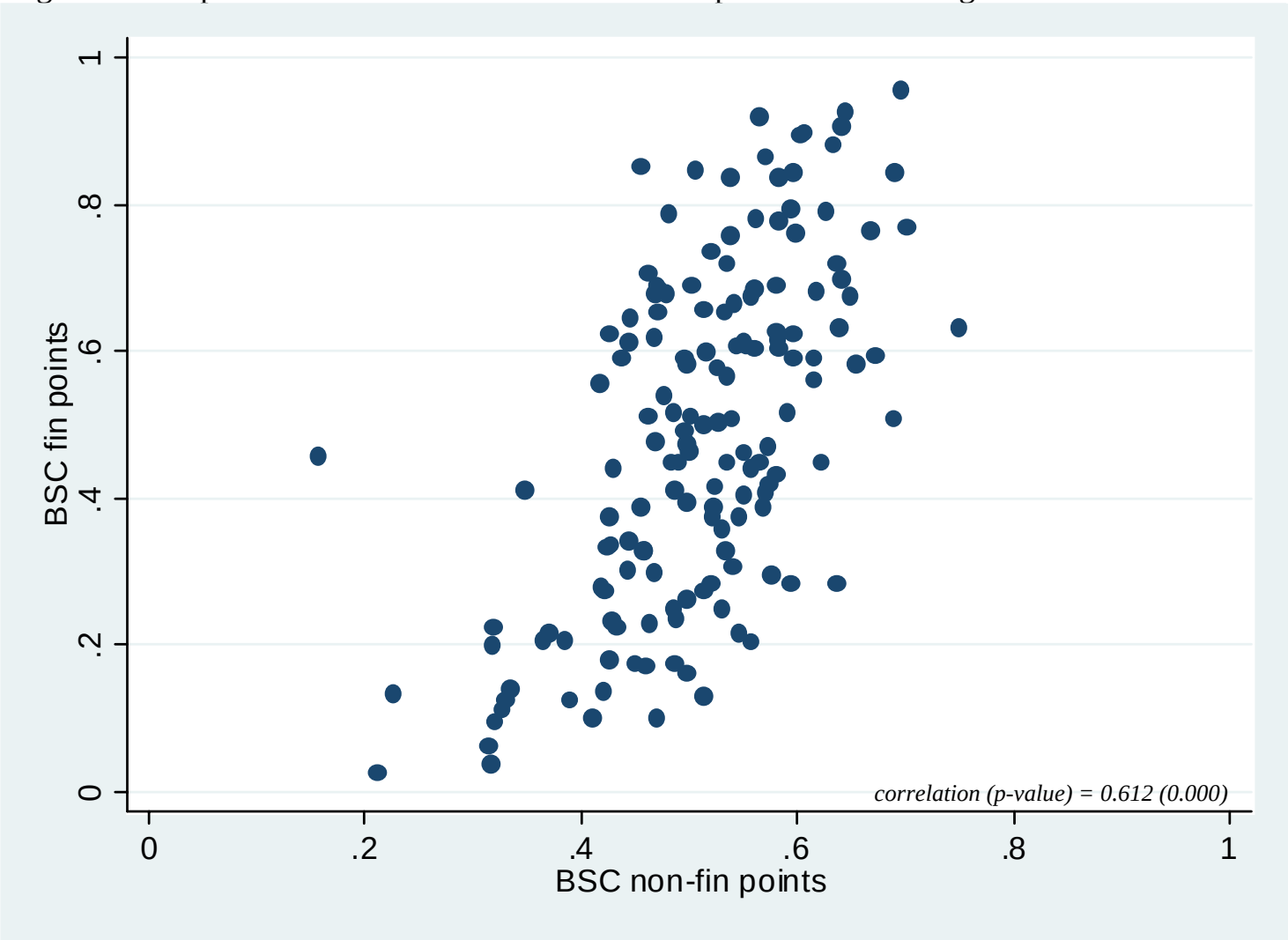
