How to predict performance management systems success? Utilization of the critical check points

Author(s): Kadak, Tarmo; Laitinen, Erkki K.

Title: How to predict performance management systems success? Utilization of the critical check points

Year: 2026

Version: Accepted manuscript

Please cite the original version:

Kadak, T., & Laitinen, E. K. (2026). How to predict performance management systems success? Utilization of the critical check points. *International Journal of Productivity and Performance Management 74*(8), 2768-2795. https://doi.org/10.1108/IJPPM-03-2024-0171

How to predict PMS success? Utilization of the critical check points

Tarmo Kadak*
Tallinn University of Technology, Estonia
tarmo.kadak@taltech.ee
Erkki K. Laitinen
University of Vaasa, Finland
erkki.k.laitinen@uwasa.fi
*Corresponding author

Abstract

Purpose – The prediction of the success of performance management systems (PMS) is difficult and may lead to contradictory results because there are many potential factors locating at different hierarchical levels of an organization and usually only reflecting a part of success. Therefore, in this study we make use of the PMS chain theory developed by Kadak & Laitinen (2016). This theory is based on fifteen check points (CP) of compatible key factors (KF) which form a logical and comprehensive chain for PMS success. The purpose of this research paper is firstly to compare the relationship between PMS success and CP variables in original (2015) and newer (2022) samples (RQ 1). Secondly, the purpose is to assess different statistical methods to predict the success of PMSs using information from these CPs (RQ 2).

Design/methodology/approach – This approach is based on the chain theory of PMS success introduced by Kadak & Laitinen (2016). For the empirical research, data have been collected from 73 Estonian companies that use PMS. This survey data contains information on 15 CPs and assessments of the success of PMS. The research compares five different statistical methods (sum of variables, linear regression, logistic regression, ridge regression, and a numerical Solver solution) to predict the success of PMS with the help of CP variables information. The success of PMS is in this study measured by the self-assessed impact on organizational performance.

Findings – The results of the study show that in the sample all 15 CP variables are highly correlated both with PMS success and with each other. The main contingencies are similar in the older and newer sample. This supports the chain theory and also RQ 1 on the similarity of the two samples. If we evaluate the performance of statistical methods in explaining PMS success with the help of ROC curve, the best results were produced by the logistic regression analysis and the numerical Solver optimization method (RQ 2). In addition, we found three CP variables that produce incremental information about PMS and affect therefore on the PMS success more than other CPs.

Practical implications – The research produces models that can be practically used to design and to predict the success of PMS. The simplest model is to calculate the sum of the CP variables, while the most effective model is based on the logit of these variables or the weighted sum of the three most important CP variables solved by a numerical optimization method. The results can be used also in constructing an efficient predictor of PMS success of companies in time.

Originality/value – The study adds to the literature a detailed and comprehensive, and multi-faceted, methodology for assessing and predicting the success of PMS. Used methodology is more detailed than the general consensus (about which elements form a PMS) and that lists measurable PMS components under each PMS element.

Keywords Performance Management Systems, performance measures, chain of check points, PMS success

Paper type Research paper

1. Introduction and motivation

Performance Management Systems (PMS) are widely used by management to improve organizational performance. In generally, PMS can be defined as a systematic process of quantifying and improving performance of an organization. However, according to common definitions of PMS, PMSs in more detail are defined by their scope, the interrelationship of their components (control mechanisms), direction, to whom PMSs are directed and where these direct the organizations. In addition, PMS includes control mechanisms covering the entire organization and organizational goals that are linked together, so that the employees are made responsible for goals, directing them to achieve the goals with the help of control mechanisms (Broadbent & Laughlin, 2009; Ferreira & Otley, 2009, Otley, 1999; Flamholtz et al., 1985).

Thus, PMSs are in essence complex, on the one hand reflecting the internal and external environments of companies, which are very company-specific, and on the other hand containing much that is common between PMSs of different companies (the structure and components of PMS). Because it is influenced by many contingency factors, the PMSs created are company-specific and designed to be tailor-made and vary in their success. Therefore, companies all around the world report mixed, either positive or negative, news of the success of PMSs (Waal, 2018).

PMSs have achieved immense popularity in academic and business communities (Lewandowski & Cirella, 2022). Their spread is supported by the various intended and unintended benefits they provide (Franco-Santos & Otley, 2018). They highlighted that the most common unintended consequences are gaming, information manipulation, selective attention and illusion of control. In these cases, PMSs cannot support companies. But if PMS is used properly, then it can strongly help management to implement the strategy of the organization and to improve its performance and provide other benefits (Cuccurullo et al., 2016; Rampho & Boon-itt, 2012; Kaplan & Norton, 1996). Waal et al. (2009) showed that organizations with fully completed PMS, gained more financial and non-financial advantages. Franco-Santos et al. (2012) in a comprehensive literature review listed the positive consequences of PMS on performance in various aspects, like positive influence on financial, stock market, non-financial, customer, team and managerial performances, also about achievement of strategic goals and outcomes. However, there is no consensus whether PMSs are beneficial or detrimental to the performances of organizations (Waal & Kourtit, 2013; Nørreklit et al., 2012; Liedtka et al., 2008; Bourne et al., 2007; Laitinen, 2003; Nørreklit, 2003, 2000; Bonner & Sprinkle, 2002). This disagreement raises the important question of when and on what basis the success of a PMS can be assessed.

The success of PMS is in fact defined by its ability to support management. If PMS helps management successfully to carry out their tasks, it can be considered successful. However, PMS is a complicated system (Okwir et al., 2018) and often may lead to dysfunctional behaviour (Franco-Santos & Otley, 2018; Cuguero-Escofte & Rosanas, 2016) and to decline of organizational performance if it is not constructed properly. In this case, PMS is considered unsuccessful. Thus, it is of importance to develop methodologies to assess the potential effect of the construction on PMS success.

There are many contradictory examples about the benefits and success of PMSs in the literature. Different consequences and performances of implemented PMSs are found mixed and contradicting (Malagueño et al., 2018; Llah et al., 2017; Bento et. al., 2016; Franco-Santos et al., 2012; Bedford et al., 2008; Yu et al., 2008; Neely, 2008; Davis & Albright, 2004; Malina & Selto, 2001). Furthermore, Hoque (2004) and Hoque and James (2000) reported positive and Jazayeri and Scapens (2008) negative relationships between PMSs and financial performance. These studies are mainly based on similar circumstances – manufacturing, large size firms and on the same broad description of PMS (usage of nonfinancial measures). Thus, even similar variables and similarly defined PMSs, may provide contradicting impacts on financial performance. In the same way, Crabtree & Debusk (2008) reported positive and Hassabelbaby et al. (2005) and Said et al. (2003) negative relationships between PMSs and financial performance. These studies are also based on similar circumstances – wide range of industry firms, and PMSs are similarly described as using nonfinancial measures. So, also in these studies similar circumstances and similarly described PMSs provided contradicting financial performances. Furthermore, these studies addressed PMS in a fragmentary manner and did not consider the content or the construction of PMS.

There is a vast amount of literature giving instructions on how to design and implement a PMS but relatively few studies on the success and failure of PMSs (Bourne, 2005). Originally, this literature included mainly PMS success stories, but during the last twenty years studies have addressed also experienced difficulties in design and implementation (Bourne et al., 2000; Bourne et al., 1999; Schneiderman, 1999; McCunn, 1998). Moreover, there studies emphasizing the negative aspects of implemented PMSs (Couturier & Sklavounos, 2019; Choong, 2014; Taticchi et al., 2012; Bourne, 2008; Franco-Santos et al., 2007; Neely, 2005; Bourne et al., 2000; Neely et al., 1995). It has even been claimed that as many as 70% of PM initiatives fail (McCunn, 1998). The failures in adaption of BSC are also reported in researches by Bento (2016), Behery et al., (2014), Othman (2006), Speckbacher et al. (2003), Malmi (2001) and Letza (1996).

In order to be useful, a PMS have to be accord in with the characteristics of the organization and its surroundings (Tangen, 2005). This obviously leads to a high variety of different PMSs in use as a response to different strategic initiatives of firms (Hope & Hope, 1995; Govindarajan & Shank, 1992; Miles & Snow, 1978).

Furthermore, the term "system" in the name of the PMS highlights the question of what exactly that system should consist of. For the question of the contents of PMS it is needed to know what is found in the composition of PMSs. Therefore, researchers have attempted to develop systematic methods to assess PMS construction (Najmi et al., 2005; Tangen, 2004; Bititci et al., 1997). The methodologies to assess the PMSs are typically based on a logical order of components and actions of PMSs to measure and improve their performance.

Regarding the fundamentals of PMSs, a consensus is emerging on what PMSs should be, what constant and permanent it should contain. Lewandowski and Cirella (2022) list five necessary fundamental consensual components constituting PMSs: 1) goals reflecting organizational expectation and delineate performance; 2) measurement comprising of the metrics used to operationalize performance; 3) a review referring to the evaluation and feedback of performance information; 4) a performance-related reward system, and 5) update of PMS. The same related components are presented by Franco-Santos et al. (2012), Olsen et al. (2007), Speckbacher et al. (2003) and Wettstein & Kueng (2002).

Franco-Santos et al. (2012) presented a four-type typology in which PMSs are defined by the components and by the consequences of PMS. In this typology, all four types include financial as well as non-financial performance measures as components implicitly or explicitly linked to strategy and used to inform managerial decision-making and to evaluate organizational performance. In addition, in the second type, PMS is showing explicit cause-and-effect relationships among the measures as a component. In the third type, PMSs are not linked with the performance evaluation results to monetary rewards. In the fourth type, PMS are linked to influences on monetary rewards. However, Olsen et al. (2007) suggested three generic criteria to assess PMS effectiveness: causality, continuous improvement, and process control.

Speckbacher et al. (2003) introduced a typology that covers BSCs. Their first type contains financial and non-financial strategic performance measures grouped into perspectives. The second type additionally employs a specific approach to describe the organization's strategy using a sequential cause-and-effect logic to link tangible and intangible assets. The third type includes an additional feature, which makes the incentive pay contingent on the performance results of the scorecard measures. Wettstein & Kueng (2002) assessed PMS success according to six different dimensions following this order: scope of measurement, data collection, storage of data, communication of performance results, use of performance measures, and quality of performance measurement processes. These studies point the way to how PMS should be assessed based on a consensual methodology.

However, there is a paucity of research that examines the success of PMSs created in the previous ways in a holistic and not fragmented manner, and under what conditions PMSs created in this way are successful. Kadak & Laitinen (2016) responded to this challenge and developed a more comprehensive approach based on the chain theory of PMS success. Success of a PMS is difficult to assess, since there are many success factors which are located at different hierarchical levels of an organization. Therefore, the essential point of their theory was to find a logical chain of key factors (KF) that in spite of different hierarchical levels makes PMS successful for an organization. Their model includes all five necessary fundamental consensual components by Lewandowski and Cirella (2022) which are measured by the KFs using information of fifteen check points (CP). In addition, their model covers life cycle of PMS from its design through implementation and use (Bourne et al. 2000).

