Balanced scorecard innovations in government departments: Adoption, use and system effectiveness

Zahirul Hoque*
Professor of Accounting
School of Accounting
La Trobe University
Bundoora, VIC 3086, Australia
Telephone (613) 9479 3433
Fax (613) 9479 2356
E-mail: z.hoque@latrobe.edu.au

And

Carol Adams
Professor of Accounting and Sustainable Development Strategy
Faculty of Law and Management
La Trobe University
Bundoora, VIC 3086, Australia
Telephone (613) 9479 2680
Fax (613) 9479 1484
E-mail: c.adams@latrobe.edu.au

*Corresponding author


Acknowledgements: We gratefully acknowledge the financial support provided for this project by CPA Australia.
Balanced scorecard innovations in government departments: Adoption, use and system effectiveness

Abstract

Recent public sector accounting literature advocates “newer” forms of performance measures such as balanced scorecard (BSC)-type performance management innovations for government organizations that are oriented towards longer-term strategies and outcomes. There is a lack of empirical evidence on whether this outcome-oriented performance management process leads to improved organizational effectiveness. In this paper, we attempt to fill this gap in prior public sector accounting research by looking at the adoption, usage and effectiveness of BSC-type performance management innovations in a sample of government departments in Australia. Drawing on management accounting innovation and new public management literatures, we explore whether, after controlling for resource (annual operating budget), employee education, and employee experience, the adoption and usage of BSC performance measures are associated with executives’ perceptions of system effectiveness. Our empirical results indicate that whilst actual use does not match adoption of BSC performance measures, there is nevertheless a direct association between BSC innovations and perceived system effectiveness. Our findings have implications for further research and those making strategic decisions for public sector entities.

Keywords: New public management; Performance measurement; Balanced scorecard; Public sector accounting; Effectiveness; Government.

Introduction and prior literature

In recent years, interest in performance management in the public sector has grown enormously, as evidenced by the voluminous literature on new public management (NPM) and performance management innovations such as results or outcome-oriented performance measures (Smith, 1993; Hood, 1995; Ittner and Larcker, 1998; Bevir et al., 2003; Cavalluzzo and Ittner, 2004) and balanced scorecards (Kaplan and Norton, 1996; Niven, 2003). NPM is a most recent paradigm change in public sector governance worldwide (Hood, 1995; Lane, 2000). In contrast to the long-standing importance of rule compliance, NPM doctrines emphasize private sector accounting norms, bottom line management, performance indicators.
and audit, strategy-oriented control system and managing for outcomes\(^1\) (for a summary of NPM components, see Hood, 1995, Table 1, p. 96). No doubt, such a greater emphasis on outputs\(^2\) and outcomes (or results) served to increase pressure on public sector organizations to improve their performance. The use of BSC-type performance measures might be expected where private sector management styles of the 1970s, 1980s and 1990s have been promoted in response to pressures to improve performance. In this paper, we attempt to shed some light on how government organizations respond to increased pressure to perform using empirical evidence concerning the adoption and use of BSC-type performance management innovations (hereafter BSC innovation) within a sample of Australian government departments. Prior empirical research in this area is limited (Geiger and Ittner, 1996; Cavalluzzo and Ittner, 2004).

The Australian government, like many others, instigated a public sector reform program, commencing in 1993, aimed at achieving a performance culture (Parker and Guthrie, 1993; Independent Committee of Inquiry, 1993; Hoque and Moll, 2001). In October 1992 a committee, headed by Professor Hilmer, was commissioned by Prime Minister, Paul Keating (Labor) to undertake an independent inquiry into the need for a national policy aimed at ‘promoting and maintaining competitive forces to increase efficiency and community welfare, while recognising other social goals’ (Independent Committee of Inquiry, 1993). In August 1993 the commission presented the proposal from its inquiry to the Federal Government, which is now referred to as the National Competition Policy (NCP). The NCP is built on the premise that increasing competition and creating a level playing field between the public and private sector will lead to an improved use of national resources

\(^1\) Outcome measures focus on whether the population is better off. Outcome measures are concerned with the results of a program in terms of how it operates and what it achieves. (Niven, 2003).

\(^2\) Output measures track the number of people served, services provided, or units produced by a program or service (Niven, 2003).
Hood (1991, 1995) argues that encouraging a performance management approach means managing results, not rules. Geiger and Ittner (1996) found, in the case of US government agencies, that although elaborate cost accounting systems are associated with the output of those systems being used to satisfy external requirements, they are not more likely to use these more elaborate systems for internal purposes. Use of cost systems for internal purposes occurred where government agencies were required to recover costs through revenues or fees. The requirement to recover costs through revenues and fees has also been a feature of public sector management in Australia (see, for example, Guthrie, 1999; O’Faircheallaigh et al., 1999; Webster and Hoque, 2005) and, therefore, in order for private sector accounting techniques such as BSC performance measures to flourish, we would expect the emphasis to be on output controls and flexibility, rather than on input\textsuperscript{3} controls and compliance. Such emphasis is designed to maintain a focus on what is being achieved or produced (outcomes and outputs) and improving transparency and financial accountability in the public sector (Parker and Guthrie, 1993; Clark and Corbett, 1999; O’Faircheallaigh et al., 1999). The relatively lower emphasis on input controls might be expected to limit the extent to which the BSC is used by managers in decision making. Given the focus of the National Competition Policy on efficiency and growth, we would also expect to find a limited focus on social and environmental performance measures consistent with prior research which indicates a lack of accountability for social and environmental performance in the Australian public sector relative to the private sector. Gibson and Guthrie (1995) analysed the contents of the 1994 annual reports of 20 public sector entities and 40 private sector entities finding that only 18% of public sector entities compared with 67% of listed companies had a separate section on environmental performance.

\textsuperscript{3} Input measures track program inputs such as staff time and budgetary resources (Niven, 2003).
None of the public sector organisations reported environmental performance against targets. Burritt and Welch (1997a) examined the annual reports of 60 Australian Commonwealth public sector organizations over the ten-year period, 1984-1993. They found that budget entities funded from taxation revenues disclosed a much larger volume of environment-related data than did non-budget entities funded by the market, and that the differential increased markedly in the 1990s. Burritt and Welch (1997b) note that whilst private sector organizations are held to account by competitive market forces, corporate legislation and shareholder pressure, the Joint Committee on Public Accounts (JCPA) has raised concerns about the lack of mechanisms to scrutinize the accountability of commercially-oriented public sector entities.

Given this background and building upon management accounting innovation and NPM literatures, in this paper we address the following research questions:

1. To what extent do government departments adopt or implement specified elements of balanced scorecard (BSC) innovations?
2. To what extent do government departments actually use BSC measures in their day-to-day operations and strategic decisions?
3. How do government departments find the BSC innovation beneficial to their organization’s program efficiency and effectiveness?
4. Is there a direct relationship between BSC innovations adoption, use of BSC measures and systems effectiveness?

Prior public sector accounting research has not systematically examined these phenomena. The first research question, concerning the adoption of specific BSC elements, will help develop an understanding of the nature of BSC innovations adopted as part of the organization’s management control systems in government departments. Addressing the second and third research questions will assist in assessing the usefulness of the normative assertion about the role of “new” performance management systems, such as BSC, in government organizations.
In the private sector context, prior management accounting research has reported mixed findings regarding the performance consequences of management accounting innovations. While some found a positive and significant association between management accounting innovation implementation and firm performance, others found no significant relationship between these two phenomena (e.g., Innes and Mitchell, 1991; Cooper et al., 1992; Scapens and Roberts, 1993; Abernethy and Lillis, 2001; Abernethy and Bouwens, 2005). In this study we make an incremental contribution to the management accounting research literature by statistically testing a relationship between BSC innovations adoption, actual usage of BSC measures, and perceived systems effectiveness in the public sector context.