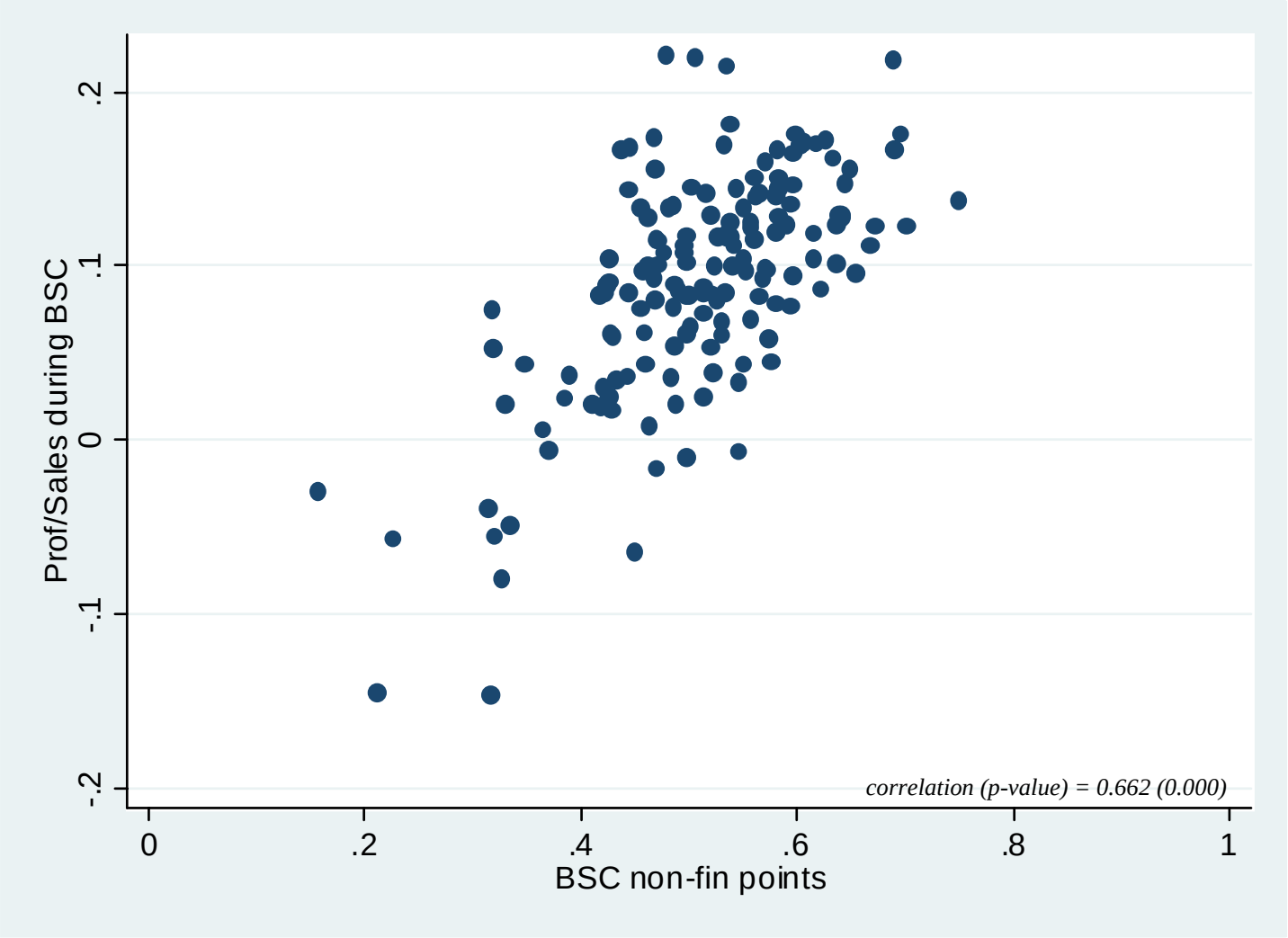