Kadak & Laitinen (2016) tested empirically the chain theory using a small sample of Estonian and Finnish firms. They showed that the sum information from the CP assessments ("strength of the chain") can be an efficient predictor of PMS success. Thus, an important characteristic for the success is the completeness of the logical chain in PMS. The efficient prediction of success is however difficult, since there are many success factors in PMS reflected by fifteen CPs which are strongly dependent on each other. This kind of situation is characterized by strong multicollinearity which makes it difficult to assess the relative importance of CPs. The sum information of CPs makes use of the hidden assumption that these CPs are equally important leading to equal weights for each CP (unity weight structure).

This study is based on a recent sample of Estonian firms and on the use of different statistical methods in this sample to estimate an efficient weight structure for CPs. The purpose of this study is to assess the usability of the check point (CP) information of the chain theory (Kadak & Laitinen, 2016) to predict the success of PMSs. For that two research questions (RQs) were set:

RQ 1: Is the relation of CPs and PMS success stable over time so that the older sample (2015) and the newer sample (2022) have similar characteristics?

RQ 2: Is the sum of CPs still an efficient predictor of PMS success? Is it possible to find out a more efficient predictor using statistical methods?

The content of the paper follows the research trajectory and is following. First, the motivation of the research and research questions are discussed in the introduction. In this section resulted by complexity of PMSs, being influenced by many contingency factors; the PMSs that have been created offer benefits but also have many unintended consequences; there is much contradicting and mixed information about when PMSs are or are not beneficial. This creates a need to assess the success of PMSs. For assessment is important to determine what PMSs consist of. Consensual components constituting PMSs are presented and the need to use a holistic and not fragmented PMS for assessing the success of PMS has been derived. In the second section, the chain rule theory of PMS and components of PMS is briefly presented. The empirical survey data and statistical methods are discussed in the third section. The recent sample is consisted of survey information from 73 Estonian firms. This information is analyzed by a set of statistical methods to find the most efficient way to extract the weight structure of CPs. The fourth section presents the empirical results. Five different statistical methods are applied in this assessment: equal weights (sum of variables), linear regression, logistic regression, ridge regression, and a Solver numeric solution. In addition, stepwise analyses of linear and logistic analyses are performed. The conclusions are presented in the final section. ~~The results of the analysis can be used in constructing an efficient predictor of PMS success.~~

2. The chain theory of PMS success

Kadak & Laitinen (2016) discuss the chain theory of PMS success in detail. Therefore, in this paper the chain theory is discussed only briefly. The purpose of the theory is to create a comprehensive and logical chain of KFs which ensures the success of PMS. In all, the chain theory is based on fourteen KFs reflected by fifteen CPs. These CPd are divided between three process stages: the performance measures, which are aligned with the organization's strategy. The second phase is implementing the measures by putting appropriate systems and procedures in place to collect and process data that allow measurements to be made. The third phase is ensuring that the measures are used as part of decision-making, while challenging the validity of measures on a regular basis. (Okwir et al., 2018; Bititci et al., 2012; Folan & Browne, 2005; Mason-Jones & Towill, 2000; Neely, 1999). The fourteen KFs are the following:

Structure of PMS

1. Clearly expressed mission/vision statement. The mission of the organization should be clear and based on the preferences of shareholders. Mission statement generally refers to the reason why organization exists (Simons, 2000). This mission statement creates a vision referring to a picture of the organization in the future.

2. Clearly expressed strategy, consistent with the mission but also with contextual factors. The strategy of the organization should be clearly expressed and aligned with the mission and contextual factors. Strategy briefly describes how an organization intends to fulfil its mission based on its potential, operating environment and considering the stakeholders interests (Kadak, 2011).

3. PMS is based on identified causal relationships between critical success factors (CSF) aligned with the strategy map or/and objectives aligned with strategy (Speckbacher et al., 2003). PMS should consist of a logical set of CSFs based on the causal relationships in the strategy map. This kind of strategy map ensures that PMS will be connected with the strategy of an organization due to the use of causal relations which exist between different components of strategy and which may be set into different areas of CSFs (Waal, 2007).

4. Clearly stated corporate organization-level objectives (CO), based on strategy and congruent with the CSFs in the strategy map. The organization should clearly express corporate-level objectives which are derived from the strategy and are congruent with the CSFs in the strategy map. These objectives make the strategy more specific and concrete for implementation (Kadak, 2011). The firm's chosen strategy has to be consistent with the chosen internal objectives and value drivers (Ittner & Larcker, 2001).

5. Organization-level performance measures that form a comprehensive set of performance measures (PMs or KPIs), measuring the attainment level of COs with the help of target values, are mutually congruent, and aligned with COs. The organization should have a comprehensive set of PMs (that covers all CSFs), measuring the attainment levels of COs in relation to target values, and forming a mutually congruent set aligned with COs. Since each CSF can be critical, each of them should be accompanied by PMs making PMS comprehensive (Laitinen, 2004).

6. Delegation of organization-level COs as goals to every hierarchical level of the organization so that congruence of the goals between and within every level is attained. Corporate objectives (COs) should be delegated as goals to every hierarchical level of the organization so that congruence of these goals between and within every hierarchical level is attained. In this way it is made sure that everyone in an organization works on the same direction to achieve the organizational objective (Waal, 2007).

7. A comprehensive set of PMs at each organizational level measuring the attainment level of goals, being congruent, and aligned with goals. The organization should have a comprehensive set of performance measures, which measure the attainment level of goals with the help of target values in every

hierarchical level of an organization. There must be congruence of the performance measures so that they are aligned with goals in every hierarchical level of an organization. These kinds of PMs are thus used at different levels in an organization to evaluate success in achieving goals, and thus satisfying the expectations of different stakeholders (Ferreira & Otley, 2009).

8. Identification of key processes (series of activities) that are critical to attain COs, delegation of these processes to attain goals in each hierarchical level of organization, and measuring performance of each process for input, process, and output. The organization should identify key processes to attain corporate objectives and delegate these processes to attain goals in every hierarchical level of an organization. In addition, the organization should have performance measures for each key process in every hierarchical level of an organization for input, process, and output control. The key processes should be derived from the measures of the goal and be critical for the achievement of the goal associated with that CSF and measure (Kadak, 2011).

User commitment to PMS

9. PMS is designed and implemented interactively with the users of PMS. The designing and implementation of PMS should be carried out interactively with the users of PMS. For the successful building and implementation of PMS it is necessary to have the involvement and commitment of top-management and members of organization from the beginning. Performance management and implementation of PMS is very important, and thus this must be top priority (Waal, 2007).

10. PMS is used both interactively and diagnostically in the way that there is created a dynamic tension. In order to be successful (to improve performance), PMS should be used both interactively and diagnostically to create dynamic tensions for the organization. The organization need to create an appropriate dynamic tension that is likely to stimulate the right mix between compliant behavior and creative search efforts necessary for organizational success (Simons, 2000).

11. PMS is intensively used by employees in charge. PMS should be intensively used by the staff in charge in order to be successful in improving the performance of the organization. This characteristic also considers PMS as an information system and refers partially to the concept of system use that is an important determinant of system success (DeLone & McLean, 2003). Intensive use of PMS by senior and operating managers must exist to ensure efficient interactive control (Bisbe et al., 2007).

Incentive system

12. PMS is associated with an incentive system. The organization should have an incentive system connected with PMS that motivates employees to perform actions consistent with the goals allocated to them (Franco-Santos et al., 2012). If PMS is not connected with an appropriate incentive system, it is not motivational to drive employees towards improved performance. Kerpershoek et al. (2016) and Jensen (2003) state, the link between performance measurement and rewards has proven to be a critical choice that can lead to negative unintended consequences. PMSs allow the design of goal-congruent incentive systems. Reward and compensation systems are designed to align individual interests with those of the organization (Simons, 1995).

Quality of information

13. PMS produces valid and reliable information to users in time and in a useful form. The value of PMS is in its ability to provide the users of the systems with valid, reliable, timely, and usable information for their work. These kinds of characteristics consider PMS as an information system and focus on information quality. Information quality can include a large number of attributes such as relevance,

understandability, accuracy, conciseness, completeness, understandability, currency, timeliness, and usability (DeLone & McLean, 2003).

Continuous updating

14. PMS is continuously updated. PMS should be continuously updated for changes in the organization and its surrounding environment. If PMS is based on an out-of-date model of organization and its environment, its value for staff in improving performance can be questioned. Environments change, organizations change, and so PMSs also need to change in order to sustain their relevance and usefulness (Ferreira & Otley, 2009).

3. Data and methods of the study

Data and questions

For the research, survey data was collected by means of a questionnaire sent to the companies asking about the background information of the companies, as well as information about the success of the PMS and about the fulfilment of requirements set by the fifteen CPs. The survey was sent to randomly selected Estonian companies, of which 73 companies answered the survey. The companies that responded to the survey are generally medium-sized, as their average number of employees is 262.7 employees. However, the size distribution of companies is typically skewed and the median number of employees is only 68.0 employees. The average turnover of the companies is 757984.5 thousand euros and the median being 16129.0 thousand euros. The average balance sheet total in the sample companies is 522,609.0 thousand euros, while the median is only 8,653.0 thousand euros. The profitability of the responding companies is satisfactory on average, as the average return on capital percentage is 11.54 the median being 7.56. Of the responding companies, 26 (35.6%) are from manufacturing industry, 23 (31.5%) from service industry, 17 (23.3%) from trade industry, 4 (5.5%) from construction industry and 1 (1.4%) from public sector. The responding companies form 38% of all companies with over 250 employees in Estonia, which allows it to be considered representative. A sample bias test was realized during follow-up calls. They showed only variances of little relevance in PMSs between the firms who responded and who did not respond. However, we think that the sample size is quite small which arises a call for future studies with larger groups of respondents.

The original questionnaire includes a number of questions about the background, performance, success of PMS, and the existence and quality of CPs (on KFs) according to the chain theory. In this paper, only the questions used to answer to the research questions are introduced. The success of PMS was measured through the self-assessed impact of PMS on organizational performance. The impact is reflected by three different measures on a Likert scale from 1 to 7 (see Appendix 1). The first measure depicts the achievement rate of company's strategic objectives (STRPER). This is important since the essence of PMS is that it supports management to achieve strategic objectives. The second variable measures the impact of PMS on non-financial performance of the company during the last three years (NFIPER). For the validity of the measure, it is important to choose a longer-run impact. In the same way, the third measure reflect the impact of PMS on financial performance during the same period (FINPER).

Each measure is self-assessed by the manager responded to the questionnaire. The Cronbach alpha between these three measures is 0.790 reflecting respectable or very good internal consistency (reliability). The Tukey's test for non-additivity is rejected on p-value 0.008. Thus, the scores of the success measures were summed up and used to classify observations into two classes with respect to the overall impact of PMS. The sum of the three measures with a score higher or equal to 15 (median) are classified as more successful PMSs (39) whereas the rest of PMSs (score less than 15) are classified as less successful PMSs (34). The score 15 refers on average to a response of 5.00 for the three measures that exceeds the average of the 1-7 Likert scales (4.00).