Our study contributes to the theoretical debate as to how the choice of performance management systems is influenced by the organization’s internal strategic choice and economic rationality (Broadbent and Guthrie, 1992; Hood, 1995; Chenhall, 2005). Prior public sector accounting studies find greater emphasis on the non-economic reasons (e.g., external legitimacy or coercive isomorphism) for implementing private sector accounting norms (e.g., Geiger and Ittner, 1996; Cavalluzzo and Ittner, 2004; Hoque et al., 2004).

Following Niven (2003), balanced scorecard innovations in our study refer to the implementation of three BSC-type performance measures: (1) input measures that track program inputs such as staff time, budgetary resources and natural resources; (2) output measures which track the number of people served, services provided, or units produced by a program or service; and, (3) outcome measures that focus on whether the target population is any better off. In the public sector context, Niven (2003) suggests that such BSC type measures can provide the performance measurement framework to move from a "deciding" phase to a "doing" phase.² We use the term ‘systems effectiveness’ to refer to management’s perceived benefit or

² For some notable contributions to the BSC research literature by accounting scholars, see Hoque and James (2000), Lipe and Salterio (2000, 2002), Ittner et al. (2003a, 2003b), Nørreklit (2003), Banker et al. (2004), and Wong-On-Wing et al. (2007),
usefulness of BSC type performance management systems for organizational efficiency, economy and effectiveness (Cavalluzzo and Ittner, 2004), that is, how the BSC has helped achieve the organization’s objectives for a particular task.

**Adoption of elements of BSC innovations**

For several decades, performance measurement has been used in the private sector as an internal informational tool to evaluate the performance of organizational operations, including program and budgetary decisions (Hopwood, 1972; Chenhall, 2005). Further, management uses performance measures to track employee performance against agreed targets. A performance measurement system allows employees at all levels of the business to assess progress in achieving targets, and to take corrective actions, if necessary. Kaplan and Norton’s (1992, 1993, 1996 and 2001) BSC concepts expand on mere financial measures and incorporate non-financial measures such as customer satisfaction, internal business processes, and employee learning and growth.

Niven (2003) argues that government organizations experience some degree of difficulty applying the original architecture of the BSC, mainly due to the overriding financial perspective in the NPM environment. Therefore, the original BSC framework has recently been modified to reflect the objectives of non-profit and government organizations. As outlined above, Niven (2003) discusses three different types of performance measures for the public sector: (a) input measures; (b) output measures; and (c) outcome measures. Inputs and outputs focus on the program or service, whereas outcomes focus on the results of the program in terms of how it operates and what it achieves.

Some authors have argued that increasingly changing public sector operating environments demand diverse performance information for managing operations effectively (Broadbent and Guthrie, 1992; Lapsley and Pettigrew, 1994; Lapsley, 1994, 1999; Parker and Gould, 1999). In the current NPM climate Government
organizations have come under seemingly continual pressures from both internal and external sources to demonstrate improvements in performance and achievement of goals (Bevir et al., 2003; Hood and Peters, 2004). As a result, they now place more emphasis on the implementation of BSC innovations than ever before (Niven, 2003; Modell, 2005). Given the NPM environment, BSC measures might be expected to focus on measures of outputs and outcomes such as better community services, but may also include input measures such as employee learning and growth, quality and natural resource consumption. This discussion leads to our first research question:

To what extent do government departments adopt specified elements of balanced scorecard (BSC)-type performance management innovations?

**Actual usage of BSC innovations**

It has been suggested that decision-making processes within public sector entities can be improved through use of private sector accounting tools and techniques (Broadbent and Guthrie, 1992). An example of such adoption in the Australian public sector is the recent introduction of accrual accounting and performance measurement processes (Parker and Guthrie, 1993; Funnell and Cooper, 1998; Hoque and Moll, 2001). It can be argued that private sector-type performance measurement can play a vital role in mapping the future direction of public sector organizations by giving decision makers and strategists information for setting strategies and ensuring that inputs, processes, and outputs/outcomes are aligned to organizational strategies and critical (or core) success factors.

One important additional potential use of performance management innovations is to communicate with external users who have a vested interest in the direction and success (outcome) of public sector entities. The BSC might do this by providing information concerning, for example: program performance, managing activities, budgets, strategic planning and rewarding staff.

---

5 External users include lenders, creditors, suppliers, ratepayers, taxpayers, and parties performing a review of oversight functions, namely parliaments, governments, regulatory agencies, analysts, media and special interest community groups.
This discussion, concerning the extent to which BSC performance measures in the NPM environment are concerned with meeting the needs of external, particularly regulatory, stakeholders rather than assisting managers in decision making, leads us to our second research question:

To what extent do government departments actually use BSC measures in their day-to-day operations and strategic decisions?

**Benefits of BSC innovations (systems effectiveness)**

In 2003, the Australian governments (states, territories and federal) introduced a performance management reporting framework, known as *Managing (or working) for Outcomes*, for public sector entities (Department of Finance and Administration, 2003a, 2003b). This framework focuses on: (a) reporting actual achievements against the approved output performance targets (quantity, quality, timeliness and unit cost), with explanations for significant variations in performance compared to targets; and, (b) reporting performance against the original budget in revised budgets (Department of the Prime Minister, 2006; Herawaty and Hoque, 2007). The framework provides government entities with the tools necessary to effectively monitor, evaluate and improve their performance in the delivery of outputs to the community (Department of the Prime Minister, 2006). This initiative of the Australian Government is similar to that introduced by the United States Government under the *Government Performance and Results Act* of 1993 which “requires managers of each government activity (i.e. project, program, or operation) to clarify their missions and strategic objectives and to measure relevant outputs, service levels, and outcomes for each activity in order to evaluate performance toward these objectives” (cited in Cavalluzzo and Ittner, 2004, p. 245). The focus is thus on outputs and outcomes, rather than inputs. Thus, we would expect the BSC to include more output and outcome measures than input measures and to promote efficient and effective agency management with value for money service delivery relative to
effective managerial accountability for public sector managers (Cavalluzzo and Ittner, 2004). This discussion leads us to our third research question:

How do government departments find the BSC innovation beneficial to their organization’s program efficiency and effectiveness?

Testing for an association between BSC innovations adoption, actual usage of BSC measures, and systems effectiveness

Since NPM doctrines prescribe private sector-type accounting systems for the public sector (Hood, 1995), we might expect there to be a direct association between BSC innovations adoption, usage of BSC measures and systems effectiveness, as measured in this study. Thus, our fourth research question is:

Is there a direct relationship between BSC innovations adoption, use of BSC measures and systems effectiveness?

Research method

Data collection

We used a mail-out questionnaire survey to collect empirical data on BSC adoption, BSC use, profile of the participants, and systems effectiveness.

The questionnaire survey was developed using previously published research on public sector management accounting (e.g., Hood, 1995; Melkers et al., 2002; Cavalluzzo and Ittner, 2004). In order to refine the design and focus the content, the survey was pilot tested with two departmental heads in charge of the budgeting and performance management before being sent out to all 109 government departments at federal and state/territory levels in August 2005.