Figure 3: Comparison of non-financial with financial performance during Balanced Scorecard



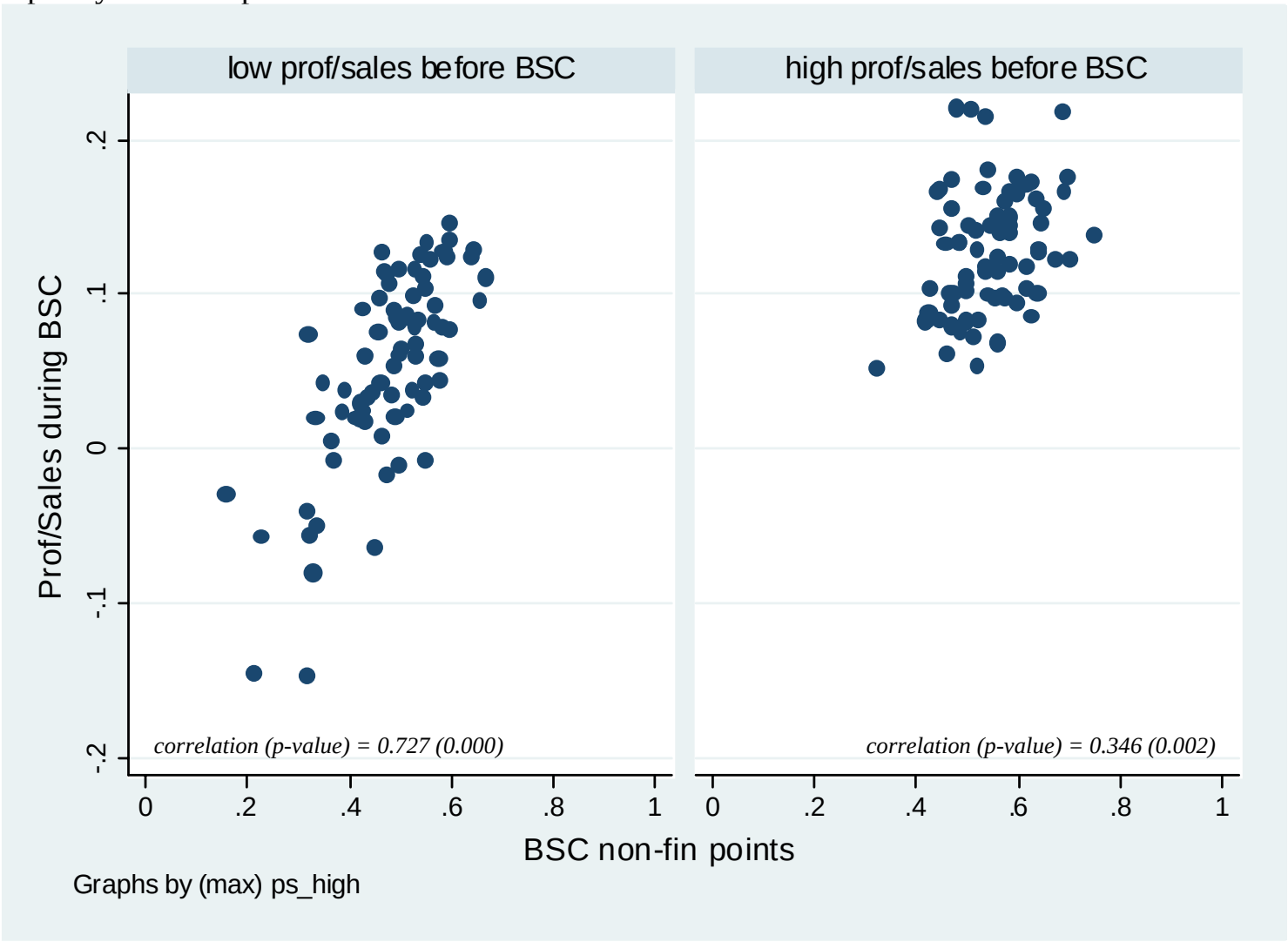
Notes: 156 observations (one for each Division 1 Brand A branch); x-axis is the share that each branch earned of the total points they could have earned on the non-financial BSC measures (customer, internal, people and supplier) over the period August 2002 - July 2004; y-axis is the share that each branch earned of the total points they could have earned on the financial BSC measures over the period August 2002 - July 2004.

Figure 3b: Comparison of non-financial with financial performance during Balanced Scorecard, using profits over sales rather than BSC measures



