Management Accounting Research 42 (2019) 1–25

Contents lists available at ScienceDirect

Management Accounting Research

journal homepage: www.elsevier.com/locate/mar

Relationships among types of use, levels of sophistication, and organizational outcomes of performance measurement systems: The crucial role of design choices

Thomas W. Guenther[a,*], Anja Heinicke[b]

[a] Chair of Business Management, esp. Management Accounting and Control, Technische Universität Dresden (TU Dresden), Faculty of Business Management and Economics, D-01062 Dresden, Germany
[b] Technische Universität Dresden (TU Dresden), Faculty of Business Management and Economics, D-01062 Dresden, Germany

ARTICLE INFO

JEL classification:
M40
L20

Keywords:
Performance measurement systems
Levers of control
Cost/benefits
Interactive use
Diagnostic use
Balanced scorecard
Organizational performance

ABSTRACT

Based on the levers of control (LoC) framework and contingency theory, this study examines the relationships among the sophistication of a firm's performance measurement system (PMS), the firm's emphasis on the interactive and diagnostic use of the PMS, and the organizational outcomes of the PMS. We also examine whether firm size moderates these relationships. Based on a survey of 276 midsized enterprises, this study provides empirical evidence of direct, positive relationships between both types of use and the benefits of a PMS. However, PMS sophistication positively moderates the relationship between interactive use and PMS benefits, but it negatively moderates the relationship between diagnostic use and PMS benefits. Thus, our study suggests that an increased emphasis on diagnostic use may reduce benefits for the firm when the PMS is more sophisticated. This result contributes to our understanding of why existing outcome effects of PMSs in the literature have been shown to be sometimes negative, positive or nonsignificant. We also find that, compared with smaller firms, larger firms benefit more from the interactive use of a PMS, whereas our results are robust for other contextual or structural variables.

1. Introduction

This study examines the relationships among the level of sophistication of a firm's performance measurement system (PMS), the firm's emphasis on the interactive and diagnostic use of the PMS, and the organizational outcomes of the PMS. PMSs, particularly the Balanced Scorecard (BSC), have reached a significant rate of adoption in practice (e.g., Crabtree and DeBusk, 2008; Neely et al., 2008; Rigby and Bilodeau, 2009; Speckbacher et al., 2003). In this paper, PMSs are defined as measurement systems encompassing both financial and nonfinancial performance measures used to operationalize strategic objectives (Franco-Santos et al., 2012; Henri, 2006b). Prominent examples of PMSs are the BSC (Kaplan and Norton, 1996), the performance prism (Neely et al., 2002), the performance pyramid (Lynch and Cross, 1991), the quantum performance measurement system (Hronec, 1993) and the French counterpart of the BSC, the tableau de bord (Bourguignon et al., 2004).

Review articles emphasize a variety of positive effects of PMSs (e.g., Franco-Santos et al., 2012; Garengo et al., 2005; Hoque, 2014). However, empirical evidence regarding PMS effects on organizational outcomes is ambiguous and sometimes contradictory (e.g., Davis and Albright, 2004; Lee and Yang, 2011; Micheli and Manzoni, 2010). Some studies find a positive relationship between PMS adoption and organizational performance (OP) (e.g., Bisbe and Malagueno, 2012; Burney and Widener, 2007; Hoque and James, 2000; Ittner and Larcker, 1998), whereas other studies find no relationship (e.g., Ittner et al., 2003a; Perera et al., 1997; Verbeeten and Boons, 2009; Yongvanich and Guthrie, 2009) or even a negative relationship (e.g., Ittner et al., 2003b). Thus, Franco-Santos et al. (2012) and Hoque (2014) call for further empirical research clarifying the relationship between PMS and performance.

One reason for the ambiguous results is the different designs and/or uses of PMSs, as shown by the high heterogeneity in the meta-analysis of Endrikat et al. (2018). Because both PMS design and type of use affect OP (Braam and Nijssen, 2004; Griffith and Neely, 2009) and explain both the costs and benefits of the PMS (Gimbert et al., 2010; Malina and Selto, 2001), design, use and organizational outcomes must be considered mutually related (Ferreira and Otley, 2009; Henri,

* Corresponding author.
E-mail addresses: thomas.guenther@tu-dresden.de (T.W. Guenther), anja.heinicke@tu-dresden.de (A. Heinicke).

https://doi.org/10.1016/j.mar.2018.07.002
Received 28 February 2015; Received in revised form 24 May 2018; Accepted 9 July 2018
Available online 31 July 2018

2006a, 2006b).

Research has focused primarily on either PMS design (e.g., Ittner et al., 2003b; Lee and Yang, 2011) or use (e.g., Grafton et al., 2010; Henri, 2006a; Koufteros et al., 2014), each of which is only partially related in research to organizational outcomes. Furthermore, research on the relationship between PMS design and use is scarce, and the few existing studies either fail to explore the impact on organizational outcomes (e.g., the case study of Agostino and Arnaboldi (2012) and the survey paper of Henri (2006b)) or focus on specific elements of PMS design (e.g., the diversity of PMS measurement in Henri, 2006b).

Thus, to the best of our knowledge, a detailed examination of the outcome effects of the fit of PMS design and use is lacking. Moreover, it is unclear whether the identified individual effects of PMS design and use on organizational outcomes are moderated by the other variable. Therefore, researchers have called for more research on the use and benefits of PMSs (Bukh and Malmi, 2005; Malina et al., 2007; Malmi, 2001; Tuomela, 2005; Wiersma, 2009).

Our study attempts to close this research gap by simultaneously exploring the relationships among the design, use and organizational outcomes of a PMS. We find that diagnostic use can have both beneficial and dysfunctional effects on PMS outcomes and that the fit of each type of use with PMS sophistication is crucial for PMS benefits. Thus, the combination of all three variables determines the extent to which PMS benefits are realized.

Empirical research on PMS design and use has relied on different conceptual frameworks. We use the levers of control (LoC) framework of Simons (1994, 1995, 2014) because it explicitly considers (interactive and diagnostic) use in addition to beliefs and boundary systems as essential management controls (MCs) of a firm. The use of performance measures is essential for the LoC framework (Simons, 1995), and in his most recent book, Simons (2014) explicitly included the BSC as a PMS. From a methodological perspective, empirical PMS research is predominantly based on contingency theory (Franco-Santos et al., 2012; Hoque, 2014), which we also apply in this paper. In this regard, the design and use of a firm's PMS are structural variables of the firm, and we examine the relationship of the fit between them on a firm's organizational outcomes. In alternative models, we examine how different contextual factors and other structural variables—such as ownership structure, environmental uncertainty, firm age, strategic orientation and size—are associated with our structural variables.

Examining costs and benefits of PMSs is especially important in our setting of midsized firms because small and medium enterprises (SMEs) often have limited resources, skills, and capital (Lohrke et al., 2006). Although SMEs play a significant role in the global economy (Cromie et al., 1995; Vachani, 2005), research often focuses on either large (Chenhall, 2003) or small companies (Lopez and Hiebl, 2015; Mitchell and Reid, 2000). However, midsized firms are an important determinant of both growth and employment in economies around the world.[1] Furthermore, midsized firms are an interesting setting since they vary considerably in size, structure, organizational forms, strategies and business systems. Thus, in our paper, we investigate midsized firms to explore how formal controls such as PMSs (often used in large organizations) are used in firms that are not large but nonetheless are sufficiently large to use formal controls (small firms often do not use sophisticated MCs, see, e.g., Sandino, 2007).

Our study contributes to the literature in two ways. First, our paper helps explain previous ambiguous results on the outcome effects of a PMS (e.g., Franco-Santos et al., 2012; Ittner et al., 2003b; Lee and Yang, 2011; Speckbacher et al., 2003) by empirically disentangling the relationships among the design of a PMS (measured as PMS sophistication level), the interactive and diagnostic use of the PMS, and the organizational outcomes of the PMS. Based on survey data for 276 midsized firms, we first provide evidence of significant positive direct relationships between both types of use and PMS benefits. Using polynomial regression analysis, we show that the fit of PMS sophistication level with interactive and diagnostic use is crucial to realize the benefits of a PMS. Thus, we show that the benefits of each type of use are conditional on PMS sophistication. The resulting interaction terms have opposite signs, indicating that a greater emphasis on diagnostic use is more beneficial with a simpler PMS, whereas a greater emphasis on interactive use requires a more sophisticated PMS. In addition, because firms use their PMSs both diagnostically and interactively, the effects overlap and exert both functional and dysfunctional influences on PMS benefits. For a low level of PMS sophistication, more emphasis on both the diagnostic and interactive use of a PMS increases PMS benefits. For higher levels of sophistication, more emphasis on interactive use increases PMS benefits, whereas more emphasis on diagnostic use is dysfunctional and decreases the benefits. This finding may help shed light on the ambiguous and contradictory results in the literature. We find significant effects on the more proximate outcome measure (i.e., PMS benefits) but not on organizational performance (OP).

Second, we contribute to the contingency literature by expanding our knowledge concerning the moderating effect of firm size on the relationships among the diagnostic and interactive use of a PMS (see the call by Chenhall, 2007), PMS sophistication level and PMS benefits. We find that, compared to smaller midsized firms, larger midsized firms design and use PMSs more like large firms do in general and benefit more from greater interactive use of a PMS.

Our results contribute to the LoC framework of Simons (1995) by finding that diagnostic use is more beneficial for a less sophisticated PMS and by showing that incentive systems, in contrast to Simons (1995), are more strongly associated with interactive than diagnostic use. Furthermore, we extend contingency theory by showing that PMS benefits depend on the fit of PMS design and its use. The positive relationship between diagnostic (interactive) use of a PMS and PMS benefits is weaker (stronger) when PMS sophistication is greater. By showing that the two variables are mutually related, our study may spur future research on PMSs to analyze the design and use of PMSs simultaneously. Our paper also suggests that overly sophisticated PMSs may be dysfunctional and that firms should consider using different PMSs for diagnostic and interactive use.

In Section 2, we present the conceptual background for the study and develop our hypotheses. After explaining our research method in the third section, we present and discuss the results in Section 4. Finally, we provide concluding remarks in Section 5.

2. Prior work and hypothesis development

2.1. Conceptual background

We employ the LoC framework of Simons (1995, 2014) as the underlying framework for our study to explain the selection of the focal variables in our research model. Simons (2014: 227) explicitly addresses the BSC as one control system. The LoC framework is particularly suitable for our study because it describes top managers' use of MCs by differentiating interactive and diagnostic use in addition to beliefs and boundary controls. The type of PMS use is regarded as being related to its design. Moreover, the LoC framework of Simons (1995, 2014) and the related literature also consider the relationship between the design of an MC system (specifically, of a PMS, as in this study) and its benefits. Finally, the LoC framework has been used in prior MC and PMS research (e.g., Bisbe and Otley, 2004; Henri, 2006a; Koufteros et al., 2014; Widener, 2007).

Following Simons (1995), *diagnostic controls* are formal information

[1] On the date our data were collected, midsized German firms accounted for 84.2% of all apprenticeship contracts, 59.4% of all employees, and 54.8% of the total value added in 2011 (see the fact book "German Mittelstand" of the German Ministry for Commerce; http://www.bmwi.de/BMWi/Redaktion/PDF/Publikationen/factbook-german-mittelstand,property=pdf,bereich=bmwi2012,sprache=de, rwb=true.pdf).

systems that constrain behavior by monitoring actions and results against preset targets, whereas *interactive controls* communicate top managers' primary concerns through the organization. Interactive controls provide information on strategic uncertainties and facilitate debate and dialogue on threats and opportunities. In our paper, we explicitly follow the revision of the LoC framework by Tessier and Otley (2012: 178), who do "not consider diagnostic and interactive controls as control systems in their own right but rather as a description of how control systems are used". Thus, we address the diagnostic and interactive *use* of a PMS as an attribute of an MC system associated with the dual role of controls as explicit managerial intentions to shape a PMS.

Notably, MCs should be understood not as a dichotomous decision to switch the control on or off but as a continuous decision between different degrees of emphasis. In addition, empirical evidence (e.g., Widener, 2007, Heinicke et al., 2016) shows that diagnostic and interactive controls—as well as beliefs and boundary controls—are mutually interrelated (Simons, 1994). Furthermore, both interactive and diagnostic use are desirable, and each of them can be both "enabling" and "constraining" (Tessier and Otley, 2012). We also adopt the view of Tessier and Otley (2012) that managers can decide whether to reward performance and connect a PMS to an incentive system.

Contingency theory is often used as an approach for MC and PMS research, as it is in our paper (Chenhall, 2003; Franco-Santos et al., 2012; Otley, 2016). We address contingency theory to explain our methodological approach. Applying the typical classification of empirical contingency approaches (e.g., Burkert et al., 2014; Drazin and Van de Ven, 1985; Gerdin and Greve, 2004, 2008), our paper follows the Cartesian approach because we measure all variables on a continuous scale. Specifically, we explore contingency fit (Gerdin and Greve, 2004), also called interaction fit (Drazin and Van de Ven, 1985), i.e., the relationship of structural variables with organizational outcomes. Thus, we assume that finding the optimal design and use of a PMS is time consuming and requires considerable financial and other resources. Therefore, we believe that most firms will not be in equilibrium, contrary to research based on congruence fit (also called selection fit), and that lower-performing firms may also survive, leading to variations in OP, as observed in our data. Chenhall (2003: 135) concludes that "if disequilibrium conditions are assumed, then it may be useful for contingency based studies to first establish the adoption and use of MCs, then to examine how they are used to enhance decision quality, and finally to investigate links with organizational performance". Following Chenhall's advice, we relate PMS design and use as structural variables of the firm to PMS benefits, viewed as a more proximate organizational outcome; however, we also examine the relationship of PMS design and use with OP in an additional model. Furthermore, various contextual factors and other structural variables—such as ownership structure, environmental uncertainty, firm age, strategic orientation and size—can be associated with both types of structural variables. Moreover, in further robustness tests, we observe the association of other MCs (e.g., budgeting and strategic planning) as structural variables with PMS design and use.

In sum, contingency theory is the basis for our empirical approach, whereas the LoC framework is used to establish the relationships among the three variables in our research model: PMS design, use and benefits.

2.2. Prior findings on the outcome effects of PMS adoption and implementation

Empirical evidence on the relationship between PMS adoption and OP is mixed, inconclusive and contradictory (Franco-Santos et al., 2012; Ittner and Larcker, 1998; Michel and Manzoni, 2010). Some studies find a positive relation between PMS adoption and OP (e.g., Bisbe and Malagueno, 2012; Burney and Widener, 2007; Hoque and James, 2000; Ittner and Larcker, 1998), whereas other studies find no relationship (e.g., Ittner et al., 2003a; Perera et al., 1997; Verbeeten and Boons, 2009; Yongvanich and Guthrie, 2009) or even a negative relationship (e.g., Ittner et al., 2003b). Schaeffer and Matlachowsky (2008) find that for three out of six case study firms, the level of sophistication of their BSCs decreased over time, and two firms even discontinued BSC use. Similarly, Kasurinen (2002) reports that his case study firm halted the adoption of a type II BSC (i.e., a BSC with cause-and-effect relationships used to break down the strategy) due to implementation problems and doubts about the future benefits of the PMS. A recent meta-analysis conducted by Endrikat et al. (2018) on the effect of PMSs on OP includes 25 empirical papers and reports a significant, positive population effect overall. However, they also find high heterogeneity in the reported results due to different types of PMS design and use. Furthermore, they find higher effect sizes when OP assessment is based on organizational capabilities[2] than when it is based on financial measures. Thus, the ways to measure the organizational outcomes of a PMS also seem crucial.

Papalexandris et al. (2004: 364) note that the cost and time invested in BSC implementation "may well outweigh improvements in organizational performance". Similarly, based on their quasiexperimental study, Griffith and Neely (2009) report significant variation in the impact of BSC implementation on branch performance in a large distribution firm. Michel and Manzoni (2010: 466) state that "if done poorly, they [PMSs] can be very expensive and not only ineffective but harmful and indeed destructive". They conclude that PMSs can be both functional and dysfunctional for firms and that PMS design and types of use determine organizational outcomes. Thus, there is a need to explore the isolated effects of the level of sophistication, diagnostic use and interactive use, as well as the fit among them, on organizational outcomes.

2.3. Design and PMS sophistication level

A PMS represents a set of metrics that help to quantify information about a firm's actions and thereby provides an overview of a firm's performance (Neely et al., 1995). Hence, the identification and selection of these metrics should reflect a firm's strategy and help to translate its strategy into day-to-day business activities (Kaplan and Norton, 2001). Different surveys show that PMS adoption, especially BSC adoption, is substantial in practice (e.g., Crabtree and DeBusk, 2008; Neely et al., 2008; Rigby and Bilodeau, 2009; Speckbacher et al., 2003). However, merely examining adoption by using a dichotomous question regarding whether a PMS is implemented seems inadequate. Furthermore, the Kaplan and Norton concept of the BSC has changed considerably over time (Hoque, 2014; Speckbacher et al., 2003). Moreover, firms claim to employ a PMS if they use a mixture of financial and nonfinancial measures, whereas Kaplan and Norton (2001) note that their concept of BSC goes far beyond the mere use of such measures. Thus, to capture differences in PMS design, it is necessary to use a continuous measure for adoption that captures the *PMS sophistication level*, defined as the extent of development of essential design features, functions or processes of a specific PMS within a firm.

The PMS sophistication level represents the technical quality of the PMS with respect to design characteristics, for example, as suggested for the BSC by Kaplan and Norton (1996) or as addressed by the PMS framework of Ferreira and Otley (2009) (e.g., multiple perspectives of financial and nonfinancial indicators, strategy maps, action plans, and connections to incentive systems). Nevertheless, PMS sophistication levels should be differentiated from the degree of embeddedness or dissemination within an organization (e.g., how many divisions or departments are using the PMS).

In the literature, PMS sophistication level has been captured by the two typologies (Franco-Santos et al., 2012; Speckbacher et al., 2003) described and compared in Fig. 1. Speckbacher et al. (2003) suggest

[2] The organizational capabilities measures analyzed in their study correspond to the different criteria for measuring PMS benefits in our study.