Based on our pilot study, the survey was addressed to heads of performance

---

6Government in Australia is three-tiered - Commonwealth (or Federal), State and Local. State parliaments are subject to the national constitution as well as their state constitutions. A federal law overrides any state law not consistent with it. In practice, the two levels of government cooperate in many areas where states and territories are formally responsible, such as education, transport, health, and law enforcement. Income tax is levied federally, and debate between the levels of governments about access to revenue and duplication of expenditure functions is a perennial feature of Australian politics. Local government bodies are created by legislation at the state and territory level. More information about Australia’s system of government is available from the Commonwealth Department of Foreign Affairs and Trade. The Australian Commonwealth Government comprises five states (New South Wales, Queensland, Tasmania, South Australia, Western Australia) and two territories (Northern Territory and Australian Capital Territory).
management of each department (our unit of analysis), as the person most likely to
provide accurate and useful data concerning our research issues. To develop a
complete mailing list, we identified the names and addresses of the government
departments from their websites. In order to personalise the mailing, we telephoned
each department’s secretary to be sure of the appropriate name of the head of
performance management within each department.

To improve the response rate, we followed Dillman’s (1983) Total Design
Method\textsuperscript{7}. Given the potential for poor responses that could arise from lengthy and
complex surveys, we gave considerable attention to refine the visual appearance of
the survey. The mail-out survey package included a cover letter explaining the
purpose of the research, a copy of the survey, and two postage-paid envelopes –
one for returning the survey, and the second to allow respondents to request a copy
of the survey results.

Forty-two of the 109 questionnaires sent out in the first mailing were returned.
A follow-up letter four weeks after the initial mail-out yielded a further nine returned
questionnaires. Ten departments wrote declining to complete the questionnaire,
citing reasons such as contravening departmental policy and staffing constraints.
Consequently, a total of 51 departments completed questionnaires, which
represented a response rate of 46.8 per cent. Analysis of the early and late
respondents’ questionnaires using t-tests indicated no differences on the basis of the
variables of interest. Similarly, t-tests showed no differences between respondents
and non-respondents on the basis of the size and organizational types or activities.

\textsuperscript{7}Dillman (1983) suggested the use of Total Design Method (TDM) which focuses on the design and the
procedures of the survey. Among the principles of TDM the design of the instrument should: avoid
questions on the cover page and last page; the first question should apply to everyone; it should be
interesting and easy to answer; and, it should place the most-interesting and topic-related questions
first, followed by potentially objectionable questions and finally demographic information. Dillman (1983)
suggests appropriate content for the cover letter with printed mailing date, individual names and
addresses and that it is signed by the researcher(s) with a blue ballpoint using sufficient pressure. He
suggested a postcard follow-up reminder be sent 1 week after first mail-out, after 3 weeks a second
cover letter and questionnaire to non-respondents, and after 7 weeks a second cover letter complete
with another cover letter and replacement questionnaire.
Table 1 provides the profile of respondents. On average, the respondents were 45.5 years old, had worked in government departments for an average of 7.9 years, and had held their present position for an average of 3.5 years. The highest qualification for 55% of respondents was a Bachelor or postgraduate degree and for 12% of the population a CPA or CA qualification.

<table>
<thead>
<tr>
<th>Variables measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adoption of BSC innovations</td>
</tr>
</tbody>
</table>

A modified version of the Melkers et al (2002) instrument reflecting recent initiatives of the Australian Government, discussed above, was used to measure the extent of the adoption of BSC innovations. Seven BSC performance dimensions were captured: (1) inputs (activities planned), (2) activities/processes, (3) outcomes, (4) outputs, (5) cost/processes efficiency and quality measures, (6) customer/community satisfaction measures, and (7) employee learning and growth measures (employee satisfaction, employee turnover, employee training and education, employee absenteeism). These measures are similar to those used by Cavalluzzo and Ittner (2004). Respondents were asked to indicate the extent to which each of the seven performance dimensions are being adopted in their department, on a five-point, fully anchored Likert scale ranging from one ('to a very little extent') to five ('to a very great extent'). Factor analysis (principal component) reveals a single factor with eigenvalues greater than one explaining 63.4 per cent of the total variance. A single construct for the organization's overall adoption of BSC innovations (hereafter $BSC_{ADOPT}$), by calculating the arithmetic average of the respondents' scores, was computed for each item within the factor. The reliability coefficient (Cronbach alpha) for this single construct is 0.74.
Actual usage of BSC measures

Based on prior public sector accounting literature (e.g. Melkers et al., 2002; Niven, 2003; Cavalluzzo and Ittner, 2004), to assess the actual usage of BSC measures in the subject organizations we asked respondents to indicate, on a five-point scale (ranging from 1 = to a little extent to 5 = to a very great extent), the extent to which BSC measures were used in various activities or areas, namely (a) measuring program performance, (b) managing an activity or program, (c) budget formulation, (d) taking actions based on results, (e) strategic planning, (f) punishing or rewarding staff, (g) satisfying legislative requirements, (h) satisfying community expectations, (i) social responsibility goals, (j) environmental goals, (k) satisfying professional associations, and (l) following others. Table 3 provides information for 12 usages under two categories – internal managerial use and external institutional use. The reliability coefficient for this construct is 0.89.

Systems effectiveness

To assess respondents' perceived BSC systems effectiveness, we used a three-item instrument, originally developed by Melkers et al. (2002). This instrument is also similar to that used by Cavalluzzo and Ittner (2004). We asked respondents to indicate, on a five-point, fully anchored scale ranging from one ('strongly disagree') to five ('strongly agree'), their agreement or disagreement with the following statements regarding their experience(s) with BSC innovations adopted: (1) BSC measures have enhanced program efficiency, (2) BSC measures have enhanced program effectiveness, and (3) the entity is “better off” since using BSC measures. Principal component analysis extracts a single factor with an eigenvalue greater than one that explains 89.5 per cent of the total variance. Therefore, we compute a single construct representing systems effectiveness (hereafter SYS_EFFECT), by calculating the arithmetic average of the respondents' scores of each item within the factor. The Cronbach coefficient alpha for the construct is 0.94.
Control variables

Three control variables were included for their potential effects on our models. The first control variable we use is \textit{TIME\_POSITN}, which is the log of the number of years the participant has held the position in the department. The second control variable is the educational background (\textit{EMP\_EDU}) of the participants. The third control variable is \textit{RESOURCE} that captures the annual fiscal budget of the department. These control variables are included to account for the organization’s potential human and financial resource constraints for system effectiveness – the depending variable in our statistical (regression) models.

Results

\textit{Research question 1:} To what extent do government departments adopt specified elements of balanced scorecard (BSC)-type performance management innovations?

Table 2 presents evidence on the extent of adoption of BSC innovations (\textit{BSC\_ADOPT}). The results indicate that BSC measures were adopted, to a very great extent in the departments who participated, in the areas of output measures (63.0 per cent; mean, 3.57), cost efficiency and quality measures (47.8 per cent; mean, 3.26), and measures of activities and processes (48.9 per cent; mean, 3.22). BSC measures adopted most commonly to little or no extent were in the areas of employee learning and growth measures (65.2 per cent; mean, 2.20), and for input measures (37.8 per cent; mean, 2.89). These results suggest that whilst the private sector is increasingly focussing on employee learning and growth in the current skills shortage climate, this is not an important focus in the public sector, perhaps putting it at a competitive disadvantage as an employer.

Although not specifically asking about use in the BSC, our study also reveals a low adoption of social and environmental sustainability-related performance
measures in the subject government departments. While not presented in a tabular form, the most commonly adopted measures, other than financial measures used in the financial accounts, were employee diversity and economic impacts. Employee diversity measures and economic impact measures were adopted by 43.1 per cent and 41.5 per cent of respondents respectively to a large or very great extent. Natural resource conservation and emission levels were reported as being adopted to a little or no extent by 34.1 per cent of respondents. Our results thus support our expectation that in the NPM context there would be a much greater focus on output and outcome measures than input measures.