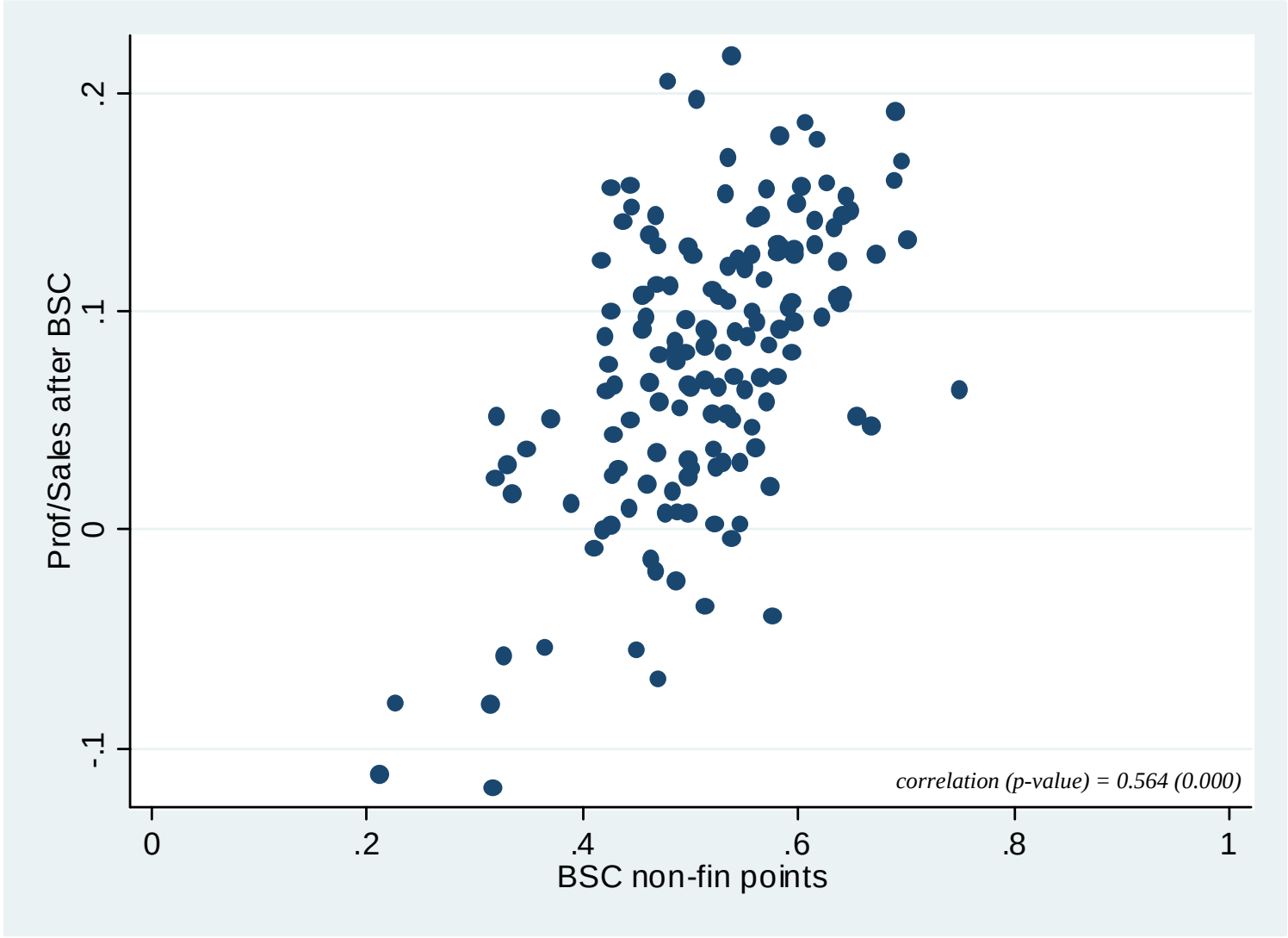
Notes: 156 observations (one for each Division 1 Brand A branch); x-axis is the share that each branch earned of the total points they could have earned on the non-financial BSC measures (customer, internal, people and supplier) over the period August 2002 - July 2004; y-axis is the average value of trading profits over sales for each branch over the period August 2002 - July 2004.

Figure 4: Comparison of non-financial with financial performance during Balanced Scorecard, split by financial performance before BSC