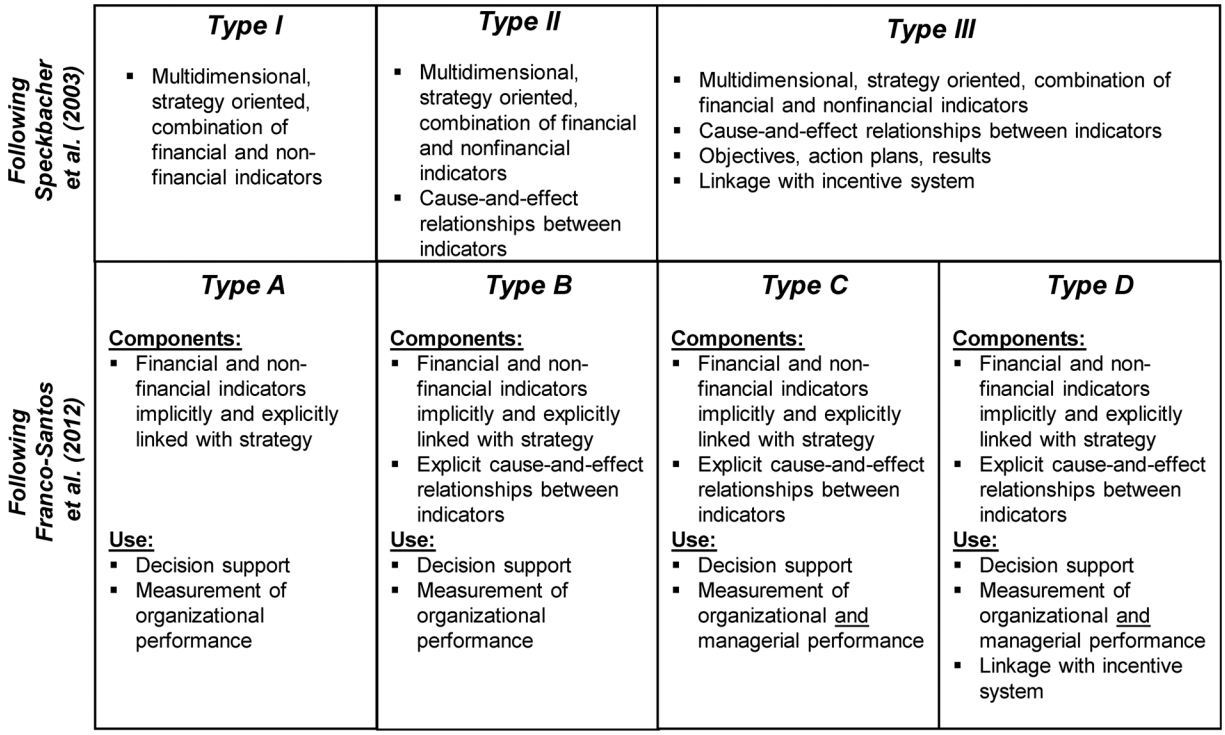


Fig. 1. Typologies of PMS sophistication levels.

that a PMS may develop over time and, hence, that research should consider the stages of PMS development because the content, implementation, and expected PMS benefits may vary based on the type of PMS.

According to Speckbacher et al. (2003), the three identified types of PMSs (see Fig. 1) reflect successive levels of sophistication in the process of implementing a PMS. In their review, Franco-Santos et al. (2012) recognize that "… Speckbacher et al.'s typology advocates that the three types of BSC are incrementally related, whilst [we] find that this might not necessarily be the case" (Franco-Santos et al., 2012: 8). Franco-Santos et al. (2012) show that research has linked financial and nonfinancial performance measures with incentive pay without necessarily considering cause-and-effect relationships between these performance measures (e.g., Ittner et al., 2003a). Furthermore, they generalize their typology beyond Speckbacher et al. (2003)'s focus on the BSC for (contemporary) PMSs in general.

The key differences in PMS sophistication are the extent to which performance measures are linked to strategy, the extent to which strategy is explicitly described through cause-and-effect relationships, and the extent to which managerial performance is linked to reward systems. Speckbacher et al. (2003) also refer to objectives and action plans that Franco-Santos et al. (2012) describe as design elements of a PMS intended "to inform managerial decision-making […] and [to] evaluate organizational performance" (p. 81). Thus, in our paper, we use a continuous multidimensional measure based on the features suggested by Speckbacher et al. (2003) and similarly presented by Franco-Santos et al. (2012).

2.4. Use of the PMS

In addition to design (i.e., PMS sophistication level), the emphasis on different *types of use* is key to PMS outcomes. In this regard, Simons' (1995) distinction between interactive and diagnostic use in the LoC framework helps to differentiate and more deeply explore the relationships among different types of use, different levels of sophistication, and PMS benefits.

Following Simons (2014: 227), *diagnostic controls* based on a cybernetic logic are characterized as any formal information system used "to (1) set a goal in advance, (2) measure outputs, (3) compute and calculate performance variances, and (4) use that variance information as feedback to alter inputs and/or processes to bring performance back in line with preset goals and standards". Simons (2014: 234) describes *interactive controls* as "formal information systems that managers use to personally involve themselves in the decision activities of subordinates", with a focus on identifying new strategies. Thus, management devotes considerable attention to interactive controls through face-to-face discussions with subordinates and peers and through continuous debates regarding data, assumptions, and action plans (see also Bisbe et al., 2007). Thus, interactive use of a PMS focuses on strategic uncertainties and facilitates the emergence of new strategies, whereas diagnostic use focuses on critical performance measures and aids the implementation of existing strategies (Simons, 1991, 1995, 2014). Simons (1995) maintains that goal achievement is linked to incentives through diagnostic controls. He regards incentive systems as essential for supporting diagnostic use because they enable the achievement of goals set by management. Moreover, any control system can be used in an interactive or diagnostic way.

The differentiation between diagnostic and interactive use is suitable for our setting because we explicitly address top management in our survey, which is also the focus of the LoC framework of Simons (1995).[3] Moreover, the LoC framework has often been applied in MC research to distinguish different types of use (e.g., Agostino and Arnaboldi, 2012; Bisbe and Otley, 2004; Henri, 2006a; Koufteros et al., 2014; Widener, 2007).

[3] Other classifications of "use"—in addition to that of Simons (1995)—describe the dual role of controls (Grafton et al., 2010; Tessier and Otley, 2012). For example, Grafton et al. (2010) differentiate "feedback control use" from "feed-forward control use" (i.e., the formulation and use of predictions), whereas Van Veen-Dirks (2010) differentiates between "decision-facilitating" and "decision-influencing" uses. Other classifications are addressed by Ahrens and Chapman (2004); Malmi (2001); Specklé and Verbeeten (2014) and Wouters (2009).

2.5. Organizational outcomes and PMS benefits

The rationale for adopting and using a PMS is to create benefits for the firm and ultimately to improve OP; these factors reflect the *organizational outcomes* of the firm (Ittner et al., 2003a). Referring to Birnberg et al. (1983) and Kren and Liao (1988), Chenhall (2003) argues "that links between MC systems, context and performance can be tenuous because they involve many factors concerning the quality of managing the production processes". Thus, in addition to the analyzed independent variables, OP is influenced by numerous other contextual and structural variables that may be difficult to control for. Therefore, focusing on a less distant outcome variable, such as *PMS benefits* in our study, seems more appropriate for investigating how firms can profit from PMS design and use.

Furthermore, with respect to MC research based on contingency theory, Chenhall (2003) notes that researchers often do not explicitly mention that firms are assumed not to be in equilibrium. Thus, in contrast to research using congruence fit or selection fit models, studies like ours—which address the interaction or contingency fit—assume that firms other than the best-performing firms can survive and that we can still observe variance in OP (see Burkert et al., 2014; Drazin and Van de Ven, 1985; Gerdin and Greve, 2004, 2008). In addition to this methodological argument, it is worthwhile to explore PMS benefits more closely from a theoretical perspective. Franco-Santos et al. (2012: 96–97) conclude that "growing consensus in the literature seems to be that [PMSs] do not automatically improve firm performance. Evidence suggests that it is the way these systems are designed, developed, and, more importantly, used that brings about performance improvements". Accordingly, they call for further research on the circumstances under which we can expect to find positive or negative consequences of PMSs. Thus, efforts to examine outcomes beyond OP and explore the source of benefits expand the current literature.

PMS benefits can be defined as the firm's satisfaction with the purposes of adopting a PMS as a control system and its intended pay-offs for users (Speckbacher et al., 2003). The literature addresses a wide list of PMS benefits (e.g., Kaplan and Norton, 1996; Malmi, 2001; Nørreklit, 2000; Wiersma, 2009). For example, PMSs have been recommended because they foster strategy implementation by strengthening the strategic focus of managers, clarifying and improving the communication of firm strategy, and thus aligning people's behavior with the strategic objectives of the firm (e.g., Ahn, 2001; Franco-Santos et al., 2012; Garengo et al., 2005; Jazayeri and Scapens, 2008; Kaplan and Norton, 2006). Furthermore, PMSs can support the strategy formulation process by allowing firms to assess whether an intended or emerging strategy is adequate and can encourage executives to understand strategy as a continuous process of rigorously reassessing and modifying strategies (Bisbe and Malagueño, 2012; Franco-Santos et al., 2012; Garengo et al., 2005; Kaplan and Norton, 2008). Speckbacher et al. (2003) summarize the potential outcomes of PMSs in a comprehensive list of 17 expected benefits, which are listed in Panel B, Table 2.

To conclude, we decided not to use OP as an outcome variable for our research model and chose to use PMS benefits as a more proximate outcome variable based on the 17 PMS benefits listed by Speckbacher et al. (2003). Nevertheless, we control for relationships with OP in addition to relationships with PMS benefits.

2.6. Hypotheses

Fig. 2 depicts the research model and the hypotheses, which are developed next.

2.6.1. Type of use and PMS benefits

As Franco-Santos et al. (2012: 96) observe, "The growing consensus in the literature seems to be that [PMSs] do not automatically improve firm performance". Franco-Santos et al. (2012) conclude that design characteristics and the type of PMS use affect firm outcomes. Firms use MCs in different ways to guide organizational behavior in the intended way (Micheli and Manzoni, 2010; Simons, 1995). Prior research indicates the importance of investigating different types of PMS use (e.g., Braam and Nijssen, 2004; Henri, 2006a; Malmi, 2001; Widener, 2007). Moreover, Bukh and Malmi (2005) note that PMS use affects PMS benefits. The literature addresses multiple benefits of PMSs (e.g., Kaplan and Norton, 1996; Malmi, 2001; Nørreklit, 2000; Wiersma, 2009), summarized in the comprehensive list of 17 benefits suggested by Speckbacher et al. (2003).[4] Although Bisbe and Malagueno (2012) conclude that the diagnostic use of a PMS calls into question the benefits of PMS implementation, an emphasis on both diagnostic and interactive use of PMSs can be valuable for a firm. Indeed, according to Simons (1995, 2014), emphasis on the diagnostic use of a PMS helps a firm to implement its intended strategy, constrain behavior and monitor whether the targets for critical performance variables are achieved using a cyclical approach of checks and balances. In contrast, an emphasis on interactive use promotes innovation and stimulates the emergence of new strategies by inspiring communication between hierarchical levels and communication with subordinates. If the use of a PMS complements a firm's strategy, the firm should expect a positive benefit (Braam and Nijssen, 2004; Micheli and Manzoni, 2010).

Prior empirical work has explored the relationship between different *types of uses* and *organizational outcomes* but has implicitly disregarded differences in PMS sophistication levels. For example, Henri (2006a) finds that a greater emphasis on diagnostic use is negatively associated with market orientation, entrepreneurship, innovativeness and organizational learning as organizational outcomes, whereas interactive use is positively associated with these outcomes. Interestingly, the dynamic tension of both interactive and diagnostic use, measured as an interaction term, has no significant effect. Henri (2006a) concludes that there is a mismatch between the diagnostic use of a PMS and the requirements for the four analyzed capabilities of his study. Grafton et al. (2010) show that the use of feedback controls is positively associated with the extent of existing organizational capabilities, which in turn is positively associated with business unit performance, whereas the use of feedforward controls is positively connected with the development of new capabilities. In contrast to Henri (2006a); Koufteros et al. (2014) find, not only for interactive but also for diagnostic PMS use, a positive relationship with three types of organizational capabilities as benefit measures (i.e., strategic management, operational and external stakeholder-related capabilities). These three benefit measures are also positively associated with different OP measures. In contrast to Henri (2006a), the authors also find a significant positive effect of the dynamic tension of both types of use. Other studies also examine the relationship between use and organizational outcomes; however, they focus on either diagnostic or interactive use or do not explicitly differentiate between the two types of use (e.g., Bisbe and Otley, 2004; Braam and Nijssen, 2004; de Geuser et al., 2009; Kihn, 2007). For example, Braam and Nijssen (2004) report a negative effect of "measurement-focused BSC use" (which is similar to diagnostic use) and a positive effect of "strategy-focused BSC use" (which is similar to interactive use) on OP, whereas Bisbe and Otley (2004) analyze the interactive use not only of different MCs but also of BSCs and find positive relationships with innovation and performance.

To conclude, prior works report ambiguous results regarding the relation between diagnostic use and PMS benefits but positive results for interactive use. Thus, we propose a direct relationship with interactive use but a null hypothesis for the relationship with diagnostic use. All the following hypotheses are based on contingency theory because we examine the relationship between a structural variable of the firm—i.e., the emphasis on PMS use—and a (proximate) outcome

[4] In the following discussion, when we refer to the "benefit" of a PMS, we consider the net benefit to be the difference between the various benefits of a PMS and the costs of its adoption and use.

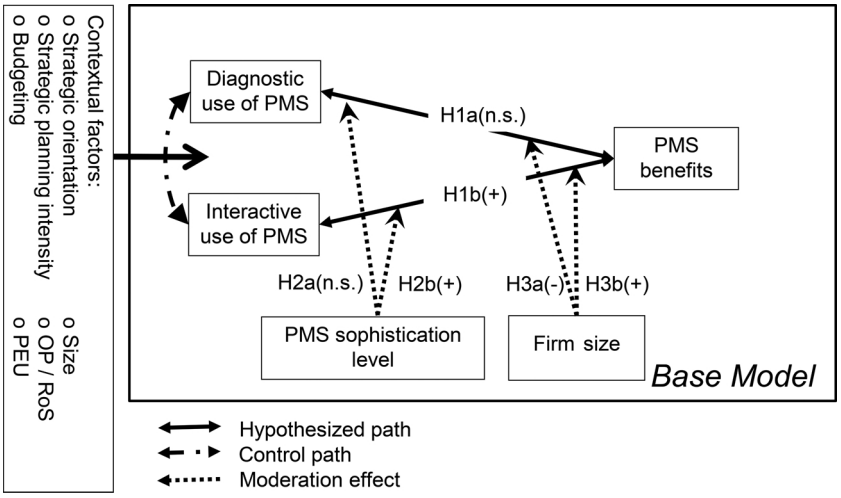


Fig. 2. Research model.

measure:

H1a. Emphasis on the diagnostic use of a PMS is not associated with PMS benefits.

H1b. Emphasis on the interactive use of a PMS is positively associated with PMS benefits.

2.6.2. Fit between PMS sophistication level and use of a PMS

In their review of strategic PMSs, Micheli and Manzoni (2010) identify two key factors that affect the benefits obtained from a firm's PMS: (1) whether the PMS design reflects the assigned purpose and (2) whether the PMS is linked to the firm's strategy formulation and implementation. Kaplan and Norton (1996) propose the concept of the BSC and argue that firms should use the BSC as a strategic management system (Kaplan and Norton, 1996). Thus, the BSC should incorporate cause-and-effect relations between the chosen measures, and these if-then statements (to use their terms) should pervade all four perspectives in which the measures are grouped in order to ensure that the PMS is used to formulate firm strategy and facilitates communication and learning, thus enabling effective control (Bisbe and Malagueno, 2012; Malina and Selto, 2001; Malina et al., 2007). Such a PMS is more sophisticated than a traditional PMS. Thus, PMS benefits can be expected to vary with PMS sophistication levels (Malmi, 2001). That said, PMS sophistication has rarely been examined in the empirical literature.

Speckbacher et al. (2003) find that firms with a more developed PMS are more satisfied with their PMS. However, a low percentage of firms in their study reported using a strategic type of BSC (type III following the classification in Fig. 1). The direct relationship between *design* and *organizational outcomes* has been examined in previous studies. For example, Ittner et al. (2003b) explore the relationship between OP and measurement system satisfaction with two different PMS design approaches: (1) the measurement and use of a diverse set of financial and nonfinancial measures (i.e., similar to type I in Speckbacher et al. (2003), see Fig. 1) and (2) the alignment of these measures with the firm's strategy and value drivers (similar to type II). Whereas measurement diversity is positively associated with satisfaction and stock market returns, Ittner et al. (2003b) find little support for the relationship of alignment with strategy and the relationship of value drivers with satisfaction and OP. Lee and Yang (2011) find that both PMS adoption (moderated by organizational structure) and the stage of a PMS using the classification of Speckbacher et al. (2003) (conditional on competition) are positively associated with OP.

Several papers argue that inadequate PMS design can promote PMS failure and dysfunction. Following Blau (1964) and Jaworski and Young (1992), we define *PMS dysfunction* as a negative consequence of the adoption and use of a PMS for the organization and define *PMS failure* as the complete abandonment of a PMS following PMS dysfunction. Dysfunctional effects are diverse, such as "ossification" (i.e., organizational paralysis due to excessively rigid PMS); the crowding out of intrinsic motivation, or "tunnel" effects (i.e., a strict managerial focus on those KPIs included in the PMS); gaming; smoothing; biasing; or creating a system that is too complex and difficult to understand, which leads to information overload, the spread of employee awareness over too many objectives or increased administrative costs (e.g., Birnberg et al., 1983; Henri, 2006a; Ittner et al., 2003a; Schick et al., 1990; Simons, 1995; Smith, 1995; Townley et al., 2003). Yongvanich and Guthrie (2009) observe no significant differences in satisfaction and perceived benefits gained from different types of BSCs. Furthermore, Gimbert et al. (2010) find no difference in the support of strategy formulation (i.e., as a PMS benefit) between having no PMS and having a "nonstrategic", merely operational PMS (i.e., a PMS with a low sophistication level). However, these studies do not explicitly account for the fact that different types of use may be related to the effects of PMS sophistication levels on organizational outcomes.