The overall mean score for \( BSC_{ADOPT} \) is 3.32, with median of 3.00 and standard deviation of 0.72. These results indicate that the subject government departments implemented BSC type performance measures to a moderate extent (a little over the average). The Kendall's coefficient of concordance (\( W \)) test for all the rankings of means was used to examine the degree of agreement in ranking among the respondents (Bryman and Cramer, 1995). The results of this exercise presented in Table 2 (\( W = 0.180; \chi^2 = 48.69; df = 6; p = 0.000 \)) indicate an observed significance level of 0.000. This suggests that the rate of adoption of the various BSC performance measures is largely consistent across our sample.

**Research question 2:** To what extent do government departments use BSC measures in their day-to-day operations and strategic decisions?

The data in Panel A of Table 3 on departmental internal use of performance management innovations, shows that BSC performance measures are least often used for punishment or reward of staff (85.7 per cent; mean, 1.50) consistent with the limited adoption of measures of employee learning and growth in the BSC (see table 2). The most common reason for adopting BSC performance measures to a large or very great extent was to measure program performance (53.3 per cent; mean, 3.53),
and to manage an activity or program (53.3 per cent; mean, 3.36). The overall mean for the internal use of BSC measures in the subject organizations is 2.93 (median = 3.00 and standard deviation = 0.84).

The data in Panel B of Table 3 indicate that 58.7 per cent of the responding departments tend to use BSC measures to satisfy external legislative requirements such as government regulation and policy. Interestingly, our findings (Panel B, Table 3) demonstrate the limited use of BSC measures to satisfy community expectations, social responsibility goals, environmental goals and professional associations (means range between 1.40 and 2.80).

The overall mean for internal use of BSC of 2.93 (Panel A, Table 3) is low compared to the BSC adoption rate of 3.32 (in Table 2). Whilst, the majority of participants (not shown in a tabular form) indicated that their performance measurement systems were adequate for most of their needs, only six and a half per cent recorded that their performance measurement systems were adequate for all of their needs and twenty-six per cent reported only half of their needs being adequately meet by their current performance measurement system. Further, reports generated from performance measurement systems were equally used for external financial reporting and internal managerial control (71.7 per cent). Over half of the participants also indicated that reports were used for budget preparation (58.7 per cent) and legal requirements (50 per cent). Our results indicate that the focus on output and outcome measures, limits the usefulness of BSC performance measures in internal decision making.

The results of the Kendall’s coefficient of concordance are: $W = 0.456; \chi^2 = 93.40; \text{df} = 5; p = 0.000$. Since the observed significance level associated with the test is small ($p = 0.000$), it can be suggested that there is no disagreement among the responding departments regarding the above uses of BSC performance measures.
Research question 3: How do government departments find the BSC innovation beneficial to their organization’s program efficiency and effectiveness?

The results presented in Table 4 indicate that the majority of participants agree that the results of using performance measures, SYS_EFFECT, have enhanced program efficiency (mean, 3.33) and program effectiveness (mean, 3.46), and that their entity is ‘better off’ since using BSC performance measures (mean, 3.71). The overall mean for SYS_EFFECT is 3.53, with median of 4.00 and the standard deviation of 1.74. The observed significance level associated with the Kendall’s test ($W = 0.456; \chi^2 = 93.40; df = 5$) is small ($p = 0.000$). Therefore, we conclude that there is a good consensus among the respondents concerning their perceived systems (BSC) effectiveness.

Research question 4: Is there a direct relationship between BSC innovations adoption, internal managerial usage of BSC measures and systems effectiveness?

This research question has two parts. First it looks at the relationship between BSC innovations adoption and system effectiveness (hereafter research question 4a). Secondly, it explores the relationship between actual usage of BSC measures and system effectiveness (hereafter research question 4b). To explain these two issues in detail we run two separate statistical analyses. We present the results in turn.

BSC innovations adoption and system effectiveness

Table 5 presents zero order (Pearson) correlation coefficients and partial correlation coefficients, after controlling for RESOURCE, EMP_EDU and LENGTH-POSTN. As expected, both zero order and partial correlation coefficients indicate a positive and significant association between system effectiveness (the dependent variable) and the adoption of BSC performance measures ($r = 0.67, p < 0.01$) and $r =$
While the control variable, \textit{RESOURCE} (annual fiscal $\text{budget}$), appears to be significantly associated with the adoption of BSC performance measures ($r = 0.31, p < 0.05$), other control variables, \textit{EMP\_EDU} (the level of employee education) and their length of services in the current position (\textit{LENGTH\_POSN}), do not appear to be significantly associated with both the adoption of BSC measures and systems effectiveness.

\textbf{INSERT TABLE 5 HERE}

Further, we run multiple regression analysis to explore the research question 4a using the following regression model:

$$SYS\_EFFECT = \beta_0 + \beta_1 BSC\_ADOPT + \beta_2 RESOURCE + \beta_3 EMP\_EDU + \beta_4 LENGTH\_POSN + e$$

\text{Model (1)}

where \textit{SYS\_EFFECT} is systems effectiveness, \textit{BSC\_ADOPT} is the extent of the adoption of BSC-type performance measures, \textit{EMP\_EDU} is the level of education of the employees (1 = graduate and above with professional qualifications, and 0 otherwise), \textit{LENGTH\_POSN} is the log of the number of years the participant has held the position in the department, \textit{RESOURCE} is the natural logarithm (log) of the annual fiscal budget of the department. $\beta_0$ is constant, and $e$ is the error term.

Table 6 presents the results of the regression analysis. The overall regression model for the independent and control variables explained 55.9 per cent (adjusted $R^2$) of the variance in the dependent variable, \textit{SYS\_EFFECT} ($F = 11.47; p = 0.00$). The data indicate that the standardized beta coefficients $\beta_1$ (\textit{BSC\_ADOPT}), $\beta_2$ (\textit{RESOURCE}) and $\beta_4$ (\textit{LENGTH\_POSN}) are all positive and significant ($\beta_1 = 0.84; p = 0.00$, two-tailed; $\beta_2 = 0.26; p = 0.06$, two-tailed; $\beta_4 = 0.18; p = 0.09$, one-tailed). Although the control variable, \textit{EMP\_EDU}, is positive, it is not statistically significant. Thus these regression results and correlations results presented in Table 5 support our prediction that greater BSC innovations adoption is associated with subject respondents' perceived system effectiveness. As shown in Table 6, multicollinearity
tests via Tolerance and VIF (variance inflation factor) detect no multicollinearity among the independent and control variables, thereby no major threat for regression analysis.

INTERNAL MANAGERIAL USAGE OF BSC MEASURES AND SYSTEMS EFFECTIVENESS

Zero-order and partial correlation analyses presented in Table 7 show that systems effectiveness (SYS_EFF) is significantly associated with the extent of the internal use of BSC performance measures (BSC_INUSE) \(r = 0.73, \ p < 0.01\); \(r = 0.76, \ p < 0.01\). Surprisingly, no control variables are significantly associated with either BSC_USE or SYS_EFF. Thus the results of this univariate analysis provide strong evidence to support our idea that actual usage of BSC innovation for day-to-day operations and strategic decisions is positively and significantly associated with SYS_EFF.

Table 7 also shows that external usage of BSC measures is positively and significantly associated with SYS_EFF (Zero-order \(r = 0.46\) and partial \(r = 0.56\)). Notice that these correlations are not significantly higher than the correlations between internal usage of BSC measures and SYS_EFF (Zero-order \(r = 0.73\) and partial \(r = 0.76\)). These results support the argument put forward by institutional theory (Meyer and Rowan, 1977; DiMaggio and Powell, 1983, 1991) that organizations sometimes tend to use an organizational system without much regard for their organizational performance – they do use them mainly for external compliance reasons. For similar observation in government performance measurement, see Geiger and Ittner (1996) and Cavalluzzo and Ittner (2004).

