Performance Measures and Intra-Firm Spillovers: Theory and Evidence

Bouwens, J.; Hofmann, C.; van Lent, L.

DOI
10.2308/jmar-51903

Publication date
2018

Document Version
Submitted manuscript

Published in
Journal of Management Accounting Research

Performance Measures and Intra-Firm Spillovers: Theory and Evidence

Jan Bouwens
Tilburg University
CentER and Department of Accountancy
PO Box 90153, 5000 LE Tilburg, The Netherlands
[T] +31 13 466 8288, [F] +31 13 466 8001
j.bouwens@uvt.nl

Christian Hofmann
LMU Munich
Chair for Managerial Accounting
80939 Munich, Germany
[T] +49 89 2180 72035, [F] + 49 89 2180 72030
hofmann@bwl.lmu.de

Laurence van Lent
Tilburg University
CentER and Department of Accountancy
PO Box 90153, 5000 LE Tilburg, The Netherlands
[T] +31 13 466 8288, [F] +31 13 466 8001
vanlent@uvt.nl

Current version: March 2013

We appreciate helpful comments from Bart Dierynck, Yan Du, Raffi Indjejikian, DJ Nanda, Richard Sansing, Jack Stecher, Ahmed Tahoun, Andreas Weiler, seminar participants at the 2010 AAA Management Accounting Section midyear meeting in Seattle, the University of Leuven, the 2010 workshop on Accounting and Economics in Vienna, the 2010 GEABA workshop in Frankfurt, and the University of Mannheim.

Performance Measures and Intra-Firm Spillovers: Theory and Evidence

Abstract

We revisit the question of how performance measures are used to evaluate business unit managers in response to intra-firm spillovers because prior studies have documented conflicting empirical evidence. Specifically, we are interested in variation in the relative incentive weightings of aggregated "above-level" measures (e.g., firm-wide net income), "own-level" business unit measures (e.g., business unit profit), and specific "below-level" measures (e.g., R&D expenses) in response to spillover arising from either the focal manager's effect on other business units' performance or the other units' effect on the focal manager's performance. Our theory highlights existence of an interaction between the two directions of spillovers. In our empirical work, we account for the interaction effect. Based on a purpose-developed survey of 122 business unit managers, we report that the incentive weighting on above-level measures increases by approximately 50 percentage points when managers face both types of spillovers (as opposed to one type of spillover).

Keywords: Contracting, Business unit performance measurement, Organizational design, Interdependencies

JEL-code: D23, L22, M12, M4

Performance Measures and Intra-Firm Spillovers: Theory and Evidence

1. Introduction

One of the most fundamental problems in the management of firms with multiple business units is to ensure that these units cope well with the interdependencies that exist between them (Roberts, 2004). A solid insight that derives from prior work in the area holds that performance measurement systems can be designed to achieve this aim (Abernethy, Bouwens and van Lent, 2004, Bushman, Indjejikian and Smith, 1995, Keating, 1997). However, these prior studies offer conflicting empirical evidence on how different performance measures deal with spillovers that arise from the presence of mutually-dependent business units. We provide a simple model that offers an explanation for why prior studies have found conflicting results on the association between spillovers and the use of performance measures. The model implies an appropriate specification to examine this association empirically and, using data from a sample of 122 business units, we reassess the existing evidence.

In their early work, Bushman et al. (1995) document a positive association between spillovers and the use of "above-level" performance measures, which summarize the performance of multiple units, in the compensation contract of business unit managers. These authors argue that such above-level measures provide incentives to "internalize" the effects of interdependencies. In contrast, Keating (1997) reports that the use of above-level performance measures increases only for a very specific *direction* of spillover. Here, direction refers to whether the spillovers arise because other units in the firm affect the performance of the focal manager or because the focal manager affects the performance of other units in the firm. Only when a given business unit manager's actions affect the performance of managers of other units within the firm are above-level measures used more. Indeed, for the reverse situation, when the spillover effect is such that the focal business unit manager's performance is "impacted" by the actions of other managers in the firm, the use of above-level performance measures *decreases*. What's more, Keating (1997) predicts that performance measures that only summarize the performance of a manager's own unit (i.e., "own-level measures", such as business unit profits) are used less regardless of the specific direction of spillovers. He finds, however, no evidence of an association between both type of spillovers and the use of own-level measures. In contrast, Abernethy et al. (2004) report that the use of own-level measures *increases* when the actions of other managers affect the performance of the business unit manager but *decreases* when the spillover direction is reversed, that is when a given manager's actions "impact" on other managers in the firm.

We argue that the empirical evidence in these prior studies is difficult to synthesize for three reasons. First, Keating (1997) and Abernethy et al. (2004) show that the direction of spillovers matter in the association with the use of performance measures. These studies disagree, however, about the question which performance measure (above-level or own-level) is associated with spillovers. Bushman et al. (1995), which predates these studies, does not account for the direction of the spillovers. Second, while Bushman et al. (1995) motivate their empirical work with a formal model, Keating (1997) and Abernethy et al. (2004) base their predictions on informal reasoning. It is a priori unclear whether a theory which reflects the direction of spillovers yields predictions that are different in substance from those in Bushman et al. (1995). Without theoretical guidance, however, the empirical findings of the former two studies might be based on misspecified models. Finally, extant theory yields predictions in terms of the association between spillovers and a *set of optimal incentive weights*, where the relation between the incentive weights can be captured by *incentive ratios* (i.e., the ratio of weights placed on different types of performance measures). Yet, the empirical evidence is not based on incentive ratios, but on constructs that confound ex ante incentive weights with expected performance (as Bushman et al. (1995) highlight in their work) or that consider the use of performance measures in isolation without modeling the fact that putting more percentage weight on, say, above-level measures implies putting less percentage weight on own-level measures (as in Keating (1997) and Abernethy et al. (2004)).

We therefore reexamine how spillovers between business units within a firm affect the incentive weightings of performance measures used to evaluate its managers.[1] In an attempt to address the

[1] While interdependencies create measurement challenges for responsibility accounting, these interdependencies can be addressed using cost allocation and transfer pricing (Zimmerman, 2003) Under a perfect transfer-pricing system,

discrepancies in prior studies, we offer two innovations. First, within a principal/two-agent LEN contracting setting, we consider the relation between the direction of spillovers and the optimal weighting of performance measures. A multi-task setting allows for divergence between spillover and local actions, i.e., variation of the extent to which a manager's actions affect the performance of other business units. We combine prior literature by proposing that spillover has a congruity effect (which is in line with Bushman et al., 1995) and a noise effect (Keating, 1997),[2] and that performance measures are differentially affected by the direction of spillovers. In addition, due to the summing-up characteristic of accounting measures, the properties of above-level measures (such as firm-wide profits) are a function of the properties of the own-level measures (such as business unit profits) that add up to these more aggregate measures. Together, spillovers have both a direct and an indirect (summing-up) impact on performance measures. Understanding the incentives provided by each type of measure on the efforts of business unit managers requires a careful analysis of the totality of these effects.

We provide comparative statics regarding the effects of intra-firm spillovers on the weightings of performance measures relative to one another. We show that one subtle consequence of our setting is that there is an interaction between the two directions of spillovers and their relation with the relative weighting on performance measures. This interaction effect is an outcome of our multi-task setting and the modeling of a joint congruity and noise effect from spillovers. The omission of such an interaction effect in empirical tests yields misspecified regressions.

Second, in our empirical analysis, we account for the full spectrum of performance measures available to the firm for contracting. These include the two types examined in prior work, namely above-level and own-level measures but also (to complete the range) below-level or specific measures, which do not summarize the overall performance of a business unit manager, but rather reflect only a single task or

---

a business unit manager's own-level measure would reflect any externalities arising from intra-firm interdependencies, making above-level measures less useful in performance evaluation (Bushman et al., 1995). We address in our paper those interdependencies that cannot be adequately accounted for by transfer-pricing schemes perhaps because of top management's lack of knowledge (Keating, 1997) or exogenous restrictions such as tax laws (Bouwens and van Lent, 2007).

[2] We thank DJ Nanda for pointing this out.

subset of tasks. We find that our theoretical predictions with regard to the existence of an interaction effect do not change in the presence of below-level measures. However, by considering the full range of performance measures, we can tailor our empirical analysis more closely to the theoretically-relevant concept of the incentive ratio; what's more, when in practice senior managers decide to place more percentage weight on one particular performance measure, it implies that remaining measures obtain lower percentage weights in the compensation contract. The weighting of performance measures is relative compared to other available measures, which needs to be explicitly accounted for in the empirical analysis.

Turning to this analysis, we test for the relations between the direction of spillovers and the relative weightings of above-level, own-level, and below-level performance measures, while controlling for the degree of decentralization and various other factors known to influence the use of performance measures (Abernethy et al., 2004, Ittner, Larcker and Rajan, 1997, Keating, 1997). We gather survey data from a sample of 122 managers of business units across a range of industries. We use this survey data to obtain proxies for spillovers *at the business unit level* as well as information about how performance measures are used to evaluate business unit managers. Because data on the direction of spillovers at the business unit level and data on the compensation contract of business unit managers is not available from publicly available sources, we gather our data using surveys (Ittner and Larcker, 2001, Lanen, 1995, Luft and Shields, 2003). We recognize that using surveys to gather data may be problematic. To address such potential problems, we implement several strategies, including the use of multiple proxies to conduct convergent validity tests for our main variables, statistical and procedural measures to mitigate concerns about common rater and other respondent biases, and simulations that explicitly account for estimation uncertainty in our regressions.