PMS sophistication level and *use of a PMS* may not be independent of each other; rather, they may be connected and may both influence PMS benefits (Bukh and Malmi, 2005). In their review of the PMS literature, Franco-Santos and Bourne (2005) identify process and contextual factors that enhance the effectiveness of PMSs. Several of these factors relate to the design (and implementation) and subsequent use of a PMS and thus signal that PMS use and sophistication level are closely related (see also Ferreira and Otley, 2009; Henri, 2006a, 2006b; Malmi, 2001; Wouters, 2009). Henri (2006b: 97) explicitly states that PMS design and use are "two closely intertwined dimensions of PMS that must be examined specifically".

Furthermore, Micheli and Manzoni (2010) argue that a PMS should be designed based on the purpose for which it is developed and that there is no "one size fits all" model. Thus, they argue that the benefits of a measurement system strongly depend on both its intended roles (i.e., type of use) and its characteristics. With respect to the LoC framework, some researchers argue that the diagnostic use of a PMS restricts it to a measurement control system, whereas interactive use renders the PMS a strategic management tool (Henri, 2006a; Kaplan and Norton, 2001). Thus, it seems worthwhile to explore whether PMS sophistication levels moderate the relationship between both types of PMS use and PMS benefits.

Simons (1995: 61) explicitly refers to profit plans and budgets as "the most pervasive diagnostic control systems in modern business firms". The list of typical diagnostic controls presented in Simons (1995) in his more recent book is expanded by the BSC concept

(Simons, 2014: 227). Furthermore, Simons (1995, 2014) explicitly links diagnostic controls to formula-based incentive systems "as a way of powering up, or motivating goal achievement" (Simons, 2014: 229). In his analysis of 12 case studies, Merchant (1989) finds that all diagnostic performance measures are linked to reward systems (Merchant, 1989), whereas nonfinancial indicators are deemed meaningless if they are not linked to rewards (Eccles, 1991). Finally, Ahn (2001) underlines the importance of explicit targets based on his case study firm.

Relating these arguments to PMS sophistication levels indicates that greater emphasis on *diagnostic use* would require a more sophisticated PMS. In addition to the measurement of financial and nonfinancial measures, explicit targets and the link to an incentive system are required. In the typology of Speckbacher et al. (2003), the inclusion of these features results in a type III PMS; Franco-Santos et al. (2012) would call it a type D PMS (see Fig. 1).

In contrast to the above argument that diagnostic use requires a high level of sophistication, Henri (2006b) postulates a negative relationship between monitoring (representing diagnostic use of a PMS) and the diversity of measurement (i.e., his measure for PMS sophistication levels). He argues that diagnostic control is strongly related to financial information and to budgetary control, which requires only a limited PMS sophistication level. Based on his case study findings, Tuomela (2005) argues that setting targets was perceived as a challenge by the case firm. Furthermore, researchers have argued that it is difficult to explicate the impact of particular nonfinancial measures on financial results (Anthony and Govindarajan, 1998) and that experience regarding the financial effects of nonfinancial factors—typically demonstrated in a PMS via cause-and-effect relationships (one of the elements of PMS sophistication levels)—is lacking. The dysfunctional effects of PMSs that are discussed in the literature are often connected to the use of a PMS as the basis for managerial incentive systems (e.g., Birnberg et al., 1983; Henri, 2006a; Simons, 1995; Smith, 1995; Townley et al., 2003). Specifically, the substitution of extrinsic motivation for intrinsic motivation (i.e., incentive systems), a focus on those KPIs in management that are the basis for rewards, and the gaming effects in incentive systems linked to a PMS raise questions about whether PMSs linked to incentive systems are beneficial for diagnostic use.

Furthermore, diagnostic use is characterized by two important features connected with mechanistic controls: first, tight control over operations and strategies and, second, highly structured channels of communication and restricted information flows (Burns and Stalker, 1961; Henri, 2006a). Regarding the first characteristic, diagnostic use is linked to action plans derived from strategies, detailed target setting, and variance analysis of the differences between planned and actual numbers. These features are also typically addressed by PMSs that incorporate nonfinancial measures, such as the BSC. Diagnostic use mainly concerns single-loop learning (Argyris and Schön, 1978) in the sense of "plan, do, check, act". However, certain features of a PMS—such as cause-and-effect relationships, incentive systems, or explicit links to strategy—are more beneficial for double-loop learning, which is needed for interactive use to innovate and develop new strategies.

Regarding the second feature of mechanistic controls, diagnostic use often follows standardized reporting processes and responsibilities (Abernethy and Brownell, 1999), whereas interactive use requires cross-functional processes and communication and a free flow of information (Kohli and Jaworski, 1990). Vandenbosch (1999) argues that diagnostic use results in corrective action at best, which is insufficient for a firm's development of new and creative ideas. At worst, diagnostic use leads to unproductive discussions about the validity of KPIs, reasons for variance, and actions to take for improvement (similarly, Henri, 2006a). The latter arguments suggest that diagnostic use would require only a limited PMS sophistication level, i.e., a PMS designed to meet the basic needs of mechanistic controls and single-loop learning. Similarly, in their literature review, Franco-Santos et al. (2012) find evidence that firms do not necessarily combine incentive systems with their PMSs. Malmi (2001) also finds that several of his 17 case firms use BSCs for informational purposes only and do not use them as management systems, which would not require target setting or a link to the incentive system; thus, these firms would require only a less sophisticated PMS for diagnostic use.

In summary, empirical findings suggest that a greater emphasis on diagnostic use may not necessarily require a higher level of sophistication and that there might be dysfunctional effects of an overly sophisticated PMS. However, the theoretical literature is ambiguous, supports contradictory arguments, and creates tensions between contrary positions. Unfortunately, the empirical evidence is scarce and limited to certain features of PMS design. Thus, based on the ambiguous theoretical arguments and contradictory empirical findings, we propose for the association of the fit between the two structural variables, emphasis on diagnostic use and PMS sophistication level, with PMS benefits the following:

H2a. The relationship between emphasis on diagnostic PMS use and PMS benefits is unaffected by PMS sophistication levels.

With respect to *interactive use*, Kaplan and Norton (1996, 2001) explicitly describe the BSC as a strategic management system, and researchers relate resulting strategic capabilities to the interactive use of a PMS (e.g., Agostino and Arnaboldi, 2012; Bisbe and Malagueno, 2012; Henri, 2006b; Specklé and Verbeeten, 2014; Tuomela, 2005; Wouters, 2009). Regarding the various design elements of a PMS, the inclusion of nonfinancial measures in addition to financial measures can facilitate the exploration of other perspectives in addition to a purely financial perspective because nonfinancial measures are assumed to be leading indicators of future financial performance (Kaplan and Norton, 1996; Tuomela, 2005). Furthermore, describing strategies based on cause-and-effect relationships, defining target values and action plans for nonfinancial performance measures, and connecting PMSs with incentive systems can contribute to the search for new strategies and therefore to the enabling role of PMSs (Ahrens and Chapman, 2004; Wouters, 2009), engendering discussions between top management and employees (Simons, 1995). Incentive systems are also necessary for interactive PMS use, although they differ from those applied for diagnostic PMS use. Specifically, Simons (2014) recommends subjective assessments of managers "to reward an individual's innovative efforts and contributions". In this regard, rewards should be based on contributions and efforts rather than on results to inspire communication and information sharing (Simons, 2014). Furthermore, Tuomela (2005) finds in his case study that targets are used implicitly as the basis for discussions in firms where it is difficult to establish explicit targets. Moreover, (re)focusing the attention of subordinates on strategic uncertainties and strategic issues is part of this interactive use of PMS (Henri, 2006b; Simons, 1995). Interactive PMS use can yield a better understanding of the business by revealing cause-and-effect relationships (Chenhall, 2005) or can lead to cross-functional and cross-hierarchical discussions by breaking down firm strategy (Abernethy and Brownell, 1999; Abernethy and Lillis, 1995), thereby overcoming a restricting and channeled information flow within internal reporting.

Regarding empirical evidence, Henri (2006b) examines the relationship between different types of PMS use and the diversity of PMS measurement, which represents PMS sophistication levels. He finds that only attention focusing and strategic decision making (which can be associated with interactive use in the LoC framework) are positively related to the diversity of measurement, but monitoring, which is a typical diagnostic use, is not. However, they do not explore other design elements of PMSs, such as cause-and-effect relationships describing firm strategy, incentive systems or action plans. Furthermore, the case study of Agostino and Arnaboldi (2012) investigates the relationship between the interactive and diagnostic uses of the BSC and its design characteristics but does not examine the consequences for organizational outcomes. Building on previous research (e.g., Speckbacher et al., 2003;

Kaplan and Norton, 1996), Agostino and Arnaboldi (2012) identify four design choices (selecting key performance measures, cascading the PMS throughout the firm, setting targets, and implementing a reward system). They find that BSCs used for interactive purposes are associated with a predominance of nonfinancial KPIs, cascading of BSCs from the corporate report, implicit targets, and links to reward systems (i.e., a higher PMS sophistication level), and vice versa for diagnostic use. The last finding conflicts with Simons' LoC framework, which suggests that reward systems are explicitly associated with diagnostic use (Simons, 1995).

To conclude, both the theoretical considerations and initial empirical findings are less ambiguous for interactive use than for diagnostic use. Thus, we expect that the fit between the two structural variables, emphasis on interactive use and PMS sophistication level, fosters PMS benefits. Hence, we propose the following:

H2b. The positive relationship between an emphasis on interactive PMS use and PMS benefits is stronger in firms with more sophisticated PMSs.

2.6.3. The role of firm size

Although the development and implementation of a PMS is a time-consuming and expensive process, many firms implement such systems to more efficiently align the firm with its strategy and to control critical performance measures. Therefore, contingency variables that affect PMS design and use and their impact on firm performance must be identified (Franco-Santos et al., 2012). There is broad consensus in the literature that firm size is an important contextual factor (e.g., Henri, 2006a; Hoque and James, 2000). In this study, we explore whether firm size (i.e., size variations within our sample of midsized firms) moderates the relationships in our research model.

The literature indicates that larger firms use more sophisticated control systems (Abdel-Kader and Luther, 2008; Chenhall, 2003) and are more likely to adopt PMSs such as the BSC (Hoque and James, 2000; Speckbacher et al., 2003). When firms grow in size, managers have to address a broader range of information, communicate such information through more and different channels, and coordinate resources and actions across a larger number of employees, more regions or more and larger legal entities (Chenhall, 2003, 2007). Thus, larger firms typically adopt tighter MCs, such as diagnostic PMS use (Baird et al., 2004). In contrast, smaller firms typically rely more on interpersonal and informal controls (Merchant, 1981; Heinicke et al., 2016), have lower coordination needs due to their smaller scale of operations (Chenhall, 2007), and communicate more directly because of their smaller staff (Merchant, 1981). Furthermore, management and strategy setters are closer to the market and to operations (Hoque and James, 2000). Bruns and Waterhouse (1975) find that larger firms are associated with greater administrative controls, whereas smaller firms are characterized by personal controls.

In his LoC framework, Simons (1995) notes that firms must manage the tension between limited attention and opportunity-seeking behavior and must use both interactive controls (to facilitate the emergence of new strategies) and diagnostic controls (to implement existing strategies, ensure goal achievement, and focus on critical performance measures). Smaller firms may benefit more from a use that allows them to compensate for their comparative shortcomings. Because smaller firms use more interpersonal controls and have lower constraints on knowledge sharing, coordination and communication, we expect that they benefit more from monitoring crucial performance measures and from tight control of goal achievement. In contrast, because larger firms face more strategic uncertainty and more complexity than smaller firms, they benefit more from interactive PMS use, which better satisfies the above-described needs of communication, coordination and attention focusing. Finally, we hypothesize a moderation approach (form of fit) of contingency theory by exploring size as a contingency factor:

H3a. The relationship between the emphasis on diagnostic PMS use and PMS benefits is weaker in larger firms than in smaller firms.

H3b. The positive relationship between the emphasis on interactive PMS use and PMS benefits is stronger in larger firms than in smaller firms.

3. Research method

3.1. Sample and data collection

To test our hypotheses, we gather empirical data from German midsized firms by using a structured questionnaire.[5] For this purpose, we use the Amadeus database to create our target sample and to collect archival data on these firms. We focus on midsized firms with sales between € 20 million and € 1 billion,[6] and we remove firms focused on public administration and financial services from our sample,[7] leaving a sample comprising production, trade, and service firms. The database contains 16,018 firms within the sales range in focus. First, we classify firms based on ownership, distinguishing between family and nonfamily firms, to ensure that the sample has broad representativeness. Because most studies use ownership and family involvement in firm management to distinguish family from nonfamily firms (e.g., Astrachan et al., 2002), we label a firms as a family firm if the family owns more than 50.00% of the firm.[8] Subsequently, we classify these firms in four sales categories (i.e., 20–39, 40–99, 100–249, and 250–1,000 million €) to produce a 2 × 4 ownership/sales matrix. We then randomly select approximately every 6th firm from each cell to generate our stratified random sample of 2500 firms. After correcting this sample for closed, moved, or merged firms, we obtain a final target sample of 2452 firms.

For the survey process, we follow the common guidelines for survey design (e.g., Dillman et al., 2014) to improve the response rate. The questionnaire was developed in 2010 and pretested in February/March 2011. In May 2011, we sent the questionnaire with a cover letter to one member of each firm's top management team (e.g., CEO, CFO). In June 2011, we sent a reminder letter to nonrespondents with an access link to the online version of the questionnaire. In total, by August 2011, we received 276 usable responses, representing a response rate of 11.26%. This response rate is comparable to that in similar studies (Bisbe and Malagueno, 2012; de Geuser et al., 2009; Widener, 2007).

We conduct a nonresponse analysis to compare respondents and nonrespondents and find no significant differences in sales or ownership (family/nonfamily) between the respondents and the target population ($p > 0.05$). Furthermore, when we compare sales, profit margin (RoS) and RoE, we do not find significant differences between our sample and the target population ($p > 0.05$). We also do not find significant differences ($p > 0.05$) between early and late respondents in terms of diagnostic and interactive use of PMSs and PMS sophistication levels. However, late respondents report significantly lower PMS

[5] The comprehensive nine-page survey instrument was also used to collect data for another paper by the authors, which examines the relationship between a flexible culture and the LoC concept and thus investigates an entirely different research question.

[6] Bisbe and Malagueno (2012) and O'Gorman and Doran (1999) use a similar sales range for midsized firms. This definition of midsized firms is confirmed by our survey, as 88.6% of the responding companies consider themselves midsized firms.

[7] Public sector organizations typically rank nonfinancial over financial objectives and have different MCs and especially PMSs. Furthermore, financial service firms (e.g., banks, insurance firms) are excluded, as OP is measured differently (e.g., RoE instead of RoA), objectives are dominated by financial targets, risk management plays a prominent role, and operational processes and strategy development are different from those in manufacturing and service firms.

[8] We obtain information related to ownership from the Amadeus database.

Table 1
Questionnaire items, descriptive statistics, confirmatory factor analysis, and reliability measures.

Emphasis on diagnostic use of the PMS (Henri, 2006)
Please rate the extent to which your top management team used a PMS to perform the following tasks in the last three years:
Scale: 1 = not at all; 5 = to a great extent

Item	Theoretical range	Mean	Standard deviation	Standardized loading (CFA)	Cronbach's alpha	Variance extracted	Individual item reliability	Composite reliability
Track progress toward goals	1–5	4.09	0.957	0.858[a]	0.861	0.72	0.737	0.869
Monitor results	1–5	4.44	0.747	0.764***			0.583	
Compare outcomes to expectations	1–5	4.07	1.010	0.830***			0.689	
Review key measures	1–5	4.02	1.028	0.703***			0.495	

Emphasis on interactive use of the PMS (Henri, 2006)
Please rate the extent to which your top management team used a PMS to perform the following tasks in the last three years:
Scale: 1 = not at all; 5 = to a great extent

Item	Theoretical range	Mean	Standard deviation	Standardized loading (CFA)	Cronbach's alpha	Variance extracted	Individual item reliability	Composite reliability
Enable discussions in meetings of superiors, subordinates, and peers	1–5	3.46	1.067	0.635***	0.899	0.62	0.404	0.897
Enable continual challenge and debate of underlying data, assumptions, and action plans	1–5	3.52	1.021	0.681***			0.463	
Provide a common view of the organization	1–5	3.43	1.080	0.790***			0.624	
Tie the organization together	1–5	2.91	1.085	0.795[a]			0.632	
Enable the organization to focus on common issues	1–5	3.12	1.148	0.781***			0.610	
Enable the organization to focus on critical success factors	1–5	3.59	1.077	0.729***			0.531	
Develop a common vocabulary in the organization	1–5	3.38	1.110	0.797***			0.635	

Note: The table reports descriptive statistics, the results of exploratory and confirmatory factor analyses, and reliability measures. All Cronbach's alpha coefficients exceed the common threshold of 0.70 (Nunnally, 1978), all factor loadings are significant ($p < 0.001$) and substantial (standardized factor loading > 0.6), all individual item reliabilities exceed the common threshold of 0.400 (Bagozzi et al., 1994), and all composite reliability measures lie above the common threshold of 0.6 (Bagozzi and Yi, 1988). The variance extracted exceeds the threshold of 0.5 for all latent constructs.
[n] = 276, chi square (42) = 147.654, p = 0.000, CMINDF = 3.516; CFI = 0.940; RMSEA = 0.096; SRMR = 0.0615.
* Significant at the 0.05 level; ** Significant at the 0.01 level; *** Significant at the 0.001 level (two-tailed).
a = Reference indicator (A change in the reference indicators reveals significant ($p < 0.001$) and substantial factor loadings even for the reference indicators.).

benefits than early respondents. Overall, the results support the generalizability of our findings and indicate a low response bias. Appendix A1 summarizes our nonresponse analyses, and Appendix A2 shows the descriptive statistics for our survey sample.