To explore the above idea further, we run the following regression model:
SYS\_EFFECT = \beta_0 + \beta_1 BSC\_INUSE + \beta_2 RESOURCE + \beta_3 EMP\_EDU + \beta_4 LENGTH\_POSITN + e \quad \text{Model (2)}

where BSC\_INUSE is the extent of the use of BSC-type performance measures for internal managerial decisions, SYS\_EFFECT, EMP\_EDU, LENGTH\_POSN, RESOURCE, \beta_0 and e, as in model 1. Table 8 shows that BSC\_USE is positively and significantly associated with SYS\_EFFECT (coefficient = 0.75, p = 0.00). Although EMP\_EDU is positively associated with SYS\_EFFECT, it is not highly significant (p = 0.07). The regression model explained 51 per cent (adjusted R^2) (F = 8.81, p = 0.00) of the variance in the dependent variable. As shown in Table 8, none of Tolerance and VIF tests detect multicollinearity among the independent and control variables. Overall, the regression results suggest a strong support for the idea that increased internal managerial usage of BSC measures is likely to result in increased systems effectiveness.

INSERT TABLE 8 HERE

Robustness tests

To further explore the above statistical relationships with regard to the regression model (1), we run multiple regression analyses using each of the seven BSC adoption dimensions, as labelled in Table 2. The results of the seven sub-models presented in Table 9 indicate that systems effectiveness is positively and significantly associated with all but one of the BSC dimensions at the level of significance, p < 0.01. The adjusted R^2’s for these six models are also strong i.e. above 10 per cent. The BSC input dimension is not significantly associated with systems effectiveness (p > 0.10). The control variable RESOURCE appears to be a significant control variable in all seven models (for reasons of space, not shown in Table 9).

INSERT TABLE 9 HERE
As for regression model 1, we also run additional analysis for the regression model 2 using each of the six internal uses of BSC performance measures. The results, presented in Table 10, indicate that SYS_EFFECT is associated significantly with all six uses of BSC performance measures. Again, RESOURCE and EMP_EDU appear to be significant control variables in all six models (for reasons of space, not shown in Table 10). Overall, these results suggest that greater use of BSC performance measures may lead to improved systems effectiveness.

Following prior management accounting research, we conducted additional robustness checks using alternative measures. To explore whether the results presented above based on average scores of the dependent and independent variables are different when using alternative measurement technique, we ran the regression models using factor scores of the measures. Additionally, we considered other control variables, namely size of the organization in terms of the number of full-time employees, annual capital budgets, and activities of the organization. While not presented, results of this exercise are consistent with those discussed above.

**Conclusions and implications**

The purpose of our study was to examine the extent and nature of the adoption of balanced scorecard (BSC)-type innovations in Australian government departments. Our findings in relation to the BSC are indicative of the extent of adoption of private sector accounting techniques and new public management (NPM) and performance management innovations in the Australian public sector. They need to be interpreted in the context of an increased focus on output and outcome measures following the introduction of the *Managing (or working) for Outcomes* framework for public entities adopted in 2003.

Responses to our survey developed from the work of Hood (1995), Melkers et al (2002) and Cavalluzzo and Ittner (2004) were received from 51 government
departments from a total population of 109, a response rate of 46.8 per cent. The survey was particularly concerned with: the extent of adoption of BSC type performance management innovations and specific measures adopted; the use of BSC performance measures in day-to-day operations and strategic decisions; and, the perceived benefits and effectiveness of BSC performance management innovations.

We found that overall, Australian government departments have implemented BSC performance measures to a moderate extent. The most common BSC performance measures adopted were measures for output, cost efficiency and quality and activities and processes, and the least common were measures for employee learning and growth and input measures. This is not surprising given the focus of the Managing (or working) for Outcomes framework. Hood (1991, 1995) argues that encouraging a performance management approach means managing results, not rules.

Drawing on the new public management literature, we tested whether, after controlling for resource, employee education, and employee experience, the adoption and internal use of BSC performance measures are associated with executives’ perceptions of BSC performance measurement system effectiveness. Our empirical results indicate a direct association between BSC innovations and perceived system effectiveness. Of the three control variables, resource appears to be a significant influential factor.

BSC performance measures were most commonly used to satisfy legislative requirements and manage an activity or program and least commonly used to punish or reward staff or to follow others. Use of BSC performance measures is perceived to enhance programme efficiency and programme effectiveness consistent with the NPM focus on output and outcome measures. The lower use of BSC performance measures compared with their adoption is consistent with their main purpose in the NPM context, i.e. to satisfy legislative requirements and increase efficiency. It is
perhaps ironic that having adopted private sector performance management tools as used towards the end of the last century, the public sector is lagging behind the private sector in recognising the importance of intellectual capital and particularly human capital management and environmental performance as measures of success and good management. Whilst the use of the BSC itself may be described as an innovation, the design and use of the BSC within the government departments surveyed could not be described as innovative, a factor which might be explained by the importance of legislative requirements in influencing its introduction.

The lower use of BSC performance measures as compared with their adoption indicates that the use of the BSC in decision-making may be improved by the inclusion of more input measures to assist managers in decision-making. The use of indicators related to human capital management and other social and environmental performance indicators was particularly low. Ironically, in the private sector this is an area gaining increasing attention.

Our findings contribute to the theoretical debate as to whether or not government organizations implement accounting innovations for economic and strategic reasons, that is, how the performance management choice is influenced by the organization’s internal strategic choice and economic rationality (Broadbent and Guthrie, 1992; Hood, 1995). Consistent with prior public sector accounting studies we found greater emphasis on external legitimacy or coercive isomorphism as reasons for implementing (perhaps not outdated) private sector accounting norms than on economic reasons (e.g., Geiger and Ittner, 1996; Cavalluzzo and Ittner, 2004; Hoque et al., 2004; Webster and Hoque, 2005). Further, we found limited focus on social and environmental performance measures. Further research might take the form of multiple case studies of similar government departments based on face-to-face interviews to shed further light on the role of performance management innovations in the changing public sector environment. This would allow a more in-depth assessment of the extent to which performance is a function of the ‘fit’ or
match between the government departments’ regulatory environment and use of the
different combinations of BSC performance measures.
Table 1
Profile of the respondents (N = 51)

<table>
<thead>
<tr>
<th>Length of service within government departments</th>
<th>Year</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Below 2</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>2 – 5</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>5 – 8</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>8 – 11</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>11 or over</td>
<td>12</td>
</tr>
<tr>
<td>Mean = 7.85 years; Median = 5.00; Std. Deviation = 7.13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Age group</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20 – 29</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>30- 39</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>40 – 49</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>50 or over</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education level</th>
<th>Degree</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Diploma</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Bachelor/Postgraduate</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>CPA/CA</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>19</td>
</tr>
</tbody>
</table>

