



10-2003


Performance Implications of Strategic Performance Measurement in Financial Services Firms

Christopher D. Ittner
University of Pennsylvania

David F. Larcker

Taylor Randall

Follow this and additional works at: https://repository.upenn.edu/accounting_papers

 Part of the [Accounting Commons](#), [Finance and Financial Management Commons](#), and the [Management Sciences and Quantitative Methods Commons](#)

Recommended Citation

Ittner, C. D., Larcker, D. F., & Randall, T. (2003). Performance Implications of Strategic Performance Measurement in Financial Services Firms. *Accounting, Organizations and Society*, 28 (7-8), 715-741. [http://dx.doi.org/10.1016/S0361-3682\(03\)00033-3](http://dx.doi.org/10.1016/S0361-3682(03)00033-3)

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/accounting_papers/42
For more information, please contact repository@pobox.upenn.edu.

Performance Implications of Strategic Performance Measurement in Financial Services Firms

Abstract

This study examines the relation between measurement system satisfaction, economic performance, and two general approaches to strategic performance measurement: greater measurement diversity and improved alignment with firm strategy and value drivers. We find consistent evidence that firms making more extensive use of a broad set of financial and (particularly) non-financial measures than firms with similar strategies or value drivers have higher measurement system satisfaction and stock market returns. However, we find little support for the alignment hypothesis that more or less extensive measurement than predicted by the firm's strategy or value drivers *adversely* affect performance. Instead, our results indicate that greater measurement emphasis and diversity than predicted by our benchmark model is associated with higher satisfaction and stock market performance. Our results also suggest that greater measurement diversity relative to firms with similar value drivers has a stronger relationship with stock market performance than greater measurement on an absolute scale. Finally, the balanced scorecard process, economic value measurement, and causal business modeling are associated with higher measurement system satisfaction, but exhibit almost no association with economic performance.

Disciplines

Accounting | Finance and Financial Management | Management Sciences and Quantitative Methods

**Performance Implications of Strategic Performance Measurement
in Financial Services Firms**

Christopher D. Ittner*

David F. Larcker*

Taylor Randall**

Revised March, 2003

* The Wharton School, University of Pennsylvania, 2400 Steinberg Hall-Dietrich Hall,
Philadelphia, PA 19104-6365

* Eccles School of Business, University of Utah, 1645 E. Campus Center, Salt Lake City,
UT 84112

We thank the Cap Gemini Ernst & Young Center for Business Innovation and Wharton Research Program for Value Creation in Organizations for their assistance in collecting the survey data used in this paper. We also thank Sudhakar Balachandran, Marlene Plumlee, participants at the 2002 AAA Management Accounting Conference, and two anonymous reviewers for their comments.

Performance Implications of Strategic Performance Measurement in Financial Services Firms

Abstract

This study examines the relation between measurement system satisfaction, economic performance, and two general approaches to strategic performance measurement: greater measurement diversity and improved alignment with firm strategy and value drivers. We find consistent evidence that firms making more extensive use of a broad set of financial and (particularly) non-financial measures than firms with similar strategies or value drivers have higher measurement system satisfaction and stock market returns. However, we find little support for the alignment hypothesis that more or less extensive measurement than predicted by the firm's strategy or value drivers *adversely* affect performance. Instead, our results indicate that greater measurement emphasis and diversity than predicted by our benchmark model is associated with higher satisfaction and stock market performance. Our results also suggest that greater measurement diversity relative to firms with similar value drivers has a stronger relationship with stock market performance than greater measurement on an absolute scale. Finally, the balanced scorecard process, economic value measurement, and causal business modeling are associated with higher measurement system satisfaction, but exhibit almost no association with economic performance.

Performance Implications of Strategic Performance Measurement in Financial Services Firms

1. Introduction

Managerial accounting is evolving to encompass a more strategic approach that emphasizes the identification, measurement, and management of the key financial and non-financial drivers of strategic success and shareholder value (International Federation of Accountants, 1998; Institute of Management Accountants, 1999). In response, many firms are adopting strategic performance measurement (SPM) systems that (1) provide information that allows the firm to identify the strategies offering the highest potential for achieving the firm's objectives, and (2) align management processes, such as target setting, decision-making, and performance evaluation, with the achievement of the chosen strategic objectives (e.g., Gates, 1999; Otley, 1999).

Proponents of strategic performance measurement advocate two general approaches for developing SPM systems. The simplest approach calls for firms to measure and use a diverse set of financial and non-financial measures. Advocates of this "measurement diversity" approach argue that a broad set of measures keeps managers from suboptimizing by ignoring relevant performance dimensions or improving one measure at the expense of others. As a result, these advocates claim that firms achieve higher performance when they place greater emphasis on a broad set of financial and non-financial performance measures (e.g., Lingle and Schiemann, 1996). A second approach is based on contingency theory, which argues that strategic performance measures must be aligned with the firm's strategy and/or value drivers (Fisher, 1995b; Langfield-Smith, 1997). Under this approach, performance theoretically is enhanced

when “measurement gaps” between the firm’s strategic priorities and measurement practices are minimized. Thus, performance is expected to be lower when the SPM system places either less *or* more emphasis on a measurement practice than the level required by the firm’s strategy and value drivers.

Closely related to the contingency perspective is the use of measurement techniques such as the balanced scorecard process, causal business modeling, and economic value measurement. Advocates argue that these techniques help companies improve the alignment between their performance measurement systems more closely to their organizational objectives (Kaplan and Norton, 1992, 1996, 2001; Gates, 1999; Stewart, 1991; Young and O’Byrne, 2001). Despite these arguments, the extent to which firms claiming to use these techniques actually link their performance measures more closely to strategic priorities is unknown.

Using data from 140 U.S. financial services firms, we extend prior research on the performance implications of strategic performance measurement along several dimensions. First, we examine a broader set of measurement system uses (goal setting, capital investment decisions, identification of improvement opportunities and development of action plans, performance evaluation, and external disclosure) and measurement capabilities than prior studies that typically focus only on performance evaluation and compensation. Second, we investigate the relation between SPM practices and actual financial outcomes (accounting and stock returns) rather than relying exclusively on *self-reported* measurement satisfaction or firm performance. Third, we examine each of the SPM approaches and compare their relative ability to explain firm performance. Fourth, we extend prior contingency research by looking at the alignment

between specific value drivers and measurement, in addition to the traditional alignment with firm or manufacturing strategy. Fifth, we provide evidence on the use and performance consequences of the three measurement alignment techniques (balanced scorecard, economic value measurement, and business modeling), an area that has received surprisingly little attention in the research literature. Finally, we examine potential lags between the implementation of performance measurement systems and economic results.

We find consistent evidence that SPM practices are associated with one- and three-year stock returns, but not with our two accounting measures (return on assets and sales growth). In particular, financial services firms that make more extensive use of a broad set of financial and (particularly) non-financial measures than those with similar strategies or value drivers earn higher stock returns. These results are even stronger in the subsample of firms with more mature performance measurement systems, suggesting that these measurement practices yield economic results with some lag.

We find little support for the hypothesis that more or less extensive measurement than predicted by the firm's strategy or value drivers *adversely* affect performance. Instead, our results indicate that *greater* measurement emphasis and diversity than predicted by our benchmark model is associated with *higher* satisfaction and stock market performance. These findings suggest that the average measurement practices of firms pursuing similar strategies or value drivers currently are not optimal in this industry.

We also find that greater measurement diversity compared to firms with similar strategies or value drivers has a stronger relationship with stock market performance than

greater overall measurement. This evidence suggests that the appropriate benchmark for assessing measurement diversity is greater measurement *relative* to competitors with similar strategies or value drivers rather than greater measurement on an *absolute* scale.

On average, firms claiming to use a balanced scorecard exhibit few differences in their emphasis on non-financial performance categories than non-users, and make little use of the causal “business models” of leading and lagging indicators that balanced scorecard advocates claim is a foundation of the scorecard process. In contrast, economic value and business model users place significantly greater emphasis on non-financial value drivers and measures than firms that do not use these practices. Although balanced scorecard, economic value, and causal business model users all rate satisfaction with their measurement systems higher than non-users, we find almost no evidence that these techniques are associated with accounting or stock market performance.

The remainder of the paper is organized as follows. The following section reviews related literature and develops our hypotheses. We then discuss our sample selection and measurement methods. The next section reports our contemporaneous performance results, followed by our analysis of lagged performance effects in the subset of more mature systems. A summary of our results and limitations to our study conclude.

2. Literature Review and Hypotheses

A study by the Conference Board defines strategic performance measurement (SPM) as a system that “translates business strategies into deliverable results. SPM systems combine financial, strategic, and operating business measures to gauge how well a company meets its targets (Gates, 1999, p. 4). Case studies by Gates (1999) and Ruddle and Feeny (2000) find that firms have adopted two general approaches to designing SPM

systems. In the following sections, we review each of these approaches and develop hypotheses regarding their performance implications.

2.1 Greater “Measurement Diversity”

One approach to strategic performance measurement is supplementing traditional financial measures with a diverse mix of non-financial measures that are expected to capture key strategic performance dimensions that are not accurately reflected in short-term accounting measures. Fisher (1995a) and Brancato (1995) indicate that many firms believe that financial measures are too historical and “backward-looking,” lack predictive ability to explain future performance, reward short-term or incorrect behavior, provide little information on root causes or solutions to problems, and give inadequate consideration to difficult to quantify “intangible” assets such as intellectual capital. As a result, many firms are supplementing financial metrics with a diverse set of non-financial performance measures that are believed to provide better information on strategic progress and success.¹

Early writings on the balanced scorecard provided considerable impetus behind this “measurement diversity” approach to SPM. Kaplan and Norton’s (1992) original balanced scorecard article argues that firms should supplement financial measures with non-financial measures focused on three other perspectives: customers, internal business processes, and learning and growth. Although the authors note that the *specific* measures within these categories should be tailored to the firm’s strategy, they also claim that these broad perspectives are common across (and should be used with) *all* strategic choices.

¹ Consistent with these claims, a number of accounting studies provide evidence that non-financial measures can be leading indicators of financial performance (e.g., Ittner and Larcker, 1997; Behn and Riley, 1999; Banker et al., 2000; Nagar and Rajan, 2001; Ittner et al., 2003). However, only Banker et al.

Other writers have extended these categories by adding additional perspectives focused on employees, partners and suppliers, and the environment (e.g., Edvinsson and Malone, 1997; Schiemann and Lingle, 1999). An assumption in these classification schemes is that measurement across all these categories is necessary regardless of the firm's strategy. Schiemann and Lingle (1999, p. 8), for example, write:

For us, measuring the “right things” entails measuring results in the six performance areas that are key to strategic success. And when we use the term “strategic measurement,” we mean measurement focused on these six perspectives or areas of performance.

Several empirical studies implicitly or explicitly draw on the measurement diversity approach in their tests. A widely-cited practitioner-oriented study by Lingle and Schiemann (1996) reports that “measurement-managed” firms (defined as those in which management updates and reviews semiannual performance measures in three or more of their six primary performance categories, and where senior management reports being in agreement on measurable criteria for determining strategic success) achieve statistically higher self-reported industry standing, financial performance relative to competitors, and progress in managing change efforts than firms that are not “measurement-managed.”²

Scott and Tiessen's (1999) academic study indicates that work teams having more diverse performance measures (i.e., both financial and non-financial measures or more categories of measures) achieve higher self-assessed performance (relative to expectations). Hoque and James (2000) also find a significant positive relation between perceived organizational performance and the use of a diverse set of performance

(2000) and Ittner et al. (2003) examine the use of these measures for performance evaluation or decision-making purposes.

² Although Lingle and Schiemann (1996) provide evidence on the extent of “measurement gaps” in their sample, they do not use these gaps in their performance tests, instead relying on the level of measurement in their performance categories.

measures related to the four balanced scorecard categories. However, all of these studies place heavy reliance on perceptual results indicators and/or simple univariate tests, making it difficult to place substantive interpretations on their results.