3.2. Variable measurement

To measure the emphasis on the *interactive use* and *diagnostic use* of a PMS, we adopt a previously used and validated instrument from Henri (2006a). The respondents were asked to indicate on a five-point Likert scale (ranging from very little to very much) the extent to which they used a PMS in their firm. Whereas the questions regarding diagnostic use capture classical measurement tasks, such as performance measures related to monitoring, reviewing, and comparisons, the questions regarding interactive use address attention focusing. An exploratory factor analysis revealed the unidimensionality of the diagnostic and interactive use constructs (Appendix A3). Moreover, a confirmatory factor analysis showed that all Cronbach's alpha values lie above 0.8 and that the factor loadings are significant and substantial. Furthermore, the variance extracted, individual item reliability, and composite reliability measures confirmed the reliability of both constructs (Table 1).

We measure *PMS sophistication level* as a formative construct based on the typology of Speckbacher et al. (2003), which has been used in

Table 2
Reliability of the constructs.

Panel A: Reliability and validity of PMS sophistication level

	CI	Theoretical range	VIF	Regression weights	Beta	Significance
PMS sophistication level (Scale: 1 = not at all; 5 = to a great extent)	10.596					
We use a multidimensional performance measurement system that combines financial and nonfinancial performance measures.		1–5	2.041	0.253	0.284	0.000
We use a multidimensional performance measurement system that additionally describes strategy by using cause-and-effect relationships.		1–5	1.772	0.237	0.266	0.000
We use a multidimensional performance measurement system that allows to define target values and action plans.		1–5	2.880	0.253	0.309	0.000
We use a multidimensional performance measurement system that is connected with incentives for management compensation.		1–5	2.173	0.241	0.319	0.000

Note: The instrument used for measuring PMS sophistication level is developed based on the classification of Speckbacher et al. (2003). The table reports the results of the assessment of multicollinearity. The VIF values for each item and the condition index (CI) for the constructs are far below the common thresholds of 10 and 30, respectively.

Panel B: Reliability and validity of PMS benefits

	CI	Theoretical range	VIF	Regression weights	Beta	Significance
PMS Benefits	22.094					
Importance of the following expected benefits of the PMS (scale: 1 = not important at all; 5 = very important) weighted with the *perceived goal achievement* of the PMS for each of the benefits (scale: 1 = strongly not agree; 5 = strongly agree)						
Developing strategy (further)		1–25	3.251	0.064	0.090	0.002
Clarifying and communicating strategy		1–25	2.782	0.043	0.060	0.023
Improved alignment of strategic objectives with actions		1–25	2.753	0.090	0.122	0.000
Focusing resources on strategy		1–25	2.381	0.005	0.007	0.776
Developing a consistent system of objectives in the company		1–25	2.208	0.066	0.096	0.000
Improving the understanding of cause-and-effect relationships in the company		1–25	2.036	0.042	0.063	0.006
Stronger consideration of nonfinancial drivers of performance		1–25	1.809	0.058	0.082	0.000
Improving company results in the long term		1–25	2.319	0.057	0.083	0.001
Alignment of strategic initiatives		1–25	3.152	0.052	0.076	0.007
Improving strategic learning (control and feedback)		1–25	2.219	0.047	0.069	0.004
Supporting the shareholder value-based management system		1–25	1.998	0.071	0.117	0.000
Building up a base for an incentive system		1–25	1.762	0.042	0.072	0.001
Better consideration of stakeholders (stakeholder: person/group/organization that shows concern in an organization, e.g., suppliers, owner, creditors)		1–25	1.963	0.077	0.112	0.000
Improved customer focus		1–25	1.970	0.068	0.102	0.000
Identifying business process reengineering opportunities		1–25	2.319	0.053	0.078	0.001
Supporting a strategy of growth		1–25	2.305	0.053	0.081	0.001
Enhancing investment in intangibles (e.g., patents)		1–25	1.674	0.064	0.083	0.000

Note: The instrument for measuring PMS benefits is developed based on Gupta and Govindarajan (1984) and Govindarajan (1988). The dimensions of the instrument are adapted based on the expected benefits derived from the study by Speckbacher et al. (2003). The table reports the results of the assessment of multicollinearity. The VIF values for each item and the condition index (CI) for the constructs are far below the common thresholds of 10 and 30, respectively. Despite the nonsignificance of the expected benefit "Focusing resources on strategy", this item remains in the construct to keep the theoretically justified scope of expected PMS benefits.

Panel C: Reliability and validity of PEU

First-order construct	CI	Theoretical range	VIF	Regression weights	Beta	Significance
Competition	16.578					
During the last three years, how intense was competition in each of the following areas for your company from your point of view? (Scale: 1 = of negligible intensity; 5 = extremely intense)						
Competition for purchases or raw materials		1–5	1.146	0.253	0.499	0.000
Competition for manpower		1–5	1.090	0.240	0.369	0.000
Price competition		1–5	1.164	0.246	0.347	0.000
Competition for quality		1–5	1.051	0.247	0.344	0.000

(*continued on next page*)

Table 2 (*continued*)

Panel C: Reliability and validity of PEU

First-order construct	CI	Theoretical range	VIF	Regression weights	Beta	Significance
Market cycles/legal constraints	14.688					
How long are the market cycles in your industry compared to other industries? (Scale: 1 = substantially shorter; 5 = substantially longer)		1–5	1.036	0.505	0.819	0.000
During the past three years, the legal constraints surrounding your firm have from your point of view (Scale: 1 = strongly decreased; 5 = strongly increased)		1–5	1.036	0.453	0.426	0.000
External environment	7.551					
During the last three years, how stable was the external environment facing your firm from your point of view? (Scale: 1 = very stable; 5 = very dynamic)						
Supplier environment		1–5	1.294	0.502	0.575	0.000
Economic environment		1–5	1.294	0.495	0.587	0.000
Technological environment/modifications	6.729					
During the last three years, how stable was the technological environment facing your firm from your point of view? (Scale: 1 = very stable; 5 = very dynamic)		1–5	1.183	0.504	0.612	0.000
How often do new scientific discoveries/modifications to existing products emerge in your industry compared to others from your point of view? (Scale: 1 = less often; 5 = more frequently)		1–5	1.183	0.489	0.582	0.000
Predictability	11.218					
During the last three years, how was the development of the predictability of the following issues from your point of view? (Scale: 1 = much easier to predict; 5 = much harder to predict) Market activities of your competitors		1–5	1.344	0.502	0.562	0.000
Preferences of your customers		1–5	1.344	0.496	0.589	0.000

Second-order construct	CI	VIF	Regression weights	Beta	Significance
Perceived environmental uncertainty	25.439				
Competition		1.133	0.169	0.253	0.000
Market cycles/legal constraints		1.016	0.193	0.332	0.000
External environment		1.312	0.209	0.418	0.000
Technological environment/modifications		1.229	0.194	0.417	0.000
Predictability		1.070	0.183	0.284	0.000

previous research by Lee and Yang (2011) and adapted by Franco-Santos et al. (2012). Bisbe et al. (2007) address the formative measurement of a BSC construct in a similar manner. In their typology, Speckbacher et al. (2003) describe three types of PMSs that build on each other. We adapted Speckbacher et al.'s typology for the purposes of this study by creating a construct based on four design elements that characterize the different types of PMSs (i.e., key performance measures in various dimensions, cause-and-effect relations, defined targets and action plans, and links to incentives). The respondents were asked to indicate the extent to which their PMS design contains these four design elements on a five-point Likert scale, where each scale represents a range between a low and a high stage of PMS development. By using a Likert scale for these items, we are able to consider PMS sophistication level as a continuous variable measured by a formative construct with equal weighting for the four scales, in contrast to the dichotomous or nominal scale (i.e., type I, II or III, see Fig. 1) originally presented by Speckbacher et al. (2003). When the construct score is higher, so is the designed PMS sophistication level. Because the four different design elements can accumulate, different levels of sophistication can be measured, where the anchors are a minimum layout, a PMS that "only" combines financial and nonfinancial measures, and a maximum layout that is a combination of all the elements of PMS design.[9]

We consider *PMS benefits* to be a measure of effectiveness. Adapting the well-established instrument developed by Gupta and Govindarajan (1984) and Govindarajan (1988) for perceived OP to our context, we measure PMS benefits as the perceived degree of goal achievement along various dimensions. This OP instrument has been used in previous empirical research by Chenhall and Langfield-Smith (1998) and Bisbe and Otley (2004). We separately asked firms to indicate the importance of and their level of satisfaction with 17 items representing the expected benefits of a PMS on a five-point Likert scale. The perceived benefit score represents a formative index indicating each firm's importance-weighted average level of satisfaction with the listed benefit items.[10] The items were adapted from the expected benefits of BSC use summarized by Speckbacher et al. (2003) based on the literature.[11] Thus, this measure is also able to consider different contingencies of a firm by weighting the importance of a specific benefit for the firm (Bedford and Speklé, 2018). The PMS benefits measure captures managerial perceptions of their firms' PMS and is thus a proxy for its "actual" benefits. Note that this measure concerns the net benefits and costs for implementing and using a PMS, which we label *PMS benefits* for simplification.

To assess the reliability and validity of both formative constructs, we examine the multicollinearity of the constructs. The variance inflation factors (VIFs) for both PMS sophistication levels and PMS benefits lie far below the common threshold of 10 (Diamantopoulos and Winklhofer, 2001; Diamantopoulos et al., 2008). Except for the item "focusing resources on strategy" within the PMS benefits construct, all items have significant regression weights that provide important contributions to the construct (Diamantopoulos and Winklhofer, 2001; Bedford and Speklé, 2018). We nevertheless keep this nonsignificant item in the PMS benefits construct to maintain the theoretically

[9] The score for PMS sophistication level ranges from 1 (i.e., four design elements with a score of 1 each, using equal weighting for the four items) to 5 (i.e., four design elements with a score of 5 each because of the five-point Likert scale).

[10] As a robustness check, we also perform a factor analysis and define a second-order construct consisting of three factors. The results are shown in the robustness section.

[11] The score theoretically ranges between 1 and 25, as it is the product of the level of importance and level of satisfaction (both measured with five-point Likert scales) equally weighted over all 17 items, as listed in Panel B, Table 2.

justified dimensions of the expected PMS benefits. Table 2 reports the reliability and validity results for both formative constructs.

We measure *Size* as the number of employees in the firms. In the contingency-based literature, the number of employees is a commonly used proxy for organizational size because financial measures may distort comparisons owing to possible differences in accounting treatments between firms (Chenhall, 2003). Moreover, research shows that the number of employees is correlated with financial measures such as sales and total assets (Hoque and James, 2000). Furthermore, the complexity of a firm's MCs seems likely to be driven by the number of employees, who must be coordinated and aligned with overall firm goals. MC systems and especially the LoCs are connected to the behavior of a firm's employees (e.g., Simons, 1995; Widener, 2007). Finally, we follow other empirical research papers and use the number of employees as a measure for organizational size (Bisbe and Malagueno, 2012; Davila et al., 2015; Henri, 2006a, 2006b).

3.3. Control variables

Following the conceptualization of Gerdin and Greve (2004) for contingency fit testing, we include three control variables as moderating variables to test the robustness of our base model. In contingency-based research, *perceived environmental uncertainty* (*PEU*) is recognized as a powerful contextual variable (Bisbe and Malagueno, 2012; Chenhall, 2003). Previous studies suggest that PEU influences how firms use their PMSs (Bisbe and Otley, 2004; Hoque, 2014; Kihn, 2007; Widener, 2007) and that PEU explains the variance in PMS sophistication levels (Abdel-Kader and Luther, 2008; Micheli and Manzoni, 2010). To measure PEU, we use the widely used instrument developed by Gordon and Narayanan (1984) and Govindarajan (1984). Moreover, because PMSs develop over time (Speckbacher et al., 2003), we include the variable *Dynamic* to test the robustness of our model with respect to changes in PMS design over the past three years because insights from the concept of accounting change (e.g., Andon et al., 2007; Quattrone and Hopper, 2001) suggest that the interplay between the PMS and the firm may be an important determinant of further PMS development. Therefore, for all items concerning PMS sophistication levels, we asked respondents to indicate the PMS sophistication level both as of today and as of three years ago. For each item, the difference in the item value between today and three years ago is calculated; these differences are then averaged over all items to obtain the value for *Dynamic*. Furthermore, we include the variable *Ownership* to control for different ownership structures. Speckbacher and Wentges (2012) find that family involvement is an important contextual factor in PMS design. Nonfamily firms tend to use formal, sophisticated PMSs more extensively than family firms (Cromie et al., 1995), and ownership is a key characteristic that distinguishes family firms from nonfamily firms (Klein, 2000). Furthermore, we integrate a mutual path with identical path coefficients between an emphasis on diagnostic and on interactive use and between an emphasis on both diagnostic and interactive use and PMS sophistication levels to control for relationships between these variables. A relationship between both types of use has been postulated by Simons (1995) and empirically proven in previous survey studies (Henri, 2006a; Widener, 2007). Finally, we measure *OP* as a latent construct of archival RoCE, RoS and RoA.

Furthermore, to test the robustness of our results, we explore the relationships of the focal variables in our base model to other independent contextual and structural variables (i.e., strategic orientation, strategic planning intensity, budgeting, size, OP, firm age) because these variables might be related to specific variables in our model (Gerdin and Greve, 2004). We describe the measurement of these variables and present the results in the robustness section.

3.4. Data analysis

Our base model focuses on the relationships among the two types of use, PMS sophistication levels and the corresponding organizational outcomes based on contingency theory. We use the mediation approach of a contingency fit analysis (see Gerdin and Greve, 2004) by using covariance-based structural equation (path) modeling (SEM) and a moderation approach when we test for the fit between PMS sophistication levels and both types of use with polynomial regression modeling and when we examine the impact of other contingency factors on our research model (i.e., the form of fit) using multigroup causal analysis (MGCA).

We analyze the data in two steps to test our hypotheses. First, we use covariance-based SEM to estimate the entire research model (H1a to H2b) simultaneously. In particular, we refer to polynomial regression modeling to examine the moderating effect of the PMS sophistication level (Edwards and Parry, 1993; Burkert et al., 2014). Thus, the benefits of a PMS can be explained by the fit between both types of use and PMS sophistication levels:

$$Benefit = ß_0 + ß_1(diagnostic\ \ use{-}PMS\ \ sophistication\ \ level)^2 + ß_2(interactive\ \ use{-}PMS\ \ sophistication\ \ level)^2 + \varepsilon$$

When taking into account all direct, interaction and squared variables in our research model, we finally obtain:

$$\begin{aligned} Benefit = {} & a_0 + a_1\ diagnostic\ use \\ & + a_2\ (diagnostic\ \ use)^2 + a_3 interactive\ \ use + a_4(interactive\ \ use)^2 \\ & + a_5 PMS\ \ sophistication\ \ level + a_6(PMS\ \ sophistication\ \ level)^2 \\ & + a_7 diagnostic\ \ use \times PMS\ \ sophistication\ \ level \\ & + a_8 interactive\ \ use \times PMS\ \ sophistication\ \ level + controls + \varepsilon \end{aligned}$$

We use AMOS 23.0 with a maximum likelihood (ML) estimation approach, which requires multivariate, normally distributed data. The absolute values for skewness and kurtosis indicate the multivariate normality of the data, as the values are far below the threshold of 3.0 for skewness or 10.0 for kurtosis (Kline, 2011).

Second, we use MGCA to test the moderating effect of Size (H3a and H3b) and of the three control variables (PEU, Dynamic, and Ownership), as recommended by Burkert et al. (2014), to test for categorical moderating effects. To investigate Size, we divide the sample into two subgroups based on a median split. Firms with fewer than 250 employees are defined as smaller firms, and firms with more than 249 employees are classified as larger firms. Then, following the procedure by Byrne et al. (2010), we assess the measurement invariance between the two groups[12] and find the measurement invariance for smaller and larger firms. Conducting the same test for the two groups divided based on PEU, Ownership, and Dynamic,[13] we find measurement invariance in the two groups for PEU and partial scalar invariance in the two groups for Ownership and Dynamic. The results of the subgroup analyses for Size are presented in Table 5; the results of the subgroup analyses for PEU, Dynamic, and Ownership are untabulated but available to readers upon request.

4. Results and discussion

Panel A, Table 3 shows the correlations among the variables of the Base Model 1. With the exception of the interaction between interactive

[12] The prerequisite for applying MGCA is measurement invariance, which ensures that the items, constructs, and structural model operate equivalently across the groups. With measurement invariance, unambiguous conclusions can be drawn from the results, whereas partial scalar invariance is regarded as sufficient for MGCA (Steenkamp and Baumgartner, 1998).

[13] To test the control variables Dynamic and PEU, we also divide the sample into two groups for each control variable by using a median split. To validate the structural model with respect to Ownership, we divide the sample into family firms (in which the family owns more than 50.00% of the firm) and nonfamily firms.

Table 3
Correlation matrix.

Panel A: Correlations between latent constructs

	Interactive use of the PMS	Diagnostic use of the PMS	Interaction between interactive use and PMS sophistication level	Interaction between diagnostic use and PMS sophistication level	PMS sophistication level	PMS benefits	Organizational performance
Interactive use of the PMS	1						
Diagnostic use of the PMS	0.605***	1					
Interaction between interactive use and PMS sophistication level	−0.093	−0.183***	1				
Interaction between diagnostic use and PMS sophistication level	−0.184***	−0.360***	0.664***	1			
PMS sophistication level	0.421***	0.319***	0.036	0.105*	1		
PMS benefits	0.747***	0.654***	0.093	−0.119**	0.494***	1	
Organizational performance	0.001	−0.021	−0.027	0.03	−0.004	0.033	1

Note: This table reports estimated Bravais-Pearson correlations between constructs.
* Significant at the 0.1 level; ** Significant at the 0.05 level; *** Significant at the 0.01 level (two-tailed).