Our empirical findings can be summarized as follows. We document that the relative weightings of the performance measures in our sample respond to spillovers, much as theory suggests. What's more, the theory's prediction that spillovers can reinforce one another and thus affect the relative weightings of the performance measures is consistent with our data. Our estimates of the economic effects are large. For example, we document that the weighting on above-level measures is on average nine percent when the spillovers from other managers on the focal manager is low (and the spillover from the focal manager to other managers is high). Increasing, however, the degree of spillovers from other managers on the focal manager to its highest level in-sample, while keeping everything else constant, increases the average weight on above-level measures to 60 percent. The accompanying reduction in weight on own-level measures is even greater as simultaneously firms place 17 percent more weight on below-level measures when the spillovers from other managers on the focal manager increase. We find very similar effects for the reverse case in which we change the degree of spillovers from the focal manager to other managers. Note that business unit managers who have to deal with both incoming and outgoing spillovers are facing great challenges in making decisions that are congruent with the firm's objective. We show empirically that when the two spillover effects reinforce each other the percentage weight placed on both above-level and below-level measures increases and less percentage weight is put on own-level measures.

Section 2 considers the theoretical framework and derives predictions regarding the impact of intra-firm spillovers on the relative weighting of above-level and own-level performance measures. Section 3 describes the sample, the measurement of the variables, and the estimation of the model. Section 4 presents our empirical findings and Section 5 concludes.

2. Theoretical Framework

2.1 Model Description

In this section, we analyze a simple agency model where the performance of business unit managers of a multi-unit firm is assessed based on business unit and firm-wide performance measures. By considering the consequences of spillover *to* and *from* other business units on the characteristics of business unit and firm-wide performance measures and the optimal incentive weights for these measures, we provide a framework for our empirical analysis.

Consider a risk-neutral principal (e.g., firm owners) who, at date t=0, hires two risk-averse agents (i.e., business unit managers). At date t=1, each agent exerts unobservable effort. At date t=2, performance measures are realized, the agents receive their compensation, and the principal obtains the net payoff. To capture the idea that intra-firm interdependencies typically relate only to a subset of a manager's responsibilities, we consider a setting where each agent performs multiple tasks.[3]

Specifically, each agent i, i=A,B, performs two tasks k, k=1,2, and the level of effort devoted to each task k is represented by $a_{ik} \in \mathbb{R}$. The output of business unit i, y_i, is defined as

$$y_i = \mu_{i1}\, a_{i1} + \mu_{i2}\, a_{i2} + m_{j1}\, a_{j1} + m_{j2}\, a_{j2} + \varepsilon_i, \qquad i,j = \text{A,B and } i \neq j,$$

where μ_{ik} is the sensitivity of y_i to agent i's effort in task k, m_{jk} is the sensitivity of y_i to the other agent j's effort in this agent's task k, and ε_i is a normally distributed random variable capturing events beyond the agents' control, $\varepsilon_i \sim N(0,\sigma_i^2)$ and $\text{Cov}[\varepsilon_A,\varepsilon_B] = \sigma_{AB}$. More specifically, m_{jk} captures interdependencies between the business units. For $m_{jk} \neq (=)\ 0$, agent j's task k is a spillover (local) action and, for $m_{jk} > (<)\ 0$, the spillover action imposes a positive (negative) externality on the performance of business unit i.

Aggregate firm output, x, is given by the sum of individual business unit outputs, i.e.,

$$x = y_A + y_B = \Sigma_{i,k}\, (\mu_{ik} + m_{ik})\, a_{ik} + \varepsilon_x,$$

where $\mu_{ik} + m_{ik} = b_{ik}$ is the marginal productivity of agent i's effort in task k and ε_x is normally distributed aggregate noise, $\varepsilon_x \sim N(0,\sigma_x^2)$, with $\sigma_x^2 = \sigma_A^2 + \sigma_B^2 + 2\,\sigma_{AB}$ and $\text{Cov}[\varepsilon_i,\varepsilon_x] = \sigma_{ix} = \sigma_i^2 + \sigma_{AB}$. For example, when business unit output represents accounting profit, the aggregate measure reflects firm profits. Unlike the output of a business unit, aggregate output captures the total productivity of an agent's action.

The principal offers agent i a contract that depends on two performance measures: aggregate firm output, x, and agent i's business unit output, y_i.[4] We restrict agent i's compensation, w_i, to be a linear function of the measures, i.e.,

$$w_i = f_i + v_{xi}\, x + v_{yi}\, y_i, \qquad i = \text{A,B},$$

[3] Business unit managers are frequently empowered with decision rights over issues like HRM, operations, marketing and investments. Arguably, decisions regarding personnel have a smaller impact or no impact at all on the outcome of other business units relative to strategic decisions and investments.

[4] Restricting the contract to these two measures is without loss in generality: Aggregate output, x, and the output of business unit i, y_i, are an equivalent statistic for the output of the other business unit, y_j, $j \neq i$.

where f_i is the fixed salary, v_{xi} is the incentive weight for aggregate firm output, and v_{yi} is the incentive weight for business unit output.

Agent i's preference is represented by a negative exponential utility function, with $u_i = -\exp[-r(w_i - \kappa_i)]$, where r is the coefficient of the agent's absolute risk aversion and κ_i denotes agent i's cost of effort. The marginal cost is positive and separable, $\kappa_i = ½ (a_{i1}^2 + a_{i2}^2)$.

Normally distributed noise terms and negative exponential utility yield a simple representation of the agent's certainty equivalent (Holmstrom and Milgrom, 1987). In particular, agent i's certainty equivalent is characterized by

$$CE_i(f_i,v_{xi},v_{yi},a_{i1},a_{i2},a_{j1},a_{j2}) = f_i + v_{xi}\,\mathrm{E}[x|a_{i1},a_{i2},a_{j1},a_{j2}] + v_{yi}\,\mathrm{E}[y_i|a_{i1},a_{i2}] - ½\,(a_{i1}^2 + a_{i2}^2)$$
$$-\, ½\, r\, \{v_{xi}^2\, \sigma_x^2 + v_{yi}^2\, \sigma_i^2 + 2\, v_{xi}\, v_{yi}\, \sigma_{ix}\}, \quad i, j = \mathrm{A,B} \text{ and and } i \neq j, \tag{1}$$

representing expected compensation net of effort cost and risk premium. Agent i conjectures agent j's actions; in equilibrium, each agent's conjecture with respect to the other agent's actions is true.

The principal selects incentive weights that maximize expected firm profit net of the agents' compensation, subject to the agents' individual rationality and incentive compatibility constraints. Specifically, the principal solves the following problem (Holmstrom and Milgrom, 1990):

$$\max_{f_i,v_{xi},v_{yi},a_{i1},a_{i2}} \mathrm{E}[x \mid a_{i1}, a_{i2}] - \mathrm{E}[w_1 + w_2 \mid a_{i1}, a_{i2}, f_i, v_{xi}, v_{yi}] \tag{2a}$$

subject to

$$CE_i(f_i,v_{xi},v_{yi},a_{i1},a_{i2},a_{j1},a_{j2}) \geq 0, \tag{2b}$$

$$\text{and } (a_{i1},a_{i2}) \text{ maximize } CE_i(f_i,v_{xi},v_{yi},a_{i1},a_{i2},a_{j1},a_{j2}), \qquad i = \mathrm{A,B}. \tag{2c}$$

Individual rationality constraints, (2b), ensure contract acceptance by the agents. Without loss in generality, we scale each agent's reservation certainty equivalent to equal zero. The incentive compatibility constraints, (2c), reflect that agent i chooses actions a_{i1} and a_{i2} that maximize his certainty equivalent.[5]

[5] We assume that the agents act non-cooperatively. Models with inter-agent negotiations include Holmström and Milgrom (1990), Itoh (1992), and Feltham and Hofmann (2007), among others.

2.2 Optimal contracts

The principal solves the problem expressed in (2a) through (2c). According to (2c), agent *i*'s effort, a_{ik}^*, follows from the first-order conditions of (1) with respect to a_{i1} and a_{i2}:

$$a_{ik}^* = b_{ik}\, v_{xi} + \mu_{ik}\, v_{yi}, \qquad i = \text{A,B and } k = 1,2. \tag{3}$$

For each task *k*, agent *i*'s effort incentives depend on the incentive weight for aggregate firm output, v_{xi}, and the incentive weight for business unit output, v_{yi}. The agents' effort choices constitute a unique dominant strategy in the agents' action-choice subgame.

To obtain the optimal incentive contract, for each agent *i,* we set (1) equal to zero (based on (2b)) and solve for f_i^*. Substituting, for each agent *i*, f_i^*, a_{i1}^*, and a_{i2}^* into (2a), differentiating with respect to the incentive weights, and solving the first-order conditions for these variables yields the optimal incentive weights. Lemma 1 describes the ratio of optimal incentive weights for aggregate firm output and business unit output, $IR_i^* = v_{xi}^* / v_{yi}^*$.[6]

Lemma 1: *Given the optimal contract offered to agent i, the relative incentive weights on x and* y_i *are given by*

$$IR_i^* = \frac{v_{xi}^*}{v_{yi}^*} = \frac{\frac{\lambda_i^2}{\sigma_i^2} + r\sum_k b_{ik}\left(b_{ik} - \frac{\sigma_{ix}}{\sigma_i^2}\mu_{ik}\right)}{r\sum_k b_{ik}\left(\mu_{ik} - \frac{\sigma_{ix}}{\sigma_x^2} b_{ik}\right)}\,\frac{\sigma_i^2}{\sigma_x^2}, \quad i = A,B, \tag{4}$$

where $\lambda_i = b_{i2}\mu_{i1} - b_{i1}\mu_{i2} = \mu_{i1}\, m_{i2} - \mu_{i2}\, m_{i1}$ *is a measure of non-congruity of* y_i *relative to x.*

Expression (4) extends the result in Bushman et al. (1995) to a multi-task setting.[7] When choosing incentive weights in multi-task settings, the principal considers a performance measure's non-congruity with her gross payoff, λ_i (Feltham and Xie 1994). Specifically, the principal chooses the

---

[6] All proofs are presented in Appendix 1.