Theoretical research on the economic benefits from greater measurement diversity is ambiguous. Nearly all of this work focuses on reward systems. Holmstrom (1979) and Banker and Datar (1989) show that more (costless) measures are preferred to fewer if the additional measures provide incremental information on dimensions of the agent's actions that the principal wishes to motivate. In addition, Feltham and Xie (1994), Hauser et al. (1994), and Hemmer (1996) demonstrate how incentives based on non-financial measures can improve contracting by incorporating information on managerial actions that is not fully captured in contemporaneous financial results. However, these theories also imply that a performance category (e.g., the four balanced scorecard perspectives) may *not* improve incentive contracting if it provides no incremental information on the manager's action, imposes too much risk on the agent, or is too costly to measure, indicating that the measurement diversity approach to SPM may have no impact on economic performance.

The limited and mixed evidence on the performance implications of this SPM approach leads us to examine our first hypothesis:

H1: Organizational performance is positively associated with the extent to which the firm measures and uses information related to a diverse set of financial and non-financial performance measures.

2.2 Alignment with Strategy and Value Drivers

A second general approach emphasizes the implementation of performance measurement systems that are more closely linked to the firm's *specific* strategy and value drivers. Contingency theory has long held that control systems must be aligned with organizational characteristics such as firm strategy (see Fisher, 1995b for a review). Similarly, economic theories contend that the optimal design of a firm's information and reward systems is a function of the firm's business strategy (e.g., Milgrom and Roberts, 1992; Brickley et al., 1997).

More recently, advocates of this SPM approach have extended these theories to argue that a key element in managing the links between strategy and performance is identifying and measuring the specific factors, or "value drivers," that actually *lead* to strategic success or firm value (Ittner and Larcker, 2001). By linking strategies to their underlying value drivers, and tying information systems, goals and objectives, resource allocation, and performance evaluation to these drivers, SPM systems are expected to improve communication of the *specific* actions required to achieve the chosen strategy, motivate performance against strategic value driver goals, and provide more rapid feedback on whether the strategy is achieving its objectives. In addition, the SPM literature increasingly argues that the value driver analysis should not only influence the design and use of measurement systems, but should also affect external disclosure requirements (e.g., Black et al., 1998; Gates, 1999; Eccles et al., 2001). This use of SPM systems is consistent with calls in the financial accounting community for greater disclosure of information on key value drivers (e.g., American Institute of Certified Public Accountants, 1994; Wallman, 1995, Financial Accounting Standards Board, 2001).

Academic research on contingency approaches to strategic performance measurement falls into three categories. Early studies focused on the influence of perceived environmental uncertainty (PEU) on management accounting systems. According to this literature, PEU relates to the extent to which the firm's competitive environment is highly dynamic and unpredictable, factors that are likely to be highly correlated with the extent to which the firm's strategy is focused on innovation and growth (Dent, 1990; Langfield-Smith, 1997). Larcker (1981) and Gordon and Narayanan (1984), for example, examine the relation between PEU and three performance measurement system attributes: focus (internal vs. external measures), quantification (financial vs. non-financial measures), and time horizon (historical vs. future-oriented), with mixed results. Larcker's (1981) study of strategic capital budgeting decisions finds no association between environmental characteristics and variations in the perceived importance of these measurement attributes. In contrast, Gordon and Narayanan (1984) find a significant positive association between PEU and the perceived importance of externally-oriented, non-financial, and ex ante information. Chenhall and Morris (1986) examine the association between PEU and the perceived usefulness of four management accounting system attributes: scope (i.e., the external, non-financial, and future-oriented attributes examined in the two previous studies), timeliness, aggregation (i.e., the level of aggregation by time period and functional area and the use of analytical or decision models), and integration (i.e., the setting of precise targets for activities and their interdependencies and the reporting of intra-sub-unit-interactions). They find significant positive associations between PEU and preferences for broader scope and more timely information, but no

associations with aggregation and integration preferences. None of these studies investigates performance consequences.

More recent studies directly examine the effects of organizational strategy on performance measurement choices, and the relation between these choices and organizational performance. These studies typically measure strategy as a continuum between firms following a “defender,” “harvest,” or “cost leadership” strategy and firms following a “prospector,” “build,” or “innovation” strategy. The majority of these studies find significant relations between the organization’s strategy and performance measurement system, with a smaller set of studies also finding higher organizational performance when measurement is more closely aligned with the chosen strategy. Simons (1987) and Govindarajan (1988), for example, find higher performance in organizations following defender or low cost strategies when bonuses are based on budget targets. Govindarajan and Gupta (1985) also find that greater reliance on non-financial compensation criteria has a stronger positive impact in organizations following a build strategy than in those following a harvest strategy. Studies by Abernethy and Guthrie (1994), Chong and Chong (1997), and Bouwens and Abernethy (2000), among others, generally support the hypothesis that broad scope performance measurement systems are associated with higher (self-reported) performance in companies following prospector or differentiation strategies.

A third set of studies provides evidence on the associations between specific manufacturing strategies or value drivers (such as quality and flexibility), the choice of performance measures, and manufacturing performance. This research finds systematic links among these choices, with an organization’s emphasis on inventory reduction or

just-in-time production, quality, or manufacturing flexibility positively associated with the provision of non-financial measures and goals and greater emphasis on non-financial measures in reward systems. However, empirical support for the hypothesized performance benefits from these measurement practices is mixed, with some studies finding positive performance effects when manufacturing strategies are aligned with measurement systems (e.g., Abernethy and Lillis, 1995), others finding mixed results depending upon the type of performance measurement attribute being examined or the extent to which other manufacturing improvement practices are implemented (e.g., Sim and Killough, 1998; Ittner and Larcker, 1997), and others finding no significant associations (e.g., Young and Selto, 1991; Perera et al., 1997).

Although the contingency-based performance measurement literature is relatively extensive, it generally looks at only one or a few strategies or value drivers at a time, does not compare results across various SPM approaches (e.g., measurement diversity vs. strategy alignment vs. value driver alignment), and relies quite heavily on respondents' self-reported performance. As a result, we extend these studies by providing further evidence on the following hypotheses:

H2: Organizational performance is positively associated with the extent to which performance measurement practices are aligned with the firm's strategy.

H3: Organizational performance is positively associated with the extent to which the performance measurement practices are aligned with the firm's value drivers.

2.3 Measurement Alignment Techniques

The performance measurement literature proposes several techniques that are claimed to improve the alignment between performance measurement systems and the firm's organizational objectives. These techniques include the balanced scorecard process,

economic value measurement, and causal business models. Kaplan and Norton's (1996, 2001) recent writings on the balanced scorecard, for example, define this process as a method for using financial and non-financial measures to communicate the multiple, *linked* objectives that a firm must achieve to satisfy its mission and reach its long-term strategic goals. According to these authors, balanced scorecard systems improve performance by translating strategy into specific objectives and measures that are linked in a causal chain of leading and lagging indicators covering the four scorecard perspectives (financial, customer, internal business process, and learning and growth).

Economic value-based SPM systems focus on the creation of long-term shareholder value through the use of residual income or cashflow-related measures. Advocates of economic value techniques (often referred to as value-based management approaches) argue that performance measurement systems should be aligned with the firm's ultimate organizational objective: improved economic performance. Stern et al. (1995), for example, maintain that effective performance measurement requires the firm to make economic value measures the cornerstone of a total management system that focuses on shareholder value enhancement for capital budgeting, goal setting, investor communication, and compensation. The resulting performance measurement system is expected to improve alignment between performance measures and strategic objectives by requiring firms to choose internal objectives that lead to shareholder value enhancement, select strategies to achieve these objectives, identify the value drivers that actually create value for the firm, and select action plans, lower-level performance

measures, and targets based on the priorities identified in the value driver analysis (Dixon and Hedley, 1993; Stern et al., 1995; Copeland et al., 1996).⁴

A third alignment technique that is stressed in both the balanced scorecard and economic value literatures is the formal development of causal “business models.” Eccles (1991), for example, argues that effective performance measurement system design requires firms to first understand the causal model linking key success factors to the ultimate objectives of the firm. Kaplan and Norton (1996, 2001) now state that a critical element of the balanced scorecard process is the development of “strategic maps” that embed strategy in a system of cause-and-effect relations that connects desired strategic outcomes with the drivers expected to lead to these outcomes. Similarly, Dixon and Hedley (1993), Copeland et al. (1996), and Young and O’Byrne (2001) stress the importance of linking financial performance measures and their non-financial value drivers to achieve the benefits from economic value measurement programs and promote value-creating behavior in the firm.⁵ Despite this emphasis on causal business models, Gates’ (1999) study indicates that most firms adopt SPM systems without articulating these causal models or maps.

Surprisingly little research has been conducted on the performance implications of the proposed measurement alignment techniques. Surveys on the use of balanced scorecards (Towers Perrin, 1996; Chenhall and Langfield-Smith, 1998; Banker et al., 2001a, Rigby,

⁴ Forty-one percent of the firms in Gates’ (1999) study describe their SPM systems as value-based, while 40 percent describe their systems as following a balanced scorecard approach. Although many firms view their economic value measurement systems as strategic, others argue that these systems do not actually represent SPM systems because they focus on a single outcome (shareholder value) rather than on the strategies used to achieve this outcome. In addition, some critics of economic value measures charge that these metrics drive managers to focus on short-term operational and financial performance, to the detriment of longer-term investments in customers, innovation, and employee capabilities (e.g., Kaplan and Norton, 2001, pp. 378-379). We examine these assertions later in the paper.

2001; Sandt et al., 2001; Kaplan and Norton, 2001, p. 357) and economic value measures (Ittner and Larcker, 1998) typically find moderately greater satisfaction with or perceived performance from these techniques than from other performance measurement practices. In contrast, Ittner et al.'s (2003) examination of a balanced scorecard bonus plan in a large bank indicates that the scorecard plan was deemed unsuccessful, leading it to be abandoned in favor of a revenue-based incentive plan. Sandt et al. (2001) and Banker et al. (2001b) also find greater satisfaction with performance measurement systems when systematic linkages among performance measures are understood and articulated. However, none of these studies directly examines the association between these techniques and actual firm performance.

Wallace (1997) and Hogan and Lewis (1999) reach conflicting conclusions regarding the performance of firms adopting compensation plans based on economic value measures. Wallace (1997) finds economic value adopters, relative to a control sample of non-adopters, decrease new investments, increase payouts to shareholders through share repurchases, and utilize assets more intensively, leading to significantly greater change in residual income. In contrast, Hogan and Lewis (1999) find no significant difference in economic value users and non-users after matching control firms on *past* performance to control for mean reversion in performance levels.

The claimed benefits from the measurement alignment techniques, together with the limited evidence to support these claims, lead to our final hypothesis:

H4: Organizational performance is positively associated with the use of balanced scorecards, economic value measures, and causal business models.

⁵ See Heskett (1994) for examples of causal business models linking employee measures, customer measures, and financial performance.

3. Data and Variables

3.1 Sample

We test these hypotheses using a sample of U.S. financial services firms that responded to a survey conducted by the authors in conjunction with the Cap Gemini Ernst & Young Center for Business Innovation. In contrast to most prior research, we restrict our sample to a single industry. An important advantage of this choice is that we can implicitly control for the myriad of confounding variables that can substantively impact any results from a multi-industry, cross-sectional study. The financial services industry was selected because our field research on performance measurement innovations indicates that financial service firms are actively debating their choice of value drivers and performance measures. Although restricting the sample to a single industry limits our ability to generalize the results, we believe that a single industry analysis has substantially higher internal validity than a multi-industry analysis.