Panel B: Correlations between both diagnostic use and interactive use and items related to the PMS sophistication level

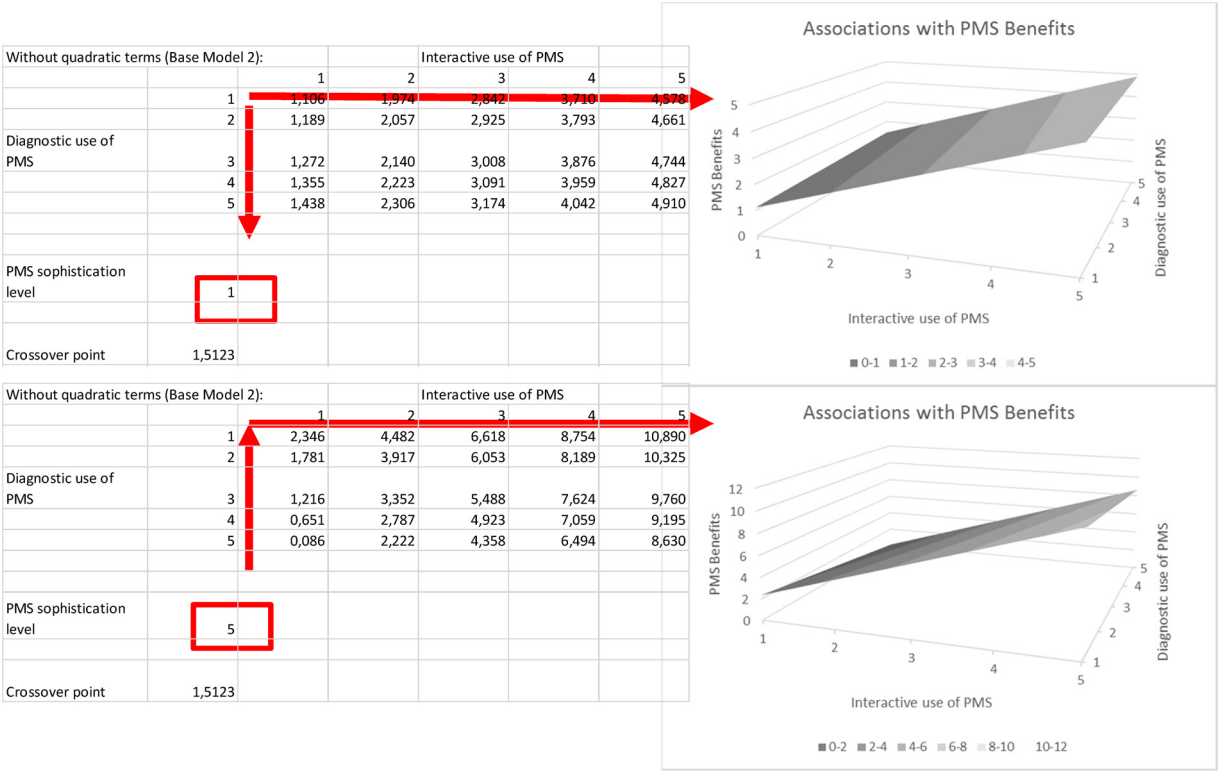
	PMS with financial and nonfinancial performance measures	PMS describing strategy by using cause-and-effect relationships	PMS allows to define target values and action plans	PMS is connected with incentives for management compensation
Diagnostic use of the PMS	0.332***	0.218***	0.273***	0.266***
Interactive use of the PMS	0.369***	0.436***	0.297***	0.344***

Note: This tables reports estimated Bravais-Pearson correlations between the latent constructs for both diagnostic use and interactive use and the single items of the PMS sophistication level measure.*Significant at the 0.1 level; ** Significant at the 0.05 level; *** Significant at the 0.01 level (two-tailed).

Table 4
Results for the structural model.

Structural model					
Independent variable	Dependent variable	Hypothesis	Exp. direct.	Base Model 1 with quadratic terms Path coefficient	Base Model 2 Path coefficient
Diagnostic use of the PMS	**PMS benefits**	**H1a**		**0.270*****	**0.245*****
Interactive use of the PMS	**PMS benefits**	**H1b**	+	**0.544*****	**0.551*****
Diagnostic use squared	PMS benefits		+	0.058	
Interactive use squared	PMS benefits		+	−0.089	
Interaction between diagnostic use and PMS sophistication level	**PMS benefits**	**H2a**		**−0.147****	**−0.162*****
Interaction between interactive use and PMS sophistication level	**PMS benefits**	**H2b**	+	**0.331*****	**0.317*****
R^2				75.4%	76.4%
(Mutual) Control paths:					
Interactive/Diagnostic use of the PMS	Diagnostic/Interactive use of the PMS			0.413***	0.415***
Diagnostic use of the PMS/PMS sophistication level	PMS sophistication level/Diagnostic use of the PMS			0.112***	0.108***
Interactive use of the PMS/ PMS sophistication level	PMS sophistication level/Interactive use of the PMS			0.257***	0.223***
PMS sophistication level	PMS benefits			0.179***	0.155***
PMS sophistication level squared	PMS benefits			0.041	
PMS benefits	Organizational performance			0.051	0.053
Model fit statistics:					
χ^2				1294.020***	522.831***
DF				650	297
CMINDF				1.991	1.760
CFI				0.911	0.951
RMSEA				0.060 [0.055; 0.064]	0.053 [0.045; 0.060]
SRMR				0.1162	0.0736

Note: This table reports the results of an ML estimation of the structural equation model depicted in Fig. 2 and two alternative models with squared constructs for interactive and diagnostic use of the PMS and of PMS sophistication level. We report the standardized coefficients. ***, **, and * indicate the significance of the p-value at < 0.01, 0.05, and 0.10, respectively. We report one-tailed p-values.
Significant path coefficients are presented in bold.



Without quadratic terms (Base Model 2):		Interactive use of PMS				
		1	2	3	4	5
Diagnostic use of PMS	1	1,100	1,974	2,842	3,710	4,578
	2	1,189	2,057	2,925	3,793	4,661
	3	1,272	2,140	3,008	3,876	4,744
	4	1,355	2,223	3,091	3,959	4,827
	5	1,438	2,306	3,174	4,042	4,910
PMS sophistication level	1					
Crossover point	1,5123					

Without quadratic terms (Base Model 2):		Interactive use of PMS				
		1	2	3	4	5
Diagnostic use of PMS	1	2,346	4,482	6,618	8,754	10,890
	2	1,781	3,917	6,053	8,189	10,325
	3	1,216	3,352	5,488	7,624	9,760
	4	0,651	2,787	4,923	7,059	9,195
	5	0,086	2,222	4,358	6,494	8,630
PMS sophistication level	5					
Crossover point	1,5123					

Fig. 3. Associations with PMS benefits using the response surface method.

use and PMS sophistication level, we find highly significant correlation coefficients for all other variables in our base model with PMS benefits ($p < 0.01$). Furthermore, despite the high correlation coefficients between PMS benefits and both interactive ($r = 0.747$) and diagnostic use of a PMS ($r = 0.654$), we find no further evidence of multicollinearity.[14] To avoid multicollinearity problems among interaction terms, we use z-standardized values for all items and for all products of item combinations to calculate interaction terms (Cortina et al., 2001; Henri, 2006a). Panel B, Table 3 reports the correlations between the single items that capture PMS sophistication level and both an emphasis on diagnostic use and an emphasis on interactive use. The results show that the correlations for interactive use are always higher than those for diagnostic use. When both types of use are compared, the difference between these correlations is particularly pronounced when the PMS describes firm strategy based on cause-and-effect relationships (e.g., in the form of a strategy map) or when the PMS is connected to incentives for management compensation. This finding indicates that a greater emphasis on interactive PMS use is associated with a higher PMS sophistication level. In particular, compared with an emphasis on diagnostic use, an emphasis on interactive use is more strongly correlated with an incentive system that is linked to a PMS. This finding is in contrast to Simonsö (1995) understanding in the LoC framework and generalizes the field study findings of Agostino and Arnaboldi (2012) for our cross-sectional sample. Nevertheless, all correlations are positive and significant but only bivariate.

4.1. Results for the structural equation model

Table 4 presents the results of the covariance-based SEM analysis for the research model. Because all quadratic terms in Base Model 1 are not significant, we focus our further analysis on the linear relationships of Base Model 2. The goodness-of-fit indices for both base models indicate that the data fit the structural model quite well.[15] However, the fit indices for the more parsimonious Base Model 2 are higher.

First, our SEM analysis results show that, despite the ambiguity in the literature, the direct relationships of both an emphasis on interactive PMS use and an emphasis on diagnostic PMS use with PMS benefits are positive and significant ($p < 0.01$) when we simultaneously control for PMS sophistication level. These results support H1b but not H1a. The path coefficient for the relationship between an emphasis on diagnostic use and PMS benefits is lower than that for the relationship between interactive use and PMS benefits (path coeff. interactive use 0.551, diagnostic use 0.245). The interactive use of a PMS stimulates the communication process and supports feedforward loops and organizational learning, whereas diagnostic use focuses on monitoring results, analyzing deviations and providing feedback. Thus, an emphasis on interactive use seems to have a stronger "impact" on PMS benefits. In both base models and all alternative models, we control for a correlation between an emphasis on interactive use and an emphasis on diagnostic use.[16]

Surprisingly, the path coefficient of the interaction term of diagnostic PMS use and PMS sophistication level on PMS benefits is significantly negative ($p < 0.05$), which rejects H2a. Thus, PMS benefits decrease as a firm's sophistication level increases when there is constant emphasis on diagnostic use. In contrast, we find a significant positive relationship between the interaction of interactive use and PMS sophistication level with PMS benefits ($p < 0.01$), which confirms H2b. This finding shows that PMS sophistication levels moderate the relationship between interactive use and PMS benefits and that the relationship is stronger when the level of sophistication is higher. Firms that use PMSs interactively are characterized by regular attention from managers throughout the firm, and the data provided by interactive PMS use are important in such firms because these data provide the basis for continual discussions with subordinates and enable communication to develop new strategies and solutions (Simon, 1995). A sophisticated PMS facilitates these features, for instance, by linking the PMS to incentives or by aligning firm strategy with the PMS, as illustrated by the high correlations in Panel B, Table 3.

Because diagnostic and interactive use are theoretically assumed (Simons, 1995) and empirically confirmed to be correlated (e.g., Widener, 2007; Heinicke et al., 2016), we use the response surface method to demonstrate their joint effect, integrating all terms determined by the three variables: interactive use, diagnostic use and PMS sophistication level. We model only the significant path coefficients of Base Model 2 because the quadratic terms for all three variables of Base Model 1 are nonsignificant. As depicted in Fig. 3, for low levels of PMS sophistication, more emphasis on either interactive or diagnostic use is associated with an increase in PMS benefits. However, for higher PMS sophistication levels, more emphasis on interactive use increases PMS benefits, whereas more emphasis on diagnostic use is dysfunctional and decreases the benefits. The level of sophistication at which the effect of diagnostic use on PMS benefits is constant can be analytically determined. For the estimated path coefficients of the SEM, the tipping point is at 1.5123, which, for a scale that ranges between 1 (low) and 5 (high), is a rather low PMS sophistication level. Given that 85.9% of all sample firms have higher scores, most firms in our sample obtain decreasing PMS benefits as their emphasis on diagnostic use increases. Note, that the mean for diagnostic use is the highest of all the analyzed constructs (mean: 4.15, theoretical range from 1 to 5). However, for more sophisticated PMSs, the decrease in benefits from more emphasis on diagnostic use is outweighed by the greater increase in benefits from more emphasis on interactive use (see the tables in Fig. 3).

Diagnostic use of a PMS is intended to monitor outcomes and to control for deviations from pre-established goals, which are necessary tasks for all firms (Simons, 1995). Our results challenge and expand Simonsö (1995) understanding based on the LoC framework that diagnostic use requires a connection to rewards and incentives for managers and employees and a link to explicit targets, which would generate benefits through a higher PMS sophistication level. We find a negative effect of diagnostic use on PMS benefits with higher levels of sophistication. Our findings generalize the earlier field research findings of Agostino and Arnaboldi (2012) and empirically demonstrate the theoretical argument of Franco-Santos et al. (2012) that a PMS can measure organizational and managerial performance without necessarily being linked to an incentive system. Moreover, Specklé and Verbeeten (2014) divide diagnostic use into operational and incentive-setting use and allocate a special role to target setting, incentives and rewards in contrast to the operational planning and process monitoring involved in operational use.

There may be several overlapping explanations for this dysfunctional interaction effect. First, diagnostic controls might work not only via the PMS but also via the budgeting system, and in particular, rewards might be linked to the achievement of mostly financial targets set within the budgeting system (see also Simons, 1995, 2014; Henri, 2006b), enabling a simpler PMS design. Furthermore, the literature reports different dysfunctional effects of a PMS when it is connected to incentive systems (e.g., ossification, the crowding out of intrinsic motivation, tunnel effects, or gaming, see Birnberg et al., 1983; Henri, 2006a; Ittner et al., 2003a; Schick et al., 1990; Simons, 1995; Smith,

[14] The highest VIF value is 2.262, which is far below the commonly accepted threshold of 10.

[15] As indicators of good model fit, following the recommendation of Hu and Bentler (1999) for samples with $n > 250$, we use the combination of the comparative fit index (CFI) and the standardized root mean square (SRMR), as well as the root mean square error of approximation (RMSEA) and the chi square divided by the model degrees of freedom (CMINDF). The common thresholds for these frequently used goodness-of-fit indices are CFI > 0.90 (Bentler, 1990, 1992), RMSEA < 0.10 (Browne and Cudeck, 1993), SRMR close to 0.080 (Hu and Bentler, 1999), and CMINDF < 5 (Wheaton et al., 1977).

[16] The estimated path coefficient for a mutual control path between an emphasis on diagnostic use and an emphasis on interactive use is 0.415 ($p < 0.01$).

Table 5
Results of the MGCA for Size.

		MGCA for Size			
Independent variable	Dependent variable	Hypothesis	Small N = 141	Large N = 135	z-score difference test
Diagnostic use of the PMS	PMS benefits	H3a (−)	0.371***	0.244**	ns
Interactive use of the PMS	PMS benefits	H3b (+)	0.486***	0.530***	p = 0.058
Interaction between diagnostic use of the PMS and PMS sophistication level	PMS benefits		−0.203***	−0.362	ns
Interaction between interactive use of the PMS and PMS sophistication level	PMS benefits		0.305***	0.531**	ns
Control paths:					
Interactive use of the PMS/ Diagnostic use of the PMS	Diagnostic use of the PMS/ Interactive use of the PMS		0.381***	0.480***	ns
Diagnostic use of the PMS/ PMS sophistication level	PMS sophistication level/ Diagnostic use of the PMS/		0.120***	0.103***	ns
Interactive use of the PMS/ PMS sophistication level	PMS sophistication level/ Interactive use of the PMS		0.152***	0.193***	ns
PMS sophistication level	PMS benefits		0.182***	0.123**	ns
PMS benefits	Organizational performance		−0.009	0.114	ns
Model comparison statistics:					
χ^2 restricted MM, unrestricted SM (fully restricted MM/SM)			1076.466 (597) 1481.780 (675)		
χ^2 difference test			P < 0.01		
DF			597		
CMINDF			1.803		
CFI			0.906		
RMSEA			0.054 [0.049; 0.059]		
SRMR			0.0969		

Note: The table reports the MGCA of the Base Model 2 for different levels for Size measured by the number of firm employees. The coefficients from the MGCA are taken from the scalar model, which constrains all factor loadings and all intercepts to be equal across the groups. The global fit indices are reported. We report standardized estimates for path coefficients. The subgroup split for Size is not precisely 50%/50%, as the basis for the split is the classification of Size in seven groups. The three lowest size groups comprise 141 firms, whereas the four largest ones comprise 135 firms.
***, **, and * indicate the significance of the *p*-value at < 0.01, 0.05, and 0.10, respectively. We report one-tailed *p*-values.

1995; Townley et al., 2003). Second, in this study, the focus of the PMS is also on nonfinancial KPIs, which might be more difficult to connect to target values. This explanation is confirmed by the case study of Tuomela (2005), wherein target setting was perceived as a challenge for the case firm. Third, it is arguably difficult to explicate the impact of specific nonfinancial measures on financial results (Anthony and Govindarajan, 1998). Fourth, there may be a lack of data on past nonfinancial performance to compare with targets or a lack of experience with the financial effects of nonfinancial KPIs, which are usually explained in a PMS through cause-and-effect relationships (one of the items constituting the PMS sophistication level). Fifth, Malmi (2001) reports that several of his 17 case firms use BSCs only as information systems, which does not require target setting or a link to the incentive system and thus a more sophisticated PMS. Finally, Yongvanich and Guthrie (2009) report that one-third of their firms did not use cause-and-effect analyses in practice.

The control paths among interactive use, diagnostic use and PMS sophistication level are all significant (p < 0.01), whereas the relationship between PMS benefits and OP is not. Thus, our results show that it makes a difference whether proximate outcomes of PMS implementation (in our case PMS benefits) or distant outcomes (in our case, OP) are examined. This analysis addresses the call by Chenhall (2003) to relate MC research to performance measures that are more proximate than OP because OP is associated with numerous antecedents that are difficult to control for. Furthermore, our results confirm previous findings of Henri (2006a), who also does not find an indirect effect of both types of use on OP but finds a direct effect of both types of use on the four organizational capabilities, which are more proximate outcomes.

In summary, PMS sophistication levels contribute positively to the relationship between an emphasis on interactive use and PMS benefits but negatively moderate the relationship between an emphasis on diagnostic use and PMS benefits.

4.2. Results for the multigroup causal analysis of size

The descriptive statistics show that all four constructs of our Base Model 2 have significantly higher values for larger midsized firms than for smaller midsized firms (Panel B, Appendix A2).[17] The results of the MGCA for Size are reported in Table 5. The goodness-of-fit indices show a good model fit. To determine whether a significant difference exists between smaller and larger firms, we compare z-score differences based on critical ratios (p < 0.1).

Table 5 shows that for both larger and smaller firms, an emphasis on diagnostic use is positively and significantly associated with PMS benefits (smaller firms coeff. = 0.371, p < 0.01; larger firms coeff. = 0.244, p < 0.05), although the path coefficient for smaller firms is higher. However, the difference is not significant (p = 0.184); thus, H3a is not supported.