<p>| Annual operating budget (A $m) | Mean = $1,044.88 | Median = $269.10 |</p>
<table>
<thead>
<tr>
<th>BSC dimension</th>
<th>Little or no extent</th>
<th>Moderate extent</th>
<th>Large and very great extent</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
<th>Observed range*</th>
<th>Factor loadings**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outputs (quantity of services provided) (BSC_OUT)</td>
<td>6 (13.1%)</td>
<td>11 (23.9%)</td>
<td>29 (63.0%)</td>
<td>3.57</td>
<td>4.00</td>
<td>1.29</td>
<td>1 – 5</td>
<td>0.84</td>
</tr>
<tr>
<td>Cost/processes efficiency and quality measures (BSC_CEQ)</td>
<td>11 (23.9%)</td>
<td>13 (28.3%)</td>
<td>22 (47.8%)</td>
<td>3.26</td>
<td>3.00</td>
<td>1.24</td>
<td>1 – 5</td>
<td>0.79</td>
</tr>
<tr>
<td>Activity/process measures (BSC_ACT)</td>
<td>10 (22.2%)</td>
<td>13 (28.9%)</td>
<td>22 (48.9%)</td>
<td>3.22</td>
<td>3.00</td>
<td>1.17</td>
<td>1 – 5</td>
<td>0.79</td>
</tr>
<tr>
<td>Customer/community satisfaction measures (BSC_COMM)</td>
<td>14 (31.2%)</td>
<td>8 (17.8%)</td>
<td>23 (51.1%)</td>
<td>3.16</td>
<td>4.00</td>
<td>1.28</td>
<td>1 – 5</td>
<td>0.71</td>
</tr>
<tr>
<td>Outcomes measures (BSC_OUTC)</td>
<td>16 (34.8%)</td>
<td>12 (26.1%)</td>
<td>18 (39.2%)</td>
<td>3.06</td>
<td>3.00</td>
<td>1.16</td>
<td>1 – 5</td>
<td>0.53</td>
</tr>
<tr>
<td>Inputs - activities planned (BSC_INP)</td>
<td>17 (37.8%)</td>
<td>9 (20.0%)</td>
<td>19 (42.2%)</td>
<td>2.89</td>
<td>3.00</td>
<td>1.30</td>
<td>1 – 5</td>
<td>0.50</td>
</tr>
<tr>
<td>Employee learning and growth measures (BSC_ELG)</td>
<td>30 (65.2%)</td>
<td>9 (19.6%)</td>
<td>7 (15.2%)</td>
<td>2.20</td>
<td>2.00</td>
<td>1.05</td>
<td>1 – 5</td>
<td>0.87</td>
</tr>
<tr>
<td>Overall BSC adoption (BSC_ADOPT)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.32</td>
<td>3.00</td>
<td>0.72</td>
<td>1 - 5</td>
<td>-</td>
</tr>
</tbody>
</table>

Kendall’s \(W\) tests: \(n = 45\); \(W = 0.180\) \(\chi^2 = 48.69\) ; df = 6 ; p = 0.000

*Theoretical Range, 1 – 5. **Principal component analysis - variance explained = 63.4 per cent; Cronbach alpha = 0.74)
Table 3 Internal and external uses of BSC performance management innovations

<table>
<thead>
<tr>
<th>Decision making areas</th>
<th>Little or no extent</th>
<th>Moderate extent</th>
<th>Large and very great extent</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
<th>Observed range*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Internal managerial use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measure program performance</td>
<td>10 (22.0%)</td>
<td>11 (24.4%)</td>
<td>24 (53.3%)</td>
<td>3.53</td>
<td>4.00</td>
<td>1.14</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Manage an activity or program</td>
<td>11 (25.0%)</td>
<td>10 (22.7%)</td>
<td>23 (53.3%)</td>
<td>3.36</td>
<td>4.00</td>
<td>1.10</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Budget formulation</td>
<td>14 (35.6%)</td>
<td>11 (24.4%)</td>
<td>18 (40.0%)</td>
<td>3.13</td>
<td>3.00</td>
<td>1.19</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Taking actions based on results</td>
<td>17 (38.6%)</td>
<td>8 (18.2%)</td>
<td>19 (43.2%)</td>
<td>3.06</td>
<td>3.00</td>
<td>1.20</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Strategic planning</td>
<td>17 (38.6%)</td>
<td>9 (20.5%)</td>
<td>18 (47.8%)</td>
<td>3.02</td>
<td>3.00</td>
<td>1.28</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Punish or reward staff</td>
<td>36 (85.7%)</td>
<td>4 (9.5%)</td>
<td>2 (4.8%)</td>
<td>1.50</td>
<td>1.00</td>
<td>0.94</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Overall internal use of BSC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.93</td>
<td>3.00</td>
<td>0.84</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Panel B: External institutional use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfy legislative requirements (law,</td>
<td>12 (26.1%)</td>
<td>7 (15.2%)</td>
<td>27 (58.7%)</td>
<td>3.46</td>
<td>4.00</td>
<td>1.36</td>
<td>1 - 5</td>
</tr>
<tr>
<td>ordinance, policy, etc)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfy community expectations</td>
<td>19 (42.2%)</td>
<td>14 (31.1%)</td>
<td>12 (26.7%)</td>
<td>2.80</td>
<td>3.00</td>
<td>1.09</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Social responsibility goals</td>
<td>22 (50.0%)</td>
<td>10 (22.7%)</td>
<td>12 (27.3%)</td>
<td>2.66</td>
<td>2.50</td>
<td>1.31</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Environmental goals</td>
<td>30 (71.4%)</td>
<td>8 (19.0%)</td>
<td>4 (7.8%)</td>
<td>1.97</td>
<td>2.00</td>
<td>1.14</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Satisfy professional associations</td>
<td>31 (76.7%)</td>
<td>6 (14.0%)</td>
<td>2 (4.8%)</td>
<td>1.77</td>
<td>1.00</td>
<td>1.08</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Follow others</td>
<td>37 (88.1%)</td>
<td>5 (11.9%)</td>
<td>0 (0.00%)</td>
<td>1.40</td>
<td>1.00</td>
<td>0.70</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Overall external institutional use of BSC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.34</td>
<td>2.50</td>
<td>1.07</td>
<td>1 - 5</td>
</tr>
</tbody>
</table>

*Theoretical Range, 1 – 5.
Table 4 Effectiveness of BSC performance management innovations (SYS_EFFECT)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Little or no extent</th>
<th>Moderate extent</th>
<th>Large very extent</th>
<th>Mean rank</th>
<th>Median</th>
<th>Standard deviation</th>
<th>Observed range</th>
<th>Factor loadings**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enhanced program efficiency</td>
<td>13 (25.2%)</td>
<td>6 (11.8%)</td>
<td>26 (51%)</td>
<td>3.72</td>
<td>4.00</td>
<td>1.10</td>
<td>1-5</td>
<td>0.91</td>
</tr>
<tr>
<td>Enhanced program effectiveness</td>
<td>11 (21.5%)</td>
<td>6 (11.8%)</td>
<td>28 (54.9%)</td>
<td>3.47</td>
<td>4.00</td>
<td>1.09</td>
<td>1-5</td>
<td>0.96</td>
</tr>
<tr>
<td>Entity is “Better Off”</td>
<td>9 (17.7%)</td>
<td>5 (9.8%)</td>
<td>31 (60.8%)</td>
<td>3.33</td>
<td>4.00</td>
<td>1.21</td>
<td>1-5</td>
<td>0.97</td>
</tr>
<tr>
<td>Overall effectiveness</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.53 (mean)</td>
<td>4.00</td>
<td>1.74</td>
<td>1-5</td>
<td>-</td>
</tr>
</tbody>
</table>

Kendall's (W) tests: n = 45; W = 0.148; χ² = 13.34; df = 2; p = 0.001

N = 45, missing n = 6 for all items

*Theoretical range, 1 – 5. **Principal component analysis - variance explained = 89.5 per cent; Cronbach alpha = 0.94)
Table 5
Zero-order and partial correlations for BSC_ADOPT with other variables