The fifteen CPs of the fourteen KFs described in the theoretical framework are in the same way measured by the questions presented in Appendix 2. In total, the pattern of questions includes fifteen questions associated with the fourteen KFs. Exceptionally, KF number 10 is reflected by two different questions referring to the diagnostic and interactive uses of PMS separately. The questions have been prepared to be as easy as possible to answer. All fifteen questions are measured on a Likert scale from 1 (do not agree at all) to 7 (agree perfectly). For presenting all CP variables into range from 0 to 1, we standardized the Likert scale by deducting unity from the value of the variable and dividing the difference by 6 as (value-1)/6.

Statistical methods

This study uses several different statistical methods to explain the success of PMS using 15 CP variables. The first and simplest method is to calculate the sum of the CP variables and use it to predict the success of the PMS (Kadak & Laitinen, 2016). The main difficulty in explaining the success of PMS with the help of CP variables is that, according to the chain theory, there are relatively many CP variables and that they are strongly correlated not only to the predictable variable but also to each other, in which case the data is strongly multicollinear. This would cause even more difficulties if the goal was to estimate the most reliable coefficients for the CP variables, which are usually biased in strongly multicollinear data and often also have the wrong sign. However, this study is concerned with predicting the success of PMS, in which case the ability to predict is more important than individual coefficients. The Cronbach alpha between the fifteen CP measures is 0.949 reflecting very good internal consistency. The Tukey's test for non-additivity is rejected on p-value <0.001. Thus, the CP variables were summed up to form a sum variable to explain the success of PMS in a consistent way.

The use of the sum variable of CPs as a predictor of PMS success means that each of the fifteen CP variables has a weight of one (unity) in the prediction. The use of a sum is a consistent with the chain theory, but by using a more efficient weight structure, it is perhaps possible to improve the quality of prediction. The usual way to estimate the weight structure for variables is to use linear regression analysis, which is the second statistical method applied in this work. In this context too, the attention is not drawn to the fact that the coefficients of the CP variables are reliable, but to the fact that the estimated function works in prediction. Linear regression analysis is a statistical analysis method in which the linear dependence of the considered response variable (PMS success) on the explanatory variables (fifteen CP variables) is estimated using here the ordinary least squares (OLS) method. In this case, the equation to be estimated passes through the origin, in which case only the weight structure is estimated, and an intercept does not appear in the equation:

$$Y = b_1X_1 + b_2X_2 + \cdots + b_{15}X_{15} + \varepsilon \quad (1)$$

where Y is the dependent variable (PMS success or SUMPERF), $X_1 \ldots X_{15}$ are independent variables (CP variables), $b_1 \ldots b_{15}$ are the regression coefficients (weight structure), and ε is the random residual. This method originally assumes that the residuals are independent, have a constant variance, have a conditional mean zero, and are normally distributed.

The coefficients of the CP variables can also be estimated using the sample classified as more successful and less successful PMSs. These coefficients can be estimated by the multiple discriminant analysis (MDA) that is however based on a set of restrictive assumptions. Therefore, as the third statistical method, the logistic regression analysis (LRA) is applied to estimate the weight structure. LRA does not require that independent variables are multivariate normal or that groups have equal covariance matrices which are basic assumptions in MDA (Hosmer & Lemeshow, 1989). In LRA, the dependent variable Y is binary so that $Y = 0$ when PMS is less successful and $Y = 1$ when PMS is considered more successful. LRA creates a score (logit) L for every observation. It is assumed that the independent variables are linearly related to L. This score or logit is used to determine the conditional probability of success P as follows:

$$P(Y = 1|X) = \frac{1}{1+e^{-L}} = \frac{1}{1+e^{-(b_1X_1+...+b_{15}X_{15})}} \qquad (2)$$

where Y is the binary dependent variable (0 or 1), $X_1 ... X_{15}$ are the fifteen independent CP variables, and $b_1 ... b_{15}$ are the coefficients. In Equation (2) the linear logit L is presented by means of the weight structure without the intercept.

The fifteen CP variables are strongly positively correlated which makes the data multicollinear. Multicollinearity can make some estimates of weight negative which, however, does not destroy the prediction ability of the model but is against the intuition. Therefore, the fourth statistical method applied in this study is the Ridge regression (RR) that is a method estimating the coefficients of the independent variables when they are highly correlated. When the independent variables are strongly correlated, the coefficients are slippery even to small changes in variables. In this kind of situation, the sum of squared errors (SSE) in RR can be lower than in OLS but the estimates are biased. In RR, bias is artificially incorporated to the regression to make the variance of the estimates lower. SSE can be presented in the following way:

$$SSE = (Y - XB)^T(Y - XB) + \lambda B^T B \qquad (3)$$

where lambda (λ) is a non-negative ridge parameter reflecting the bias. RR estimates are biased as lambda increases but may give more precise estimates than OLS. If lambda = 0, RR returns to OLS. If lambda increases towards infinity, all the regression coefficients shrink towards zero. There are several techniques how to select the lambda parameter (Lukman & Ayinde, 2016; Duzan & Shariff, 2015). In this study, we simply solve the standardized coefficients (without an intercept) increasing lambda until finally none of the estimated coefficients (weights of CP variables) is negative (λ=463).

Finally, the coefficients of the CP variables are estimated using a Solver model (Excel). When the number of positively correlated independent variables is high, the set of variables can include redundant information. In this kind of situation, a smaller set of independent variables (CP variables) can lead to improved results. Therefore, the last statistical methods applied in this study is a Solver model where the weight structure is solved searching for the best non-negative coefficients for the CP variables. This model makes it possible that in the optimal solution the coefficients of one or more CP variables are equal to zero. Thus, the model searches for the best set of coefficients which are either zero or positive. The objective of the Solver model is set to maximize the coefficient of determination between the fifteen independent variables and the sum variable of PMS success for non-negative coefficients. The solution for the optimization is searched using the Generalized Reduced Gradient (GRG) method (Lasdon et al., 1973). This method looks at the gradient or slope of the objective function as the input values change and determines that it has reached an optimum solution when the partial derivatives equal zero. GRG can handle equality constraints and also inequality constraints which are converted to equalities by the use of slack variables. GRG uses a combination of the gradient of the objective function and a pseudo-gradient derived from the equality constraints. It is an iterative method using a search procedure where the search direction is found in the way that any active constraint remains precisely active for some small move in this direction.

Thus, the coefficients of the CP variables are solved in this study using five different statistical methods which are the following: unity vector (sum of CP variables), linear regression (OLS), logistic regression (LR), ridge regression (RR), and Solver model. Moreover, the linear and logistic models are simplified using a stepwise procedure, where the most significant CP variables are one after the other entered into the model, until a pre-specified stopping rule is reached. The purpose is to compare the ability of the five methods in predicting PMS success by means of the information from the fifteen CP variables. For practical purposes, the ability is firstly assessed comparing the correlation between the prediction and the actual value of the sum variable of PMS success. However, the ability is mainly assessed by the receiver operating characteristic curve (ROC) that shows graphically the performance of a classification model for all classification thresholds. The ROC curve plots parameters true positive rate and false positive rate. The overall classification performance of the methods is numerically compared on the basis of the area under the ROC curve (AUC). AUC can get values from 0 (all classifications are incorrect) to 1 (all classifications are correct). It measures the quality of the predictions irrespective of what classification threshold is chosen. In addition, this comparison is made using the accuracy ratio (AR) that is defined as AR = 2 AUC-1. This measure AR is constructed in the way that the value 0.5 refers to the accuracy of an average classification model. This value corresponds to the value of AUC = 0.75.

4. Empirical results

Descriptive statistics from 2015 and 2022

Table 1 compares the distributions of the three PMS success variables STRPER, NFIPER, and FINPER in the older sample from 2015 (Panel 1) and in the current sample from 2022 (Panel 2). In the older sample, where the chain theory was originally developed, companies are divided into groups such that in the group "more successful PMSs" are companies whose PMS success sum variable SUMPERF exceeds the value 14 (30 firms), as other companies (32 firms) are in the lower group “less successful PMSs”. The values of the impact variables are higher in the new sample, whereby firms whose sum of the three variables exceeds 15 (39 firms) are now located in the group "more successful PMSs", while the rest of the firms (34 firms) are in the lower group. In the newer sample, the median of all three impact variables in the group "more successful PMSs" is 6 referring to a very high success. In the new sample, the values of the variable FINPER (impact on financial performance) in the group "more successful PMSs” are clearly higher than in the old sample but in the group "less successful PMSs" there has been less improvement. In both samples, the values of all three success variables in the sum variable are approximately at the same level, and thus none of them dominates the sum variable.

(Table 1 here)

Table 2 shows the distributions of the fifteen CP variables in the older (Panel 1) and newer (Panel 2) data. The results show that although the values of the success variables have clearly increased, the values of the CP variables have generally decreased in both the "more successful PMSs" and "less successful PMSs" groups. If the change in value is measured by the decrease in the average, the values have decreased the most in both groups in the variables ORGKPI and UNKPI. The F-test shows that in the new sample the differences between the groups are clearly larger than in the older sample. In the new sample, the group differences of all fifteen CP variables are statistically significant, while in the old sample the differences were not as significant. In this respect, the samples are not completely comparable, as the limit based on the sum variable for the groups "more successful PMSs" and "less successful PMS" has risen. In any case, the comparison shows that despite the fact that the values of the CP variables have decreased, the values of the impact (success) variables have increased. As the correlations between CP variables and the PMS success variables show that there is a strong dependence between them in the same way as in the older sample, this implicates that the chain theory works in practice. The dependence between the variables thus remains unchanged, so in this sense the relationship is stable over time, answering so to the first research question (RQ 1).

(Table 2 here)

Appendix 3 shows the mutual Pearson and Spearman (rank) correlation coefficients of all fifteen CP variables in the new sample (Panel 1 and 2). All correlation coefficients are high and statistically highly significant. This means that when CP variables are used as explanatory variables in the model, the data will be strongly multicollinear. At the same time, the results support the chain theory, because the CP variables form a parallel set of variables that are related to each other. High positive correlations can also mean that the CP variables have significantly overlapping information, in which case they do not all clearly produce incremental information for prediction over one another. The last rows of the tables also show the correlation of SUMPERF with the CP variables. These correlations are positive and high showing that there is a strong positive relationship. The variable SUMPERF has particularly high Pearson and Spearman correlation with the CP variables IMPUSE (0.542 and 0.516), RESP (0.561 and 0.563), QUALINF (0.507 and 0.535) and ADJ (0.596 and 0.616). Thus, these variables can be important predictors and are able to produce alone a rough prediction of PMS success.