[7] For $k = 1$, expression (4) simplifies to

$$IR_i^* = \frac{b_{i1} - \frac{\sigma_{ix}}{\sigma_i^2}\mu_{i1}}{\mu_{i1} - \frac{\sigma_{ix}}{\sigma_i^2} b_{i1}}\,\frac{\sigma_i^2}{\sigma_x^2},$$

which is essentially identical to expression (6) in Bushman et al. (1995).

weights on the performance measures to minimize the incongruity of the overall performance measure and the agent's risk premium (Datar et al. 2001).

When business unit output is perfectly congruent with aggregate firm output, λ_i=0, the relative incentive weights follow from the ratio of the sensitivity times precision of aggregate firm output and business unit output (Banker and Datar 1989). The terms in square brackets represent, for each task *k*, the adjusted sensitivity of aggregate firm output (numerator) and business unit output (denominator), respectively. The adjustment reflects that, with correlated signals, $\sigma_{ix}\neq 0$, some information on task *k* is contained in the other signal. To reflect differences in the principal's preferences for the agent's tasks, the adjusted sensitivities are weighted by the task's marginal productivity, b_{ik}. Finally, the last term in expression (4) is the precision of aggregate firm output, $1/\sigma_x^2$, divided by the precision of business unit output, $1/\sigma_i^2$.

More generally, when business unit output is not congruent with aggregate firm output, $\lambda_i\neq 0$, the degree of non-congruency, λ_i, and the agent's coefficient of risk aversion, *r*, affect the relative incentive weights. Given a non-congruent performance measure, varying the incentive weights generally changes the agent's allocation of effort across tasks and the risk imposed on the agent, i.e., the agent's risk premium.

To illustrate the impact of intra-firm interdependencies on the relative incentive weights, in line with Bushman et al. (1995), it is instructive to consider the special case where business unit outputs are uncorrelated, $\sigma_{AB} = 0$, implying $\sigma_{ix} = \sigma_i^2$. Then, expression (4) simplifies to

$$IR_i^* = \frac{\frac{\lambda_i^2}{\sigma_i^2} + r\ \sum_k b_{ik} m_{ik}}{r\ \sum_k b_{ik}\left(\mu_{ik} - \frac{\sigma_i^2}{\sigma_j^2} m_{ik}\right)}\ \frac{\sigma_i^2}{\sigma_j^2}, \quad i,j = \text{A,B and } i\neq j. \qquad (5)$$

Following Keating (1997), intra-firm interdependency can take on two forms: *spillover to the other business unit*, i.e., the effect the focal agent's business unit has on the other unit, and *spillover from the other business unit*, i.e., the effect the other agent's business unit has on the focal agent's unit. For illustration, let δ_i represent the spillover from business unit *i* to business unit *j*, and let δ_j represent the spillover from business unit j to business unit i, i,j = A,B and $i \neq j$. We next discuss (i), how performance measure sensitivity and noise may vary with the two forms of spillover and (ii), the implications for the relative weights placed on performance measures, IR_i^*, according to expression (5).

First, Bushman et al. (1995) argue that spillover creates externalities that manifest in the sensitivity of a business unit's output to the effort of the agent from another business unit. Specifically, from the perspective of focal business unit i, for a larger spillover δ_i to the other business unit, the marginal impact on the other business unit, m_{ik}, increases, $\partial m_{ik}/\partial \delta_i > 0$. In turn, from expression (5), $\partial IR_i^*/\partial m_{ik} \gtreqless 0$, i.e., in a multi-task setting, the relation between the relative weight placed on the aggregate performance measure and the impact on another business unit is generally ambiguous (Datar et al. 2001). For an increasing marginal impact of task k on the other business unit, the principal would like to motivate agent i to provide more effort on task k. While placing a larger weight v_{xi}^* on aggregate output lets agent i provide more effort on task k, it also motivates the agent to provide more effort on his other task $n \neq k$ (if $b_{in} > 0$). However, the associated increase in the agent's risk premium can be quite costly for the principal. Alternatively, to motivate more effort on task k, the principal can also increase the weight on business unit output, v_{yi}^*, thus decreasing the relative incentive weight IR_i^*. Now, increasing v_{yi}^* also motivates agent i to provide more effort on his other task $n \neq k$ (if $\mu_{in} > 0$) and increases the agent's risk premium. Overall, while the relation is generally ambiguous, the relative weight placed on aggregate firm output decreases in the marginal impact on another business unit if business unit output is sufficiently precise.[8]

Combining both observations, (i), the relation between spillover to the other business unit and the marginal impact on the other business unit and (ii), the relation between the marginal impact on the other business unit and the relative weight placed on aggregate firm output, we find that the relation between spillover to the other business unit and the relative weight placed on aggregate firm output is ambiguous,

[8] From expression (5), the numerator increases in m_{ik}. Moreover, the denominator decreases in m_{ik} if $\sigma_i^2/\sigma_j^2 > \mu_{in}/(\mu_{in} + 2\, m_{in})$. Hence, a sufficient condition for a positive relation between the relative weight placed on aggregate output and the marginal impact on another business unit is that the focal business unit's output is sufficiently noisy, $\sigma_i^2/\sigma_j^2 > \mu_{in}/(\mu_{in} + 2\, m_{in})$. We formalize this result and further comparative statics subsequently in Proposition 1.

$\partial IR_i^*/\partial\delta_i \gtreqless 0$. This is in stark contrast to a single-task setting, where the relative weight placed on aggregate firm output increases in the marginal impact on another business unit (Bushman et al. 1995, Observation 1 (ii)). Intuitively, in a single-task setting, increasing the weight on aggregate firm output increases the agent's effort in the single task but not in any further tasks.

By symmetry, for a larger spillover δ_j from the other business unit, the marginal impact of the other business unit on the focal business unit, m_{jk}, increases. However, from expression (5), $\partial IR_i^*/\partial m_{jk} = 0$, i.e., the marginal impact of the other business unit on the focal business unit does not affect the relative weight placed on aggregate output. While these externalities vary both, the focal business unit's expected output and expected aggregate firm output, it is only the focal agent's fixed salary that corrects for the associated mean effect. Hence, it is straightforward that varying the spillover from the other business unit does not affect the relative weight placed on aggregate firm output, $\partial IR_i^*/\partial\delta_j = 0$.

Second, Keating (1997) argues that spillover creates noise in business unit performance.[9] Specifically, for a larger spillover δ_i to the other business unit, the noise in the other business unit's output, σ_j^2, increases, $\partial\sigma_j^2/\partial\delta_i > 0$. From expression (5), $\partial IR_i^*/\partial\sigma_j^2 < 0$, i.e., the relative weight placed on aggregate firm output decreases when the other business unit's output becomes noisier. Note that any variation in the precision of a business unit's output yields a similar variation in the precision of aggregate firm output. Now, for a noisier output of the other business unit, the precision of aggregate firm output relative to the precision of the focal business unit's output decreases and the principal chooses a relatively smaller weight for the relatively less precise aggregate firm output as compared to business unit output. Overall, increasing the spillover to the other business unit decreases the relative weight placed on aggregate firm output, $\partial IR_i^*/\partial\delta_i < 0$.

Reiterating, by symmetry, for a larger spillover δ_j from the other business unit, the noise in the focal business unit's output, σ_i^2, increases. Now, for a noisier output of the focal business unit, the

[9] The "greater a manager's impact on divisions other than his own, the greater the noise in division accounting metrics …" (Keating, 1997, p. 247). While Keating assumes that a larger spillover to the other business unit reduces the noise in the aggregate performance measure, in our setting, due to the aggregation of business unit output to firm-wide output, the noise of the aggregate signal increases in the spillover to the other business unit.

precision of aggregate firm output relative to the precision of the focal business unit's output increases and the principal chooses a relatively larger weight for the relatively more precise aggregate firm output as compared to business unit output, $\partial IR_i^*/\partial\sigma_i^2 > 0$. Summarizing, increasing the spillover from the other business unit increases the relative weight placed on aggregate firm output, $\partial IR_i^*/\partial\delta_j > 0$.

Proposition 1 summarizes the key comparative statics regarding performance measure sensitivity and noise.

Proposition 1: *Consider the multi-task setting where business unit outputs are uncorrelated, $\sigma_{AB} = 0$; the relative incentive weights of the focal business unit, IR_i^*,*

(i) can increase or decrease in task k's marginal impact, m_{ik}, on the other business unit. The relative incentive weights increase (decrease) in m_{ik} if business unit output is sufficiently noisy (precise), assuming equally precise business outputs and a minor impact of task $n\neq k$ on the output of the other business unit, $\partial IR_i^*/\partial m_{ik} > (<) 0$ at $m_{ik}=0$ if $r\sigma_i^2 > (<) 2\mu_{in}m_{in}$ when $\sigma_A^2=\sigma_B^2$ and $m_{in}^2<\mu_{i1}^2+\mu_{i2}^2$;

(ii) are unaffected by the marginal impact of the other business unit on the focal business unit, $\partial IR_i^*/\partial m_{jk} = 0$;

(iii) decrease in the noisiness of the other business unit's output, $\partial IR_i^*/\partial\sigma_j^2 < 0$;

(iv) increase in the noisiness of the focal business unit's output, $\partial IR_i^*/\partial\sigma_i^2 > 0$.