H3b predicts that the positive relationship between an emphasis on interactive use and PMS benefits is more pronounced in larger firms than in smaller firms. The individual path coefficients are significant and positive for both groups and, as expected, higher for larger firms than for smaller firms (coeff. for smaller firms = 0.486; coeff. for larger firms = 0.530, p < 0.01). Furthermore, the two coefficients are significantly different (p = 0.058). Hence, H3b is supported. Larger firms therefore seem to benefit more from focusing on communication processes and creative searches for innovation, which are key characteristics of an emphasis on interactive use. Moreover, larger firms benefit more because they likely face greater complexity and diversity in their

[17] Because we explore midsized firms in general, when we speak of smaller and larger firms, we mean smaller and larger midsized firms within the size range of our study.

Table 6
Modifications of Base Model 2.

Additional variable used as independent variables (IV)		Base Model 2	Model 1 Base Model 2 with PMS benefits as a second-order construct	Model 2 Base Model 2 w/o strategy items of PMS benefits	Model 3 Base Model 2 w/o strategy items for PMS sophistication level	Model 4 Base Model 2 w/o strategy items for PMS benefits and PMS sophistication level	Model 5 Base Model 2 with beliefs and boundary systems	Model 6 Base Model 2 with the adoption of a PMS instead of PMS sophistication level	Model 7 Base Model 2 stead of the PMS sophi-stication level
Independent variable (IV)	*Dependent variable (DV)*								
Diagnostic use of the PMS	**PMS benefits**	**0.245*****	**0.277*****	**0.231*****	**0.235*****	**0.221*****	**0.237*****	**0.251*****	**0.242*****
Interactive use of the PMS	**PMS benefits**	**0.551*****	**0.565*****	**0.504*****	**0.564*****	**0.516*****	**0.519*****	**0.631*****	**0.584*****
Interaction between diagnostic use of the PMS and PMS sophistication level	**PMS benefits**	**−0.162*****	**−0.142****	**−0.141****	**−0.167*****	**−0.139****	**−0.168*****	−0.013	−0.042
Interaction between interactive use of the PMS and PMS sophistication level	**PMS benefits**	**0.317*****	**0.320*****	**0.327*****	**0.318*****	**0.315*****	**0.310*****	0.096	0.121
Control paths:									
Interactive use of the PMS/ Diagnostic use of the PMS	Diagnostic use of the PMS/ Interactive use of the PMS	**0.415*****	**0.418*****	**0.422*****	**0.418*****	**0.425*****	**0.385*****	**0.459*****	**0.442*****
Diagnostic use of the PMS/ PMS sophistication level	PMS sophistication level/ Diagnostic use of the PMS	**0.108*****	**0.109*****	**0.110*****	**0.127*****	**0.128*****	**0.106*****	0.034	0.064
Interactive use of the PMS/ PMS sophistication level	PMS sophistication level/ Interactive use of the PMS	**0.223*****	**0.225*****	**0.225*****	**0.185*****	**0.187*****	**0.201*****	**0.090****	**0.211*****
PMS sophistication level	PMS benefits	**0.155*****	**0.170*****	**0.161*****	**0.148*****	**0.161*****	**0.157*****	0.020	**0.124*****
PMS benefits	Organizational performance	0.053	0.025	0.058	0.053	0.059	0.053	0.052	0.052
Beliefs system	PMS sophistication level						0.076		
Boundary system	PMS sophistication level						−0.067		
Beliefs system	PMS benefits						**0.078***		
Boundary system	PMS benefits						0.017		
Model fit:									
X^2		522.8***	1414.6***	514.8***	529.2***	521.2***	879.7***	643.7***	603.7***
DF		297	829	297	297	297	524	297	297
CMINDF		1.760	1.706	1.733	1.782	1.755	1.679	2.167	2.033
CFI		0.951	0.921	0.952	0.949	0.950	0.934	0.919	0.934
RMSEA		0.053	0.051	0.052	0.053	0.052	0.050	0.065	0.061
SRMR		0.0736	0.0676	0.0732	0.0808	0.0803	0.0663	0.0667	0.0561

Note: This table reports alternative models to the Base Model 2. Model 1 is the Base Model 2 with PMS benefits as a second-order construct. Model 2 is the Base Model 2 with the PMS benefits variable excluding all strategy items, whereas Model 3 is the Base Model 2 with the PMS sophistication level variable excluding the cause-and-effect items describing firm strategy (see Panel A, Table 2). Model 4 excludes all strategic items from both PMS sophistication level and PMS benefits. Finally, Model 5 includes all four LoCs (beliefs system, boundary system, interactive use and diagnostic use of the PMS). Model 6 uses the adoption of a BSC, performance pyramid or quantum PMS instead of PMS sophistication level, and Model 7 uses the level of sophistication of MCs instead of PMS sophistication level. For all models, the path coefficient with significance levels and goodness-of-fit indices are reported. ***, **, and * indicate the significance of the p-value at < 0.01, 0.05, and 0.10. We report one-tailed p-values for all directional hypotheses and two-tailed values for the others. Significant path coefficients are presented in bold.

Table 7
Alternative models: Base model with additional independent and dependent variables.

Additional variable used as independent variables		Base Model	Model A Strategic orientation (archival data)	Model B Strategic orientation (latent)	Model C Strategic planning intensity	Model D Budgeting	Model E PEU
Independent Variable (IV)	*Dependent Variable (DV)*						
Diagnostic use of the PMS	**PMS benefits**	**0.245*****	**0.247*****	**0.254*****	**0.244*****	**0.245*****	**0.244*****
Interactive use of the PMS	**PMS benefits**	**0.551*****	**0.553*****	**0.539*****	**0.549*****	**0.545*****	**0.552*****
Interaction between diagnostic use of the PMS and PMS sophistication level	**PMS benefits**	**−0.162*****	**−0.164*****	**−0.164*****	**−0.164*****	**−0.160*****	**−0.162*****
Interaction between interactive use of the PMS and PMS sophistication level	**PMS benefits**	**0.317*****	**0.318*****	**0.315*****	**0.317*****	**0.315*****	**0.317*****
Control paths:							
Interactive use of the PMS/ Diagnostic use of the PMS	Diagnostic use of the PMS/ Interactive use of the PMS	**0.415*****	**0.412*****	**0.433*****	**0.416*****	**0.408*****	**0.419*****
Diagnostic use of the PMS/ PMS sophistication level	PMS sophistication level/ Diagnostic use of the PMS/	**0.108*****	**0.107*****	**0.129*****	**0.104*****	**0.099*****	**0.115*****
Interactive use of the PMS/ PMS sophistication level	PMS sophistication level/ Interactive use of the PMS	**0.223*****	**0.222*****	**0.186*****	**0.207*****	**0.201*****	**0.212*****
PMS sophistication level	PMS benefits	**0.155*****	**0.157*****	**0.147*****	**0.147*****	**0.149*****	**0.155*****
PMS benefits	Organizational performance	0.053	0.053	0.053	0.053	0.053	0.053
Additional Variable	*Associated with …*						
IV (see row heading)	PMS sophistication level		0.051	0.221***	0.277***	0.228***	0.152***
IV (see row heading)	PMS benefits		−0.050	0.033	0.024	0.035	−0.005
IV (see row heading)	Diagnostic use of the PMS		0.052	−0.089	0.055	0.076	−0.036
IV (see row heading)	Interactive use of the PMS		0.043	0.169**	0.071	0.144***	0.075
Model fit:							
X^2		522.8***	538.8***	594.9***	544.1***	549.9***	550.0***
DF		297	320	346	320	320	320
CMINDF		1.760	1.684	1.719	1.700	1.781	1.719
CFI		0.951	0.953	0.948	0.952	0.951	0.951
RMSEA		0.053	0.050	0.051	0.079	0.051	0.051
SRMR		0.0736	0.0716	0.0722	0.0715	0.0723	0.0726

Additional variable used as independent variables		Base Model	Model F Size (Employees)	Model G Size (Sales)	Model H Organizational performance (latent)	Model I RoS	Model J Firm age
Independent variable (IV)	*Dependent variable (DV)*						
Diagnostic use of the PMS	**PMS benefits**	**0.245*****	**0.246*****	**0.245*****	**0.243*****	**0.246*****	**0.244*****
Interactive use of the PMS	**PMS benefits**	**0.551*****	**0.547*****	**0.554*****	**0.551*****	**0.551*****	**0.551*****
Interaction between diagnostic use of the PMS and PMS sophistication level	**PMS benefits**	**−0.162*****	**−0.165*****	**−0.164*****	**−0.170*****	**−0.165*****	**−0.162*****
Interaction between interactive use of the PMS and PMS sophistication level	**PMS benefits**	**0.317*****	**0.320*****	**0.319*****	**0.324*****	**0.321*****	**0.317*****
Control paths:							
Interactive use of the PMS/ Diagnostic use of the PMS	Diagnostic use of the PMS/ Interactive use of the PMS	**0.415*****	**0.417*****	**0.413*****	**0.417*****	**0.415*****	**0.415*****
Diagnostic use of the PMS/ PMS sophistication level	PMS sophistication level/ Diagnostic use of the PMS	**0.108*****	**0.108*****	**0.104*****	**0.108*****	**0.107*****	**0.108*****
Interactive use of the PMS/ PMS sophistication level	PMS sophistication level/ Interactive use of the PMS	**0.223*****	**0.202*****	**0.215*****	**0.223*****	**0.223*****	**0.224*****
PMS sophistication level	PMS benefits	**0.155*****	**0.145*****	**0.161*****	**0.156*****	**0.157*****	**0.155*****
PMS benefits	Organizational performance	0.053	0.053	0.053			0.053
Additional Variable	*Associated with …*						
IV (see row heading)	PMS sophistication level		0.283***	0.168***	−0.011	−0.049	0.013
IV (see row heading)	PMS benefits		0.029	−0.032	0.073	0.044	−0.003
IV (see row heading)	Diagnostic use of the PMS		0.040	0.044	0.009	−0.027	−0.018
IV (see row heading)	Interactive use of the PMS		0.084	0.076	−0.009	−0.015	−0.045
Model fit:							
X^2		522.8***	550.1***	544.4***	519.3***	484.0***	548.9***
DF		297	320	320	294	246	320
CMINDF		1.760	1.719	1.701	1.766	1.967	1.715
CFI		0.951	0.951	0.952	0.951	0.946	0.951
RMSEA		0.053	0.051	0.050	0.053	0.059	0.051
SRMR		0.0736	0.0717	0.0719	0.0729	0.0769	0.0719

Note: This table reports alternative models to the Base Model 2. Model A is the Base Model 2 with an additional path of the strategic orientation of the firm (measured by archival data, following Bentley et al. (2013)) on all four constructs of the base model, whereas Model B uses strategic orientation measured by a latent construct based on our survey as an additional antecedent. Model C includes strategic planning intensity (measured by a latent construct based on our survey data), and Model D includes budgeting (also measured by latent constructs based on our survey data) as additional independent variables. Model E also considers PEU (measured by a latent construct). Models F and G use Size as measured by the number of employees and by net sales. Furthermore, Model H uses OP as measured by a latent construct of archival RoCE, RoS, and RoA, and Model I uses archival return on sales (RoS) as additional antecedents. Model J considers firm age as an additional independent variable. For all models, the path coefficient with significance levels and goodness-of-fit indices are reported. We report one-tailed p-values for all directional hypotheses and two-tailed values for the others. ***, **, and * indicate the significance of the p-value at < 0.01, 0.05, and 0.10.
Significant path coefficients are presented in bold.

internal and external environment than smaller firms. Thus, large midsized firms seem to behave like large firms in general, on which MC research has predominantly focused (Chenhall, 2003). Even if all firms, regardless of their size, actually use diagnostic controls to monitor outcomes and deviations from pre-established goals based on measurement and goal setting, a stronger emphasis on diagnostic use may not sufficiently meet the needs of larger firms in light of their previously explained specific situation.

4.3. Robustness tests

To test the robustness of our results, we analyze different modifications of our Base Model 2 (Table 6). Model 1 is Base Model 2 with PMS benefits as a second-order construct (see Wiersma, 2009) based on a factor analysis of the weighted benefit items, resulting in three first-order constructs.[18] Inferences from the results of the Base Model 2 remain unchanged. Model 2 eliminates the strategy items from the PMS benefits measure (see Panel B, Table 2), and Model 3 removes the cause-and-effect item that describes firm strategy from the PMS sophistication level measure (see Panel A, Table 2). Finally, Model 4 eliminates these strategy items from both constructs. For all models, although the model fit is slightly lower, the results for Base Model 2 are generally confirmed. As expected, when we remove the strategic items, the negative relationship of the interaction between diagnostic use and PMS sophistication level with PMS benefits becomes weaker in Models 2 and 4. However, the strategic focus of the PMS (e.g., Kaplan and Norton, 1996) is essential for its design and use and thus for the benefit to the firm (Speckbacher et al., 2003). Hence, removing these strategic items is not supported by theoretical considerations. Model 5 includes all four LoCs (beliefs, boundary, interactive, and diagnostic controls) because Simons (1994) states that all four levers work together, which has been empirically confirmed (e.g., Heinicke et al., 2016; Widener, 2007). Again, our results remain robust. The only significant additional path that we find is the relationship between the beliefs system and PMS benefits ($p < 0.10$).

Finally, to explore the face validity of our measure for PMS sophistication level, we develop alternative models in which we replace PMS sophistication level with two measures constructed from our survey. As a proxy, we measure the adoption of often-discussed PMSs and measure the existence of a BSC (Kaplan and Norton, 1996), a performance pyramid (Lynch and Cross, 1991) or the quantum PMS (Hronec, 1993) (Model 6). Model 7 uses a measure for the level of sophistication of MCs in general based on the adoption of up to 27 potential MCs taken from our survey (Model 7), which results in a scale ranging from 0 to 27 with one point given for each MC used. The correlations between PMS sophistication level and both PMS existence ($r = 0.308$; $p < 0.01$) and MC sophistication level in general ($r = 0.462$; $p < 0.01$) are significant; however, consistent with Cohen (1988), they are only moderate, indicating that all three variables measure different aspects. In Model 6, the paths between PMS sophistication level and both interaction terms with PMS benefits are no longer significant. In Model 7, neither interaction term is significant. However, from the perspective of theoretical face validity, each alternative measures a different variable: the alternative in Model 6 measures only a dichotomous decision on adoption but not a continuous variable for PMS sophistication level, and the alternative in Model 7 measures MCs in general but not PMSs specifically.

To assess the robustness of the structural model, we conduct an MGCA for each of the control variables (i.e., PEU, Dynamic, and Ownership). In all of these unreported MGCAs, the restricted model is not significantly different from the unrestricted model, indicating that the two groups divided by each control variable are not significantly different from each other. For Ownership, we do not obtain any significant differences in path coefficients. For PEU, we find that the path coefficient between diagnostic use and PMS benefits is significantly higher for a low PEU. We obtain the same result for the relationship between PMS benefits and OP. These results confirm the theory that the assessment of predefined goals, variance analysis and feedback are more beneficial in a less hostile environment than in a constantly changing environment. The MGCA for Dynamic delivers several interesting insights. In general, our main results are confirmed. However, firms whose PMS sophistication level increased over the last three years (i.e., firms with high Dynamic values) have significantly stronger relationships between PMS sophistication level and PMS benefits (low Dynamic: coeff. = 0.081, $p > 0.1$; high Dynamic coeff. = 0.202, $p < 0.01$). Furthermore, for both interactive use (low Dynamic: coeff. = 0.505, $p < 0.01$; high Dynamic: coeff. = 0.058, $p > 0.1$) and diagnostic use (low Dynamic: coeff. = -0.311, $p < 0.01$; high Dynamic: coeff. = 0.106, $p > 0.1$), the interaction terms are no longer significant. This result shows that when PMSs are adjusted over time, neither type of use is conditional on the level of sophistication and that PMS sophistication level has an overall stronger relationship with PMS benefits. Thus, the dysfunctional effects of a misfit between diagnostic use and PMS sophistication level can be resolved. Overall, we conclude that our results are robust to variations in PEU, ownership structure, and PMS development over time.

Using MGCA, we also test whether our Base Model 2 is moderated by the type of respondent, i.e., CEOs vs. CFOs and other top managers in finance and accounting. Finance and accounting top managers might be likely to exhibit an “ownership” bias (e.g., for ABC adoption Shields, 1995) because they are often the main sponsors of PMS adoption in firms. However, the untabulated results show that the inferences remain robust, confirming our Base Model 2 and rejecting a potential response bias. Furthermore, to test the robustness of our MGCA for Size, we also estimate the moderation effects of Size as a continuous variable by using the PROCESS syntax of Hayes (2013) based on single regression modeling. Our results confirm a moderating effect of Size on the relationship between an emphasis on interactive use and PMS benefits (coeff. of the interaction term 0.514; $p < 0.01$). However, the moderation effect of Size on the relationship between an emphasis on diagnostic use and PMS benefits is not significant (coeff. 0.115, $p > 0.1$). Furthermore, in untabulated results, we find that a MGCA with a median split of revenues rather than the number of employees leads to qualitatively similar inferences for the hypothesized relationships.

In Table 7, we examine additional dependent variables in alternative models, as other contextual or structural variables may be associated with either diagnostic use or interactive use. Models A and B examine the relationship between a firm’s strategic orientation and all four constructs of Base Model 2. A firm’s strategy is a contextual variable often addressed in MC research (Chenhall and Langfield-Smith, 1998; Chenhall, 2003; Langfield-Smith, 1997) and specifically associated with PMS (e.g., Ahn, 2001; Ittner et al., 2003b; Jazayeri and Scapens, 2008; Perera et al., 1997; Van der Stede et al., 2006). We draw on the typology of Miles and Snow (1978), who propose that prospector, analyzer, and defender strategies simultaneously exist as viable strategies within an industry and exclude the reactor strategy as being unsustainable. This distinction allows firms to be positioned along a continuum with prospectors and defenders at either end and analyzers in between (Hambrick, 1983). Whereas prospectors strive to be innovative market leaders in different domains by rapidly adjusting their product/market mix, defenders maintain a narrow range of product/market combinations and strive to excel in price, service or quality (Miles and Snow, 1978). In Model A, we use an archival measure validated by Bentley et al. (2013).[19] In Model B, we use a variable based

[18] The first-order constructs capture three dimensions: (1) support for strategy implementation, (2) support for stakeholder communication and (3) support for operational processes.