A random sample of 600 firms was solicited to participate in the survey. A marketing research firm telephoned senior executives from each of these firms to request participation. Those agreeing to participate were sent a survey or guided to a web site containing the questionnaire. Executives from 140 firms (23.3%) completed usable surveys during November of 1999. The respondents represent a variety of financial service sectors, including regional banks (33.3% of the sample), insurance companies (21.4%), diversified financial firms (17.1%), savings and loans (14.3%), money center banks (5.7%), and others (e.g., consumer finance, investment banking, etc.) (7.1%). Relative to the population of financial services firms on Compustat, the resulting sample

contains a larger percentage of diversified financial firms, life insurance companies, and banks, and a smaller percentage of investment banks and savings and loans.

3.2 *Strategy*

We asked respondents to evaluate 12 aspects of the company's organizational strategy and corporate environment that are commonly used to measure strategy and perceived environmental uncertainty. Principal components analysis of these questions (with oblique rotation) reveals three factors with eigenvalues greater than one. The factors capture the extent to which the firm's strategy focuses on (1) *innovation*, (2) *flexibility* in changing its product and service offerings and responding to market demands, and (3) *maintaining* current relationships and product offerings by pursuing existing customers and markets in stable environments.

The three strategy constructs represent the average standardized response to each question loading greater than 0.40 on these factors.⁶ *Flexible* equals the average answer to four questions asking the respondent's agreement with the statements, "We respond rapidly to early signals of opportunity in our market," "We have greater flexibility to respond to changes in our environment than our competitors," "We have the ability to adjust capacity within a short period of time," and "We have the ability to change product or service offerings rapidly" (scales ranging from 1 = strongly disagree to 6 = strongly agree).

Innovate equals the average standardized response to four questions asking the respondent's agreement with the statements, "We offer a more expanded range of products and services than our competitors," "We are first to market with new products or services," "We respond rapidly to early signals of opportunity in our market," and "We expect most

⁶ One question asking whether the firm is more cost efficient than its competitors did not load greater than 0.40 on any of the factors. This question is dropped from the analysis.

of our future growth in profits to come from our new product and service offerings.”

Maintain equals the average standardized response to three questions asking the respondent’s agreement with the statements, “We are most active in developing the markets we currently serve, rather than entering new markets with our products or services,” “We operate in markets for our products or services that are highly predictable,” and “It is easy to forecast how actions of competitors will affect the performance of our organization.” Coefficient alphas are 0.66 for *Innovate*, 0.75 for *Flexible*, and 0.46 for *Maintain*. The Pearson (Spearman) correlation between *Innovate* and *Flexible* is 0.49 (0.50), between *Innovate* and *Maintain* is 0.18 (0.15), and between *Flexible* and *Maintain* is 0.17 (0.15).⁷

3.3 Value Drivers

Respondents were also asked the extent to which various performance categories are important drivers of their firms’ *long-term* organizational success, on a scale ranging from 1 (not at all important or not applicable to their organization) to 6 (extremely important). The ten value driver categories include short-term financial performance (e.g., annual earnings, return on assets, cost reduction), customer relations (e.g., market share, customer satisfaction, customer retention), employee relations (e.g., employee satisfaction, turnover, workforce capabilities), supplier relations (e.g., on-time delivery, input into product/service design), operational performance (e.g., productivity, safety, cycle time), product and service quality (e.g., defect rates, quality awards), alliances with other organizations (e.g., joint marketing or product design, joint ventures), environmental performance (e.g., government citations, environmental compliance or

⁷ Despite the significant correlations, in no case does the Variance Inflation Factor (VIF) exceed 2.5 in any of our regression models, suggesting no serious problems with multicollinearity in our tests. Since one question cross-loaded on the *Innovate* and *Flexible* factors, we repeated our analyses after dropping this question from the strategy constructs. This change had no substantive impact on our results.

certification), product and service innovation (e.g., new product or service development success, development cycle time), and community (e.g., public image, community involvement). These ten categories are drawn from value driver discussions in the balanced scorecard, intangible asset, intellectual capital, and value-based management literatures (e.g., Kaplan and Norton, 1996, 2001; Edvinsson and Malone, 1997; Schiemann and Lingle, 1999).

Table 1 provides the mean importance score given to each of the value driver categories. Short-term financial performance ranks only fourth most important, behind customer relations, product and service quality, and operational performance. Innovation, community relations, and employee relations also receive relatively high importance scores, with environmental performance and supplier relations believed to be relatively unimportant in this industry.

Spearman correlations among the strategy and value driver measures are shown in Table 2. Correlations between the strategy constructs and individual value drivers are generally small (mean = 0.132, median = 0.125). Two value driver categories do not vary significantly with any of the strategies: short-term financial performance and operational results. *Innovate* is positively correlated ($p < 0.10$, two-tailed) with all of the remaining value drivers except community, with a mean (median) correlation of 0.188 (0.145). *Flexible* is positively correlated with the importance placed on quality, alliances, suppliers, and innovation, but not with the other performance categories. Firms following a *Maintain* strategy place greater emphasis on community and the environment, but this strategy is not significantly associated with the other value drivers categories. Overall, the relatively small correlations in Table 2 suggest that the firm's

chosen strategies are not synonymous with the value drivers used to achieve these strategies. For example, the evidence suggests that some firms attempt to implement an *Innovate* strategy by placing greater reliance on supplier expertise, while others focus on improving innovation by developing their own, internal employee capabilities. These results suggest that measuring a firm's overall strategy without considering its chosen value drivers provides an incomplete representation of strategic attributes.

3.4 Performance Measurement Practices

We asked respondents the extent to which their firms use each of the ten value driver categories for: (1) identifying problems and improvement opportunities and developing action plans, (2) evaluating major capital investments, (3) evaluating managerial performance, and (4) disclosing information to external parties (e.g., via fact books, analyst meetings, conference calls, press releases, company internet websites, and one-on-one meetings). Scales for these questions range from 1 (not at all important or not used at all) to 6 (used extensively or complete external disclosure). We also asked respondents to rate how well their organizations measure information in these categories (1 = extremely poor quality of measurement to 6 = high quality of measurement) and the extent to which goals are set (1 = not applicable or no goals set to 6 = explicit goals set).⁸

Table 1 provides information on the consistency between the perceived importance of the individual value driver categories and the corresponding use and quality of

⁸ Although the survey did not ask respondents the factors they considered when assessing measurement quality, Cavalluzzo and Ittner (2003) find that managers' perceptions of measurement system quality are a function of data limitations (i.e., the ability of existing information systems to provide valid, reliable, and timely data in a cost effective manner) and difficulties selecting and interpreting appropriate performance measures. Thus, responses to this question are likely to reflect performance measurement attributes other than scope, such as timeliness and information system effectiveness, that have been identified in prior studies (e.g., Chenhall and Morris, 1986; Gates, 1999). The goal setting question, in turn, is consistent with the target setting component of Chenhall and Morris' (1986) system integration construct, as well as Simons' (1987) examination of the emphasis on budgetary and output goals in different strategic contexts.

performance measures related to these categories. With the exception of short-term financial performance, the importance scores for each performance category are lower than the scores provided for the use and quality of related performance measures. These differences are consistent with the “measurement gaps” identified in other studies (e.g., Dixon et al., 1990; Lingle and Schiemann, 1996; Stivers et al., 1998). The differences vary across uses, indicating that extensive use of performance measures for one purpose does not necessarily imply that the measures are used for other purposes. The largest differences relate to the external disclosure of customer and quality information, the use of employee information for evaluating capital investments, and the quality of customer-related measures. Differences related to identifying problems and developing action plans generally are smaller than those associated with other uses.

3.5 Use of Measurement Alignment Techniques

The use of performance measurement alignment techniques is assessed using three questions on the implementation of balanced scorecards and economic value measures (e.g., economic value added or cash flow return on investment) and the extent of *formal* reliance on a “business model” or “theory of business” that causally links performance drivers to performance outcomes. Following Krumwiede (1998), a six point scale is used to measure the implementation of balanced scorecards or economic value measures, where 1 = not considered, 2 = implemented and abandoned, 3 = considering, 4 = implementing now, 5 = used, and 6 = used extensively. For our analyses, we code scorecard or economic value measure use one if the respondent stated that the firm uses or extensively uses that method, and zero otherwise.⁹

⁹ We do not distinguish between firms that have not considered, are considering, or are implementing these techniques because none of these firms actually used these techniques at the time of the survey, making it

As shown in Panel A of Table 3, implementation of economic value measures is more frequent than balanced scorecard implementation. More than one-third of the firms (36.7%) use economic value measures to some extent and another 7.9% are implementing them now. In contrast, only 20% use balanced scorecards, with an additional 10.7% of scorecard systems undergoing implementation. Roughly 50% of the firms have not considered implementing a balanced scorecard, versus 31.4% for economic value measures.

Business model reliance ranges from 1 = not at all to 6 = completely. Nearly thirty percent of the firms place no reliance on a formal, causal business model or theory of the business (Table 3, Panel B), and only 34.7% make substantial to complete use of business models (four or greater on the six point scale). Extensive reliance on business models is coded one for responses of four or greater to this question, and zero otherwise.

Although Kaplan and Norton (1996, 2001) now argue that causal business models are an integral component of the balanced scorecard concept, 76.9% of the firms claiming to use a balanced scorecard place little or no reliance on business models (not reported in the table). Similarly, 79.2% of economic value users make little or no use of business models, despite claims that greater understanding of the causal model linking financial and non-financial measures to economic results can enhance the benefits from the adoption of economic value measures.

3.6 Performance Variables

impossible to detect any performance differences due to these different responses. Very few firms have implemented and abandoned these practices, preventing us from examining the performance consequences of this decision. We also repeated the analyses using separate dichotomous variables for firms making “extensive” use of these practices and firms making “very extensive” use. These results are reported later in the paper.

We assess the performance implications of strategic performance measurement using two sets of variables: (1) managers' responses regarding their satisfaction with the performance measurement system, and (2) publicly-available information on the firm's accounting and stock market performance. We include the satisfaction measure to allow comparisons of our results to other performance measurement studies using satisfaction as their dependent variable (e.g., Banker et al., 2001a, b; Rigby, 2001; Sandt et al., 2001). The accounting and market measures provide more direct tests of the influence of measurement practices on economic performance, which most SPM advocates argue is the ultimate objective of these systems.

Three questions are used to measure a firm's satisfaction with its measurement system: (1) how well the system meets expectations (1 = has not met expectations to 6 = exceeded expectations); (2) how well the system compares to the respondent's concept of an "ideal" system (1 = not at all ideal to 6 = very close to ideal); and (3) overall satisfaction with the system (1 = not at all satisfied to 6 = completely satisfied). Actual responses to all three questions range from 1 to 6, and indicate that these financial services firms, on average, are moderately satisfied with their measurement systems. Mean (median) scores are 3.54 (4.00) relative to expectations and 3.13 (3.00) relative to ideal. Mean (median) overall satisfaction is 3.42 (4.00), with 20.0% of the respondents rating their satisfaction 5 or 6 and 37.2% rating it 1 or 2. The three satisfaction questions load on a single factor with a coefficient alpha of 0.91. Our satisfaction construct represents the average standardized response to these three questions.

We evaluate economic performance using several measures that are commonly employed to assess financial results. These include two publicly-available accounting

measures (return on assets and three-year sales growth) and two stock return measures (one-year and three-year returns). The accounting and stock return performance measures, obtained from Compustat and CRSP, are measured contemporaneously with the date of the survey. The one-year measures are for fiscal 1999 and the three-year measures are for the time period covering fiscal 1997 to fiscal 1999. For firms in our sample, mean (median) sales growth is 22.5% (16.3%), and return on assets (ROA) is 1.8% (1.2%). Mean (median) one-year stock returns are -12.9% (-13.5%), while three-year stock returns are 1.6% (1.9%).¹⁰

We include several additional variables in the economic performance tests to control for other potential determinants of accounting and stock price performance. To account for the well-known effects of organizational size and growth opportunities on firm performance, we include the log of assets (denoted *Size*) and the ratio of the book value of assets to the market value of equity (denoted *BTOM*, an *inverse* measure of growth opportunities) in the models. We also include the median performance of other firms in the same four-digit SIC code to control for the effects of industry sector on financial performance.¹¹

4. Performance Tests

We test our hypotheses by examining the relations between our firm performance variables and four different measurement system characteristics: (1) overall

¹⁰ Compared to all financial services firms on Compustat, the mean firm in our sample has significantly lower sales growth and stock returns. Return on assets is not statistically different.