Table 1 summarizes the comparative statics of the relative incentive weights when varying the spillover to and from the other business unit.[10] We label these consequences *congruity effect*, when spillover affects the sensitivity and, thus, the congruity of business unit output with aggregate firm output and *noise effect*, when spillover affects the noisiness of business unit output.

---------------- Insert Table 1 about here. -----------------

Three observations emerge from Table 1. First, there exists a non-trivial relation between the direction of the spillover and the relative weights placed on the performance measures. For example,

[10] Our model comprises just two business units. Extending the analysis to an arbitrary number of business units yields insights that are basically identical to the case of just two business units.

considering the noise effect, while the relative weight placed on the aggregate performance measure decreases if there is more spillover to the other business unit, the relative weight placed on the aggregate performance measure increases if there is more spillover from the other business unit.

Second, in real applications, both the congruity and the noise effect are most likely present. Then, the overall consequences of variations in the spillover depend on the relative importance of either effect. For example, when the congruity effect dominates and the focal business unit's output is sufficiently noisy, increasing the spillover to the other business unit increases the relative weight placed on aggregate firm output. In contrast, when the noise effect dominates, increasing the spillover to the other business unit decreases the relative weight placed on aggregate firm output.

Third, the two directions of spillover interact in determining the optimal weights placed on the performance measures.[11] Specifically, the spillover from the other business unit affects the noise in the focal business unit's output which, in turn, affects the relation between the spillover to the other business unit and the incentive weights. For example, for a small spillover from the other business unit, the focal business unit's output can be sufficiently precise, such that increasing the spillover to the other business unit decreases the relative weight placed on aggregate firm output; however, for a large spillover from the other business unit, the focal business unit's output can be sufficiently noisy, implying that increasing the spillover to the other business unit increases the relative weight placed on aggregate firm output. Thus, the model predicts an interaction between the spillover to the other business unit and the spillover from the other business unit, which would affect the relative weightings of the performance measures.

The interaction of the two directions of spillover in determining the optimal weights placed on the performance measures is a key insight of our analysis. It is a subtle consequence of (i), a multi-task setting, (ii), the differentiation between the two directions of spillover, and (iii), the differentiation

[11] Technically, the cross partials are generally non-zero, $\partial^2 IR_i^*/(\partial\delta_i\,\partial\delta_j) \neq 0$, and the sign of the derivative regarding one direction of spillover varies with the magnitude of the other direction of spillover, $\mathrm{sgn}[\partial IR_i^*/\partial\delta_i]$ depends on δ_j.

between a congruity and a noise effect that are caused by intra-firm spillover.[12] What's more, while sample selection ensures that business unit managers have responsibility over substantive multiple tasks, the significance of (ii) and (iii) lends itself directly to empirical testing. To this undertaking, we turn next.

3. Sample, variable measurement, and model estimation

3.1 Data collection and sample

We use the client list of a major audit firm to obtain our sample of business unit managers from publicly listed firms. Using the names and addresses provided on the client list, we directly contacted a random selection of these business unit managers and invited them to participate in the survey. Although using only a single audit firm's client list may bias our sample, we feel that, in our setting, this effect is likely to be small and would not outweigh the advantages of making our first contact with our respondents through the audit firm and thereby securing their willingness to participate. To be included in the sample, firms must have more than one business unit. From our initial sample of 240 business units, 122 managers agreed to participate, yielding a response rate of about 50%, which is in line with recent survey-based studies in accounting. We use this sample to collect data specifically for the purpose of the current study (and thus our data have not been used in other projects).

Surveys were administered by phone. Prior to the phone call, respondents received a package containing the questionnaire and a cover letter that explained the general aims of the research.[13] Compared with mail surveys, phone interviews offer many of the same advantages as site visits while being less costly. The researcher can verify the respondent's understanding of the questions and ensure that all questions are answered. Phone interviews also ensure that the intended respondent, and not, for example, the respondent's assistant, answers the questions. Common method bias is a valid concern in survey-based research. We use both procedural and statistical remedies to mitigate the adverse effects of

---

[12] Note that each element is necessary for the interaction effect to occur. In particular, in a single-task setting, the relative weight placed on aggregate output increases in the spillover to another business unit, implying that the sign of the comparative static is independent of the spillover from the other business unit.

[13] Appendix 2 reproduces the questionnaire items we used to construct our variables.

this bias (following Podsakoff, MacKenzie, Lee and Podsakoff, 2003). We separate the measurement of dependent and independent variables by placing these questionnaire items as far apart as possible and by using different response formats. We protected respondent anonymity and reduced evaluation apprehension by assuring respondents that there are no right or wrong answers and that they should answer questions honestly. These procedures are designed "to make people less likely to edit their responses to be socially desirable, lenient, acquiescent, and consistent with how they think the researcher wants them to respond," (Podsakoff et al., 2003, p. 888) thus reducing common method bias. We also heeded the warning that some scale items are more prone to common method bias than others. For this reason, we avoided as much as possible for our key dependent and independent variables the use of Likert scales with similar end points and formats as these similarities are likely to cause common method bias and to have anchoring effects. In addition, we conduct the single factor test in Harman (1967) to evaluate the extent to which common method bias is present in the data. If it is present, then either only a single factor will emerge or only one of several factors will account for the majority of covariance among the variables (see also, Abernethy et al., 2004). This test clearly rejects the hypothesis that common method bias is driving our results (chi-squared = 274.7; df = 77, p-value < 0.001) (Podsakoff et al., 2003).

Table 2, Panel A provides details about the characteristics of the firms in our sample. About 27% of the business units are in the service sector. Most business units have been part of their current parent company for a substantial period of time (mean = 17.50 years), although approximately 14% of the units were acquired within the past year (untabulated). The business units vary significantly in size; while the median number of employees is 290, the mean is much higher at 1,005. Approximately 10% of the business units have fewer than 25 employees (untabulated). On average, the units are responsible for 12.6% of total firm sales.

----------------- Insert Table 2 about here. -----------------

As Table 2, Panel B shows, the typical respondent is 43.85 years old and has been in their current job (business unit) for an average of 3.16 (5.93) years. The respondents are somewhat less experienced than their superiors, both in the industry (mean = –3.73 years) and in the firm (mean = –4.75 years).

3.2 Variable measurement

3.2.1. Dependent variables

Weighting of type of performance measure. In this section, we describe the dependent variables which proxy for the theoretical construct v_x^*/v_y^* analyzed in Section 2. Following earlier studies (Abernethy et al., 2004, Bouwens and van Lent, 2007), we provide respondents with a list of performance measures and ask them to indicate, in percentage terms, their superior's weighting of each measure in evaluating their annual performance. These measures include stock-price related, firm-wide, group, own-level (i.e., business unit), and below-level (i.e., specific) performance measures.[14]

We combined the separate weightings of stock-price related,[15] firm-wide, and group measures into a single "weighting of above-level measures." We have two reasons for doing so. First, our theory focuses on the use of above-level measures versus the use of own-level measures and does not further distinguish between different types of above-level measures. Without more detailed theoretical guidance, any exploration of the use of individual above-level measures would merely be ad hoc. Second, depending on the firm's structure, business units are not necessarily part of a group, and thus we found that many respondents assigned zero weight to group measures. Analyzing group measures separately from firm-wide measures would confound issues of firm structure with those of performance measurement. In addition, as we describe below, "zero-inflation" could cause problems in our estimation procedures, and combining categories to reduce the number of zeros is a recommended way of dealing with such inflation (Fry, Fry and McLaren, 1996).

Below-level measures are defined as non-summary performance measures of specific activities within the respondent's unit. Earlier work has proposed the idea that below-level measures can signal

[14] Our theoretical model in Section 2 does not include below-level measures, i.e., non-summary performance measures of specific activities within the respondent's unit (see below for more details). However, an augmented version of our model that admits these below-level measures yields similar inferences. For example, in a setting where below-level measures are aligned with business unit output but not affected by spillovers, the comparative statics of Table 1 continue to apply. More importantly, our result of an interaction of the two directions of spillover continues to hold in the augmented model. Detailed results are available upon request.

[15] Stock-price related measures are by definition associated with firm-wide performance and should therefore be treated as "above-level" measures. In addition, prior work has shown that stock-price related measures respond to interdependencies in very much the same way as other firm-wide measures (Bouwens and van Lent, 2007, Keating, 1997).

whether managers are acting in a cooperative fashion (Baiman and Baldenius (2009)) and that they can be used in an integrated way to show how managers are achieving the goals of the company (Bouwens and van Lent, 2007, Hirsch, 1994); excluding these measures could potentially bias our inference. We ask respondents about the full range of possible performance measures (ranging from above-level to below-level) to improve scale reliability. Our scale allows respondents to map actual weights into answer categories without having to rescale due to the omission of any one performance measure type. We also allowed respondents to fill in any other measures that they judged were not well represented by any of these categories (Fowler Jr., 1995).[16] Details are reported in Table 3, Panel A.