The high correlations between the independent CP variables imply that there are not many distinct hidden dimensions behind the variables which potentially have an impact on PMS success. The hidden dimensions (factors) in the data were assessed by the factor analysis based on Varimax rotation which makes the hidden dimensions linearly independent. The scree plot of the factor analysis showed that there behind the fifteen CP variables are only two hidden dimensions which explain respectively 62.6% and 9.2% of the variance of the fifteen CP variables. The highest loadings on the first factor have got by QUALINF and ACTUSE whereas STR, MIS, and STROBJ got the highest loadings on the second factor. Thus, the results of the factor analysis imply that there are only two hidden dimensions so that only a small part of the CP variables can be simultaneously effective in the estimation of the weight structure. This factually means that a very simple model with a couple of CP variables can be as efficient as a more complicated model with a larger set of variables. Therefore, in addition to the full models, it is useful to test the prediction ability of the stepwise models leading to a smaller set of predictors.

Statistical results

Table 3 shows the coefficients (weight structure) of the fifteen CP variables estimated with different statistical methods. In the sum variable of CP variables, the weight of all variables is one. The weight structure estimated by linear regression includes six negative coefficients, which is against the intuition and is a result of multicollinearity. MIS, CSF, STROBJ, UNOBJ, UNKPI and KPR have negative coefficients although their correlation to SUMPERF is strongly positive. At the p-level of 5%, only STR, STROBJ, ORGKPI, and ACTUSE have statistically significant coefficients. The coefficient of determination of the regression is relatively high (0.929), and the regression is very significant (p-value = 0.000). In this linear regression solution ORGKPI, STR and ACTUSE have received the greatest weights. In terms of absolute values, the variables STROBJ and UNKPI have the largest negative coefficients. Table 4 presents the results of the linear stepwise model. The forward stepwise procedure leads to a simplified model structure with only three CP variables which all have a positive coefficient. These CP variables are ORGKPI (7.536), ACTUSE (8.222) and QUALINF (6.282). The coefficient of determination for the linear regression is 0.905, that is not much lower than for the full model of fifteen independent variables. The VIF values especially in the entire model refer to a very strong multicollinearity. For the stepwise model the values of VIF are lower but may still refer to multicollinearity.

(Table 3 here)

(Table 4 here)

Table 3 shows that the solution produced by the logistic regression analysis (logit) has eight CP variables with a negative coefficient (MIS, CSF, ORGKPI, UNOBJ, UNKPI, ACTUSE, DGNUSE, and INSYS). At the p-level of 5%, only the variables IMPUSE, DGNUSE, and RESP have received statistically significant coefficients. The variables IMPUSE, RESP, and, ADJ have the highest positive coefficients. In absolute terms, the CP variables DGNUSE and ORGKPI have the largest negative coefficients. The logistic regression is statistically very significant, and the Cox & Snell R square is quite high (0.356) as is Nagelkerke R Square (0.475).

Table 4 also shows that the stepwise logistic model has only three CP variables (IMPUSE, DGNUSE and ADJ). From the three variables, DGNUSE has the most significant coefficient, but the coefficient is negative. The Cox & Snell R square for the stepwise model is 0.236 whereas the Nagelkerke R Square is 0.316. Table 3 presents the positive coefficients estimated by the Ridge regression. In this regression, the lambda was set equal to 463 which was just enough to make all the coefficients positive. The highest coefficients have got by the CP variables ADJ, RESP, and IMPUSE. For the original goal set VIF=1, lambda was solved as 15.399, which led to RR with the coefficient of determination 0.625 and negative coefficients for nine variables. Table 3 also shows the Solver solution that for several sets of initial values produced positive (non-zero) coefficients only for three CP variables (ADJ, RESP, and IMPUSE). Thus, it shows that there are only a few CP variables which produce additional information about PMS success more than other CP variables.

Table 5 presents the Pearson (Panel 1) and Spearman (Panel 2) correlations between SUMPERF and the predictions given by statistical methods. All the correlations in both panels of the table are statistically very significant. The highest correlations to SUMPERF are found for logistic regression logit followed by Solver solution and linear regression. The differences in the correlations between the methods are relatively small. The table also includes the results for the logistic regression probability (logistic transformation) which are quite similar as for the logistic logit. The Spearman rank correlations are identical for the logit and the probability due to the type of transformation. The correlations of the stepwise models are lower that the full models of fifteen CP variables but still higher than for the sum of CPs. The Solver solution based only on three CP variables has got a clearly higher correlation with the sum variable than the stepwise models.

(Table 5 here)

Table 6 presents descriptive statistics for the predictors of SUMPERF in the categories of "more successful PMSs" and "less successful PMSs". The measures produced by the predictors are different and are not directly comparable. The highest F-test has got by logistic regression probability followed by the Solver solution. Table 7 presents the AUC and AR values for the estimation methods. Each predictor has got a higher value of AR than 0.5 showing s better classification performance than an average model. The highest AR ratio has got by the logistic regression (logit and probability give the identical result) followed by Solver solution, Ridge regression, and linear regression. In summary, the sum of CPs will give only a rough estimate of PMS success. The simple Solver model only based on three CP variables seems to be an efficient predictor and clearly beats the stepwise models. Figures 1-7 show the ROC curves for the statistical methods. Table 8 shows the misclassification rates for the different methods. Linear regression, logistic regression, and the Solver solution lead to the lowest rates. The differences in the rates between the prediction methods are not very significant. In summary, the empirical results showed that other statistical methods worked more effectively than the sum of CP variables in predicting the success of PMS, answering so to the second research question (RQ 2).

(Table 6 here)

(Table 7 here)

(Figures 1-7 here)

(Table 8 here)

5. Conclusions

The chain theory of PMS success developed by Kadak & Laitinen (2016) is currently the only comprehensive method for ensuring and predicting the success of PMS. This method is in line with the generally agreed upon components that a PMS should contain. The chain theory is based on a total of fourteen KFs, which are measured using 15 CP variables. Kadak & Laitinen (2016) showed in their study that securing all fifteen CPs is strongly correlated with the success of PMS. For this reason, the sum of CP variables works well in predicting the success of PMS. The results published by Kadak & Laitinen in 2016 were however based on a small sample of both Estonian and Finnish companies. Therefore, they emphasized the preliminary nature of the results.

Thus, the first research question (RQ 1) of this current study was whether there is a similar dependence between CP variables and PMS success in the newer (2022) data as well. If this relationship is further valid, it can be assumed that the chain theory works over time and generally in other samples as well. This study showed that the level of CP variables, measured with the scale used and self-assessment, has decreased, but despite this, the impact of PMS on organizational success, measured with the help of self-assessment, has increased. This means that the level of the relationship between CP variables and PMS success is not completely stable, but the impacts of CP variables on PMS success have increased during the eight years. However, the correlations between CP variables and PMS success show that there is a strong dependence between them in the same way as in the older sample, which shows that the chain theory works in practice. The dependence between the variables thus remains unchanged, so that in this sense the relationship is stable over time.

The potential validity of the chain theory implies that the success of PMS depends on all CP variables which is clearly supported by high correlations between these variables. Since all CP variables thus change into the same direction with respect to PMS success, it follows that there is strong multicollinearity between CP variables. This multicollinearity makes it difficult to assess the contribution of individual variables. Since the CP variables are strongly correlated with each other, the sum of the CP variables can alone reasonably predict the success of the PMS. However, the second research question of this study (RQ 2) was whether the importance of individual CP variables can be evaluated and the prediction that the sum variable gives can be improved by estimating the weight structure for the CP variables using statistical methods.

In this study, five different methods were applied to estimate the weight structure, and, in addition, stepwise analysis was used for two of the statistical methods. Since the CP variables are strongly correlated with each other, it is possible that a reliable prediction can be obtained with a model constructed only by a few variables, and additional variables do not bring incremental information to the analysis. The results showed that other methods worked at least slightly more effectively than the sum of CP variables in predicting the success of PMS. Logistic regression analysis and an optimization model based on a Solver solution worked best for predicting. Logistic regression analysis, however, produced due to multicollinearity negative weights for CP variables, which do not reflect their separate importance. The Solver model gave positive weights to only three CP variables, but nevertheless it works well in prediction. The model is based on the variables IMPUSE, RESP, and ADJ. Thus, the model emphasizes that it is primarily essential for the success of a PMS that users participate in the design and construction of the PMS (IMPUSE), use it intensively (RESP), and update it systematically (ADJ).
The results have given an implication that models can be practically used to design and to predict the success of PMS and can be used in the design of new PMSs. The findings about some few CP variables can produce additional information about PMS success more than other CPs. Thus, the findings about high multicollinearity between CPs, direct us to investigate more this phenomenon in the future, such as to clarify the reasons for multicollinearity and to investigate possibilities to construct a multi-layered chain model. The multicollinearity refers on the overlapping and it opens new direction to investigate size of the buffer inside PMS. According to each CP, how much does each CP contribute directly to the success of the PMS and how much does he contribute to the so-called safety buffer in case the CP next to him does not perform his tasks? In addition, if now to rely on the initial tool for assessing and predicting the success of PMSs, further research could focus on cases where low PMS success has been initially detected and, after improvements, examine changes in the success of PMSs as a whole and by component.

This study is limited by a small sample size. In the future, larger samples should be used to confirm the findings. Therefore, the results of this study should be considered still preliminary.

References

Bedford, D., Brown, D.A., Malmi, T. and Sivabalan, P. (2008). Balanced Scorecard Design and Performance Impacts: Some Australian Evidence, Journal of Applied Management Accounting Research, Summer, pp. 17–36.

Behery, M., Jabeen, F. and Parakandi, M. (2014). Adopting a contemporary performance management system: A fast-growth small-to-medium enterprise (FGSME) in the UAE, International Journal of Productivity and Performance Management, 63(1), pp. 22-43. https://doi.org/10.1108/IJPPM-07-2012-0076

Bento, R.F., Mertins, L. and White, L.F. (2017). Ideology and the balanced scorecard: An empirical exploration of the tension between shareholder value maximization and corporate social responsibility, Journal of Business Ethics, 142(4), pp. 769-789. https://doi:10.1007/s10551-016-3053-6

Bisbe, J., Batista-Foguet, J.-M. and Chenhall, R. (2007). Defining management accounting constructs: a methodological note on the risks of conceptual misspecification, Accounting, Organizations and Society 32, pp. 789–820.

Bititci, U.S., Garengo, P., Dörfler, V. and Nudurupati, S. (2012). Performance measurement: challenges for tomorrow, International Journal of Management Reviews, 14, pp. 305–327.

Bititci, U.S., Carrie, A.S. and McDevitt, L. (1997). Integrated performance measurement systems: a development guide, International Journal of Operations and Production Management, 17(5/6), pp. 522–534.

Bonner, S.E. and Sprinkle, G.B. (2002). The effects of monetary incentives on effort and task performance: theories, evidence, and a framework for research, Accounting, Organizations and Society, 27(4–5), pp. 303–345.

Bourne, M.C.S. (2008). Performance measurement: learning from the past and projecting the future, 12(4), pp. 67-72.

Bourne, M.C.S., Melnyk. S. and Faull, N. (2007). The impact of performance measurement on performance, International Journal Operation and Production Management, 27(8).