<table>
<thead>
<tr>
<th>Control Variables</th>
<th>BSC_ADOPT</th>
<th>SYS_EFFECT</th>
<th>RESOURCE</th>
<th>EMPL_EDU</th>
<th>LENGTH_POSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>None*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSC_ADOPT</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYS_EFFECT</td>
<td>0.67**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RESOURCE</td>
<td>0.31*</td>
<td>-0.09</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMPL_EDU</td>
<td>-0.13</td>
<td>-0.13</td>
<td>0.13</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>LENGTH_POSN</td>
<td>-0.15</td>
<td>-0.08</td>
<td>0.14</td>
<td>0.02</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RESOURCE, EMPL_EDU &amp; LENGTH_POSN</th>
<th>BSC_ADOPT</th>
<th>SYS_EFFECT</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SYS_EFFECT</td>
<td>0.78**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Cells contain zero-order (Pearson) correlations
**p < 0.01; * p < 0.05 (Two-tailed)

Definition of variables:
BSC_ADOPT = The adoption of Balanced Scorecard (BSC) type performance management
SYS_EFFECT = Systems Effectiveness
RESOURCE = Annual fiscal (operating) budget
EMPL_EDU = Employee education
LENGTH_POSN = Length of employment in the current position
Table 6
Multiple regression results (H1)
Dependent variable = SYS\_EFFECT against BSC\_ADOPT

SYS\_EFFECT = $\beta_0 + \beta_1 BSC\_ADOPT + \beta_2 RESOURCE + \beta_3 EMP\_EDU + \beta_4 LENGTH\_POSN + e$

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Estimate</th>
<th>Standard error</th>
<th>t-value</th>
<th>Sig. (p)</th>
<th>Tolerance</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$ Intercepts</td>
<td>$\alpha_0$</td>
<td>-3.89</td>
<td>2.95</td>
<td>-1.32</td>
<td>0.19</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>$X_1$ BSC_ADOPT</td>
<td>$\beta_1$</td>
<td>0.84</td>
<td>0.08</td>
<td>6.74</td>
<td>0.00</td>
<td>0.87</td>
<td>1.15</td>
</tr>
<tr>
<td>$X_2$ RESOURCE</td>
<td>$\beta_2$</td>
<td>0.26</td>
<td>0.31</td>
<td>1.95</td>
<td>0.06</td>
<td>0.75</td>
<td>1.34</td>
</tr>
<tr>
<td>$X_3$ EMP_EDU</td>
<td>$\beta_3$</td>
<td>0.05</td>
<td>0.99</td>
<td>0.39</td>
<td>n.s.</td>
<td>0.96</td>
<td>1.04</td>
</tr>
<tr>
<td>$X_4$ LENGTH_POSN</td>
<td>$\beta_4$</td>
<td>0.18</td>
<td>0.45</td>
<td>1.40</td>
<td>0.09</td>
<td>0.78</td>
<td>1.28</td>
</tr>
</tbody>
</table>

R² = 0.613; Adjusted R² = 0.559; F [4, 29] = 11.47; p = 0.00

Variable Definitions: SYS\_EFFECT = Systems Effectiveness; BSC\_ADOPT = Balanced Scorecard (BSC) Adoption; RESOURCE = Log of the annual budget; EMP\_EDU = Level of employee education; LENGTH\_POSN = Length of position the participants held (log); n.s = Not significant; n/a = Not available
### Table 7
Additional regression analysis for **SYS_EFFECT** (the dependent variable) with seven BSC dimensions with

<table>
<thead>
<tr>
<th>Regression statistics</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
<th>Model 7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BSC_INP</td>
<td>BSC_CEQ</td>
<td>BSC_ACT</td>
<td>BSC_COMM</td>
<td>BSC_OUTPUT</td>
<td>BSC_OUTCOME</td>
<td>BSC_ELG</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.49</td>
<td>0.35</td>
<td>0.47</td>
<td>0.43</td>
<td>0.49</td>
<td>0.45</td>
<td>0.51</td>
</tr>
<tr>
<td>Coefficient</td>
<td>0.15</td>
<td>0.70</td>
<td>0.59</td>
<td>0.64</td>
<td>0.72</td>
<td>0.47</td>
<td>0.52</td>
</tr>
<tr>
<td>T-value</td>
<td>0.78</td>
<td>4.98</td>
<td>3.41</td>
<td>3.80</td>
<td>5.08</td>
<td>2.80</td>
<td>3.09</td>
</tr>
<tr>
<td>Sig.</td>
<td>0.44</td>
<td>0.000</td>
<td>0.002</td>
<td>0.001</td>
<td>0.000</td>
<td>0.009</td>
<td>0.004</td>
</tr>
<tr>
<td>R²</td>
<td>0.03</td>
<td>0.464</td>
<td>0.291</td>
<td>0.335</td>
<td>0.474</td>
<td>0.217</td>
<td>0.251</td>
</tr>
<tr>
<td>Adj. R²</td>
<td>-0.11</td>
<td>0.390</td>
<td>0.193</td>
<td>0.244</td>
<td>0.401</td>
<td>0.109</td>
<td>0.148</td>
</tr>
<tr>
<td>F-value</td>
<td>0.19</td>
<td>6.27</td>
<td>2.97</td>
<td>3.66</td>
<td>6.53</td>
<td>2.01</td>
<td>2.44</td>
</tr>
<tr>
<td>Sig.</td>
<td>0.94</td>
<td>0.001</td>
<td>0.036</td>
<td>0.016</td>
<td>0.001</td>
<td>0.120</td>
<td>0.070</td>
</tr>
</tbody>
</table>

Model 1: \( SYS_{\text{EFFECT}} = \beta_1 \text{BSC}_\text{INP} + \beta_2 \text{RESOURCE} + \beta_3 \text{EMP}_\text{EDU} + \beta_4 \text{LENGTH}_\text{POSN} + e \)
Model 2: \( SYS_{\text{EFFECT}} = \beta_1 \text{BSC}_\text{CEQ} + \beta_2 \text{RESOURCE} + \beta_3 \text{EMP}_\text{EDU} + \beta_4 \text{LENGTH}_\text{POSN} + e \)
Model 3: \( SYS_{\text{EFFECT}} = \beta_1 \text{BSC}_\text{ACT} + \beta_2 \text{RESOURCE} + \beta_3 \text{EMP}_\text{EDU} + \beta_4 \text{LENGTH}_\text{POSN} + e \)
Model 4: \( SYS_{\text{EFFECT}} = \beta_1 \text{BSC}_\text{COMM} + \beta_2 \text{RESOURCE} + \beta_3 \text{EMP}_\text{EDU} + \beta_4 \text{LENGTH}_\text{POSN} + e \)
Model 5: \( SYS_{\text{EFFECT}} = \beta_1 \text{BSC}_\text{OUTPUT} + \beta_2 \text{RESOURCE} + \beta_3 \text{EMP}_\text{EDU} + \beta_4 \text{LENGTH}_\text{POSN} + e \)
Model 6: \( SYS_{\text{EFFECT}} = \beta_1 \text{BSC}_\text{OUTCOME} + \beta_2 \text{RESOURCE} + \beta_3 \text{EMP}_\text{EDU} + \beta_4 \text{LENGTH}_\text{POSN} + e \)
Model 7: \( SYS_{\text{EFFECT}} = \beta_1 \text{BSC}_\text{ELG} + \beta_2 \text{RESOURCE} + \beta_3 \text{EMP}_\text{EDU} + \beta_4 \text{LENGTH}_\text{POSN} + e \)

**Variable definitions:**
- **SYS_EFFECT** = Systems effectiveness
- **BSC_INP** = Inputs – activities planned
- **BSC_CEQ** = Cost/process efficiency and quality measures
- **BSC_ACT** = Activities/process measures
- **BSC_COMM** = Customer/community satisfaction measures
- **BSC_OUTPUT** = Outputs (quantity of services provided)
- **BSC_OUTCOME** = Outcomes measures
- **BSC_ELG** = Employee learning and growth measures
- **RESOURCE** = Log of the annual budget
- **EMP_EDU** = Level of employee education
- **LENGTH_POSN** = Length of position the participants held (log).
Table 8  
Zero-order and partial correlations for BSC_USE with other variables