---------------- Insert Table 3 about here. -----------------

The annual performance of a business unit manager is given by: *above* · above-level performance + *own* · own-level performance + *below* · below-level performance, where the combined weighting of all three categories of performance measures (after reassigning the answers to the open-ended "remainder" category) in Table 3 Panel A must sum to 100 percent, *above* + *own* + *below* = 100%. The ratio of our variables yield proxies for our theoretical constructs; for example, *above/own* proxies for our theoretical construct of the incentive ratio v_x^*/v_y^* for the weighting on above-level and own-level performance (for observations with *below* = 0, theoretical model and empirical setting coincide).[17]

One empirical implication of the 100%-requirement is that the weightings of any two of the categories fully describe the distribution of weight among all three; that is, the weighting of any one of the categories of performance measures cannot be varied independently of the other two. In short, our data is "compositional." Statistical analysis of this kind of data requires that we explicitly incorporate the constant sum constraint into the model, which can be accomplished by using log-ratios of the proportions (Abernethy, Bouwens and van Lent, 2013, Aitchison, 1986).We use *Log(above/own)*, the natural logarithm of the weighting of above-level measures divided by the weighting of own-level measures, and

---

[16] These answers were reviewed separately by at least two of the authors and re-classified if possible to the remaining categories.

[17] Different to Bushman et al. (1995), our survey question for the percentage weighting of each performance measure yields proxies that are not confounded with expected performance.

Log(below/own), the natural logarithm of the weighting of below-level measures divided by the weighting of own-level measures, as the dependent variables in our tests. We rely on Aitchison's (1986) zero-replacement procedure to avoid the problem of dividing by zero and/or taking the logarithm of zero.

Our survey instrument has been tested in several earlier studies and is known to have good psychometric properties and construct validity (Abernethy et al., 2004, Bouwens and van Lent, 2007). Some of its more salient benefits are that the weightings correlate strongly with the use of measures in formula-based bonus contracts, which reduces the possibility that we are merely capturing "softer" perceptions; the measures can have equal weightings and do not have to be rank-ordered; and Likert scales, which tend to elicit how respondents feel about an issue rather than what they actually do, can be avoided.

3.2.2. Test variables

Spillovers. In this section, we describe the independent variables which proxy for the theoretical constructs, δ_i and δ_j, used in Section 2. Our theory suggests that the direction of spillovers is important. We can therefore not follow prior work (Bushman et al., 1995, Christie, Joye and Watts, 2003) which has constructed proxies for spillovers *within* the firm from somewhat coarse data taken from publicly available datasets as these proxies are not sufficiently refined to capture the direction of the spillover. Instead, we rely on Keating (1997) who uses single-item survey instruments to obtain a measure of the focal unit's effect on other units and of the impact of other units on the focal unit (see also, e.g., Abernethy et al., 2004). This choice also allows us to heed Luft and Shields' (2003) advice to use the same instruments across studies to improve our ability to foster understanding of the role of accounting in organizations. In addition, the Keating instrument corresponds very closely to the kind of spillovers we analyze in the theoretical part of this paper.

Specifically, we construct *Spillovers to other managers* from a survey question which asks respondents to indicate to what extent their unit's actions impact on work carried out in *other* organizational units of their firm; this variable proxies for our theoretical construct δ_i of the spillover from the focal manager's business unit *i* to business unit *j*. Similarly, *Spillovers from other managers* uses answers to the question to what extent actions of managers of other units within the firm impact work carried out in the respondent's unit; this variable proxies for our theoretical construct δ_j of the spillover from business unit *j* to the focal manager's business unit *i*. Respondents answer on a Likert-type scale that ranges from 1 (no impact at all) to 7 (a very significant impact); details are in Table 3, Panel B.

We use three other questions from the survey to test convergent validity of our test variables (following Abernethy et al., 2004). Details about the correlations underlying these validity tests are in Table 3, Panel C. First, we use the percentage of the focal unit's total production delivered to other units in the firm to assess the validity of *Spillovers to other managers*. The Spearman correlation between the two measures is 0.34 (p-value<0.01). We then use the percentage of the focal unit's total production that uses inputs sourced from other units within the firm as a validity test of *Spillovers from other managers* and find a correlation of 0.48 (p-value<0.01). Finally, we correlate both *Spillovers to other managers* and *Spillovers from other managers* with a survey question which asks how much time (as a percentage of total working time) the respondent spends on meetings with managers from other units in the firm (corr. with *Spillovers to other managers* is 0.48, p-value<0.01; corr. with *Spillovers from other managers* is 0.54, p-value<0.01) and with a question that asks to what extent the focal unit could operate as an *Independent business* (corr. with *Spillovers to other managers* is −0.47, p-value<0.01; corr. with *Spillovers from other managers* is –0.36, p-value<0.01). Together these tests suggest that our spillover metrics are valid.

Our theory shows that the two types of spillovers are likely to interact. We therefore mean-center *Spillovers to other managers* and *Spillovers from other managers* before multiplying them to construct the interaction term *To×From*. In our regressions, the coefficient on *Spillovers to other managers* (*Spillovers from other managers*) represents the effect of *Spillovers to other managers* (*Spillovers from other managers*) on the dependent variable, holding constant *Spillovers from other managers* (*Spillover to other managers)* at its sample average (Jaccard and Turrisi, 2003, Wooldridge, 2000).

Finally, we also consider a measure of spillovers that does not distinguish the direction of the effect. We use this *Unsigned spillovers* proxy to examine more closely our claim that it is essential to sign the direction of spillovers when trying to understand their effect on the weights placed on different types of performance measures. We construct *Unsigned spillovers* by summing the mean-centered values of *Spillovers to other managers* and *Spillovers from other managers* (Table 3, Panel B reports the descriptives of the non-mean centered variable). *Unsigned Spillovers* has a Cronbach's alpha, which is equal to 0.74.

3.2.3. Control variables

To define our set of controls, we follow earlier studies on the determinants of the use of performance measures in business units (Abernethy et al., 2004, Bouwens and van Lent, 2007, Bushman et al., 1995, Ittner and Larcker, 2001, Ittner et al., 1997, Keating, 1997). Table 4 presents summary statistics on these variables. *Decentralization* measures the difference in decision making authority between the respondent and his superior in five key areas: strategy, investments, marketing, internal operations, and human resources. The instrument is described in Abernethy et al. (2004) and Bouwens and van Lent (2007) and follows from an earlier proposal in Gordon and Narayanan (1984). A measure based on a composite of the above five items is correlated with five additional questions that in a yes/no format ask for detailed information about the respondent's decision making authority regarding investments.[18] Bouwens and van Lent (2007) suggest that these investment-decision questions can be used to check the validity of the *Decentralization* construct. Four out of five of the correlation coefficients between *Decentralization* and the five investment questions are positive and significant at the 5 percent level or better, and one coefficient has a *p*-value of 0.14, which can be interpreted as evidence of convergent validity.

[18] These questions are as follows. If my business unit needs a new building, I can decide to purchase, rent, or build one without my boss's prior consent. If my unit needs to replace durable equipment, I can do so without asking my boss for permission. If I need to extend the production capacity in my unit, I can decide to do so without asking my boss for permission. I can decide on the level of R&D activities in my unit. When developing new products, I can make my own investment decisions.

We use a two-item instrument taken from Abernethy et al. (2004) to capture the growth opportunities of the respondent's business unit and of the industry in which they compete (here, labeled *Growth opportunities*). We use the four-item instrument in Khandwalla (1972) to capture the competitive environment of the business unit (here, labeled *Competition*). *Varcomp* is the maximum percentage of total pay that is available as performance-dependent incentive pay (which can include cash bonuses, equity grants or stock options).

---------------- Insert Table 4 about here. ----------------

Size is the natural logarithm of the number of full time employees of a given business unit (see Table 2). Finally, we include an indicator variable (here, labeled *Service*) that takes the value of unity when the business unit operates in a service industry.

3.3 Model estimation

Our estimation has two steps. First, we use factor analysis to construct a weighted composite measure for each theoretical construct (Chenhall, 2005, Hair, Anderson, Tatham and Black, 1998). We submit all indicator variables to the factor analysis simultaneously to ensure that we obtain a clean factor score; the (untabulated) results suggest that the constructs exhibit good reliability and construct validity.[19]

Second, we estimate seemingly unrelated regressions using the log-ratios (above/own) and (below/own) of the three performance measures as the dependent variables. As noted above, the compositional nature of our data (i.e., the fact that the weightings of the three categories of performance measures must sum to unity) means that the percentage weight on any one performance measure cannot be varied independently of the weights placed on other measures. We model this dependency explicitly by taking log-ratios and by allowing the residuals from the two log-ratio regression equations to be correlated (Aitchison, 1986).[20]

[19] Our results, however, are not sensitive to this estimation choice. When we adopt a latent variable approach to deal explicitly with measurement error and provide evidence on construct validity as suggested by Ittner and Larcker (2001), neither the sign nor significance of our variables is affected. Specifically, we use Partial Least Squares to obtain estimates of the weightings used to create the latent variables scores and of the loadings that connect the latent variables with their associated manifest indicators (Chin and Newsted, 1999).
[20] The estimated correlation between the residuals of the two equations is 0.49.

To correct departures from normality and improve the size of the test, we use bootstrapping to compute the standard errors (i.e., 1,000 replications with replacement where all samples have the same size as the original sample) (Moon and Perron, 2008).

4. Empirical findings

We first discuss the summary statistics for our main variables as reported in Tables 3 and 4 as well as their Pearson correlations, which are reported in Table 5. Table 6 presents the estimation results of the seemingly unrelated regressions that examine the association between interdependencies and the use of above-level and below-level measures relative to own-level measures, and Table 7 presents evidence from a simulation of first differences (i.e., the difference between the expected value of the weight placed on above-, own-, and below-level performance measures for minimum and maximum spillover effects, respectively) for the purpose of exploring the nature of the interaction effect.

4.1 Summary statistics

We report descriptive statistics on the weightings of the above-, own-, and below-level performance measures in Table 3, Panel A. By far, the own-level measures receive the most weight on average (mean = 0.58). Above-level measures, on the other hand, are used frequently in our sample, and their average weighting in the performance evaluation of business unit managers is 0.30. In contrast, below-level, or specific, measures are not used in 54% of the companies (untabulated) and for that reason obtain an average weighting of 0.13. However, when we consider only those companies that do use below-level measures, the average weighting is much higher (mean = 0.28).