¹¹ Since firm strategy can also affect economic performance, we repeated our performance tests using the three strategy constructs as additional control variables. This addition had virtually no impact on our other results.

measurement diversity, (2) alignment with firm strategy, (3) alignment with firm value drivers, and (4) use of measurement alignment techniques.

4.1 Overall Measurement Diversity

Our first set of tests investigates whether greater “measurement diversity” is associated with superior performance (Hypothesis 1). Consistent with claims in the performance measurement literature, we assume that greater measurement diversity is characterized by more extensive use of a broad set of financial and non-financial measures for performance evaluation and decision-making purposes, *independent* of the firm’s strategy and value drivers. *Overall Measurement Diversity* equals the average standardized rating for each of the ten value driver categories across all uses (problem identification, capital investments, performance evaluation, and external disclosure), goal setting, and measurement quality.¹² A higher value for this variable means that the firm uses all of the measures to a greater extent, sets more extensive goals, and has greater measurement quality for these categories.

Given the recent emphasis on the use of non-financial measures, we also examine the relative importance of financial vs. non-financial measurement (denoted *Financial Measurement Focus* and *Non-Financial Measurement Focus*, respectively).¹³ If greater non-financial measurement is more beneficial than greater financial measurement, the association between performance and the use of non-financial measures will be stronger than that with financial measures.

¹² We use a single measurement focus construct rather than separate constructs for usage, goal setting, and measurement quality because of high correlations among these practices, which creates problems with multicollinearity. When separate constructs are computed for each of these three measurement characteristics, the correlation between usage and goal setting is 0.79, between usage and measurement quality is 0.83, and between goal setting and measurement quality is 0.74.

¹³ Financial focus is measured by the values assigned to short-term financial results and non-financial focus is measured by the average values assigned to the remaining nine performance categories.

Evidence on the influence of *Overall Measurement Diversity* on performance is presented in Panel A of Table 4.¹⁴ When satisfaction with the measurement system is the dependent variable, the coefficient on *Overall Measurement Diversity* is positive and highly significant ($p < 0.01$, two-tailed), with an adjusted R^2 of 0.215. As seen in Panel B, this result is driven by greater use and improved measurement of non-financial performance measures, which is significantly positive when overall measurement diversity is decomposed into financial and non-financial components. In contrast, greater emphasis on financial measures is not significantly associated with measurement satisfaction. This evidence is consistent with prior studies that find greater measurement satisfaction when more emphasis is placed on non-financial measures (e.g., Sandt et al., 2001), but does not support claims that greater emphasis on traditional budgetary or financial control uses of performance measures leads to lower satisfaction (Banker et al., 2001b).

Although overall and non-financial measurement focus are positively associated with satisfaction, their relation with economic performance is mixed. Neither variable is significantly associated with ROA, sales growth, or three-year stock returns. However, both variables are positively related to one-year stock returns ($p < 0.05$, two-tailed). These findings lend weak support to the hypothesis that greater measurement diversity is associated with higher performance.

With respect to the control variables, firm performance is strongly related to sector performance, while ROA is positively related to the firm's growth opportunities (i.e.,

¹⁴ We delete observations from our performance tests when studentized residuals are greater than four standard deviations from the mean in order to mitigate the impact of outliers on our results. For each regression, fewer than three observations were removed.

negatively related to its book-to-market ratio). Organizational size is not significant in any of the models.

4.2 Alignment Between Strategy and Measurement Practices

The preceding tests assume that greater measurement diversity affects performance, independent of the firm's strategies. Contingency and economic theories, on the other hand, contend that measurement practices must be aligned with the firm's strategy. We investigate these theories by examining whether performance is enhanced when strategy and measurement are more closely aligned (Hypothesis 2). We first develop benchmark models for assessing the extent of alignment by regressing each combination of measurement characteristics (use, measurement quality, and goals) and type of measure (e.g., short-term financial, customer, employee, etc.) on the three strategy constructs (*Flexible*, *Innovate*, and *Maintain*). This yields 60 individual benchmark regressions (i.e., ten types of measures multiplied by six measurement characteristics).

Our proxies for measurement system alignment are then computed by averaging the standardized residuals from these models. This approach assumes that firms, on average, have correctly chosen their performance measurement systems, and that the estimated models capture the appropriate benchmark for system characteristics given the firm's strategy (Van de Ven and Drazin, 1985). If the benchmark models represent optimal measurement practices, then any deviations from the estimated models (i.e., either too little or too much measurement emphasis) should be negatively associated with satisfaction or performance. These deviations are analogous to the "measurement gaps" discussed in the performance measurement literature in that gaps (or deviations) between

the perceived importance of a performance criteria and the criteria's use in the performance measurement system are assumed to be detrimental.

To allow for potential differences between measurement practices that are less or more extensive than predicted, we compute separate variables for negative and positive residuals and then take their absolute values (denoted *Negative Overall Measurement Residual* and *Positive Overall Measurement Residual*, respectively).¹⁵ We also subdivide each of these variables into financial and non-financial components to investigate whether this measurement distinction influences performance.

The resulting performance tests are presented in Table 5. The adjusted R²s for the satisfaction models (0.117 and 0.124) are smaller than those using the measurement diversity scores in Table 4 (0.215 and 0.210). The differences in explanatory power imply that respondents are more likely to rate measurement satisfaction in terms of the *overall* diversity and extent of measurement rather than relative to the requirements of their chosen strategies. Positive deviations from the benchmark model have a significant *positive* association with measurement satisfaction, indicating that firms with more extensive measurement than competitors following similar strategies rate their systems more highly. However, negative deviations from the benchmark models show no differences in satisfaction.

When the overall residuals are subdivided into financial and non-financial components, *Positive Non-Financial Measurement Residual* has a highly significant

¹⁵ For example, assume that three firms have average residuals of 0, 3, and -3 after averaging the residuals across all 60 benchmark models. *Negative Overall Measurement Residual* and *Positive Overall Measurement Residual* equal (0, 0) for the first firm, (0, 3) for the second firm, and (3, 0) for the third firm). Larger values (i.e., larger deviations in absolute value from the benchmark models) are hypothesized to be negatively associated with performance in these tests. If positive and negative deviations from the benchmark models have equal effects on performance, the coefficients on *Negative Overall Measurement Residual* and *Positive Overall Measurement Residual* should be equal.

positive association with satisfaction ($p < 0.01$, two-tailed), again suggesting that satisfaction is related to greater use of non-financial measures than firms with similar strategies. *Negative Non-Financial Measurement Residual*, in contrast, is not significant in the satisfaction model.

Turning to the economic performance tests, we continue to find no evidence that measurement characteristics are associated with accounting performance. We also find no evidence that lower than predicted measurement influences performance. Greater than predicted measurement (*Positive Overall Measurement Residual*), on the other hand, has positive and significant associations with both one-year stock returns ($p < 0.05$, two-tailed) and three-year stock returns ($p < 0.10$, two-tailed). When we subdivide the overall residuals into financial and non-financial components, the individual financial and non-financial residual measures are not significantly associated with stock market performance at the 10 percent level. This is somewhat surprising given the earlier significant association between *Positive Overall Measurement Residual* and the two stock return measures. Further analysis indicates that the differing results are due to positive non-financial residuals being marginally insignificant ($p < 0.12$, two-tailed) in both stock return models, and positive financial residuals being marginally insignificant in the three-year stock return model ($p < 0.14$, two-tailed). Although these variables are individually insignificant at conventional statistical levels, the two variables have a significant *combined* effect on stock market performance. These results suggest that financial and non-financial measurement deviations have a stronger *joint* effect on performance than either category individually.

In sum, the evidence in Table 5 provides little support for the hypothesis that deviations from our strategy alignment benchmark model (or “measurement gaps”) are detrimental to performance. Instead, *greater* than predicted combined measurement of financial and non-financial measures is associated with higher stock returns, while lower than predicted measurement appears to have little influence on performance. These results support Hypothesis 1 that greater measurement diversity improves performance. However, the significant associations between three-year stock returns and positive measurement residuals in Table 5, together with the insignificant three-year stock market results in Table 4, indicate that the appropriate benchmark may be greater measurement relative to competitors following similar strategies rather than greater measurement on an absolute scale.

4.3 Alignment Between Value Drivers and Measurement Practices

We next examine the association between firm performance and the alignment between perceived value drivers and measurement practices (Hypothesis 3). We measure value driver alignment using a method analogous to that used to compute the alignment between strategy and measurement practices. Each measurement practice using a particular type of measure is regressed on the score given to the corresponding value driver (e.g., use, measurement quality, and goal setting for customer measures are each regressed on the customer value driver score). Our alignment measures are based on the average standardized residuals from these models. We again allow for differences in lower and greater than predicted measurement by computing separate variables for negative and positive residuals and then taking their respective absolute values. We also

subdivide each of these variables into financial and non-financial components to examine their respective performance implications.

The results are presented in Table 6. As before, measurement system satisfaction is positively associated with greater than predicted overall measurement ($p < 0.01$, two-tailed). This is primarily due to the positive association with *Positive Non-Financial Measurement Residual*. In addition, lower than predicted overall and financial measurement now have significant negative associations with satisfaction. The *Negative Overall Measurement Residual* result differs from those in the earlier satisfaction tests and, together with the *Negative Financial Measurement Residual* result, implies that satisfaction is lower when financial measurement practices are less extensive than that required by the firm's perceived value drivers. Relative to satisfaction models using the strategy alignment variables (Table 5), the adjusted R^2 s in the Table 6 satisfaction models are larger (0.117 and 0.124 vs. 0.199 and 0.195), suggesting that managers' satisfaction with the measurement system is more closely related to the alignment with value drivers than to the alignment with strategy. However, the explanatory power using either set of alignment variables is still somewhat lower than that using the overall measurement diversity variables in Table 4 (0.215 and 0.210).

Similar to the earlier results, the value driver alignment variables are insignificant in the ROA models. However, greater than predicted overall and non-financial measurement now exhibit significant *negative* relations with sales growth, indicating that more extensive measurement and use of non-financial measures than firms with similar perceived value drivers is associated with *lower* sales growth. The stock return results are similar to those using the strategy alignment variables, but the models' explanatory

power and the significance levels of the coefficients are generally higher using the alignment with value drivers. Greater than predicted overall measurement is positively associated with one- and three-year stock returns ($p < 0.05$, two-tailed), while lower than predicted overall measurement remains insignificant. More importantly, the adjusted R^2 s in the overall measurement models increase from 0.091 (one-year return) and 0.126 (three-year return) using the strategy alignment variables to 0.110 and 0.146 using value driver alignment.

When the overall measurement residual is broken into financial and non-financial components, *Positive Non-Financial Measurement Residual* is again positive in both stock return models (though marginally insignificant in the three-year return model), with the coefficient's two-tailed confidence level in the one-year return model increasing from 0.88 using the strategy alignment measures to 0.95 using the value driver alignment measures. Similarly, the two-tailed confidence level for the positive coefficient on *Positive Financial Measurement Residual* in the three-year return model increases from 0.86 to 0.97. The models' explanatory power also improves, with the adjusted R^2 in the one-year (three-year) stock return model increasing from 0.082 (0.127) in Table 5 to 0.099 (0.163) in Table 6.