Bourne, M.C.S., Mills, J.F., Wilcox M., Neely A.D. and Platts, K.W. (2000). Designing, implementing and updating performance measurement systems, International Journal of Production and Operations Management, 20(7), pp. 754–771.

Bourne, M.C.S., Mills, J.F., Bicheno, J., Hamblin, D.J., Wilcox M., Neely, A.D. and Platts, K.W. (1999). Performance measurement system design: testing a process approach in manufacturing companies, International Journal of Business Performance Management, 1(2), pp. 154-0.

Broadbent, J. and Laughlin, R. (2009). Performance management systems: A conceptual model, Management Accounting Research, 20(4), pp. 283–295.

Choong, K.K. (2014). The fundamentals of performance measurement systems: A systematic approach to theory and a research agenda", International Journal of Productivity and Performance Management, 63(7), pp. 879-922.

Couturier, J. and Sklavounos, N. (2019). Performance dialogue: A framework to enhance the effectiveness of performance measurement systems, International Journal of Productivity and Performance Management, 68(4), pp. 699-720. https:// doi:10.1108/IJPPM-09-2017-0238

Crabtree, A. D. and DeBusk, G. K. (2008). The effects of adopting the balanced scorecard on shareholder returns, Advances in Accounting, 24(1), pp. 8–15.

Cuccurullo, C., Aria, M. and Sarto, F. (2016). Foundations and trends in performance management. A twenty-five years bibliometric analysis in business and public administration domains, Scientometrics, 108, 595–611. https://doi.org/10.1007/s11192-016-1948-8

Cuguero-Escofet, N. and Rosanas, J.M. (2016). The ethics of metrics: overcoming the dysfunctional effects of performance measurements through justice, Journal of Business Ethics, 140, pp. 615–631.

Davis, S. and Albright, T. (2004). An investigation of the effect of balanced scorecard implementation on financial performance, Management Accounting Research, 15(2), pp. 135-153. https://doi:10.1016/j.mar.2003.11.001

DeLone, W.H. and McLean, E.R. (2003). The DeLone and McLean Model of Information Systems Success: A Ten-Year Update, Journal of Management Information Systems, 19(4), pp. 9–30.

Duzan, H. and Shariff, N.S.B.M. (2015). Ridge Regression for solving the multicollinearity problem: Review of methods and models, Journal of Applied Sciences, 15, pp. 392-404. https://doi:10.3923/jas.2015.392.404

Ferreira, A. and Otley, D. (2009). The Design and Use of Management Control Systems: An Extended Framework for Analysis, Management Accounting Research, 20, pp. 263–282.

Flamholtz, E.G., Das, T.K. and Tsui, A.S. (1985). Toward an integrative framework of organizational Control, Accounting, Organizations and Society, 10(1), pp. 35-50.

Folan, P. and Browne, J. (2005). A review of performance measurement: towards performance management, Computers in Industry, 56, pp. 663–680.

Franco-Santos, M., Lucianetti, L. and Bourne, M. (2012). Contemporary performance measurement systems: A review of their consequences and a framework for research, Management Accounting Research, 23, pp. 79–119.

Govindarajan, V. and Shank, J. K. (1992). Strategic cost management: Tailoring controls to strategies, Journal of Cost Management, 6(3), 14–25.

HassabElnaby, H. R., Said, A. A. and Wier, B. (2005). The retention of nonfinancial performance measures in compensation contracts, Journal of Management Accounting Research, 17, pp. 23–42.

Hosmer, D. and Lemeshow, S. (1989). Applied Logistic Regression, Wiley Interscience, Prentice Hall International, New York.

Hoque, Z. (2004). A contingency model of the association between strategy, environmental uncertainty and performance measurement: impact on organizational performance, International Business Review, 13, pp. 485–502.

Hoque, Z. and James, W. (2000). Linking balanced scorecard measures to size and market factors: impact Measuring Business Excellence on organizational performance, Journal of Management Accounting Research, 12, pp. 1–17.

Hope, T. and Hope, J. (1995). Transforming the bottom line. Nicholas Brealey Publishing.

Ittner, C. D. and Larcker, D. F. (2001). Assessing empirical research in managerial accounting: a valuebased management perspective, Journal of accounting and economics, 32(1), pp. 349-410.

Jazayeri, M. and Scapens, R.W. (2008). The business value scorecard within BAE systems: The evolution of a performance measurement system, British Accounting Review, 40(1), pp. 48-70.

Jensen, M.C. (2003). Paying people to lie: the truth about the budgeting process, European Financial Management, 9, pp. 379–406.

Kadak, T. and Laitinen, E. K. (2016). What matters with PMS? Critical check points in the success of PMS, Studies in Managerial and Financial Accounting, Performance Measurement and Management Control: Contemporary Issues, 31, pp. 111-140.

Kadak, T. (2011). Creation of a Supportive Model for Designing and Improving the Performance Management System of an Organisation. Case studies. Tallinn University of Technology. School of Economics and Business Administration, Department of Accounting, Tallinn, Estonia.

Kaplan, R.S. and Norton, D.P. (1996). The Balanced Scorecard, Harvard, Harvard Business School Press, Boston.

Laitinen, E. K. (2004). Performance Measurement & Management: Strategic View. University of Vaasa, Levon-institute. Strategic Management Accounting Research Unit. Publication No 114, Vaasa, Finland.

Laitinen, E. K. (2003). Future-based management accounting: a new approach with survey evidence, Critical Perspectives on Accounting, 14(3), pp. 293–323.

Lasdon, L. S., Fox, R. L. and Ratner, M. W. (1973). Nonlinear Optimization Using the Generalized Reduced Gradient Method, Technical Memorandum No. 325, Department of Operations Research, Case Western Reserve University. http://www.numdam.org/item/RO_1974__8_3_73_0.pdf

Kerpershoek, E., Groenleer, M. and de Bruijn, H. (2016). Unintended responses to performance management in Dutch hospital care: bringing together the managerial and professional perspectives, Public Management Review, 18, pp. 417–436.

Letza, S.R. (1996). The design and implementation of the balanced business scorecard: An analysis of three companies in practice, Business Process Re-engineering and Management Journal, 2(3), pp. 54-76. https://doi.org/10.1108/14637159610151217

Lewandowski, R.A. and Cirella, G.T. (2023). Performance management systems: Trade-off between implementation and strategy development, Operations Management Research, 16, pp. 280–295 https://doi.org/10.1007/s12063-022-00305-4

Liedtka, S.L., Church, B.K. and Ray, M.R. (2008). Performance Variability, Ambiguity Intolerance, and Balanced Scorecard-Based Performance Assessments, Behavioral Research in Accounting, 20(2), pp. 73–88.

Llach, J., Bagur, L., Perramon, J. and Marimon, F. (2017). Creating value through the balanced scorecard: how does it work? Management Decision, 55(10), pp. 2181-2199. https://doi.org/10.1108/MD-11-2016-0812

Lukman, A.F. and Ayinde, K. (2016). Review and Classifications of the Ridge Parameter Estimation Techniques, Hacettepe Journal of Mathematics and Statistics, 46, pp. 1-26.

Malagueño, R., Lopez-Valeiras, E. and Gomez-Conde, J. (2018). Balanced scorecard in SMEs: effects on innovation and financial performance, Small Business Economics, 51(1), pp. 221-244. https://doi:10.1007/s11187-017-9921-3

Malina, M.A. and Selto, F.H. (2001). Communicating and controlling strategy: an empirical study of the effectiveness of the balanced scorecard, Journal of Management Accounting Research, 13, pp. 47–90.

Malmi, T. (2001). Balanced scorecards in Finnish companies: a research note, Management Accounting Research, 12(2), pp. 207-220. https://doi.org/10.1006/mare.2000.0154

Mason-Jones, R. and Towill, D.R. (2000). Designing, implementing and updating performance measurement system, International Journal of Operations and Production Management, 20, pp. 754–771.

McCunn, P. (1998). The balanced scorecard: the eleventh commandment, Management Accounting: Magazine for Chartered Management Account, December, 76(11), pp. 34–36.

Miles, R. E. and Snow, C. C. (1978). Organizational strategy, structure and process, McGraw-Hill. https://doi.org/10.2307/257544

Najmi, M., Rigas, J. and Fan, I.S. (2005). A framework to review performance measurement systems, Business Process Management Journal, 11(2), pp. 109–22.

Neely, A.D. (2008). Does the balanced scorecard work: An empirical investigation, Cranfield University, School of Management, Research paper, 1/08. Cranfield.

Neely, A.D. (2005). The evolution of performance measurement research: developments in the last decade and a research agenda for the next, International Journal of Operations and Production Management, 25(12), pp. 1264-1277.

Neely, A. (1999). The performance measurement revolution: why now and what next? International Journal of Operations & Production Management, 19, pp. 205–228.

Neely, A.D., Gregory, M.J. and Platts, K. (1995). Performance measurement system design: a literature review and research agenda, International Journal of Operations and Production Management, 15(4), pp. 80-116.

Nørreklit, H., Nørreklit, L., Mitchell, F. and Bjørnenak, T. (2012). The rise of the balanced scorecard! Relevance regained? Journal of Accounting & Organizational Change, 8(4), pp. 490–510.

Nørreklit, H. (2003). The Balanced Scorecard: what is the score? A rhetorical analysis of the Balanced Scorecard, Accounting, Organizations and Society, 28(6), pp. 591–619.

Nørreklit, H. (2000). The balance on the balanced scorecard a critical analysis of some of its assumptions, Management Accounting Research, 11(1), pp. 65–88.

Okwir, S., Nudurupati, S.S., Ginieis, M. and Angelis, J. (2018). Performance Measurement and Management Systems: A Perspective from Complexity Theory, International Journal of Management Reviews, 20(3), pp. 731-754. https://doi.org/10.1111/ijmr.12184

Olsen, E.O., Zhou, H., Lee, D., Ng, Y., Chong, C.C. and Padunchwit, P. (2007). Performance measurement system and relationships with performance results: a case analysis of a continuous improvement approach to PMS design, International Journal of Productivity and Performance Management, 56(7), pp. 559–582.

Otley, D. (1999). Performance management: a framework for management control systems research, Management Accounting Research, 10(4), pp. 363-382.

Rompho, N. and Boon-itt, S. (2012). "Measuring the success of a performance measurement system in Thai firms", International Journal of Productivity and Performance Management, 61(5), pp. 548–562.

Said, A.A., HassabElnaby, H. R. and Wier, B. (2003). An empirical investigation of the performance consequences of nonfinancial measures, Journal of Management Accounting Research, 15, pp. 193–223.

Schneiderman, A. (1999). Why balanced scorecards fail! Journal of Strategic Performance Measurement, Special edition, January, pp. 6–11.

Simons, R. L. (2000). Performance measurement and control systems for implementing strategy. Upper Saddle River: Pearson.

Simons, R. (1995). Levers of control: How managers use innovative control systems to drive strategic renewal, Harvard Business School Press.