[19] This measure is a continuous scale with prospectors and defenders as the

on our survey items that measure the emphasis on an environmental management system and on a quality management system because both types of systems can be associated with a prospector strategy. Table 7 shows that our main findings remain unchanged. Although the archival measure for strategic orientation is not connected to our variables, the survey measure is positively associated with an emphasis on interactive use ($p < 0.05$) and PMS sophistication level ($p < 0.01$).

Models C and D focus on two other MC mechanisms in addition to a PMS. Strategic planning intensity is constructed based on four items from our survey instrument that relate to the emphasis on strategic planning and on PMSs such as the BSC, the performance pyramid and the quantum PMS. We measure an emphasis on the budgeting system by using five of our survey items regarding the emphasis on static capital budgeting, dynamic capital budgeting, midterm financial planning, cash forecasts, and forecasts for the balance sheet and income statement. We include both constructs as additional independent variables. Table 7 shows that strategic planning intensity and budgeting emphasis are positively associated with PMS sophistication levels. Furthermore, budgeting is associated with an emphasis on interactive PMS use. For both alternative models, the results of our Base Model 2 remain unchanged.

To control for two frequently referenced contingency factors (e.g., Chenhall, 2003) and for potentially omitted variables, we also examine the relationship of PEU and Size (measured by the number of employees and by sales) as independent variables with our focal variables in Models E, F and G. Both PEU and Size are positively associated with PMS sophistication level, indicating that larger firms and firms with higher PEU have more sophisticated PMSs. However, the statistical inferences of our hypothesized results again remain robust. Furthermore, as additional antecedents of our Base Model 2, Model H uses OP as measured by a latent construct of archival RoCE, RoS and RoA, and Model I uses OP as measured by archival RoS. Neither latent OP nor RoS is significantly associated with any variable in our Base Model 2, and the robustness of our inferences is again confirmed. Model J addresses firm age as an additional hand-collected independent variable because the firm life cycle has been found to be associated with the emphasis placed on MCs (e.g., Davila and Foster, 2005; Davila et al., 2015). However, we find that firm age is not statistically significantly associated with the constructs of our Base Model 2, and the robustness of our model is supported once again.

5. Conclusion

This study contributes to the literature in two ways. First, our paper explains the inconclusive and ambiguous outcome effects of PMS reported thus far in the literature (e.g., Franco-Santos et al., 2012; Ittner et al., 2003b; Lee and Yang, 2011; Speckbacher et al., 2003) by empirically disentangling the relationships among PMS design (i.e., PMS sophistication level), interactive and diagnostic use, and the organizational outcomes of a PMS. Based on survey data from 276 midsized firms, we find evidence that the direct relationships of both types of use with PMS benefits are significant and positive. Using polynomial regression analysis, we show that the fit between PMS sophistication level and both interactive and diagnostic use is crucial for PMS benefits. Thus, the benefits of the emphasis on both types of use are conditional on the PMS sophistication level. The resulting interaction terms have opposite signs, indicating that more emphasis on diagnostic use is more beneficial with a simpler PMS but dysfunctional with a more sophisticated PMS, whereas more emphasis on interactive use is more beneficial with a more sophisticated PMS. In addition, because firms use their PMSs both diagnostically and interactively, the two effects overlap and produce both functional and dysfunctional effects on PMS benefits, which may explain the ambiguous and contradictory results in the literature. For example, both case study and survey results may be different when the case firm or the majority of the sample firms have a high or low level of the frequently neglected variable of PMS sophistication. Furthermore, we find significant effects only for the more proximate outcome measure of PMS benefits but not for organizational performance (OP).

Our study expands the survey paper of Henri (2006b) and the case studies of Agostino and Arnaboldi (2012), who examine the relationship between different types of PMS use and PMS sophistication levels. In addition to Henri (2006b), we examine other design elements of PMSs by applying the typologies of Speckbacher et al. (2003) and Franco-Santos et al. (2012). With respect to Agostino and Arnaboldi (2012), we answer their call for cross-sectional studies that generalize their case study results. Expanding both papers, we also include organizational outcomes in our research model. Our results also confirm the finding of Franco-Santos et al. (2012) that firms do not necessarily need to connect their PMSs to their incentive systems. This finding is inconsistent with the LoC framework because Simons (1995) assumes that diagnostic use includes a linkage to incentive systems (i.e., a more advanced PMS sophistication level), which is not the case for PMSs in practice, according to our results. In conclusion, for future research, design and types of use should always be regarded as a “system” and should not be analyzed in isolation from each other. Design and use are two structural variables of the PMS accounting choices made by managers, but they are conditional on one another.

Second, this study contributes to the PMS literature (e.g., Speckbacher et al., 2003; Franco-Santos et al., 2012, Hoque, 2014) by providing a more comprehensive picture of the moderation effect of size on the relationships among both types of use and PMS outcomes. Specifically, we find that larger midsized firms can gain greater benefits from their PMS than smaller firms when they emphasize interactive use. Thus, we observe a shift in benefits from more technical to more social controls (Tessier and Otley, 2012) as size increases. Explanations of this finding may be that larger firms are more complex and diverse; thus, they require controls that allow them to manage uncertainty and find new strategies. In contrast, smaller firms typically use more interpersonal controls and are more closely connected to the markets and to the environment, which reduces communication and coordination problems but creates challenges related to the successful implementation and monitoring of intended and emerging strategies (Chenhall, 2003, 2007; Simons, 1995). Finally, the investigated relationships are surprisingly robust when we consider different measurements and other contextual and structural variables.

Our study results should be considered in light of its limitations. We use empirical data for a specific geographical area and conduct our analyses with a sample comprising 276 firms, and although we find no concerns regarding the representativeness of our sample, the generalizability of the results outside our population may nonetheless be limited. Future research may aim to confirm our results by using larger sample sizes and different settings. We use cross-sectional data collected with a survey-based method. Thus, although we carefully develop the investigated relationships based on underlying theory, we cannot empirically demonstrate the causality of the relationships.

Our results also have interesting implications for practice.

(*footnote continued*)
two anchors and is based on additional data collected over the past five years for six variables for all 17,261 firms in the Amadeus database in 2011 that are within the size range and industry definition of our sample. The six variables are 1) R&D to Sales, 2) Employees to Sales, 3) Sales Growth, 4) Ratio of Marketing Expenses to Sales, 5) Employee Fluctuation (standard deviation in the number of employees over five years), and 6) Capital Intensity (fixed assets to total assets). We average the six variables over five years and categorize each variable into quintiles per industry using two-digit NACE industry codes. A firm is given a score of 5 when the firm is positioned in the top quintile of its industry, a score of 4 for the second highest quintile, and so forth. The scores for the six measures are added for each firm-year. Thus, a firm can reach a maximum score of 30 (associated with a prospector type of strategy) and a minimum score of 6 (associated with a defender strategy).

Specifically, our study shows that the benefits of a PMS are driven by both PMS sophistication level and the two types of use, whereas the fit between PMS sophistication and the emphasis on both types of PMS use is crucial but has opposite effects for diagnostic and interactive use. This finding may help practitioners avoid misspecifications in PMS design, as stated in one case in Agostino and Arnaboldi (2012). Managers have to consider the fit of both types of use with PMS sophistication levels to create benefits for firms. Practitioners should also understand that more sophisticated PMSs might by dysfunctional for diagnostic use, which suggests the need for a simpler PMS design or different designs for diagnostic and interactive PMS use. Furthermore, practitioners should be aware that the importance and benefits of interactive PMS use increase with firm size.

Acknowledgments

We would like to thank *Jan Endrikat*, *Isabelle Grabner* and *Frances Kennedy* for their feedback on earlier versions of this paper. Earlier drafts of this paper were presented at the *9*th *Conference on New Directions in Management Accounting* in Brussels, Belgium, 2014; at *the 77*th *Conference of the German Academic Association for Business Research* in Vienna, Austria, 2015; and at the *European Accounting Association Conference* in Glasgow, United Kingdom, 2015. We thank all the participants for their helpful comments and suggestions. A very early, but now completely changed version of the paper has been part of the cumulative dissertation thesis of the second author. Furthermore, we are grateful for the very fruitful comments of the editor, *Wim Van der Stede*, and two anonymous reviewers during the review process of the *Management Accounting Review*. In addition, *Jan Endrikat*, *Matthias Mahlendorf* and *Sally K. Widener* gave us guidance on and recommendations for improving our revisions.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.mar.2018.07.002.

Appendix A1. Generalizability of the obtained sample

Panel A: Representativeness of the obtained sample

Sales in m€			Ownership		
	Received questionnaires	Expected questionnaires		Received questionnaires	Expected questionnaires
20–39	88	96.0	family firms	74	67.6
40–99	95	98.9	nonfamily firms	202	208.4
100–249	58	51.4			
250–1.000	35	29.6			
Total	276		Total	276	
Chi-square test statistic	2.655			0.797	
Degrees of freedom	3			1	
p-value	0.448			0.372	

Note: The table reports the results of the chi-square statistics for the test of distributional adequacy of the received sample. No significant differences are reported ($p > 0.05$).

Panel B: Nonresponse analysis for the financial characteristics of the firms

Variable	Respondents (n = 276)	Addressed nonrespondents (n = 2224)	Survey population (n = 2500)	Mann-Whitney U test
Sales category (1: 20-39 m€; 2: 40-99 m€; 3: 100-249 m €, 4: 250-1,000 m€)	2.14	2.04	2.05	Z = −1.603 (p = 0.109)
Profit margin	3.91	3.54	3.58	Z = −1.487 (p = 0.137)
RoE	13.78	10.17	10.57	Z = −1.849 (p = 0.064)

Note: The table reports variable means and the results of a Mann-Whitney U test for a comparison of variable means between respondents and addressed nonrespondents. No significant differences ($p > 0.05$) are reported.

Panel C: Comparison of constructs for early and late respondents

Construct	Early respondents (before follow-up procedure) mean rank	Late respondents (after follow-up procedure) mean rank	Mann-Whitney U test
Diagnostic use of the PMS	139.74 (n = 210)	134.57 (n = 66)	Z = −0.463, p = 0.643
Interactive use of the PMS	140.83 (n = 210)	131.08 (n = 66)	Z = −0.867, p = 0.386
PMS sophistication level	142.25 (n = 210)	126.58 (n = 66)	Z = −1.392, p = 0.164
PMS benefits	144.55 (n = 210)	119.27 (n = 66)	Z = −2.244, p = 0.025

Note: The table reports the results of a comparison of early and late respondents based on the Mann-Whitney U test. A significant difference is reported only for PMS benefits. For the other constructs, no significant differences (p > 0.05) are found.

Appendix A2. Descriptive statistics

Panel A: Sales and number of employees of our sample firms

	Sales in EUR (n = 276)	Number of employees (n = 246)
Mean	118,688.36	496.54
Standard deviation	154,281.73	753.247
10% quantile	23,956.50	53.70
90% quantile	320,584.50	1,100.20

Note: The table reports the characteristics of the sample in terms of the mean, standard deviation, 10% quantile, and 90% quantile sales (in EUR) and number of employees.

Panel B: Construct means, minimum and maximum of the study variables split by firm size

	Smaller firms n = 141			Larger firms n = 135			
	Construct mean	Construct minimum	Construct maximum	Construct mean	Construct minimum	Construct maximum	Mann-Whitney U test
Diagnostic use of the PMS	4.0711	1	5	4.2414	1	5	Z = −2.554, p = 0.011
Interactive use of the PMS	3.2556	1	5	3.4355	1	5	Z = −1.977, p = 0.046
PMS sophistication level	2.5937	1	5	3.2207	1	5	Z = −4.483, p < 0.01
PMS benefits	10.52	1	19	11.49	1	25	Z = −1.917, p = 0.055
Budgeting	3.51	1	5	3.84	0	4	Z = −2.180, p = 0.029
Strategic planning intensity	1.11	0	4	1.27	0	4	Z = −2.921, p = 0.003

Note: The table reports construct means, minima and maxima of the study variables, split by firm size using a medium split. Budgeting is measured by five items from our survey regarding the use of static capital budgeting, dynamic capital budgeting, midterm financial planning, cash forecasts, and forecasts for the balance sheet and income statement in the responding firms (theoretical minimum of 0 and maximum of 5 for the scale). Strategic planning intensity is measured by four items in our survey instrument that are related to firm emphasis on strategic planning and on PMSs such as the BSC, performance pyramid and quantum PMS (theoretical minimum of 0 and maximum of 4). Emphases on diagnostic use and on interactive use are latent constructs based on five-point Likert scales (theoretical minimum of 1 and maximum of 5 each). The score for PMS sophistication level ranges between 1 and 5, whereas the score for PMS benefits ranges between 1 and 25. The subgroup split for Size is not precisely 50%/50%, as the basis for the split is the classification of Size in seven groups in the survey. The three lowest size groups comprise 141 firms, whereas the four largest ones comprise 135 firms.

Appendix A3 Exploratory factor analysis

	Factor 1 (interactive use of the PMS)	Factor 2 (diagnostic use of the PMS)
Track progress toward goals		0.867
Monitor results		0.942
Compare outcomes to expectations		0.809
Review key measures		0.737
Enable discussions in meetings of superiors, subordinates, and peers	0.624	
Enable continual challenge and debate of underlying data, assumptions, and action plans	0.614	
Provide a common view of the organization	0.690	
Tie the organization together	0.978	
Enable the organization to focus on common issues	0.894	
Enable the organization to focus on critical success factors	0.640	
Develop a common vocabulary in the organization	0.834	

Note: The table reports the results of an exploratory factor analysis of the questions in our questionnaire for the reflective constructs, i.e., diagnostic and interactive use of the PMS. The results for the formative constructs (PMS sophistication level and PEU) are reported separately. We use principal components with promax rotation to extract all factors with an eigenvalue > 1. The results of the explorative factor analysis confirm that the latent constructs are unidimensional. No item loads below 0.3.

References

Abdel-Kader, M., Luther, R., 2008. The impact of firm characteristics on management accounting practices: a UK-based empirical analysis. Br. Acc. Rev. 40 (1), 2–27.
Abernethy, M.A., Brownell, P., 1999. The role of budgets in organizations facing strategic change: an exploratory study. Acc., Org. Soc. 24 (3), 189–204.
Abernethy, M.A., Lillis, A.M., 1995. The impact of manufacturing flexibility on management control system design. Acc., Org. Soc. 20 (4), 241–258.
Agostino, D., Arnaboldi, M., 2012. Design issues in balanced scorecards: the "what" and "how" of control. Eur. Manage. J. 30 (4), 327–339.
Ahn, H., 2001. Applying the balanced scorecard concept: an experience report. Long Range Plann. 34 (4), 441–461.
Ahrens, Th., Chapman, C.S., 2004. Accounting for flexibility and efficiency: a field study of management control systems in a restaurant chain. Contemp. Acc. Res. 21 (2), 271–301.
Andon, P., Baxter, J., Chua, W.F., 2007. Accounting change as relational drifting: a field study of experiments with performance measurement. Manage. Acc. Res. 18 (2), 273–308.
Anthony, R.N., Govindarajan, V., 1998. Management Control Systems. Irwin, Boston, MA.
Argyris, C., Schön, D.A., 1978. Organizational Learning. Addison-Wesley, Reading, MA.
Astrachan, J.H., Klein, S.B., Smyrnios, K.X., 2002. The F-PEC scale of family influence: a proposal for solving the family business definition problem. Family Bus. Rev. 15 (1), 45–58.
Bagozzi, R.P., Yi, Y., 1988. On the evaluation of structrual equation models. J. Acad. Mark. Sci. 16 (1), 74–94.
Bagozzi, R.P., Baumgartner, H., 1994. The evaluation of structural equation models and hypotheses testing. In: Bagozzi, R.P. (Ed.), Principles of Marketing Research. Blackwell Business, Cambridge, MA, pp. 386–422.
Baird, K.M., Harrison, G.L., Reeve, R.C., 2004. Adoption of activity management practices: a note on the extent of adoption and the influence of organizational and cultural factors. Manage. Acc. Res. 15 (4), 383–399.
Bedford, D.S., Speklé, R.F., 2018. Construct validity in survey-based management accounting and control research. J. Manage. Acc. Res., in print.
Bentler, P.M., 1990. Comparative fit indexes in structural models. Psychol. Bull. 107 (2), 238–246.
Bentler, P.M., 1992. On the fit of models to covariances and methodology to the bulletin. Psychol. Bull. 112 (3), 400–404.
Bentley, K.A., Omer, T.C., Sharp, N.Y., 2013. Business strategy, financial reporting irregularities, and audit effort. Contemp. Acc. Res. 30 (2), 780–817.
Birnberg, J.G., Turopolec, L., Young, S.M., 1983. The organizational context of accounting. Acc., Org. Soc. 8 (2/3), 111–129.
Bisbe, J., Malagueno, R., 2012. Using strategic performance measurement systems for strategy formulation: does it work in dynamic environments? Manage. Acc. Res. 23 (4), 296–311.
Bisbe, J., Otley, D., 2004. The effects of the interactive use of management control systems on product innovation. Acc., Org. Soc. 29 (8), 709–737.
Bisbe, J., Batista-Foguet, J.M., Chenhall, R., 2007. Defining management accounting constructs: a methodological note on the risks of conceptual misspecification. Acc., Org. Soc. 32 (7), 789–820.
Blau, P.M., 1964. The Dynamics of Bureaucracy: A Study of Interpersonal Relations in Two Government Agencies, 2nd ed. The University of Chicago Press, Chicago.
Bourguignon, A., Malleret, V., Nørreklit, H., 2004. The American balanced scorecard versus the French tableau de bord: the ideological dimension. Manage. Acc. Res. 15 (2), 107–134.
Braam, G.J.M., Nijssen, E.J., 2004. Performance effects of using the balanced scorecard: a note on the Dutch experience. Long Range Plann. 37 (4), 335–349.
Browne, M.W., Cudeck, R., 1993. Alternative ways of assessing model fit. In: Bollen, K.A., Long, J.S. (Eds.), Testing Structural Equation Models. Sage, Newbury Park, CA, pp. 136–162.
Bruns, W.J., Waterhouse, J.H., 1975. Budgetary control and organization structure. J. Account. Res. 13 (2), 177–203.
Bukh, N., Malmi, T., 2005. Re-examining the cause-and-effect principle of the balanced scorecard. In: Jönsson, S., Mouritsen, J. (Eds.), Accounting in Scandinavia – The Northern Lights. Liber & Copenhagen Business School Press, Malmö pp. 87–113.
Burkert, M., Davila, A., Mehta, K., Oyon, D., 2014. Relating alternative forms of contingency fit to the appropriate methods to test them. Manage. Acc. Res. 25 (1), 6–29.
Burney, L., Widener, S.K., 2007. Strategic performance measurement systems, job-relevant information, and managerial behavioral responses-role stress and performance. Behav. Res. Acc. 19, 43–69.
Burns, T., Stalker, G.M., 1961. The Management of Innovation. Tavistock, London.
Byrne, B.M., 2010. Structural Equation Modeling With AMOS: Basic Concepts, Applications, and proGramming. Taylor & Francis, New York, London et al.
Chenhall, R.H., 2003. Management control systems design within its organizational context: findings from contingency-based research and directions for the future. Acc., Org. Soc. 28 (2), 127–168.
Chenhall, R.H., 2005. Integrative strategic performance measurement systems, strategic alignment of manufacturing, learning and strategic outcomes: an exploratory study. Acc., Org. Soc. 30 (5), 395–422.
Chenhall, R.H., 2007. Theorizing contingencies in management control systems research. In: Chapman, C.S., Hopwood, A.G., Shields, M.D. (Eds.), Handbook of Management Accounting Research. Elsevier, Oxford, UK.
Chenhall, R.H., Langfield-Smith, K., 1998. The relationship between strategic priorities, management techniques and management accounting: an empirical investigation using a systems approach. Acc., Org. Soc. 23 (3), 243–264.
Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Lawrence Erlbaum, Mahwah, NJ.
Cortina, J.M., Chen, G., Dunlap, W.P., 2001. Testing interaction effects in Lisrel: examination and illustration of available procedures. Org. Res. Methods 4 (4), 324–360.
Crabtree, A.D., DeBusk, G.K., 2008. The effects of adopting the balanced scorecard on shareholder returns. Adv. Acc. 24 (1), 8–15.
Cromie, S., Stephenson, B., Monteith, D., 1995. The management of family firms: an empirical investigation. Int. Small Bus. J. 13 (4), 11–34.
Davila, A., Foster, G., 2005. Management accounting systems adoption decisions: evidence and performance implications from early-stage/startup companies. Acc. Rev. 80 (4), 1039–1068.
Davila, A., Foster, G., Jia, N., 2015. The valuation of management control systems in start-up companies: international field-based evidence. Eur. Acc. Rev. 24 (2), 207–239.
Davis, S., Albright, T., 2004. An investigation of the effect of balanced scorecard implementation on financial performance. Manage. Acc. Res. 15 (2), 135–153.
de Geuser, F., Mooraj, S., Oyon, D., 2009. Does the balanced scorecard add value? Empirical evidence on its effects on performance. Eur. Acc. Rev. 18 (1), 93–122.
Diamantopoulos, A., Winklhofer, H.M., 2001. Index construction with formative indicators: an alternative to scale development. J. Mark. Res. 38 (2), 269–277.
Diamantopoulos, A., Riefler, P., Roth, K.P., 2008. Advancing formative measurement models. J. Bus. Res. 61 (12), 1203–1218.
Dillman, D.A., Smyth, J.D., Christian, L.M., 2014. Internet, Mail, and Mixed-Mode Surveys: The tailored Design Method, 4th ed. John Wiley & Sons, Hoboken.
Drazin, R., Van de Ven, A.H., 1985. Alternative forms of fit in contingency theory. Adm. Sci. Q. 30, 514–539.
Eccles, R., 1991. The performance measurement manifesto. Harv. Bus. Rev. 63 (1), 131–137.