Overall, the analyses in Table 6 provide only limited support for the hypothesis that performance is enhanced when deviations or "measurement gaps" between a firm's performance measurement system and its value drivers are minimized. Rather, performance is more strongly associated with measurement systems that are more extensive and diverse than those of competitors with similar value drivers. Measurement system satisfaction is enhanced when measurement is more extensive than predicted and

reduced when measurement is less extensive than predicted. Consistent with the earlier results, we again find that stock returns are higher when financial *and* non-financial measurement are more extensive than predicted by the value driver benchmark model. In addition, the greater statistical significance and explanatory power in these stock return models, relative to those in earlier tests, suggest that assessing measurement practices relative to the firm's value drivers is an important input into the choice of performance measures.

4.4 Alignment Techniques, Value Drivers, and Measurement System Characteristics

We next examine the measurement alignment techniques discussed in the strategic performance measurement literature. We first consider whether the perceived value drivers and types of information used for different purposes are related to the adoption of the balanced scorecard, economic value measures, and causal business modeling.¹⁶ We conduct these analyses to validate responses to these questions and to verify or refute claims about the performance measurement emphases of firms adopting these practices.

Table 7 summarizes the results. The table lists the statistically significant mean differences ($p < 0.10$, two-tailed) between users and non-users of these three techniques (each coded 1 if that technique is used extensively, and 0 otherwise). The fewest differences relate to the use or non-use of balanced scorecards. Despite the emphasis on non-financial customer, innovation, and internal business process measures in the balanced scorecard literature, we find little evidence that firms claiming to use scorecards place greater emphasis on these performance categories than non-users. Mean responses

¹⁶ We also examined mean differences in strategies between users and non-users of these measurement practices. Strategy is not significantly different between balanced scorecard users and non-users. Extensive users of economic value measures and causal business models, on the other hand, tend to

regarding the firm's value drivers and the measures used in problem identification are not significantly different in the two groups. Two categories that do show significant differences are supplier and employee measures, even though these categories are not explicitly included in Kaplan and Norton's (1996) four balanced scorecard perspectives. Scorecard users set supplier goals more extensively and disclose more information on supplier performance (along with community and alliance performance) to external parties than non-users. Employee measures are also used more extensively for evaluating capital investments and managerial performance and have higher measurement quality in scorecard adopters. However, the many insignificant differences in the two groups bring into question whether firms claiming to have balanced scorecards are actually *using* the information, or have merely implemented measurement systems that capture information corresponding to the scorecard categories without making changes in the information used for decision-making and performance evaluation.

Economic value users rate operating performance a more important value driver than non-users, consistent with claims that these measures promote an operational focus (Kaplan and Norton, 2001, pp. 378-379). This focus is also evident in users' greater emphasis on operating performance information for each measurement characteristic. Environmental performance is considered a more important value driver in economic value users, but short-term financial performance is not. However, short-term financial performance (which can include economic value) has higher measurement quality in economic value adopters and is used more extensively for problem identification and the development of action plans.

emphasize flexibility and innovation to a greater extent than non-users, but the emphasis on *Maintain* strategies is not significantly different.

The greater emphasis on operational performance and short-term financials in economic value users does *not* come at the expense of non-financial measures. On average, economic value users make equal or greater use of information related to all of the non-financial categories. For example, users report that employee measures are far more important for problem identification and the development of action plans, are externally disclosed to a greater extent, have higher measurement quality, and more likely to have goals. Supplier measures, in turn, are used more extensively for goal setting and the evaluation of capital investments. Goals related to the performance of alliances are set to a greater extent, alliance performance measurement quality is higher, and greater use is made of alliance information for performance evaluation. In contrast to the balanced scorecard comparisons, economic value users have innovation measures with higher measurement quality than non-users, and place greater emphasis on innovation measures for problem identification and action plans. Information on customers is also disclosed more extensively, along with information on environmental performance. In no case is less emphasis placed on non-financial measures by economic value users than by non-users. The equal or greater emphasis on many of the non-financial measures by economic value users is consistent with surveys finding that economic value measures are typically used in conjunction with other performance measures (Ittner and Larcker, 1998; Answerthink and Stern Stewart, 2001), and is inconsistent with claims that economic value measures cause companies to overlook longer-term non-financial performance information.

Reliance on a formal, causal business model or “theory of the business” exhibits the largest number of significant differences in the importance, measurement quality, and

use of the various performance categories. Business model users perceive operational, supplier, and environmental performance to be more important long-term value drivers than firms making little or no use of causal business models. The greater importance placed on supplier and environmental information for each measurement characteristic supports this perception. The importance of operational performance leads business model users to set more extensive operational goals. Business modeling is also positively associated with the use of quality and innovation measures for goal setting, problem identification, performance evaluation, external disclosure, and measurement quality.¹⁷

Overall, the evidence in Table 7 indicates that differences in value drivers and performance measurement characteristics tend to be greater between users and non-users of causal business models than between users and non-users of balanced scorecards or economic value measurement. The evidence also indicates that many common perceptions about the value drivers and measurement characteristics of balanced scorecard and economic value users are incorrect, with economic value users placing significantly more emphasis on non-financial measures than non-users and balanced scorecard adopters exhibiting few differences from non-adopters.

4.5 Measurement Alignment Techniques and Firm Performance

The performance implications of the measurement alignment techniques (Hypothesis 4) are examined in Table 8. Although all three practices have strong, positive associations with measurement system satisfaction ($p < 0.01$, two-tailed), the economic

¹⁷ To examine claims that these three techniques lead to closer alignment between the firm's strategies and value drivers and its performance measurement system, we also estimated models regressing the absolute values of the residuals from our strategy or value driver benchmark models on indicator variables for the use of the three techniques. If these techniques lead to closer alignment, they should be negatively associated with the absolute values of the residuals (i.e., use of these techniques leads to smaller deviations from the benchmark model). None of the indicator variables was significant in any of the models, providing no support for claims that these techniques result in closer performance measure alignment.

results are weak and inconsistent. Only causal business modeling is positively associated with any of the accounting and stock return measures, and then only with ROA.

Moreover, extensive use of balanced scorecards is *negatively* associated with ROA.

None of these techniques is statistically related to sales growth or either stock return measure.

As an alternative specification (not reported in the tables), we added interactions between business modeling and both balanced scorecard and economic value use as additional predictor variables. Balanced scorecard and economic value advocates suggest that these techniques are more beneficial when they are used in conjunction with causal business models. The interactions were not significant, with the other results changing little from those reported in Table 8.

We also repeated the tests after replacing the single balanced scorecard and single economic value variables with two variables each: the first coded one if the respondent stated that the technique is used (5 on the survey scale) and zero otherwise, and the second coded one if the respondent stated that it is used extensively (6 on the survey scale) and zero otherwise. Only the second variables capturing very extensive use were statistically significant in the satisfaction model. No significant associations with economic performance were found using any of these variables. Results also remained virtually identical to those reported in Table 8 when the dichotomous business model variable was replaced by a business model variable ranging from 1 (not at all) to 6 (completely).

Taken as a whole, these results provide little support for the hypothesis that the use of these measurement alignment techniques influences economic performance.

Furthermore, the contrasting satisfaction and financial results raise serious questions about the validity of using satisfaction (or other perceived performance measures) as an indicator of management accounting system success.¹⁸

5. Performance Results in More Mature Systems

The preceding tests make no distinction between measurement systems that have been in place for some time and those that have recently been modified. If the benefits from improved strategic performance measurement take some time to materialize, these benefits may not be observed in firms with recently modified systems. As a result, our contemporaneous performance tests may understate the performance consequences of strategic performance measurement practices. Consequently, we repeat the analyses using the 94 firms (67.1% of our sample) that report no major changes in their performance measurement systems in the past two years.¹⁹

The results from these tests are summarized in Table 9. The performance implications are quite similar to those using the entire sample. Measurement satisfaction continues to have consistently positive associations with greater overall measurement diversity, and particularly with greater non-financial measurement. Although the coefficients on *Negative Overall Residual* and *Negative Non-Financial Residual* are significantly positive in a few of the strategy alignment tests using this subsample, we find little consistent evidence that lower than predicted measurement influences performance.

¹⁸ As an additional robustness check of our financial performance results, we reran the tests using measurement system satisfaction as an independent variable in place of the other proxies for system characteristics and uses. This allows us to examine whether any financial benefits from strategic performance measurement system characteristics are driven by users' satisfaction with these systems. System satisfaction was *negatively* associated with sales growth, and was not statistically significant in the other accounting or stock return models.

Accounting results remain limited and mixed. Both positive and negative financial measurement residuals are positively associated with ROA in the value driver alignment tests, indicating that either more or less financial measurement than firms with similar value drivers is related to higher accounting performance. *Positive Overall Residual* is negatively associated with sales growth using the alignment with either strategy or value drivers. These findings are consistent with earlier results using the full sample in Table 6. *Positive Non-Financial Measurement Residual* also has a significant negative association with sales growth in the value driver alignment model.

The strongest and most consistent results relate to the association between strategic measurement practices and stock returns. *Positive Overall Residual* and *Positive Non-Financial Measurement Residual* have significant, positive associations with both one- and three-year stock returns using either the strategy alignment or value driver alignment variables. These results are not only consistent with those using the full sample, but are more significant. In particular, one-year stock returns using *Positive Non-Financial Measurement Residual* (computed using either strategy or value driver alignment) are now significant at the 5% level or better (two-tailed) using this subsample, while three-year stock returns (using either alignment variable) are now significant at the 10% level (two-tailed). *Positive Financial Measurement Residual* is also positively associated with three-year stock returns using the value driver alignment variable (and is only marginally insignificant [$p < 0.12$, two-tailed] using the strategy alignment variable). In no case is lower than predicted measurement significantly associated with stock market performance.

¹⁹ An additional advantage to using this subsample is that the three-year sales growth and three-year stock return measures correspond to periods in which the current measurement system was in place.

Business modeling remains strongly associated with measurement satisfaction ($p < 0.01$, two-tailed). In contrast, economic value and balanced scorecard use are no longer significantly associated with satisfaction, indicating that economic value and scorecard users with more mature measurement system are *not* more satisfied than non-users. Further analysis finds that the significant positive association between economic value and scorecard use and satisfaction in the full sample is driven by higher satisfaction in the subset of firms that have recently made major changes in their measurement system. This result is particularly intriguing because it suggests that the higher satisfaction associated with economic value and balanced scorecard systems may be short-lived. Finally, none of the measurement alignment techniques is statistically related to firm accounting or stock price performance, suggesting that our earlier insignificant results for these techniques are not due to the inclusion of new measurement systems that have not had time to yield significant performance improvements.

6. Conclusions

Using data from U.S. financial services firms, we examine the relative ability of various strategic performance measurement approaches to explain firm performance. These approaches include greater measurement diversity, improved alignment with firm strategy and value drivers, and the use of performance measurement alignment techniques including the balanced scorecard, economic value measures, and causal business modeling. We find that a variation of the measurement diversity approach has the strongest association with stock market performance. In particular, firms that make more extensive use of a broad set of financial and (particularly) non-financial measures than those with similar strategies or value drivers earn higher stock returns. These results

are even stronger when the performance measurement system has undergone no significant changes in the past three years.

In contrast, we find little evidence that SPM practices are associated with accounting measures (ROA and sales growth). One potential explanation for these contrasting results is that the performance implications of SPM systems are more likely to be captured in forward-looking stock market measures than in short-term, historical accounting measures. Finally, the measurement alignment techniques proposed in the performance measurement literature are positively associated with measurement system satisfaction, but exhibit almost no association with economic performance.