Speckbacher, G., Bischof, J. and Pfeiffer, T. (2003). A descriptive analysis on the implementation of balanced scorecards in German-speaking countries, Management Accounting Research, 14(4), pp. 361-388. https://doi.org/10.1016/j.mar.2003.10.001

Strauss, E. and Zecher, C. (2013). Management control systems; a review, Journal on Management Control, 23, pp. 233–268.

Tangen, S. (2005). Insights from practice Analysing the requirements of performance measurement systems, Measuring Business Excellence, 9(4), 46–54. https://doi.org/10.1108/13683040510634835

Tangen, S. (2004). Evaluation and Revision of Performance Measurement Systems, doctoral thesis, Woxe'n Centrum, Department of Production Engineering Royal Institute of Technology Stockholm.

Taticchi, P., Balachandran, K. and Tonelli, F. (2012). Performance measurement and management systems: state of the art, guidelines for design and challenges, Measuring Business Excellence, 16(2), pp. 1-19.

Waal, A. A. (2018). Defining and measuring high performance organization. Retreived January 4, 2025, from https://www.hpocenter.nl/wp-content/uploads/2018/07/Clute-2018-Defining-and-measuring-the-high-organizational-organization-De-Waal.pdf

Waal, A. A. and Kourtit, K. (2013). Performance measurement and management in practice: Advantages, disadvantages and reasons for use, International Journal of Productivity and Performance Management, 62(5), pp. 446–473.

Waal, A., Kourtit, K. and Nijkamp, P. (2009). The relationship between the level of completeness of a strategic performance management system and perceived advantages and disadvantages, International Journal of Operations & Production Management, 29(12), 1242–1265.

Waal, A. A. (2007). Strategic Performance Management. A Managerial and Behavioural Approach, Palgrave Macmillan, New York.

Wettstein, T. and Kueng, P. (2002). A maturity model for performance measurement systems, in Brebbia, C.A., Pascolo, P. (Eds), Management Information Systems 2002: GIS and Remote Sensing, WIT Press, Southampton.

Othman, R. (2006). Balanced scorecard and causal model development: preliminary findings, Management Decision, 44(5), pp. 690-702. https://doi.org/10.1108/00251740610668923

Yu, L., Perera, S. and Crowe, S. (2008). Effectiveness of the Balanced Scorecard: The Impact of Strategy and Causal Links, Journal of Applied Management Accounting Research, 6(2), pp. 37–56.

Appendices

Appendix 1. PMS success measures. Questions about the impact of PMS on organizational performance.

1. STRPER: How was the achievement rate of the latest strategic objectives assessed on a scale from 1 (extremely poor) to 7 (extremely good)?

2. NFIPER: How much do you consider the performance management system has improved the non-financial performance (productivity, efficiency, effectiveness) in the three last years on a scale from 1 (extremely little) to 7 (extremely much)?

3. FINPER: How much do you consider the performance management system has improved the financial performance (profitability, margins, turnover ratios) in the three last years on a scale from 1 (extremely little) to 7 (extremely much)?

Appendix 2. Questions about the 15 check points (CPs) of PMS.

1. MIS: Our organization has a stated mission (it is documented in a written form) (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

2. STR: Our organization has a document describing corporate strategy (it is documented in a written form) (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

3. CSF: Our organization’s strategy document (map) includes descriptions of causal relationships between Critical Success Factors (CSFs) (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

4. STROBJ: Our corporate strategy document includes a description of corporate organization-level objectives (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

5. ORGKPI: Our corporate performance management system includes a set of organization-level performance measures (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

6. UNOBJ: Our organization has goals set on different hierarchical levels of the organization (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

7. UNKPI: Our organization has a set of performance measures on different levels of organization (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

8. KPR: Our organization has defined key processes (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

9. IMPUSE: Our Performance Management System of Organization was built and designed with the users of this system (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

10.1. ACTUSE: Our Performance Management System of Organization is used interactively where managers are constantly and personally discussing with subordinates to learn strategic uncertainties and involved subordinates decisions (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

10.2. DNGUSE: Our Performance Management System of Organization is used diagnostically where managers monitor organizational outcomes and correct deviations from present standards of performance (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

11. RESP: Our Performance Management System of Organization is intensively used by employees in charge (users of the system) (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

12. INSYS: Our Performance Management System of Organization is associated with an incentive system to motivate employees towards achievement of objectives (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

13. QUALINF: Our Performance Management System of Organization produces valid, reliable, timely and usable information for managerial work (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

14. ADJ: Our Performance Management System of Organization is systematically updated when changes in organization or environment are emerged (assess on a scale from 1 = does not agree at all to 7 = agree perfectly).

Appendix 3. Correlation coefficients between the fifteen checkpoint variables.

Panel 1. Pearson correlations

Variables	MIS	STR	CSF	STROBJ	ORGKPI	UNOBJ	UNKPI	KPR	IMPUSE	ACTUSE	DGNUSE	RESP	INSYS	QUALINF	ADJ
MIS	1,000	0,826	0,663	0,679	0,571	0,553	0,510	0,651	0,422	0,260	0,461	0,331	0,406	0,447	0,427
STR	0,826	1,000	0,735	0,861	0,684	0,671	0,623	0,728	0,552	0,399	0,494	0,430	0,565	0,599	0,568
CSF	0,663	0,735	1,000	0,788	0,636	0,612	0,520	0,735	0,568	0,367	0,531	0,498	0,572	0,646	0,532
STROBJ	0,679	0,861	0,788	1,000	0,789	0,678	0,646	0,754	0,604	0,403	0,563	0,497	0,613	0,629	0,619
ORGKPI	0,571	0,684	0,636	0,789	1,000	0,730	0,751	0,658	0,616	0,347	0,449	0,584	0,489	0,603	0,687
UNOBJ	0,553	0,671	0,612	0,678	0,730	1,000	0,849	0,776	0,628	0,498	0,581	0,635	0,581	0,668	0,643
UNKPI	0,510	0,623	0,520	0,646	0,751	0,849	1,000	0,729	0,605	0,471	0,574	0,628	0,540	0,613	0,644
KPR	0,651	0,728	0,735	0,754	0,658	0,776	0,729	1,000	0,625	0,412	0,728	0,578	0,629	0,651	0,513
IMPUSE	0,422	0,552	0,568	0,604	0,616	0,628	0,605	0,625	1,000	0,652	0,592	0,658	0,619	0,682	0,652
ACTUSE	0,260	0,399	0,367	0,403	0,347	0,498	0,471	0,412	0,652	1,000	0,459	0,429	0,571	0,598	0,603
DGNUSE	0,461	0,494	0,531	0,563	0,449	0,581	0,574	0,728	0,592	0,459	1,000	0,634	0,645	0,629	0,486
RESP	0,331	0,430	0,498	0,497	0,584	0,635	0,628	0,578	0,658	0,429	0,634	1,000	0,539	0,650	0,598
INSYS	0,406	0,565	0,572	0,613	0,489	0,581	0,540	0,629	0,619	0,571	0,645	0,539	1,000	0,803	0,559
QUALINF	0,447	0,599	0,646	0,629	0,603	0,668	0,613	0,651	0,682	0,598	0,629	0,650	0,803	1,000	0,809
ADJ	0,427	0,568	0,532	0,619	0,687	0,643	0,644	0,513	0,652	0,603	0,486	0,598	0,559	0,809	1,000
SUMPERF	0,165	0,346	0,310	0,314	0,471	0,339	0,305	0,239	0,542	0,344	0,199	0,561	0,265	0,507	0,596
p-value	0,163	0,003	0,008	0,007	<,001	0,003	0,009	0,042	<,001	0,003	0,091	<,001	0,023	<,001	<,001

Panel 2. Spearman rank correlations

Variables	MIS	STR	CSF	STROBJ	ORGKPI	UNOBJ	UNKPI	KPR	IMPUSE	ACTUSE	DGNUSE	RESP	INSYS	QUALINF	ADJ
MIS	1,000	0,829	0,643	0,692	0,565	0,582	0,534	0,633	0,412	0,245	0,441	0,358	0,424	0,407	0,434
STR	0,829	1,000	0,752	0,855	0,667	0,637	0,597	0,700	0,535	0,372	0,449	0,440	0,538	0,538	0,534
CSF	0,643	0,752	1,000	0,788	0,621	0,580	0,522	0,745	0,548	0,345	0,510	0,490	0,562	0,629	0,530
STROBJ	0,692	0,855	0,788	1,000	0,782	0,641	0,625	0,744	0,561	0,342	0,515	0,471	0,568	0,577	0,578
ORGKPI	0,565	0,667	0,621	0,782	1,000	0,733	0,739	0,653	0,584	0,338	0,436	0,572	0,483	0,577	0,671

UNOBJ	0,582	0,637	0,580	0,641	0,733	1,000	0,856	0,737	0,592	0,426	0,557	0,636	0,577	0,635	0,613
UNKPI	0,534	0,597	0,522	0,625	0,739	0,856	1,000	0,689	0,597	0,419	0,574	0,660	0,512	0,573	0,611
KPR	0,633	0,700	0,745	0,744	0,653	0,737	0,689	1,000	0,613	0,376	0,711	0,584	0,614	0,607	0,478
IMPUSE	0,412	0,535	0,548	0,561	0,584	0,592	0,597	0,613	1,000	0,632	0,639	0,636	0,610	0,668	0,628
ACTUSE	0,245	0,372	0,345	0,342	0,338	0,426	0,419	0,376	0,632	1,000	0,443	0,406	0,520	0,583	0,597
DGNUSE	0,441	0,449	0,510	0,515	0,436	0,557	0,574	0,711	0,639	0,443	1,000	0,643	0,666	0,650	0,485
RESP	0,358	0,440	0,490	0,471	0,572	0,636	0,660	0,584	0,636	0,406	0,643	1,000	0,578	0,652	0,583
INSYS	0,424	0,538	0,562	0,568	0,483	0,577	0,512	0,614	0,610	0,520	0,666	0,578	1,000	0,815	0,598
QUALINF	0,407	0,538	0,629	0,577	0,577	0,635	0,573	0,607	0,668	0,583	0,650	0,652	0,815	1,000	0,813
ADJ	0,434	0,534	0,530	0,578	0,671	0,613	0,611	0,478	0,628	0,597	0,485	0,583	0,598	0,813	1,000
SUMPERF	0,186	0,350	0,290	0,308	0,474	0,373	0,358	0,266	0,516	0,340	0,260	0,563	0,357	0,535	0,616
p-value	0,114	0,002	0,013	0,008	<,001	0,001	0,002	0,023	<,001	0,003	0,026	<,001	0,002	<,001	<,001

Tables

Table 1. Descriptive statistics of the PMS success measures in two samples.
Panel 1: Sample 2015

	More Successful PMSs (n=30)			Less successful PMSs (n=32)		
Variable	Mean	Median	St. dev.	Mean	Median	St. dev.
STPER	5,600	5,692	0,724	3,906	3,920	0,734
NFIPER	5,300	5,333	0,794	3,563	3,609	0,982
FINPER	5,000	5,000	0,910	3,375	3,462	0,871
Sum of variables	15,900	15,909	1,539	10,844	11,143	1,986

Legend: If the sum of success variables exceeds 14, PMS is considered more successful

Panel 2: Sample 2022

	More successful PMS (n=39)			Less successful PMS (n=34)		
Variable	Mean	Median	St. dev.	Mean	Median	St. dev.
STRPER	6,000	6,000	0,562	4,529	5,000	1,261
NFIPER	5,744	6,000	0,677	4,029	4,000	1,267
FINPER	5,795	6,000	0,833	3,735	4,000	1,310
Sum of variables	17,539	18,000	1,393	12,294	13,000	2,725

Legend: If the sum of success variables exceeds 15, PMS is considered more successful
See Appendix 1 for PMS success measures.