<table>
<thead>
<tr>
<th>Control Variables</th>
<th>BSC_INUSE</th>
<th>BSC_EXUSE</th>
<th>SYS_EFFECT</th>
<th>RESOURCE</th>
<th>EMPL_EDU</th>
<th>LENGTH_POSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>None*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSC_INUSE</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSC_EXUSE</td>
<td>0.53**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYS_EFFECT</td>
<td>0.73**</td>
<td>0.46**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RESOURCES</td>
<td>0.12</td>
<td>0.05</td>
<td>-0.01</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMPL_EDU</td>
<td>0.08</td>
<td>0.23</td>
<td>-0.17</td>
<td>0.16</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>TIME_POSITN</td>
<td>-0.16</td>
<td>-0.23</td>
<td>0.01</td>
<td>-0.40*</td>
<td>-0.08</td>
<td>1.00</td>
</tr>
<tr>
<td>RESOURCES, EMPL_EDU &amp; TIME_POSITN</td>
<td>BSC_INUSE</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SYS_EFFECT</td>
<td>0.76**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BSC_EXUSE</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SYS_EFFECT</td>
<td>0.56**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Cells contain zero-order (Pearson) correlations  
**p < 0.01; * p < 0.05 (Two-tailed)

**Definition of Variables:**  
BSC_INUSE = Internal usage of balanced scorecard (BSC) performance measures  
BSC_EXUSE = External institutional usage of balanced scorecard (BSC) performance measures  
SYS_EFFECT = Systems effectiveness  
RESOURCE = Annual fiscal (operating) budget  
EMPL_EDU = Employee education  
LENGTH_POSN = Length of employment in the current position
Table 9
Multiple regression results (H2)
Dependent variable = SYS_EFFECT against BSC_INUSE

SYS_EFFECT = β₀ + β₁BSC_INUSE + β₂RESOURCE + β₃EMPL_EDU + β₄LENGTH_POSN + e

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Estimate</th>
<th>Standard error</th>
<th>t-value</th>
<th>Sig. (p)</th>
<th>Tolerance</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>α₀ Intercept</td>
<td>α₀</td>
<td>3.77</td>
<td>2.37</td>
<td>1.59</td>
<td>0.06</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>X₁ BSC_INUSE</td>
<td>β₁</td>
<td>0.75</td>
<td>0.09</td>
<td>5.78</td>
<td>0.00</td>
<td>0.96</td>
<td>1.04</td>
</tr>
<tr>
<td>X₂ RESOURCE</td>
<td>β₂</td>
<td>-0.01</td>
<td>0.31</td>
<td>-0.06</td>
<td>n.s.</td>
<td>0.82</td>
<td>1.23</td>
</tr>
<tr>
<td>X₃ EMP_EDU</td>
<td>β₃</td>
<td>0.25</td>
<td>1.08</td>
<td>1.88</td>
<td>0.07</td>
<td>0.97</td>
<td>1.03</td>
</tr>
<tr>
<td>X₄ LENGTH_POSN</td>
<td>β₄</td>
<td>0.11</td>
<td>0.47</td>
<td>0.76</td>
<td>n.s.</td>
<td>0.78</td>
<td>1.28</td>
</tr>
</tbody>
</table>

R² = 0.575; Adjusted R² = 0.510; F [4, 26] = 8.81; p = 0.00

**Variable Definitions:**
- **SYS_EFFECT** = Systems effectiveness
- **BSC_INUSE** = Internal usage of balanced scorecard (BSC) measures
- **RESOURCE** = Log of the annual budget
- **EMPL_EDU** = Level of employee education
- **LENGTH_POSN** = Length of position the participants held (log)

n.s = Not significant; n/a = Not available
Table 10
Additional regression analysis for SYS_EFFECT (the dependent variable) with six internal usages of BSC performance measures

<table>
<thead>
<tr>
<th>Regression statistics</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BSC_INUSE_PRO</strong></td>
<td>0.37</td>
<td>0.43</td>
<td>0.45</td>
<td>0.38</td>
<td>0.34</td>
<td>0.52</td>
</tr>
<tr>
<td><strong>BSC_INUSE_ACT</strong></td>
<td>0.70</td>
<td>0.65</td>
<td>0.60</td>
<td>0.71</td>
<td>0.72</td>
<td>0.53</td>
</tr>
<tr>
<td><strong>BSC_INUSE_BDG</strong></td>
<td>5.16</td>
<td>4.53</td>
<td>3.46</td>
<td>5.13</td>
<td>5.20</td>
<td>3.104</td>
</tr>
<tr>
<td><strong>BSC_INUSE_ACTION</strong></td>
<td>0.481</td>
<td>0.425</td>
<td>0.302</td>
<td>0.498</td>
<td>0.494</td>
<td>0.292</td>
</tr>
<tr>
<td><strong>BSC_INUSE_SPLAN</strong></td>
<td>0.409</td>
<td>0.343</td>
<td>0.202</td>
<td>0.423</td>
<td>0.421</td>
<td>0.183</td>
</tr>
<tr>
<td><strong>BSC_INUSE_RP</strong></td>
<td>6.72</td>
<td>5.18</td>
<td>3.03</td>
<td>6.69</td>
<td>6.83</td>
<td>2.68</td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
<td>0.005</td>
</tr>
<tr>
<td><strong>Adj. R²</strong></td>
<td>0.401</td>
<td>0.343</td>
<td>0.202</td>
<td>0.423</td>
<td>0.421</td>
<td>0.183</td>
</tr>
<tr>
<td><strong>F-value</strong></td>
<td>6.72</td>
<td>5.18</td>
<td>3.03</td>
<td>6.69</td>
<td>6.83</td>
<td>2.68</td>
</tr>
<tr>
<td><strong>Sig.</strong></td>
<td>0.001</td>
<td>0.003</td>
<td>0.034</td>
<td>0.001</td>
<td>0.001</td>
<td>0.054</td>
</tr>
</tbody>
</table>

Model 1: SYS_EFFECT = β₁BSC_INUSE_PRO + β₂RESOURCE + β₃EMP_EDU + β₄LENGTH_POSN + e
Model 2: SYS_EFFECT = β₁BSC_INUSE_ACT + β₂RESOURCE + β₃EMP_EDU + β₄LENGTH_POSN + e
Model 3: SYS_EFFECT = β₁BSC_INUSE_BDG + β₂RESOURCE + β₃EMP_EDU + β₄LENGTH_POSN + e
Model 4: SYS_EFFECT = β₁BSC_INUSE_ACTION + β₂RESOURCE + β₃EMP_EDU + β₄LENGTH_POSN + e
Model 5: SYS_EFFECT = β₁BSC_INUSE_SPLAN + β₂RESOURCE + β₃EMP_EDU + β₄LENGTH_POSN + e
Model 6: SYS_EFFECT = β₁BSC_INUSE_RP + β₂RESOURCE + β₃EMP_EDU + β₄LENGTH_POSN + e

Variable Definitions:
SYS_EFFECT = Systems effectiveness; BSC_INUSE_PRO = Measure program performance; BSC_INUSE_ACT = Manage an activity or program; BSC_INUSE_BDG = Budget formulation; BSC_INUSE_ACTION = Taking actions based on results; BSC_INUSE_SPLAN = Strategic planning; BSC_INUSE_RP = Punish or reward staff; RESOURCES = Log of the annual budget EMP_EDU = Level of employee education; TIME_POSITN = Length of position the participants held (log).
References