As with any study of this type, our results are subject to a number of limitations, including potential response biases, endogeneity of our predictor variables, model specification, and the difficulty in using a survey instrument to obtain factual, detailed information on exactly how firms measure performance. In addition, our results may not generalize to other industries or competitive settings. Future studies can also extend our analyses by examining a broader set of performance measurement system attributes, such as the level of aggregation and integration, that other studies have found to be determinants of the perceived usefulness of measurement systems (e.g., Chenhall and Morris, 1986) but are not included in our survey. Finally, prior research on management accounting and performance measurement system innovations indicates that technical and organizational factors can play an important role in the perceived success of system implementation (e.g., Shields, 1995; Anderson and Young, 1999; Cavalluzzo and Ittner, 2003). Although we do not examine these factors in our tests due to the absence of data on these issues, future studies can make a significant contribution by examining how

these factors interact with system design choices to influence *actual* performance outcomes.

Despite these limitations, our findings carry a number of implications for research and practice. First, the differing results for measurement satisfaction and stock market performance raise important questions about the validity of using perceptual satisfaction or outcome measures to evaluate measurement system success. Although recent practitioner publications promote the benefits of balanced scorecards and economic value measures by citing surveys on measurement system satisfaction (e.g., Kaplan and Norton, 2001, pp. 356-357), we find no evidence that these higher satisfaction levels translate into improved financial performance.

Second, the descriptive statistics on the use of the various performance measure categories and business modeling practices by adopters and non-adopters of balanced scorecards indicate that many firms that claim to have implemented this technique have not fully adopted Kaplan and Norton's (1996, 2001) prescriptions. Future studies will need to devise improved methods for eliciting what firms mean by a "balanced scorecard" and how this information is actually being used.

Third, the significant performance implications of positive deviations from our benchmark models indicate that (subject to the validity of these models) average measurement practices of firms pursuing similar strategies or value drivers currently are not optimal in this industry. Instead, our results indicate that greater measurement emphasis and diversity than similar firms is associated with higher stock market performance. The results also indicate that the appropriate benchmark for assessing

measurement diversity is greater measurement *relative* to competitors with similar strategies or value drivers rather than greater measurement on an *absolute* scale.

Finally, the greater explanatory power of stock return models using value driver alignment as predictor variables suggest that researchers and practitioners should go beyond the alignment of measurement practices with organizational strategy to investigate and measure the specific value drivers underlying strategic success.

References

- Abernethy, M., Lillis, A., 1995. The impact of manufacturing flexibility on management control system design. *Accounting, Organizations and Society* 20, 241-258.
- Abernethy, M., Guthrie, C., 1994. An empirical assessment of “fit” between strategy and management information system design. *Accounting and Finance*, 49-66.
- American Institute of Certified Public Accountants, 1994. *Improving Business Reporting -- A Customer Focus*. American Institute of Certified Public Accountants, Inc., New York, NY.
- Anderson, S., Young, S.M., 1999. The impact of contextual and procedural factors on the evaluation of activity based costing systems. *Accounting, Organization and Society*, 525-559.
- Answerthink and Stern Stewart & Co. 2001. *Shareowner Alignment Index Report*. Hackett Benchmarking & Research and Stern Stewart & Co., New York, NY.
- Banker, R.D., S.M. Datar. 1989. Sensitivity, Precision, and Linear Aggregation of Signals for Performance Evaluation. *Journal of Accounting Research*: 21-39.
- Banker, R.D., Potter, G., Srinivasan, D., 2000. An empirical investigation of an incentive plan that includes nonfinancial performance measures. *The Accounting Review*, 65-92.
- Banker, R.D., S.N. Janakiraman, C. Konstans. 2001a. *Balanced Scorecard: Linking Strategy to Performance*. New York: Financial Executives International.
- Banker, R.D., S.N. Janakiraman, C. Konstans, M.J. Pizzini. 2001b. *Determinants of Chief Financial Officers' Satisfaction with Performance Measurement Systems*. Working paper. University of Texas at Dallas.
- Behn, B., Riley, R., 1999. Using nonfinancial information to predict financial performance: the case of the U.S. airline industry. *Journal of Accounting, Auditing, and Finance*, 29-56.
- Black, A., Wright, J., Bachman, J., Makall, M., Wright, P., 1998. *In Search of Shareholder Value: Managing the Drivers of Performance*. Pitman Publishing, London.
- Bouwens, J, Abernethy, M., 2000. The consequences of customization on management accounting system design. *Accounting, Organizations and Society*, 241-258.
- Brancato, C.K. 1995. *New Performance Measures – A Research Report*. New York, NY: The Conference Board, Report Number 1118-95-RR.

Brickley, J., Smith, C., Zimmerman, J., 1997. *Managerial Economics and Organizational Architecture*. Richard D. Irwin, Burr Ridge, IL.

Cavalluzzo, K., Ittner, C., 2003. Implementing performance measurement innovations: evidence from government. *Accounting, Organizations and Society*, in press.

Chenhall, R., Langfield-Smith, K., 1998. The relationship between strategic priorities, management techniques and management accounting: an empirical investigation using a systems approach. *Accounting, Organizations and Society* 23, 243-264.

Chenhall, R., Morris, D., 1986. The impact of structure, environment, and interdependence on the perceived usefulness of management accounting systems. *The Accounting Review*, 16-36.

Chong, V., Chong, K., 1997. Strategic choices, environmental uncertainty and SBU performance: a note on the intervening role of management accounting systems. *Accounting and Business Research*, 268-276.

Copeland, T., Koller, T., Murrin, J., 1996. *Valuation: Measuring and Managing the Value of Companies*. John Wiley & Sons, New York, NY.

Dent, J., 1990. Strategy, organization and control: some possibilities for accounting research. *Accounting, Organizations, and Society* 15, 3-25.

Dixon, J.R., Nanni, A.J., Vollmann, T.E. 1990. *The New Performance Challenge: Measuring Operations for World-Class Companies*. Dow Jones-Irwin: Homewood, IL.

Dixon, P., Hedley, B., 1997. *Managing for Value*. Braxton Associates, Boston, MA.

Eccles, R. 1991. The performance measurement manifesto. *Harvard Business Review* 69, 131-137.

Eccles, R., Herz, R., Keegan, E., Phillips, D. 2001. *The Value Reporting Revolution: Moving Beyond the Earnings Game*. John Wiley & Sons, New York, NY.

Edvinsson, L., Malone, M. 1997. *Intellectual Capital: Realizing Your Companies True Value by Finding its Hidden Brainpower*. Harper Business, New York, NY.

Feltham, G. and J. Xie. 1994. Performance Measure Congruity and Diversity in Multi-task Principal/Agent Relations. *The Accounting Review* 69: 429-453.

Financial Accounting Standards Board, 2001. *Improving Business Reporting: Insights into Enhancing Voluntary Disclosures*. Financial Accounting Standards Board.

Fisher. 1995a. Use of Nonfinancial Performance Measures. In *Readings in Management Accounting*, edited by S.M. Young, 329-335. Englewood Cliffs, NJ: Prentice Hall.

- Fisher, J., 1995b. Contingency-based research on management control systems: categorization by level of complexity. *Journal of Accounting Literature* , 24-53.
- Gates, S., 1999. *Aligning Strategic Performance Measures and Results*. The Conference Board, New York, NY.
- Gordon, L., Narayanan, V., 1983. Management accounting systems, perceived environmental uncertainty and organizational structure: an empirical investigation. *Accounting, Organizations and Society*, 33-47.
- Govindarajan, V., A contingency approach to strategy implementation at the business-unit level: integrating administrative mechanisms with strategy. *Academy of Management Journal*, 828-853.
- Govindarajan, V, Gupta, A., 1985. Linking control systems to business unit strategy: impact on performance. *Accounting, Organizations and Society*, 51-66.
- Hauser, J.R., D. Siemester, B. Wernerfelt. 1994. Customer Satisfaction Incentives. *Marketing Science*, 327-350.
- Hemmer, T. 1996. On the Design and Choice of “Modern” Management Accounting Measures. *Journal of Management Accounting Research*: 87-116.
- Heneman, R.L., G. Ledford, M. Gresham. 1999. *The Effects of Changes in the Nature of Work on Compensation*. Working paper. Ohio State University.
- Heskett, J., 1994. Putting the service-profit chain to work. *Harvard Business Review* (March/April): 164-174.
- Hogan, C., Lewis, C., 1999. *The long-run performance of firms adopting compensation plans based on economic profits*. Unpublished working paper, Vanderbilt University.
- Holmstrom, B. 1979. Moral Hazard and Observability. *Bell Journal of Economics*: 74-91.
- Hoque, Z., James, W. 2000. Linking balanced scorecard measures to size and market factors: impact on organizational performance. *Journal of Management Accounting Research* 12, 1-17.
- Institute of Management Accountants, 1999. *Counting more, counting less. Transformations in the management accounting profession*. IMA Publications, Montvale, N.J.
- International Federation of Accountants, 1998. *International Management Accounting Practice Statement: Management Accounting Concepts*. International Federation of Accountants, New York, NY.

Ittner, C., Larcker, D., 1995. Total quality management and the choice of information and reward systems. *Journal of Accounting Research*, 1-34.

Ittner, C., Larcker, D., 1997. Are non-financial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research*, 1-35.

Ittner, C., Larcker, D., 1998. Innovations in performance measurement: trends and research implications. *Journal of Management Accounting Research* 6, 205-238.

Ittner, C., Larcker, D., 2001. Assessing empirical research in managerial accounting: a value-based management perspective. *Journal of Accounting and Economics*, 349-410.

Ittner, C., Larcker, D., and Meyer, M., 2003. Subjectivity and the weighting of performance measures: evidence from a balanced scorecard. *The Accounting Review*, in press.

Kaplan, R., Norton, D., 1992. The balanced scorecard: measures that drive performance. *Harvard Business Review*, 71-80.

Kaplan, R., Norton, D., 1996. *The Balanced Scorecard: Translating Strategy into Action*. Harvard Business School Press, Boston, MA.

Kaplan, R., Norton, D., 2001. *The Strategy Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Harvard Business School Press, Boston, MA.

Krumwiede, K., 1998. The implementation stages of activity-based costing and the impact of contextual and organizational factors. *Journal of Management Accounting Research* 10, 239-277.

Langfield-Smith, K., 1997. Management control systems and strategy: a critical review. *Accounting, Organizations and Society*, 207-232.

Larcker, D., 1981. The perceived importance of selected information characteristics for strategic capital budgeting decisions. *The Accounting Review*, 519-535.

Lingle, J., Schiemann, W., 1996. From balanced scorecard to strategic gauges: is measurement worth it?. *Management Review* 85, 56-61.

Milgrom, P., Roberts, J., 1992. *Economics, Organization & Management*. Prentice Hall, Englewood Cliffs, NJ.

Nagar, V., Rajan, M., 2001. The revenue implications of financial and operational measures of product quality. *The Accounting Review*, 495-513.

- Otley, D., 1999. Performance management: a framework for management control system design. *Management Accounting Research* 10, 363-382.
- Pererra, S., Harrison, G., Poole, M., 1997. Customer-focused manufacturing strategy and the use of operations-based non-financial performance measures: a research note. *Accounting, Organizations and Society* 22, 557-572.
- Rigby, D. 2001. Management tools and techniques: a survey. *California Management Review*, 139-160.
- Ruddle, K., D. Feeny. 2000. *Transforming the Organization: New Approaches to Management, Measurement and Leadership*. Oxford, England: Templeton College.
- Sandt, J., U. Schaeffer, J. Weber. 2001. Balanced Performance Measurement Systems and Manager Satisfaction—Empirical Evidence from a German Study. Working paper. WHU—Otto Beisheim Graduate School of Management.
- Schick, A., L. Gordon, and S. Haka. 1990. Information Overload: A Temporal Approach. *Accounting, Organizations and Society*: 199-220.
- Schiemann, W.A., J.H. Lingle. 1999. *Bullseye!: Hitting Your Strategic Targets Through High-Impact Measurement*. New York: The Free Press.
- Scott, T., Tiessen, P., 1999. Performance measurement and managerial teams. *Accounting, Organizations and Society*, 263-285.
- Shields, M., 1995. An empirical analysis of firms implementation experiences with activity-based costing. *Journal of Management Accounting Research*, 1-61.
- Sim, K., Killough, L., 1998. The performance effects of complementarities between manufacturing practices and management accounting systems. *Journal of Management Accounting Research*, 325-345.
- Simons, R., 1987. Accounting control systems and business strategy: an empirical analysis. *Accounting, Organizations and Society*, 49-62.
- Stern, J., G. Stewart, D. Chew. 1995. The EVA financial management system. *Journal of Applied Corporate Finance* 8, 32-46.
- Stewart, G.B. 1991, *The Quest for Value*. New York: HarperBusiness.
- Stivers, B., Covin, J., Hall, N., Smalt, S. 1998. How nonfinancial performance measures are used. *Management Accounting* (February), 44-49.
- Towers Perrin. 1996. *Compuscan Report*.