Table 2. Descriptive statistics of the fifteen checkpoints (CP) in two samples.
Panel 1: Sample 2015

	More successful PMSs (n=30)			Less successful PMSs (n=32)				
Variables	Mean	Median	St. Dev.	Mean	Median	St. Dev.	F-test	p-value
MIS	0,933	1,000	0,254	0,781	1,000	0,420	2,929	0,092
STR	0,933	1,000	0,254	0,688	1,000	0,471	6,423	0,014
CSF	0,500	0,500	0,509	0,219	0,000	0,420	5,667	0,020

STROBJ	0,867	1,000	0,346	0,625	1,000	0,492	4,948	0,030
ORGKPI	0,967	1,000	0,183	0,781	1,000	0,420	4,963	0,030
UNOBJ	0,833	1,000	0,379	0,781	1,000	0,420	0,262	0,611
UNKPI	0,900	1,000	0,305	0,781	1,000	0,420	1,604	0,210
KPR	0,900	1,000	0,305	0,719	1,000	0,457	3,329	0,073
IMPL	0,639	0,667	0,259	0,531	0,583	0,269	2,566	0,114
ACTUSE	0,678	0,667	0,219	0,563	0,667	0,235	3,991	0,050
DNGUSE	0,694	0,667	0,224	0,583	0,667	0,216	3,964	0,051
RESP	0,678	0,667	0,243	0,526	0,500	0,258	5,658	0,021
INSYS	0,930	1,000	0,254	0,750	1,000	0,440	3,969	0,051
QUALINF	0,717	0,750	0,219	0,583	0,667	0,220	5,701	0,020
ADJ	0,867	1,000	0,346	0,656	1,000	0,483	3,849	0,054
Total	12,036	12,918		9,567	12,084			

Panel 2: Sample 2022

	More successful PMSs (n=39)			Less successful PMSs (n=34)				
Variables	Mean	Median	St. dev.	Mean	Median	St. dev.	F-test	p-value
MIS	0,739	0,833	0,335	0,544	0,583	0,365	5,6662	0,0200
STR	0,769	0,833	0,285	0,495	0,500	0,354	13,4089	0,0005
CSF	0,568	0,667	0,326	0,363	0,333	0,344	6,8552	0,0108
STROBJ	0,705	0,833	0,309	0,441	0,333	0,357	11,4426	0,0012
ORGKPI	0,731	0,833	0,285	0,480	0,500	0,304	13,1841	0,0005
UNOBJ	0,761	0,833	0,291	0,529	0,583	0,332	10,0729	0,0022
UNKPI	0,705	0,833	0,302	0,490	0,500	0,307	9,0572	0,0036
KPR	0,709	0,833	0,298	0,510	0,500	0,326	7,4789	0,0079
IMPL	0,718	0,667	0,233	0,402	0,500	0,288	26,8674	0,0000
ACTUSE	0,722	0,833	0,227	0,510	0,500	0,301	11,7330	0,0010
DNGUSE	0,680	0,667	0,258	0,549	0,500	0,268	4,4947	0,0375
RESP	0,688	0,667	0,265	0,382	0,333	0,289	22,2001	0,0000
INSYS	0,769	0,833	0,261	0,549	0,667	0,303	11,1298	0,0014
QUALINF	0,765	0,833	0,219	0,471	0,500	0,300	23,3392	0,0000
ADJ	0,752	0,833	0,244	0,407	0,417	0,313	27,9362	0,0000
Total	10,782	11,833		7,123	7,250			

See Appendix 2 for checkpoint variables.

Table 3. The weights of the checkpoint variables for different methods of prediction.

Checkpoint	Sum of CPs	Linear regression	p-value	VIF	Logistic regression	p-value	Ridge regression	Solver solution
MIS	1	-0,0694	0,9811	15,922	-1,0240	0,6250	0,0000	0
STR	1	8,9218	0,0430	33,712	2,5670	0,3760	0,0211	0
CSF	1	-2,9072	0,3325	10,401	-2,3280	0,2840	0,0160	0
STROBJ	1	-9,8748	0,0212	27,624	1,7170	0,5030	0,0138	0

ORGKPI	1	12,1335	0,0011	20,556	-2,6810	0,2140	0,0374	0
UNOBJ	1	-0,6803	0,8542	24,853	-1,5130	0,5200	0,0165	0
UNKPI	1	-5,7904	0,1225	21,879	-0,6820	0,7510	0,0124	0
KPR	1	-1,6694	0,6731	25,716	0,9480	0,7160	0,0034	0
IMPUSE	1	1,0598	0,7375	14,133	4,6790	0,0240	0,0472	1,7487
ACTUSE	1	6,1007	0,0327	12,462	-2,7580	0,1590	0,0242	0
DGNUSE	1	4,9513	0,1538	18,261	-6,1180	0,0210	0,0010	0
RESP	1	4,3931	0,1265	10,878	4,0630	0,0360	0,0524	2,9610
INSYS	1	0,5204	0,8840	23,047	-1,2200	0,6130	0,0086	0
QUALINF	1	3,4554	0,4869	40,237	2,3790	0,4610	0,0404	0
ADJ	1	0,9893	0,7907	21,454	3,2410	0,1570	0,0547	3,8317

Legend: For Ridge regression lambda is set equal to 463 to make all coefficients positive.

Table 4. The weights of the checkpoint variables for stepwise methods of prediction.

	Linear stepwise regression			Logistic stepwise regression	
Checkpoint	Coefficient	p-value	VIF	Coefficient	p-value
MIS					
STR					
CSF					
STROBJ					
ORGKPI	7,536	<,001	6,802		
UNOBJ					
UNKPI					
KPR					
IMPUSE				3,314	0,028
ACTUSE	8,222	<,001	7,484		
DGNUSE				-4,241	0,003
RESP					
INSYS					
QUALINF	6,282	0,028	11,641		
ADJ				2,038	0,062

Table 5. The correlation coefficients between the sum variable of PMS success and different predictions.

Panel 1. Pearson correlation coefficients

Methods	SUM PERF	Sum of' CPs	Linear regression	Linear stepwise	Logistic logit	Logistic probability	Logistic stepwise	Ridge regression	Solver solution
SUMPERF	1,000	0,462	0,632	0,533	0,677	0,673	0,577	0,576	0,656
Sum of CPs	0,462	1,000	0,819	0,921	0,399	0,448	0,389	0,975	0,878
Linear regression	0,632	0,819	1,000	0,896	0,485	0,505	0,460	0,876	0,873

Linear stepwise	0,533	0,921	0,896	1,000	0,433	0,468	0,530	0,952	0,886
Logistic logit	0,677	0,399	0,485	0,433	1,000	0,952	0,798	0,539	0,657
Logistic probability	0,673	0,448	0,505	0,468	0,952	1,000	0,766	0,572	0,678
Logistic stepwise	0,577	0,389	0,460	0,530	0,798	0,766	1,000	0,518	0,572
Ridge regression	0,576	0,975	0,876	0,952	0,539	0,572	0,518	1,000	0,955
Solver solution	0,656	0,878	0,873	0,886	0,657	0,678	0,572	0,955	1,000

Panel 2. Spearman correlation coefficients

Methods	SUM PERF	Sum of' CPs	Linear regression	Linear stepwise	Logistic logit	Logistic probability	Logistic stepwise	Ridge regression	Solver solution
SUMPERF	1,000	0,480	0,634	0,537	0,665	0,665	0,540	0,590	0,643
Sum of CPs	0,480	1,000	0,770	0,887	0,440	0,440	0,412	0,965	0,849
Linear regression	0,634	0,770	1,000	0,877	0,509	0,509	0,433	0,847	0,847
Linear stepwise	0,537	0,887	0,877	1,000	0,456	0,456	0,504	0,929	0,863
Logistic logit	0,665	0,440	0,509	0,456	1,000	1,000	0,750	0,575	0,690
Logistic probability	0,665	0,440	0,509	0,456	1,000	1,000	0,750	0,575	0,690
Logistic stepwise	0,540	0,412	0,433	0,504	0,750	0,750	1,000	0,528	0,545
Ridge regression	0,590	0,965	0,847	0,929	0,575	0,575	0,528	1,000	0,940
Solver solution	0,643	0,849	0,847	0,863	0,690	0,690	0,545	0,940	1,000

Table 6. Descriptive statistics of the predictors of PMS success.

	More successful PMSs (n=39)			Less successful PMSs (n=34)				
	Mean	Median	St. dev.	Mean	Median	St. dev.	F-test	p-value
Sum_of_CPs	10,782	11,542	2,919	7,123	7,333	3,768	21,802	<,001
Linear_regression	16,620	16,837	3,612	10,996	10,453	5,043	30,547	<,001
Linear stepwise	16,251	17,142	3,801	10,768	11,762	5,509	25,001	<,001
Logit	1,511	1,360	1,510	-0,876	-0,220	1,900	35,712	<,001
Logistic_probability	0,744	0,796	0,211	0,392	0,445	0,268	39,365	<,001
Logistic stepwise	1,030	1,065	0,992	-0,167	0,003	1,186	22,062	<,001
Ridge_regression	0,253	0,266	0,064	0,155	0,161	0,086	30,928	<,001
Solver_solution	6,175	6,567	1,709	3,394	3,705	2,154	37,775	<,001

Table 7. Area under ROC curve (AUC) and accuracy ratio (AR) for PMS success predictors.

			95% confidence interval		Accuracy ratio (AR)
	AUC	St. error	Lower bound	Upper bound	
Sum of CPs	0,790	0,054	0,684	0,896	0,580
Linear regression	0,817	0,051	0,718	0,917	0,634
Linear stepwise	0,791	0,055	0,684	0,898	0,582
Logit	0,849	0,044	0,763	0,935	0,698
Logistic probability	0,849	0,044	0,763	0,935	0,698

Logistic stepwise	0,784	0,054	0,679	0,890	0,568
Ridge regression	0,833	0,048	0,738	0,928	0,666
Solver solution	0,844	0,046	0,754	0,934	0,688

Table 8. Misclassification rates of PMS success predictors.