Panel B documents the summary statistics for our key interdependency variables *Spillovers to other managers* and *Spillovers from other managers*, as well as for the variables we use to establish convergent validity (*Supply to other managers, Supply from other managers, %Time-Meet,* and *Independent business*). The test variables span the full theoretical range of one to seven and show on average substantial intra-firm interdependency (i.e., the mean is approximately four for both spillover variables, signifying that spillovers "to some extent" affect the functioning of the business unit). The descriptive statistics for *Unsigned spillover* show that spillovers in firms vary from a virtual absence of any kind of interdependencies to the existence of reciprocal relations between units (the observed minimum and maximum values equals the theoretical range). We find similar variation in the two *Supply* variables, which measure the percentage of a business unit's total production that is delivered to other units within the firm. In addition, business unit managers spend between 0 and 40 percent of their time meeting with managers from other units in the firm (mean = 9.97). About 50% of the respondents report that their firm can operate as an independent business (outside of the current parent company) for a substantial part of their activities (median = 6).

As shown in Table 5, which presents the correlations among the variables in our study, the weightings of the three performance-measure categories (i.e., above-level, own-level, and below-level) are, as we would expect given that they are constrained to sum to unity, negatively associated. More importantly, we find that weightings of above-level measures are significantly positively associated with *Spillovers from other managers*, but not with *Spillovers to other managers*. Conversely, weightings of below-level measures are significantly positively correlated with *Spillovers to other managers*, but not with *Spillovers from other managers.* In line with these results, weightings of own-level measures are negatively associated with both types of interdependencies (albeit only weakly with *Spillovers from other managers*). Taken together, it appears that below-level and above-level measures are both used in response to increasing intra-firm dependencies, but that each type of measure is adapted to the different kinds of demands that arise from spillovers that affect, on the one hand, the focal unit and, on the other, the other unit. As our theory suggests a non-linear relation between spillovers and optimal incentive weights, however, these linear associations should be interpreted with some care.

----------------- Insert Table 5 about here. -----------------

4.2 Regression results

The empirical specification of the model is based on our theoretical framework. Specifically, our model predicts that the relative weightings (i.e., ratio of optimal incentive weights) of the above-level versus the own-level measures depend on the effect the focal unit has on other business units as well as on the effect the other business units have on the focal unit. In addition, because both types of spillover can reinforce one another, our model motivates the inclusion of an interaction term. Recall that we examine the use of above-level measures in relation with below-level measures for two reasons. First, our measurement scale obtains information about the weight on above-level measures relative to the weight on own-level and below-level measures. We wish to preserve this information in the estimations. Second, prior literature suggests that below-level measures might be helpful in dealing with spillovers (Baiman and Baldenius, 2009, Bouwens and van Lent, 2007).

Table 6 presents our seemingly unrelated regression results. First, we estimate a regression (presented in the Column with heading "Model (1)") which includes the *Unsigned spillovers* proxy together with a comprehensive set of control variables. The dependent variables are the log-ratios (above/own) and (below/own) that together describe the full spectrum of performance measure choices. This model is motivated by our desire to examine the empirical importance of the distinction we have drawn in our theory between spillovers *to* others (δ_i) and spillovers *from* others (δ_j). Consistent with the idea that it is essential to allow different directions of spillovers to affect the choice of performance measures differently, we find no significant association between *Unsigned spillovers* and either the log-ratio (above/own) or the log-ratio (below/own).[21]

----------------- Insert Table 6 about here. -----------------

In addition, we find that some of our control variables have explanatory power. Increasing the proportion of performance-dependent incentive pay in a manager's total compensation package reduces the weightings of below-level performance measures relative to the weightings of own-level measures. We also find that the evaluation of managers in more decentralized business units tends to be based on own-level measures more than on above-level measures. We find qualitatively similar results for the control variables in the remaining models, described below, which we will not discuss again.

[21] To reconcile this finding with Bushman et al. (1995), we regress the weight on above-level measures on *Unsigned spillovers* (that is, we ignore the compositional nature of our dependent variable). Consistent with this prior study's findings, we obtain a positive and marginally significant coefficient, which implies that higher levels of (unsigned) spillovers are associated with more weight on above-level measures.

Next, we estimate a regression that includes the two signed spillover proxies (i.e., *Spillovers to other managers* and *Spillovers from other managers*) but does not allow their effect to be non-linear (as we exclude their interaction). Note that if the true data-generating process is non-linear, as is suggested by our theoretical analysis, then this regression is misspecified and its findings should be interpreted with care.[22] The results are presented in Table 6 in the columns with the heading "Model (2)".

We find no significant association between either of the two signed spillover proxies and the log-ratio (above/own). However, we do find that *Spillovers from other managers* is positively and significantly associated with the log-ratio (below/own), implying that when the actions of managers of other business units spillover to the focal manager's business unit, more weight is placed in the focal manager's performance assessment on below-level measures. We also find that *Spillovers to other managers* is negatively associated with the log-ratio (below/own), albeit marginally (p-value = 0.11). Thus, as the focal manager's actions increasingly affect others in the firm, the weight on below-level measures in the former manager's performance evaluation decreases relative to the weight on own-level measures. Together the results of Models (1) and (2) suggest that lumping the two types of spillovers into one proxy might confound inferences. Indeed, the two types of spillover carry opposite signs in the association with the relative weight on own-level and below-level performance measures.

We estimate the full model as suggested by our theory in the final columns of Table 6 ("Model (3)"); in these regressions, we include both signed spillover proxies and their interaction to allow for non-linear relations. The interaction term *To × From* is strongly significant in the regression with the log-ratio (above/own) as the dependent variable. At the same time, the signed spillover proxies are not significant (as in Model (2)). The significant interaction term validates the idea that the relation between weight on performance measures and spillovers is non-linear. In the current parameterization (i.e., with mean centered spillover proxies), the findings imply that holding *Spillovers to other managers* (*Spillovers from*

[22] The bias in the estimated coefficients, however, might be small. This is due to the mean-centering of *Spillovers to other managers* and *Spillovers from other managers*. If the "true" model is $Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 X \times Z + \varepsilon$ and a researcher estimates mistakenly the following model, $Y = \gamma_0 + \gamma_1 X + \gamma_2 Z + \varepsilon$, then the estimated coefficient $\hat{\gamma}_1$is typically close to $\widehat{\beta_1} + \widehat{\beta_3}\bar{Z}$, where $\bar{Z}$ is the sample average of Z (Sorensen and Ozer-Balli, 2011). Mean centering implies that $\bar{Z} = 0$, so $\hat{\gamma}_1 \approx \widehat{\beta_1}$.

other managers) constant at its sample average, *Spillovers from other managers* (*Spillovers to other managers)* is not associated with the weight placed on above-level measures relative to the weight on own-level measures.[23] While there is no significant relation between the spillover proxies and the log-ratio (above/own) when considering average values of spillovers, as the relation is non-linear, a completely different picture might emerge when considering other values of the empirical distribution of spillovers. We will return to this possibility when we discuss our simulation results below.

We also find a strongly significant interaction term *To* × *From* in the regression with the log-ratio (below/own) as the dependent variable. At the same time, as before in Model (2), we find that *Spillovers from other managers* (*Spillovers to other managers*) is positively (marginally negatively) related to the relative weight placed on below-level versus own-level measures. Thus, we find again that the relation between spillovers and the use of performance measures is non-linear. While evaluated at the sample mean, each type of spillover is significantly associated with the log-ratio, the non-linearity of the relation calls for a closer inspection of what happens at different values of the empirical distribution of spillovers. As mentioned before, we do so in the simulations described below.

Overall, our results in Table 6 underscore the importance of separating the two directions of spillover and allowing for an interaction between the two directions of spillover in understanding the role of different types of performance measures. Table 6 also provides a possible reconciliation of the mixed results in prior work on the association between spillovers and the weightings of performance measures. Regressions that do not separate the two directions of spillover (as in Bushman et al., 1995) or that do not allow for an interaction between the two directions of spillover (as in Keating, 1997 and Abernethy et al., 2004), are potentially misspecified and should be interpreted with care. In turn, these results highlight the usefulness of an analytical model to provide a framework for the empirical analysis.

---

[23] Specifically, suppose we have the following equation: $Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 X \times Z + \Sigma\beta_j \text{Controls}_j + \epsilon$. The marginal effect of X on Y when Z is held at its sample minimum ($\underline{Z}$) is defined as $\partial Y/\partial X|_{Z=\underline{Z}}$, which equals $\beta_1 + \beta_3\underline{Z}$. Clearly, owing to the interaction term, β_1 only describes the marginal effect of X on Y when Z = 0, which equals the sample mean of Z due to our mean-centering procedure.

Despite the support of our theory, the estimation results in Table 6 are hard to interpret regarding what happens to the (relative) weighting of each performance measure as intra-firm spillovers vary. For example, the estimation results in Table 6, at first glance, appear to suggest that *Spillovers from other managers* and *Spillovers to other managers* do not affect the weighting placed on above-level measures directly, but only through the interaction term. This inference, however, would be incorrect. While our regressions are econometrically justified, three complications somewhat obfuscate the intuition behind them. First, we estimate a system of inter-related equations, which implies that the effects demonstrated in any one equation cannot be evaluated in isolation. Second, we model an interaction term that allows *Spillovers to other managers* and *Spillovers from other managers* to affect the relative use of each performance measure both directly and indirectly. Third, our dependent variable is a log-ratio and not the percentage weights placed on the performance measures.