Van de Vin, A.H., Drazin, R. 1985. Alternative forms of fit in contingency theory. *Administrative Science Quarterly*, 514-539.

Wallace, J., 1997. Adopting residual income-based compensation plans: do you get what you pay for?. *Journal of Accounting and Economics* 24, 275-300.

Wallman, S., 1995. The future of accounting and disclosure in an evolving world: the need for dramatic change. *Accounting Horizons* 9, 81-91.

Young, S., O'Byrne, S. 2001. *EVA and Value-Based Management*. McGraw-Hill, New York, NY.

Table 1

Mean Survey Responses on the Importance of Performance Categories to Long-Term Organizational Success and Their Use in Performance Measurement and Decision-Making

	Importance to Long-Term Success ^a	Extent Goals Set ^b	<i>Extent Measures Related to These Categories Are Used For the Following Purposes:</i>				
			Problem Identification ^c	Capital Investments ^c	Performance Evaluation ^c	External Disclosure ^d	Measurement Quality ^e
<i>Performance Category:</i> ^f							
Short-Term Financial Results	4.572	5.369	5.190	4.581	5.138	5.428	5.465
Customer Relations	5.511	4.285	4.378	3.917	4.236	2.722	3.724
Employee Relations	3.862	2.853	2.775	2.109	2.913	1.871	2.465
Operational Performance	5.020	4.600	4.573	4.431	4.747	3.889	4.453
Quality	5.031	4.018	3.992	3.876	3.723	2.941	3.588
Alliances	3.060	2.223	2.204	2.122	1.938	2.262	1.979
Supplier Relations	2.875	2.051	2.063	2.150	1.904	1.664	2.001
Environmental Performance	2.079	1.740	1.594	1.695	1.485	1.572	1.634
Innovation	4.114	3.407	3.294	3.598	2.941	3.005	2.706
Community	4.066	3.492	3.053	2.641	2.814	3.127	2.804

^a Scale: 1 = measure not applicable or not at all important to 6 = extremely important

^b Scale: 1 = measure not applicable or no goals established to 6 = explicit goals established

^c Scale: 1 = measure not applicable or not used at all to 6 = used extensively

^d Scale: 1 = measure not applicable or no external disclosure to 6 = complete external disclosure

^e Scale: 1 = measure not applicable or extremely poor quality of measurement to 6 = high quality of measurement

^f Performance categories are defined as short-term financial performance (e.g., annual earnings, return on assets, cost reduction), customer relations (e.g., market share, customer satisfaction, customer retention), employee relations (e.g., employee satisfaction, turnover, workforce capabilities), supplier relations (e.g., on-time delivery, input into product/service design), operational performance (e.g., productivity, safety, cycle time), product and service quality (e.g., defect rates, quality awards), alliances with other organizations (e.g., joint marketing or product design, joint ventures), environmental performance (e.g., government citations, environmental compliance or certification), product and service innovation (e.g., new product or service development success, development cycle time), and community (e.g., public image, community involvement).

Table 2

Spearman Correlations Among Strategy Variables and Value Driver Importance Scores

	Flexible	Innovate	Maintain	Financial	Customer	Employee	Operations	Quality	Alliances	Suppliers	Environment	Innovation
Innovate	0.50 ^{***}											
Maintain	0.15 [*]	0.16 [*]										
Financial	0.03	0.04	0.07									
Customer	0.03	0.16 [*]	0.12	0.13								
Employee	0.06	0.25 ^{***}	0.13	-0.04	0.32 ^{***}							
Operations	0.06	0.13	0.10	0.12	0.32 ^{***}	0.09						
Quality	0.17 ^{**}	0.20 ^{**}	0.12	-0.11	0.35 ^{***}	0.10	0.40 ^{***}					
Alliances	0.21 ^{**}	0.28 ^{***}	0.05	-0.01	0.23 ^{***}	0.17 ^{**}	0.19 ^{**}	0.33 ^{***}				
Suppliers	0.15 [*]	0.28 ^{***}	0.03	0.09	0.16 [*]	0.20 ^{**}	0.11	0.25 ^{***}	0.42 ^{***}			
Environment	0.11	0.16 [*]	0.20 ^{**}	-0.12	0.03	0.10	0.13	0.19 ^{**}	0.26 ^{***}	0.45 ^{***}		
Innovation	0.16 [*]	0.31 ^{***}	0.01	0.11	0.41 ^{***}	0.13	0.25 ^{***}	0.27 ^{***}	0.31 ^{***}	0.27 ^{***}	0.22 ^{**}	
Community	0.07	0.07	0.21 ^{**}	0.03	0.44 ^{***}	0.10	0.17 ^{**}	0.30 ^{***}	0.25 ^{***}	0.10	0.10	0.32 ^{***}

Flexible = the extent to which the firm follows a strategy focused on flexibility; Innovate = the extent to which the firm follows an innovation-oriented strategy; Maintain = the extent to which the firm follows a strategy focused on maintaining existing relationships in stable and predictable markets. See table 1 for definitions of the remaining value driver performance categories.

***, **, * indicate p-value < 0.01, 0.05, 0.10 respectively (two-tailed).

Table 3
The Use of new Performance Measurement Techniques
in Financial Services Firms

Panel A. Use of Balanced Scorecard and Economic Value Measures (n = 140)

	Balanced Scorecard	Economic Value Measures
Not considered	50.7%	31.4%
Implemented and abandoned	1.4%	2.1%
Considering	17.1%	22.1%
Implementing now	10.7%	7.9%
Use	15.0%	24.5%
Use extensively	5.0%	12.1%

Panel B. Formal Reliance on a “Theory of the Business” or Business Model That Causally Links Performance Drivers to Performance Outcomes (n = 136)

1 = Not at all	29.7%
2	16.7%
3	18.8%
4	19.6%
5	10.8%
6 = Completely	4.3%

Table 4**Association Between Measurement Diversity and Performance in Financial Services Firms****Panel A: Association Between Overall Measurement Diversity and Performance**

	Satisfaction	ROA	Sales Growth	One-year Stock Returns	Three-year Stock Returns
Intercept	0.012 (0.176)	0.012 (1.645)	0.150** (1.892)	0.045 (0.401)	-0.305** (-2.131)
Overall Measurement Diversity	0.873*** (6.255)	0.001 (0.471)	-0.031 (-1.224)	0.087** (2.033)	0.064 (1.326)
Sector Performance		0.597*** (7.082)	0.674*** (4.609)	0.847*** (3.474)	1.219*** (3.852)
SIZE		0.0006 (0.800)	0.0069 (0.360)	-0.0117 (-0.918)	0.0236 (1.565)
BTOM		-0.012*** (-3.056)	-0.176*** (-3.159)	0.0096 (0.160)	0.0632 (0.633)
Adj. R ²	0.215	0.321	0.259	0.095	0.118

Panel B: Association Between Financial and Non-Financial Measurement Focus and Performance

	Satisfaction	ROA	Sales Growth	One-year Stock Returns	Three-year Stock Returns
Intercept	0.012 (0.176)	0.0118 (1.634)	0.150* (1.884)	0.045 (0.395)	-0.310** (-2.153)
Financial Measurement Focus	0.072 (0.682)	-0.008 (-0.375)	-0.0041 (-0.219)	-0.0032 (-0.102)	0.0213 (0.595)
Non-Financial Measurement Focus	0.793*** (5.830)	0.0018 (0.607)	-0.0279 (-1.119)	0.0843** (2.035)	0.0505 (1.093)
Sector Performance		0.603*** (7.051)	0.675*** (4.527)	0.853*** (3.479)	1.230*** (3.858)
SIZE		0.006 (0.431)	0.0069 (0.914)	-0.0115 (-0.903)	0.0236 (1.562)
BTOM		-0.0124*** (-3.041)	-0.177*** (-3.132)	0.0102 (0.170)	0.0695 (0.686)
Adj. R ²	0.210	0.317	0.252	0.088	0.111

***, **, * indicates p-value < 0.01, 0.05, 0.10 respectively (two-tailed test)

Overall Measurement Diversity = the average extent to which measures related to the ten performance categories are used for performance evaluation and decision-making, have goals set, and have high quality of measurement. Financial Measurement Focus = the average extent to which short-term accounting measures are used for performance evaluation and decision-making, have goals set, and have high quality of measurement. Non-Financial Measurement Focus = the average extent to which measures related to the nine non-financial performance categories are used for performance evaluation and decision-making, have goals set, and have high quality of measurement. Sector Performance = the median performance of other firms in the same four-digit SIC code. SIZE = the log of assets. BTOM = the ratio of book value of assets to the market value of the firm (an inverse measure of growth opportunities).

Table 5
Association Between the Performance of Financial Services Firms and the Alignment Between Measurement Practices and Firm Strategy

Panel A: Association Between Overall Measurement Alignment and Performance

	Satisfaction	ROA	Sales Growth	One-year Stock Returns	Three-year Stock Returns
Intercept	-0.727*** (-2.415)	0.0058 (0.664)	0.199** (2.166)	-0.0994 (-0.732)	-0.460*** (-2.824)
Positive Overall Measurement Residual	1.530*** (3.875)	0.0085 (1.089)	-0.0945 (-1.408)	0.225** (2.046)	0.240* (1.941)
Negative Overall Measurement Residual	0.328 (0.797)	0.0070 (0.884)	-0.0156 (-0.223)	0.0854 (0.735)	0.150 (1.146)
Sector Performance		0.592*** (6.917)	0.663*** (4.492)	0.837*** (3.407)	1.218*** (3.867)
SIZE		0.0007 (0.835)	0.0064 (0.856)	-0.0084 (-0.670)	0.0245 (1.650)
BTOM		-0.0128*** (-3.133)	-0.177*** (-3.170)	0.0041 (0.068)	0.0531 (0.533)
Adj. R ²	0.117	0.322	0.260	0.091	0.126

Panel B: Association between Financial and Non-Financial Measurement Alignment and Performance

	Satisfaction	ROA	Sales Growth	One-year Stock Returns	Three-year Stock Returns
Intercept	-0.712** (-2.234)	0.0060 (0.656)	0.195* (1.910)	-0.133 (-0.938)	-0.543*** (-3.157)
Positive Financial Measurement Residual	-0.112 (-0.273)	0.0017 (0.225)	-0.0627 (-0.833)	0.114 (0.998)	0.193 (1.473)
Negative Financial Measurement Residual	-0.285 (1.458)	0.0026 (0.699)	-0.0041 (-0.116)	0.0527 (0.935)	0.0456 (0.730)
Positive Non-Financial Measurement Residual	1.526*** (4.074)	0.0068 (0.871)	-0.0598 (-0.851)	0.170 (1.580)	0.177 (1.481)
Negative Non-Financial Measurement Residual	0.651 (1.549)	0.0074 (0.440)	0.0227 (0.294)	0.0435 (0.372)	0.123 (0.932)
Sector Performance		0.605*** (6.824)	0.778*** (4.928)	0.859*** (3.450)	1.228*** (3.897)
SIZE		0.0007 (0.819)	0.0055 (0.686)	-0.0069 (-0.543)	0.0264* (1.767)
BTOM		-0.0128*** (-3.085)	-0.189*** (-3.102)	0.0082 (0.135)	0.0766 (0.754)
Adj. R ²	0.124	0.312	0.262	0.082	0.127

Notes to Table 5

***, **, * indicate p-value < 0.01, 0.05, 0.10 respectively (two-tailed test)

The measurement residual variables represent the average standardized residuals from prediction models regressing measurement uses, goal setting, and quality on the three strategy variables (*Flexible*, *Innovate*, and *Maintain*). The positive measurement residual variables equal the average residual if this value is positive, and zero if it is negative. The negative measurement residual variables equal the absolute value of the average residual if the average residual is negative, and zero if it is positive. See Table 4 for other variable definitions.