Edwards, J.R., Parry, M.E., 1993. On the use of polynomial regression equations as an alternative to difference scores in organizational research. Acad. Manage. J. 36, 1577–1613.
Endrikat, J., Guenther, T.W., Titus, R., 2018. A Meta-analysis of the Performance Effects of Contemporary Performance Measurement Systems, Working paper. TU Dresden.
Ferreira, A., Otley, D., 2009. The design and use of performance management systems: an extended framework for analysis. Manage. Acc. Res. 20 (4), 263–282.
Franco-Santos, M., Bourne, M., 2005. An examination of the literature relating to issues affecting how companies manage through measures. Prod. Plann. Control 16 (2), 114–124.
Franco-Santos, M., Lucianetti, L., Bourne, M., 2012. Contemporary performance measurement systems: a review of their consequences and a framework for research. Manage. Acc. Res. 23 (2), 79–119.
Garengo, P., Biazzo, S., Bititci, U.S., 2005. Performance measurement systems in SMEs: a review for a research agenda. Int. J. Manage. Rev. 7 (1), 25–47.
Gerdin, J., Greve, J., 2004. Forms of contingency fit in management accounting research - a critical review. Acc., Org. Soc. 29 (3), 303–326.
Gerdin, J., Greve, J., 2008. The appropriateness of statistical methods for testing contingency hypotheses in management accounting research. Acc., Org. Soc. 33 (7), 995–1009.
Gimbert, X., Bisbe, J., Mendoza, X., 2010. The role of performance measurement systems in strategy formulation processes. Long Range Plann. 43 (4), 477–497.
Gordon, L.A., Narayanan, V.K., 1984. Management accounting systems, perceived environmental uncertainty and organization structure: an empirical investigation. Acc., Org. Soc. 9 (1), 33–47.
Govindarajan, V., 1984. Appropriateness of accounting data in performance evaluation: an empirical examination of environmental uncertainty as an intervening variable. Acc., Org. Soc. 9 (2), 125–135.
Govindarajan, V., 1988. A contingency approach to strategy implementation at the business unit level. Integrating administrative mechanisms with strategy. Acad. Manage. J. 31 (4), 828–853.
Grafton, J., Lillis, A.M., Widener, S.K., 2010. The role of performance measurement and evaluation in building organizational capabilities and performance. Account. Org. Soc. 35 (7), 689–706.
Griffith, R., Neely, A., 2009. Performance pay and managerial experience in multitask teams: evidence from within a firm. J. Labor Econ. 27 (1), 49–82.
Gupta, A.K., Govindarajan, V., 1984. Business unit strategy, managerial characteristics, and business unit effectiveness at strategy implementation. Acad. Manage. J. 27 (1), 25–41.
Hambrick, D.C., 1983. Some tests of the effectiveness and functional attributes of Miles and Snow's strategic types. Acad. Manage. J. 26 (1), 5–26.
Hayes, A.F., 2013. Introduction to Mediation, Moderation, and Conditional Process Analysis, A Regression-Based Approach. Guilford, New York.
Heinicke, A., Guenther, T.W., Widener, S.K., 2016. An examination of the relationship between the extent of a flexible culture and the levers of control system: the key role of beliefs control. Manage. Acc. Res.h 33 (1), 25–41.
Henri, J., 2006a. Management control systems and strategy: a resource-based perspective. Acc., Org. Soc. 31 (6), 529–558.
Henri, J., 2006b. Organizational culture and performance measurement systems. Acc., Org. Soc. 31 (1), 77–103.
Hoque, Z., 2014. 20 years of studies on the balanced scorecard: trends, accomplishments, gaps and opportunities for future research. Br. Acc. Rev. 46 (1), 33–59.
Hoque, Z., James, W., 2000. Linking balanced scorecard measures to size and market factors: impact on organizational performance. J. Manage. Acc. Res. 12, 1–17.
Hronec, S.M., 1993. Vital Signs: Using Quality, Time and Cost Performance Measurements to Chart Your Company's Future. Amacom, New York.
Hu, L., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria verus new alternatives. Struct. Equ. Model. 6 (1), 1–55.
Ittner, C.D., Larcker, D.F., 1998. Innovations in performance measurement: trends and research implications. J. Manage. Acc. Res. 10, 205–238.
Ittner, C.D., Larcker, D.F., Meyer, M.W., 2003a. Subjectivity and the weighting of performance measures: evidence from a balanced scorecard. Account. Rev. 78 (3), 725–758.
Ittner, C.D., Larcker, D.F., Randall, T., 2003b. Performance implications of strategic performance measurement in financial services firms. Acc., Org. Soc. 28 (7-8), 715–741.
Jaworski, B.J., Young, S.M., 1992. Dysfunctional behavior and management control: an empirical study of marketing managers. Acc., Org. Soc. 17 (1), 17–35.
Jazayeri, M., Scapens, R.W., 2008. The business values scorecard within BAE systems: the evolution of a performance measurement system. Br. Account. Rev. 40 (1), 48–70.
Kaplan, R.S., Norton, D.P., 1996. Using the balanced scorecard as a strategic management system. Harv. Bus. Rev. 74 (1), 75–85.
Kaplan, R.S., Norton, D.P., 2001. Transforming the balanced scorecard from performance measurement to strategic management: part I and II. Acc. Horizons 15 (1) 87–104 and 15(2), 147–160.
Kaplan, R.S., Norton, D.P., 2006. Alignment: Using the Balanced Scorecard to Create Corporate Synergies. Harvard Business School Press, Boston.
Kaplan, R.S., Norton, D.P., 2008. Mastering the management system. Harv. Bus. Rev. 86 (1), 63–77.
Kasurinen, T., 2002. Exploring management accounting change: the case of balanced scorecard implementation. Manage. Acc. Res. 13 (3), 323–343.
Kihn, L.A., 2007. Financial consequences in foreign subsidiary manager performance evaluations. Eur. Acc. Rev. 16 (3), 531–554.
Klein, S.B., 2000. Family businesses in Germany: significance and structure. Family Bus. Rev. 13 (3), 157–181.
Kline, R.B., 2011. Principles and Practices of Structural Equation Modeling. The Guilford Press, New York, NY.
Kohli, A.K., Jaworski, B.J., 1990. Market orientation: the construct, research propositions, and managerial implications. J. Mark. 54 (2), 1–18.
Koufteros, X., Verghese, A., Lucianetti, L., 2014. The effect of performance measurement systems on firm performance: a cross-sectional and a longitudinal study. J. Oper. Manage. 32 (6), 313–336.
Kren, L., Liao, W.M., 1988. The role of accounting information in the control of organizations: a review of the evidence. J. Acc. Lit. 7 (1), 280–309.
Langfield-Smith, K., 1997. Management control systems and strategy: a critical review. Acc., Org. Soc. 22 (2), 207–232.
Lee, C., Yang, H., 2011. Organization structure, competition and performance measurement systems and their joint effects on performance. Manage. Acc. Res. 22 (2), 84–104.
Lohrke, F.T., Kreiser, P.M., Weaver, K.M., 2006. The influence of current firm performance on future SME alliance formation intentions: a six-country study. J. Bus. Res. 59 (1), 19–27.
Lopez, O.L., Hiebl, M.R.W., 2015. Management accounting in small and medium-sized enterprises: current knowledge and avenues for further research. J. Manage. Acc. Res. 27 (1), 81–119.
Lynch, R.L., Cross, K.F., 1991. Measure up - Yardsticks for Continuous Improvement. Basil Blackwell, Cambridge, MA.
Malina, M.A., Selto, F.H., 2001. Communicating and controlling strategy: an empirical study of the effectiveness of the balanced scorecard. J. Manage. Acc. Res. 13, 47–90.
Malina, M.A., Nørreklit, H.S.O., Selto, F.H., 2007. Relations among measures, climate of control, and performance measurement models. Contemp. Acc. Res. 24 (3), 935–982.
Malmi, T., 2001. Balanced scorecards in Finnish companies: a research note. Manage. Acc. Res. 12 (2), 207–220.
Merchant, K.A., 1981. The design of the corporate budgeting system: influences on managerial behavior and performance. Acc. Rev. 56 (4), 813–829.
Merchant, K.A., 1989. Rewarding Results: Motivating Profit Center Managers. Harvard Business School Press, Boston, MA.
Micheli, P., Manzoni, J., 2010. Strategic performance measurement: benefits, limitations and paradoxes. Long Range Plann. 43 (4), 465–476.
Miles, R.E., Snow, C.C., 1978. Organizational Strategy, Structure and Process. McGraw-Hill, New York, NY.
Mitchell, F., Reid, G.C., 2000. Problems, challenges and opportunities: the small business as a setting for management accounting research. Manage. Acc. Res. 11 (3), 385–390.
Neely, A., Gregory, M., Platts, K., 1995. Performance measurement system design. A literature review and research agenda. Int. J. Oper. Prod. Manage. 25 (12), 1228–1263.
Neely, A.D., Adams, C., Kennerley, M., 2002. The Performance Prism: The Scorecard for Measuring and Managing Business Success. Pearson Education, London.
Neely, A.D., Yaghi, B., Youell, N., 2008. Enterprise Performance Management:The Global State of the Art. Oracle and Cranfield School of Management Publication, Cranfield.
Nørreklit, H., 2000. The balance on the balanced scorecard, a critical analysis of some of its assumptions. Manage. Acc. Res. 11 (1), 65–88.
Nunnally, J.C., 1978. Psychometric Theory, Second ed. McGraw-Hill Book Company, New York, NY.
O'Gorman, C., Doran, R., 1999. Mission statements in small and medium-sized businesses. J. Small Bus. Manage. 37 (4), 59–66.
Otley, D., 2016. The contingency theory of management accou nting and control: 1980–2014. Manage. Acc. Res. 31 (1), 45–62.
Papalexandris, A., Ioannou, G., Prastacos, G.P., 2004. Implementing the balanced scorecard in Greece: a software firm's experience. Long Range Plann. 37 (4), 351–366.
Perera, S., Harrison, G., Poole, M., 1997. Customer-focused manufacturing strategy and the use of operations-based non-financial performance measures: a research note. Acc. Org. Soc. 22 (6), 557–572.
Quattrone, P., Hopper, T., 2001. What does organizational change mean? Speculations on a taken for granted category. Manage. Acc. Res. 12 (4), 403–435.
Rigby, D., Bilodeau, B., 2009. Management Tools and Trends. Bains and Co., Inc.
Sandino, T., 2007. Introducing the first management control systems: evidence from the retail sector. Acc. Rev. 82 (1), 265–293.
Schaeffer, U., Matlachowsky, P., 2008. Warum die balanced scorecard nur selten als strategisches Managementsystem genutzt wird (Why the balanced scorecard is seldomly used as strategic management systems). Zeitschrift für Planung & Unternehmenssteuerung 19 (2), 207–232.
Schick, A., Gordon, L., Haka, S., 1990. Information overload: a temporal approach. Acc., Org. Soc. 15 (3), 199–220.
Shields, M.D., 1995. An empirical analysis of firms' implementation experiences with activity-based costing. J. Manage. Acc. Res. 7, 148–166.
Simons, R., 1991. Strategic orientation and top management attention to control systems. Strateg. Manage. J. 12 (1), 49–62.
Simons, R., 1994. How new top managers use control systems as levers of strategic renewal. Strateg. Manage. J. 15 (3), 169–189.
Simons, R., 1995. Levers of Control: How Managers Use Innovative Control Systems to Drive Strategic Renewal. Harvard Business School Press, Boston, MA.
Simons, R., 2014. Performance Measurement and Control Systems for Implementing Strategy. Pearson, Harlow.
Smith, P., 1995. On the unintended consequences of publishing performance data in the public sector. Int. J. Public Admin. 18 (2/3), 277–310.
Speckbacher, G., Wentges, P., 2012. The impact of family control on the use of performance measures in strategic target setting and incentive compensation: a research note. Manage. Acc. Res. 23 (1), 34–46.
Speckbacher, G., Bischof, J., Pfeiffer, T., 2003. A descriptive analysis on the implementation of balanced scorecards in German-speaking countries. Manage. Acc. Res. 14 (4), 361–387.
Specklé, R.F., Verbeeten, F.H.M., 2014. The use of performance measurement systems in the public sector: effects on performance. Manage. Acc. Res. 25 (2), 131–146.
Steenkamp, J.E.M., Baumgartner, H., 1998. Assessing measurement invariance in cross-national consumer research. J. Consum. Res. 25 (1), 78–90.
Tessier, S., Otley, D., 2012. A conceptual development of Simons' levers of control framework. Manage. Acc. Res. 23 (3), 171–185.
Townley, B., Cooper, D.J., Oakes, L., 2003. Performance measures and the rationalization of organizations. Org. Stud. 24 (7), 1045–1071.

Tuomela, T., 2005. The interplay of different levers of control: a case study of introducing a new performance measurement system. Manage. Acc. Res. 16 (3), 293–320.
Vachani, S., 2005. Problems of foreign subsidiaries of SMEs compared with large companies. Int. Bus. Rev. 14, 415–439.
Van der Stede, W.A., Chow, C.W., Lin, T.W., 2006. Strategy, choice of performance measures, and performance. Behav. Res. Acc. 18 (1), 185–205.
Van Veen-Dirks, P., 2010. Different uses of performance measures: the evaluation versus reward of production managers. Acc., Org. Soc. 35 (2), 141–164.
Vandenbosch, B., 1999. An empirical analysis of the association between the use of executive support systems and perceived organizational competitiveness. Acc. Org. Soc. 24 (1), 77–92.
Verbeeten, F.H.M., Boons, A.N.A.M., 2009. Strategic priorities, performance measures and performance: an empirical analysis in Dutch firms. Eur. Manage. J. 27 (2), 113–128.
Wheaton, B., Muthen, B., Alwin, D.F., Summers, G.F., 1977. Assessing reliability and stability in panel models. In: Heise, D.R. (Ed.), Sociological Methodology. Jossey-Bass, San Francisco, CA, pp. 84–136.
Widener, S.K., 2007. An empirical analysis of the levers of control framework. Acc., Org. Soc. 32 (7-8), 757–788.
Wiersma, E., 2009. For which purposes do managers use balanced scorecards? An empirical study. Manage. Acc. Res. 20 (4), 239–251.
Wouters, M., 2009. A developmental approach to performance measures – results from a longitudinal case study. Eur. Manage. J. 27, 64–78.
Yongvanich, K., Guthrie, J., 2009. Balanced scorecard practices amongst Thai companies: performance effects. Pacific Acc. Rev. 21 (2), 132–149.