In other words, using our regression results to answer questions of practical relevance—such as, what happens to the weighting of above-level measures when the firm wants to increase collaboration between units—is not straightforward. Indeed, consider the implications of increasing *Spillovers to other managers* in the regression equations. Not only would this affect both log-ratios directly, but through the interaction term, it would also affect them indirectly. Then, given the compositional nature of the dependent variables, changes in one log-ratio (say, (below/own)) will have consequences for the other log-ratio (in this case, (above/own)). Taking all of these distinct effects into account, it is hard to foresee how increasing *Spillovers to other managers* would ultimately affect the weight placed on the above-level measures. We aim to provide a more intuitive interpretation of our estimation results in the simulations, to which we turn our attention next.

4.3 Simulation results

Ultimately, we are interested in reconciling our empirical findings in Table 6 with how the percentage weighting of each performance measure changes as intra-firm spillovers vary. These relations cannot be simply inferred from the results in Table 6, as, e.g., the non-linear specification implies that the partial derivative of the log-ratio (above/own) with respect to *Spillovers to other managers* (*Spillovers* *from other managers*) depends on the level of *Spillovers from other managers* (*Spillovers to other managers*). Indeed, both sign and significance of the partial derivative may change as one considers different levels of the interacting variable (Jaccard and Turrisi, 2003).

As a result, we follow Abernethy et al. (2013) and conduct simulations of first differences, that is, the difference in the expected value of the weight placed on a performance measure (i.e., above-level, own-level, or below-level) when the value of *Spillovers to other managers* (*Spillovers from other managers*) is changed from its sample minimum to its sample maximum, while holding *Spillovers from other managers* (*Spillovers to other managers*) constant at its sample maximum value.[24] All other explanatory variables are set at their mean. Simulations explicitly deal with the estimation uncertainty present in our empirical model and provide an intuitive approach to statistical interpretation (King et al., 2000). We assess the significance of the first difference using its empirical distribution from the simulation.

Table 7 summarizes the findings from the analysis of first differences. Cell entries represent the average (computed over 1,000 simulations) of the weight placed on the associated performance measure for minimum and maximum values of *Spillovers to other managers* (*Spillovers from other managers*), respectively. The rows labeled "First diff." present the statistic of interest, i.e., the first difference.

----------------- Insert Table 7 about here. -----------------

[24] Specifically, the simulation involves taking *M* draws from the multivariate normal distribution with mean $\hat{\beta}$, the coefficient matrix from the seemingly unrelated regression model, and variance matrix $V(\hat{\beta})$, the estimated variance-covariance matrix for the coefficients in the model. We simulate first differences by setting the value for *Spillovers to other managers* (*Spillovers from other managers*) to its sample minimum, and *Spillovers from other managers* (*Spillovers to other managers*) to its sample maximum. We use the maximum value for *Spillovers from other managers* (*Spillovers to other managers*) to ensure that the potential effect of having two types of spillovers (to and from) is captured in the simulations. We keep all other variables constant at their means. We then generate the expected value of the transformed outcome variable (the log-ratio) conditional on these starting values for the explanatory variables by taking one draw from the normal distribution. We use the inverse logistic function to transform the log-ratio into the original scale of percentage weights. Next, we set the value for *Spillover from other managers* (*Spillovers to other managers*) to its sample maximum and *Spillovers to other managers* (*Spillovers from other managers*) to its sample maximum. Again, all other variables are set at their means. We generate the expected value of the transformed log-ratio conditional on these ending values for the explanatory variables. The first difference is simply the difference between these two expected values of the transformed log-ratio. We then repeat this procedure 1,000 times to approximate the distribution of first differences (see, Abernethy et al., 2013, King, Tomz and Wittenberg, 2000, Zelner, 2009).

Turning first to the performance measures considered in prior studies (Bushman et al., 1995; Keating, 1997; Abernethy et al., 2004), i.e., above-level and own-level measures, we find that spillovers have economically very meaningful consequences for incentive weight choices. Specifically, changing *Spillovers from other managers* from its minimum to its maximum value increases the weight placed on above-level measures from 9 to 60 percent. The first difference (52 percent) is significant at the 1 percent level. This observation is consistent with the argument that a noise effect will produce a positive relation between the spillover from the other business unit and the relative weight placed on aggregate firm output.

A similar change in *Spillovers to other managers* increases the weight placed on above-level measures from 11 to 60 percent. This first difference (equal to 49 percent) is also significant at the 1 percent level. Based on the model in Section 2, for this type of spillover, the argument can support both an increase and a decrease in the optimal weight on aggregate firm output relative to business unit output. The noise effect yields a negative association between the spillover to the other business unit and the optimal incentive ratio, whereas the congruity effect can produce a negative as well as a positive association, and the association is positive when the spillover from the other business unit is large to create a sufficiently noisy business unit output. Our observation of a strong positive association is consistent with the case in which congruity has a positive effect on the optimal incentive ratio that is sufficiently large to dominate the negative effect arising from having a noisy aggregate firm output.[25,26]

The model in Section 2 does not provide theoretical predictions for what should happen with regard to below-level measures. We find, however, that the weight placed on below-level measures increases when spillovers grow. Specifically, changing *Spillovers from other managers* from its minimum to its maximum value increases the weight on below-level measures by 17 percent (p-value < 0.01). We

[25] It is instructive to compare this finding with Keating's (1997) study, which also allows signed spillover effects. Keating finds that spillovers from other managers reduce the use of above-level measures whereas spillovers on other managers increase the use of above-level measures). In contrast, based on the estimation derived from our full model, both types of spillovers increase the use of above-level measures.

[26] Abernethy et al. (2004) report that own-level measures increase with the *Spillovers from other managers* and decrease with the *Spillovers to other managers* (using our construct labels). We find, based on our full model, that own-level measures receive lower weightings only when both types of spillover are high.

also find an increase of the weight placed on below-level measures when considering the same change in *Spillovers to other managers*; this increase (10 percent), however, does not attain significance at conventional levels.

Taken together, the simulation results suggest a clear picture of how firms cope with spillovers between managers. Above-level measures are used to a greater extent when either type of spillovers (i.e., *from* or *to* other managers) becomes more prevalent. To a lesser extent, firms use below-level measures, especially when coping with increased *Spillovers from other managers*.

It is worthwhile to point out, however, that our simulations focus on what happens if both spillover effects are strong at the same time. Table 7 also documents that own-level measures are given very significant weights if the spillover effect is only strong in one direction. For example, when *Spillovers to other managers* are at their sample maximum value (and *Spillovers from other managers* is at the sample minimum), the expected value of the weighting on own-level measures is 91 percent. This finding suggests that senior management does not employ noisy (and thus costly) above-level measures unless spillovers create significant issues in the cooperation between units in the firm.

4.4 Additional analyses

Some recent studies suggest that information asymmetry is an important determinant of which performance measures are used to evaluate managers (Abernethy et al., 2004, Bouwens and van Lent, 2007, Hwang, Erkens and Evans, 2009, Raith, 2008). Given that these earlier studies document a high correlation between decentralization and information asymmetry,[27] we include only *Decentralization* in our regressions. Nevertheless, when we include *Information Asymmetry*, which we measure using the instrument in Dunk (1993), in the regressions, the findings remain virtually unchanged and none of our test variables change in either sign or significance.[28]

---

[27] In their Table 5, Bouwens and van Lent (2007), for example, report that the correlation coefficient between decentralization and information asymmetry is 0.43.

[28] The instrument in Dunk (1993) asks respondents to compare their information about the work carried out in the business unit to that of their superior and consists of five questions, which are reproduced in Appendix 2.

5. Discussion and conclusions

Roberts (2004) describes the problem of how firms encourage cooperation between managers as one of the foremost issues in organizational design. He also points out that designing incentive systems that motivate managers to do the right thing is no trivial task. In firms with multiple business units, doing the right thing requires managers to cooperate with each other and, specifically to be aware of how one's actions influence the performance of other managers in the firm. Performance measures can play an important role in highlighting such spillover effects. We provide theory and empirical evidence on the existence of two types of spillover and show that each type of spillover can affect performance measures differently. For this reason, we expect the weightings of these performance measures to depend on the degree to which each type of interdependency is present in the firm and, more importantly, on whether these interdependencies are present at the same time. Indeed, as expected, we find that the association between each type of spillover and the weighting of a performance measure depends critically on the magnitude of the effect of the other type of spillover. In addition, we find that both above-level and below-level measures receive more weight than own-level measures when the focal manager's actions are affected by other units in the firm. Above-level measures are also used more in response to higher levels of spillovers deriving from the actions of the focal managers which impact on other managers. Based on our theory, this suggests that for these spillovers (i.e., originating from the focal managers and affecting others in the firm), having a congruous performance measure (which promotes that managers work in the best interest of the whole firm) is more important than having a precise, i.e., less risky measure.

The role of below-level measures in dealing with spillovers has not been documented before (although some prior studies suggest that below-level measures can be effective in responding to knowledge asymmetry problems, see, e.g., Raith (2008) and Hwang, Erkens and Evans (2009)), but has important implications for designing performance evaluation systems in practice.

We draw our empirical evidence from a survey of business unit managers, which supplies us with new data appropriate to the level of analysis required to conduct research into intra-firm interdependencies. We use information gathered directly from business unit managers who are affected by these interdependencies and/or are encouraged to work with other managers to overcome the problems such interdependencies cause. These respondents also provide us with data about the performance measures their seniors use when evaluating their performance. As such, in our setting, the benefits of using survey data outweigh the potential validity-related costs. To offset these potential costs and further validate our findings, we employ several safeguards when selecting the sample, phrasing the survey questions, administering the survey, and analyzing the respondents' answers.