Table 6
Association Between the Performance of Financial Services Firms and the Alignment Between Measurement Practices and Firm Value Drivers

Panel A: Association Between Overall Measurement Alignment and Performance

	Satisfaction	ROA	Sales Growth	One-year Stock Returns	Three-year Stock Returns
Intercept	-0.124 (0.616)	0.0061 (0.709)	0.192** (2.230)	-0.0359 (-0.280)	-0.347** (-2.194)
Positive Overall Measurement Residual	1.153*** (3.276)	0.0090 (1.147)	-0.140** (-2.153)	0.239** (2.263)	0.244** (2.049)
Negative Overall Measurement Residual	-0.792** (-2.240)	0.0071 (0.995)	-0.0244 (-0.372)	0.0071 (0.944)	0.0002 (0.002)
Sector Performance		0.603*** (7.146)	0.673*** (4.672)	0.821*** (3.403)	1.170*** (3.742)
SIZE		0.0006 (0.807)	0.0086 (1.151)	-0.0129 (-1.025)	0.0189 (1.269)
BTOM		-0.0129*** (-3.155)	-0.165*** (-2.997)	0.0004 (0.008)	0.0429 (0.437)
Adj. R ²	0.199	0.323	0.280	0.110	0.146

Panel B: Association between Financial and Non-Financial Measurement Alignment and Performance

	Satisfaction	ROA	Sales Growth	One-year Stock Returns	Three-year Stock Returns
Intercept	-0.037 (-0.138)	0.0050 (0.586)	0.194** (2.252)	-0.0503 (-0.385)	-0.392** (-2.466)
Positive Financial Measurement Residual	-0.153 (-0.493)	0.0115* (1.858)	0.0660 (1.213)	0.0870 (0.963)	0.222** (2.248)
Negative Financial Measurement Residual	-0.293* (-1.668)	0.0063* (1.854)	0.0360 (1.174)	0.0246 (0.478)	0.0491 (0.860)
Positive Non-Financial Measurement Residual	1.100*** (0.317)	0.0040 (0.536)	-0.155** (-2.498)	0.197* (1.955)	0.160 (1.487)
Negative Non-Financial Measurement Residual	-0.564 (-1.643)	0.0027 (0.406)	-0.0475 (-0.754)	-0.0025 (-0.025)	-0.0148 (-0.129)
Sector Performance		0.597*** (7.054)	0.651*** (4.405)	0.816*** (3.353)	1.232*** (3.950)
SIZE		0.0005 (0.651)	0.0073 (0.973)	-0.0131 (-1.036)	0.0184 (1.244)
BTOM		-0.0144*** (-3.488)	-0.188*** (-3.302)	-0.0094 (-0.153)	0.0091 (0.091)
Adj. R ²	0.195	0.333	0.284	0.099	0.163

Notes to Table 6

***, **, * indicate p-value < 0.01, 0.05, 0.10 respectively (two-tailed test)

The measurement residual variables represent the average standardized residuals from prediction models regressing measurement uses, goal setting, and quality on the ten value driver categories. The positive measurement residual variables equal the average residual if this value is positive, and zero if it is negative. The negative measurement residual variables equal the absolute value of the average residual if the average residual is negative, and zero if it is positive. See Table 4 for other variable definitions.

Table 7

Significant Mean Differences in Value Drivers and Measurement Characteristics Between Users and Non-Users of Performance Measurement Alignment Techniques

	<u>Balanced Scorecard</u>				<u>Economic Value Added</u>				<u>Business Modeling</u>			
	<u>Performance Category</u>	<u>Use</u>	<u>Not Use</u>	<u>Difference</u>	<u>Performance Category</u>	<u>Use</u>	<u>Not Use</u>	<u>Difference</u>	<u>Performance Category</u>	<u>Use</u>	<u>Not Use</u>	<u>Difference</u>
Drivers of Long-Term Organizational Success	No Differences				Customers	5.67	5.42	0.25	Operational Performance	5.45	4.95	0.50
					Operational Performance	5.27	4.87	0.40	Suppliers	3.37	2.79	0.58
					Environmental	2.39	1.90	0.49	Environmental	3.37	2.74	0.63
Extent Used in Problem Identification and Developing Action Plans	No Differences				Short- Term Financial	5.42	5.06	0.36	Employee	3.42	2.65	0.77
					Employee	4.56	4.27	0.29	Quality	5.07	3.81	1.26
					Operational Performance	4.84	4.42	0.42	Suppliers	2.71	1.49	1.22
					Innovation	3.61	3.11	0.50	Environment	2.15	1.49	0.66
									Innovation	4.50	3.09	1.41
Extent Used to Evaluate Capital Investments	Employee	2.57	1.99	0.58	Operational Performance	4.80	4.21	0.59				
					Suppliers	2.48	1.96	0.52				
Extent Used in Performance Measurement	Employee	3.44	2.78	0.66	Employee	3.44	2.61	0.83	Employee	3.53	2.81	0.72
					Operational Performance	5.02	4.59	0.43	Quality	4.55	3.60	0.95
					Alliances	2.21	1.77	0.44	Alliances	2.45	1.88	0.57
									Suppliers	2.39	1.84	0.55
									Environment	1.97	1.41	0.56
Extent Used in External Disclosure	Alliance	2.72	2.15	0.57	Customer	3.08	2.52	0.56	Quality	3.68	2.80	0.88
					Supplier	2.05	1.57	0.48	Alliances	3.00	2.13	0.87
					Community	3.65	2.99	0.66	Operational Performance	4.49	3.54	0.95
									Suppliers	2.13	1.59	0.54
									Environmental	1.86	1.40	0.46
Measurement Quality	Employees	3.05	2.31	0.74	Short-Term Financial	5.70	5.33	0.37	Quality	4.42	3.44	0.98
					Employees	2.90	2.22	0.68	Alliances	2.47	1.91	0.56
					Operational Performance	4.89	4.20	0.69	Suppliers	2.53	1.92	0.61
					Innovation	2.95	2.57	0.38	Environmental	2.13	1.52	0.61
									Innovation	3.45	2.57	0.88
Extent Goals Set	Suppliers	2.51	1.93	0.58	Employee	3.29	2.6	0.69	Customer	4.79	4.20	0.59
					Operational Performance	4.95	4.39	0.56	Operational Performance	5.08	4.54	0.54
					Alliances	2.48	2.07	0.41	Quality	4.92	3.87	1.05
					Suppliers	2.30	1.90	0.40	Suppliers	2.71	1.94	0.77
									Environment	2.34	1.64	0.70
								Innovation	4.36	3.27	1.09	

Cells report significant differences in mean practices ($p < 0.10$, two-tailed t-tests) between users and non-users of the measurement alignment techniques (coded 1 if used extensively, else 0)

Table 8
The Relation Between Financial Performance and Performance Measurement Alignment Techniques

	Satisfaction	ROA	Sales Growth	One-year Stock Returns	Three-year Stock Returns
Intercept	-0.353 ^{***} (-3.757)	0.0105 (1.341)	0.192 ^{**} (2.301)	0.0109 (0.091)	-0.301 [*] (-1.923)
Economic Value	0.454 ^{***} (3.026)	0.0037 (1.165)	-0.0427 (-1.516)	-0.0345 (-0.739)	0.0259 (0.493)
Balanced Scorecard	0.537 ^{***} (2.957)	-0.0117 ^{***} (-2.787)	-0.0371 (-0.995)	-0.0616 (-0.968)	-0.0225 (0.283)
Business Modeling	0.676 ^{***} (3.250)	0.0078 [*] (1.67)	0.0477 (1.119)	0.0271 (0.407)	0.0348 (0.428)
Sector Performance		0.354 ^{***} (5.055)	0.660 ^{***} (4.351)	0.869 ^{***} (3.378)	1.223 ^{***} (3.677)
SIZE		0.0018 (1.256)	0.0059 (0.696)	-0.0034 (-0.229)	0.0232 (1.329)
BTOM		-0.0115 ^{***} (-2.643)	-0.204 ^{***} (-3.437)	0.0033 (0.053)	0.0366 (0.339)
Adj. R ²	0.187	0.263	0.258	0.057	0.078

***, **, * indicate p-value < 0.01, 0.05, 0.10 respectively (two-tailed test)

Economic Value = 1 if the firm uses economic value measures, 0 otherwise; Balanced Scorecard = 1 if the firm uses a balanced scorecard, 0 otherwise; Business Modeling = 1 if the firm makes extensive use of formal, causal business modeling, 0 otherwise. See Table 4 for other variable definitions.

Table 9
Coefficients on the Performance Measurement Constructs in the Subsample of Firms
With More Mature Measurement Systems
(No Major Changes in Past Two Years)

	Satisfaction	ROA	Sales Growth	One-year Stock Returns	Three-year Stock Returns
<i><u>Overall Measurement Practices (Hypothesis 1)</u></i>					
Overall Measurement Diversity	0.892 ^{***}	-0.001	-0.042	0.060	0.055
Financial Measurement Focus	-0.022	-0.000	0.004	-0.013	-0.003
Non-Financial Measurement Focus	0.850 ^{***}	-0.001	-0.041	0.064	0.054
<i><u>Alignment With Strategy (Hypothesis 2)</u></i>					
Positive Overall Residual	2.382 ^{***}	0.009	-0.191 ^{**}	0.257 ^{**}	0.319 ^{**}
Negative Overall Residual	0.953 [*]	0.010	-0.072	0.118	0.212
Positive Financial Measurement Residual	0.326	-0.001	-0.117	-0.093	0.133
Negative Financial Measurement Residual	0.088	0.001	-0.031	-0.008	0.063
Positive Non-Financial Measurement Residual	2.125 ^{***}	0.009	-0.115	0.280 ^{**}	0.255 [*]
Negative Non-Financial Measurement Residual	0.889 [*]	0.009	0.010	0.135	0.155
<i><u>Alignment With Value Drivers (Hypothesis 3)</u></i>					
Positive Overall Residual	1.682 ^{***}	0.007	-0.203 ^{**}	0.307 ^{***}	0.320 ^{**}
Negative Overall Residual	-0.334	0.003	-0.069	0.158	0.098
Positive Financial Measurement Residual	0.282	0.016 ^{**}	0.075	0.015	0.237 [*]
Negative Financial Measurement Residual	0.112	0.007 [*]	0.027	0.000	0.086
Positive Non-Financial Measurement Residual	1.479 ^{***}	-0.000	-0.212 ^{***}	0.284 ^{***}	0.244 [*]
Negative Non-Financial Measurement Residual	-0.380	-0.002	-0.083	0.152	0.064
<i><u>Measurement Alignment Techniques (Hypothesis 4)</u></i>					
Economic Value	0.322	0.002	-0.009	-0.008	0.043
Balanced Scorecard	0.444	0.002	0.059	-0.080	-0.050
Business Modeling	1.014 ^{***}	0.001	-0.019	-0.027	0.061

***, **, * indicate p-value < 0.01, 0.05, 0.10 respectively (two-tailed test)

See Tables 4 to 7 for variable definitions.