	More successful PMS	Less successful PMS	All PMS
Sum of CPs	0,3077	0,1765	0,2466
Linear regression	0,2821	0,1765	0,2329
Linear stepwise	0,2051	0,2353	0,2192
Logit	0,3077	0,1176	0,2192
Logistic probability	0,3077	0,1176	0,2192
Logistic stepwise	0,2308	0,2647	0,2466
Ridge regression	0,2564	0,2059	0,2329
Solver solution	0,2821	0,1176	0,2055

Figures

Figure 1. ROC curve for the sum of CP variables.

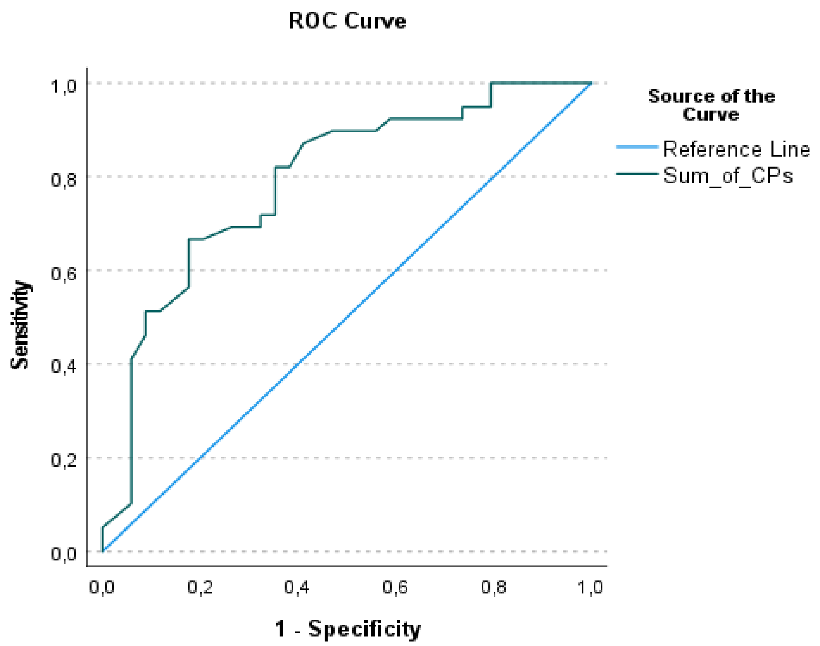

Figure 2. ROC curve for the linear regression.

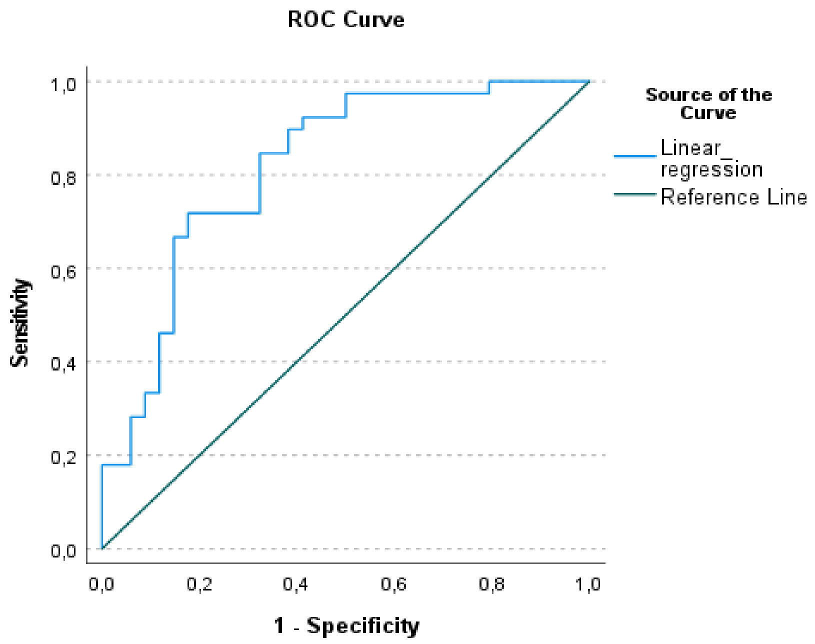

Figure 3. ROC curve for the linear stepwise model.

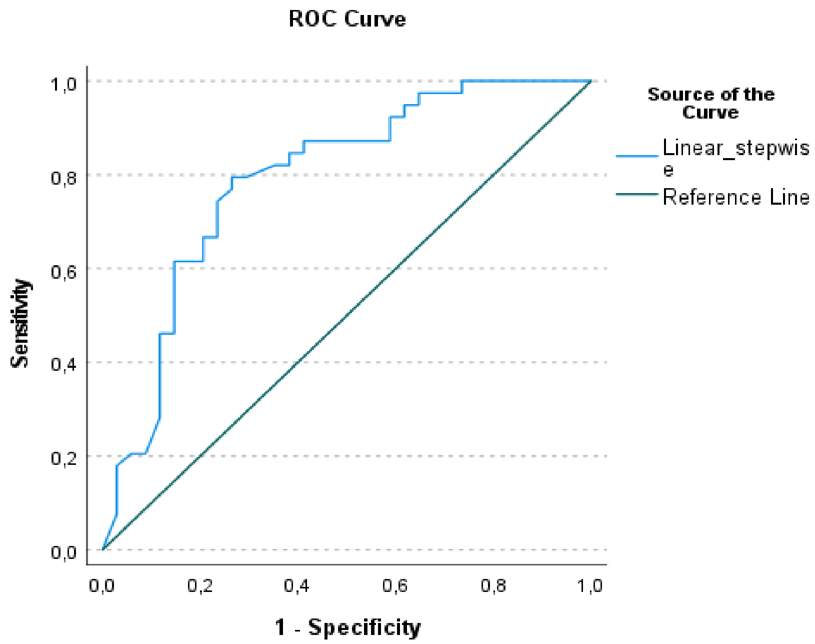

Figure 4. ROC curve for the logistic regression logit.

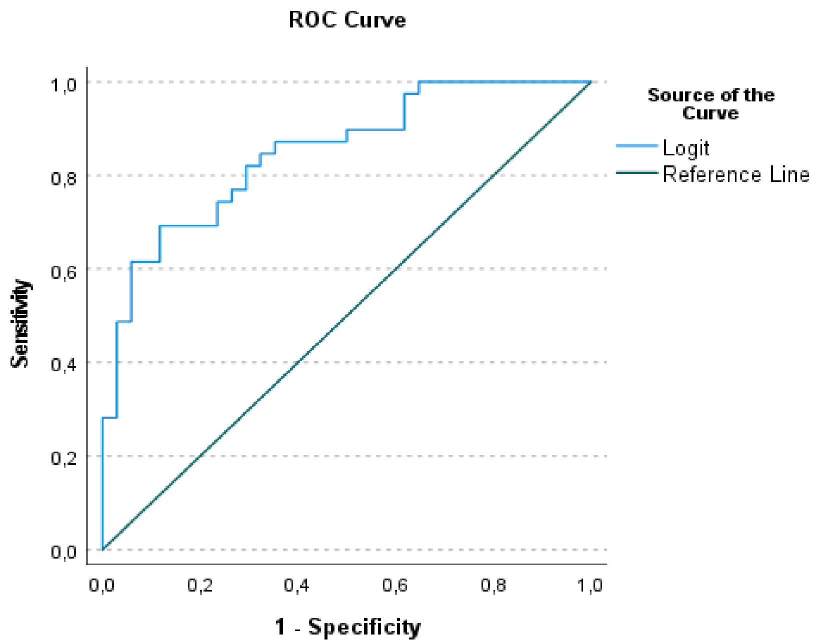

Figure 5. ROC curve for the logistic stepwise model.

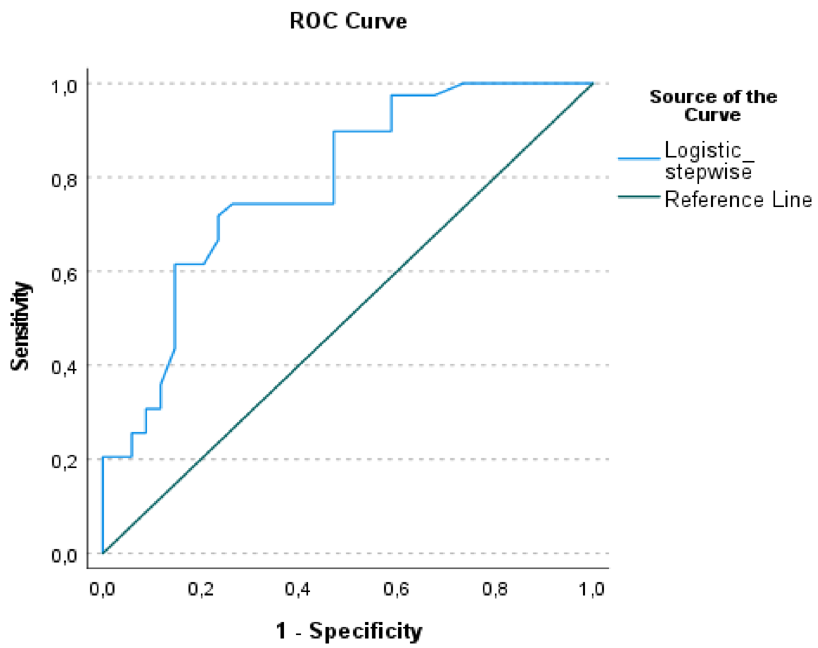

Figure 6. ROC curve for the ridge regression.

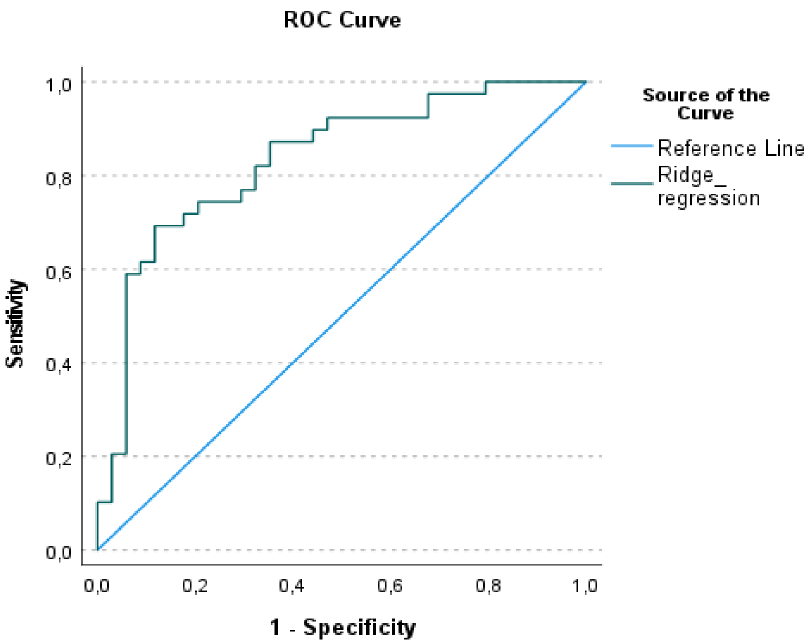

Figure 7. ROC curve for the Solver solution.