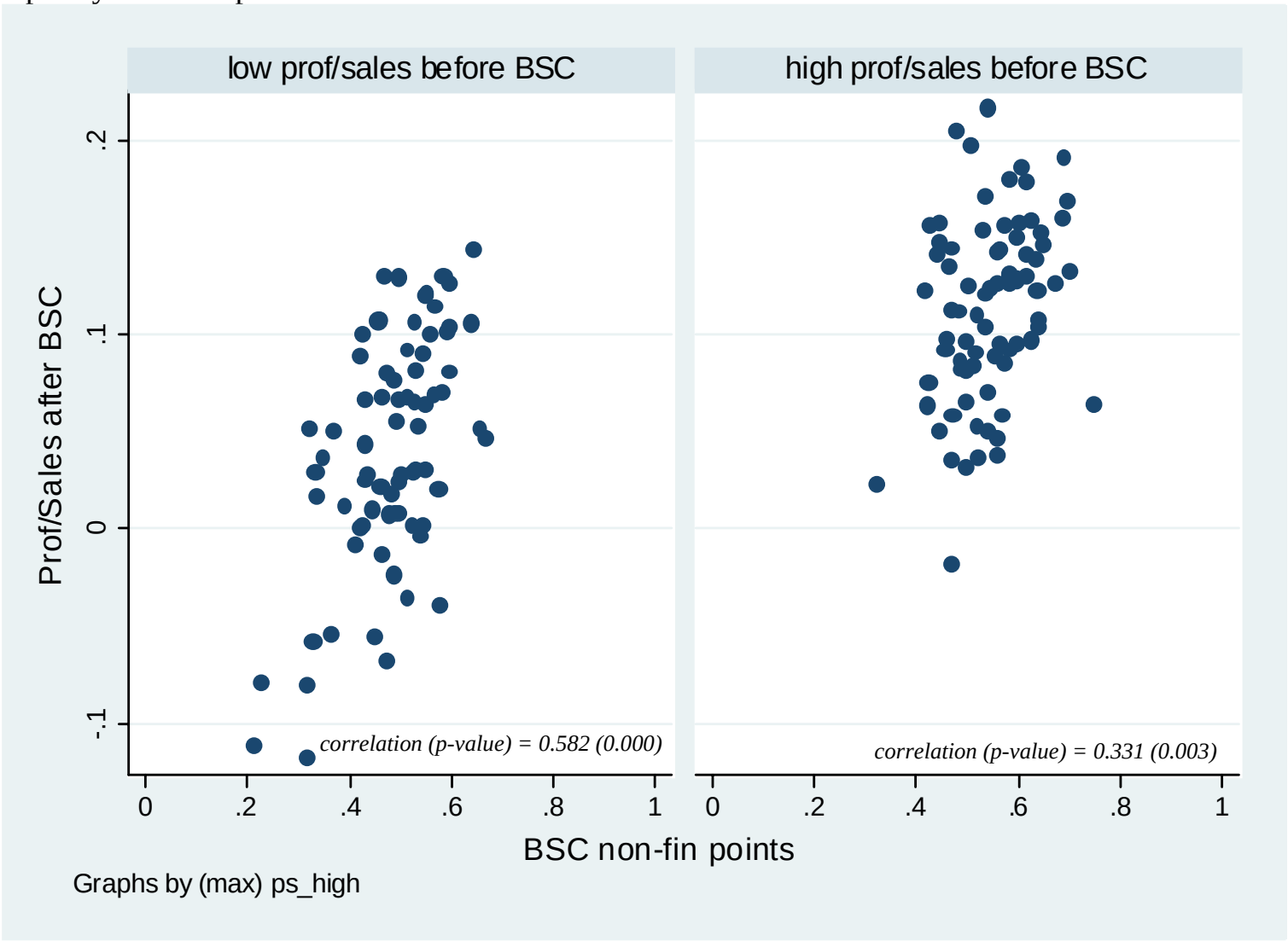
Notes: 156 observations (one for each Division 1 Brand A branch); x-axis is the share that each branch earned of the total points they could have earned on the non-financial BSC measures (customer, internal, people and supplier) over the period August 2002 - July 2004; y-axis is the average value of trading profits over sales for each branch over the period August 2002 - July 2004; left-hand panel are those branches that had below median ratio of profit to sales over the period August 1999 - July 2002.

Figure 5: Comparison of non-financial with financial performance after Balanced Scorecard



Notes: 156 observations (one for each Division 1 Brand A branch); x-axis is the share that each branch earned of the total points they could have earned on the non-financial BSC measures (customer, internal, people and supplier) over the period August 2002 - July 2004; y-axis is the average value of trading profits over sales for each branch over the period August 2004 - July 2005.

Figure 6: Comparison of non-financial with financial performance after Balanced Scorecard, split by financial performance before BSC



Notes: 156 observations (one for each Division 1 Brand A branch); x-axis is the share that each branch earned of the total points they could have earned on the non-financial BSC measures (customer, internal, people and supplier) over the period August 2002 - July 2004; y-axis is the average value of trading profits over sales for each branch over the period August 2004 - July 2005; left-hand panel are those branches that had below median ratio of profit to sales over the period August 1999 - July 2002.