Overall, our study attempts to understand why some important prior work has documented conflicting empirical evidence on how spillovers affect the use of performance measures in incentive systems. Our answer highlights that spillovers are "directional"; that is, for the purpose of performance measurement, it matters a great deal whether spillovers affect a manager or whether it is the manager's actions, which "impact" on the performance of others in the firm. What's more, spillovers have both direct and indirect effects on accounting performance measures, as above-level measures are computed by summing up own-level measures. Thus, any spillover effects on own-level measures can be compounded once the own-level measures of multiple business units are aggregated into a single above-level measure.

Empirically nor in theory do we address the possibility that firms use non-financial "above-level" measures to deal with spillover effects. Non-financial measures potentially do not face the "summing up" property of accounting performance measures and could therefore respond differently to spillovers. Exploring this possibility is left for future research.

References

M.A. Abernethy, Bouwens, J., van Lent, L. 2004. Determinants of control system design in divisionalized firms. *The Accounting Review*. **79**(3) 545-570.

M.A. Abernethy, Bouwens, J., van Lent, L. 2013. The role of performance measures in the intertemporal decisions of business unit managers. *Contemporary Accounting Research*. **in press**.

J. Aitchison. 1986. *The statistical analysis of compositional data*. Chapman and Hall, London.

S. Baiman, Baldenius, T. 2009. Non-financial performance measures as coordination devices. *The Accounting Review*. **84**(2) 299-330.

J. Bouwens, van Lent, L. 2007. Assessing the performance of business unit managers. *Journal of Accounting Research*. **45**(4) 667-697.

R.M. Bushman, Indjejikian, R.J., Smith, A. 1995. Aggregate performance measures in business unit manager compensation: The role of intrafirm interdependencies. *Journal of Accounting Research*. **33**(Supplement) 101-127.

R.H. Chenhall. 2005. Integrative strategic performance measurement systems, strategic alignment of manufacturing, learning and strategic outcomes: an exploratory study. *Accounting, Organizations and Society*. **30**(5) 395-422.

W.W. Chin, Newsted, P.R. 1999. *Structural equation modeling analysis with small samples using partial least squares*. Sage, City.

A.A. Christie, Joye, M., Watts, R. 2003. Decentralization of the firm: theory and evidence. *Journal of Corporate Finance*. **9**(1) 3-36.

A. Dunk. 1993. The effects of budget emphasis and information asymmetry on the relation between budgetary participation and slack. *The Accounting Review*. **68**(2) 400-410.

G. Feltham, Hofmann, C. 2007. Limited commitment in multi-agent contracting. *Contemporary Accounting Research*. **24** 345-375.

F.J. Fowler Jr. 1995. *Improving survey questions*. Sage, Thousand Oaks.

J.M. Fry, Fry, T.R.L., McLaren, K.R. 1996. Compositional data analysis and zeros in micro data. Working paper, Monash University.

L. Gordon, Narayanan, V. 1984. Management accounting systems, perceived environmental uncertainty and organizational structure: an empirical investigation. *Accounting, Organizations and Society*. **9**(1) 33-47.

J.F. Hair, Anderson, R.E., Tatham, R.L., Black, W.C. 1998. *Multivariate data analysis*. Prentice Hall, Upper Saddle River, NJ.

H.H. Harman. 1967. *Modern factor analysis*. University of Chicago Press, Chicago, IL.

M.L. Hirsch. 1994. *Advanced management accounting*. South-Western Publishing, Cincinnati, Ohio.

B. Holmstrom, Milgrom, P. 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*. **55**(2) 303-328.

B. Holmstrom, Milgrom, P. 1990. Regulating trade among agents. *Journal of Institutional and Theoretical Economics*. **146** 85-105.

Y.C. Hwang, Erkens, D.H., Evans, J.H. 2009. Knowledge Sharing and Incentive Design in Production Environments: Theory and Evidence. *Accounting Review*. **84**(4) 1145-1170.

H. Itoh. 1992. Cooperation in hierarchical organizations: an incentive perspective. *Journal of Law, Economics and Organization*. **146** 321-345.

C.D. Ittner, Larcker, D.F. 2001. Assessing empirical research in managerial accounting: a value-based management perspective. *Journal of Accounting & Economics*. **32**(1-3) 349-410.

C.D. Ittner, Larcker, D.F., Rajan, M.V. 1997. The choice of performance measures in annual bonus contracts. *Accounting Review*. **72**(2) 231-255.
J. Jaccard, Turrisi, R. 2003. *Interaction effects in multiple regression*. Sage Publications, Thousand Oaks.
A.S. Keating. 1997. Determinants of divisional performance evaluation practices. *Journal of Accounting and Economics*. **24** 243-274.
P. Khandwalla. 1972. The effect of different types of competition on the use of management controls. *Journal of Accounting Research*. **10**(2) 275-285.
G. King, Tomz, M., Wittenberg, J. 2000. Making the most of statistical analysis: Improving interpretation and presentation. *American Journal of Political Science*. **44**(2) 341-355.
W.N. Lanen. 1995. Discussion of aggregate performance measures in business unit manager compensation: The role of intrafirm dependencies. *Journal of Accounting Research*. **33**(Supplement) 129-134.
J. Luft, Shields, M. 2003. Mapping management accounting research: graphics and guidelines for theory-consistent empirical research. *Accounting, Organizations and Society*. **28**(2-3) 169-250.
H.R. Moon, Perron, B. 2008. *Seemingly unrelated regressions*. Palgrave Macmillan, City.
J.C. Nunnally, Bernstein, I.H. 1994. *Psychometric theory*. McGraw-Hill, New York.
P.M. Podsakoff, MacKenzie, S.B., Lee, J.Y., Podsakoff, N.P. 2003. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*. **88**(5) 879-903.
M. Raith. 2008. Specific knowledge and performance measurement. *RAND Journal of Economics*. **39**(4) 1059-1079.
J. Roberts. 2004. *The modern firm: Organizational design for performance and growth*. Oxford University Press, Oxford.
B.E. Sorensen, Ozer-Balli, H. 2011. Interaction effects in econometrics. Working-paper, University of Houston
J.M. Wooldridge. 2000. *Introductory econometrics: a modern approach*. South-Western College Publishing.
B.A. Zelner. 2009. Using simulation to interpret results from logit, probit, and other nonlinear models. *Strategic Management Journal*. **30**(12) 1335-1348.
J.L. Zimmerman. 2003. *Accounting for decision making and control*. Irwin, Homewood.

APPENDIX 1

Proof of Lemma 1:

The certainty equivalent of agent i is

$$CE_i(f_i, v_{xi}, v_{yi}, a_{i1}, a_{i2}, a_{j1}, a_{j2}) = f_i + v_{xi}\, \mathrm{E}[x|a_{i1}, a_{i2}, a_{j1}, a_{j2}] + v_{yi}\, \mathrm{E}[y_i|a_{i1}, a_{i2}] - ½\, (a_{i1}^2 + a_{i2}^2)$$

$$- ½\, r\, \{v_{xi}^2\, \sigma_x^2 + v_{yi}^2\, \sigma_i^2 + 2\, v_{xi}\, v_{yi}\, \sigma_{ix}\}, \quad i,j = \text{A, B and A} \neq \text{B}. \qquad (A.1)$$

The first order conditions on agent i's actions, a_{i1} and a_{i2}, are

$$b_{ik}\, v_{xi} + \mu_{ik}\, v_{yi} - a_{ik} = 0, \quad k = 1,2. \qquad (A.2)$$

Solving equations (A.2) for a_{ik} gives (3).

Setting equation (A.1) equal to zero (based on (2b)) and solving for f_i gives the optimal fixed salary. Substituting, for each agent i, f_i^* plus a_{i1}^* and a_{i2}^* into the objective function (2a) yields the principal's unconstrained decision problem. Taking first-order conditions with respect to v_{xi} and v_{yi} and solving the equations simultaneously for these variables gives the optimal weightings of the performance measures, i.e.,

$$v_{xi}^* = D^{-1}\, [\lambda_i^2 + r\, \Sigma_k\, b_{ik}\, (b_{ik} - \sigma_{ix} / \sigma_i^2\, \mu_{ik})\, \sigma_i^2], \text{ and} \qquad (A.3a)$$

$$v_{yi}^* = D^{-1}\, r\, \Sigma_k\, b_{ik}\, (\mu_{ik} - \sigma_{ix} / \sigma_x^2\, b_{ik})\, \sigma_x^2, \qquad (A.3b)$$

where $D = \lambda_i^2 + r\, [\Sigma_k\, b_{ik}\, (b_{ik} - \sigma_{ix} / \sigma_i^2\, \mu_{ik})\sigma_i^2 + \Sigma_k\, \mu_{ik}\, (\mu_{ik} - \sigma_{ix} / \sigma_x^2\, b_{ik})\, \sigma_x^2 + r\, (\sigma_i^2\, \sigma_x^2 - \sigma_{ix}^2)]$. The optimal incentive ratio (4) follows from substituting (A.3a) and (A.3b) into $IR_i^* = v_{xi}^*/v_{yi}^*$.

Proof of Proposition 1:

The following comparative statics hold (with i,j=A,B, $i \neq j$, and k = 1,2):

(i) $$\frac{\partial IR_i^*}{\partial m_{ik}} = \frac{\left(-2\frac{\lambda_i}{\sigma_i^2}\mu_{in} + r\,(b_{ik} + m_{ik})\right)}{r\left(\Sigma_t\, b_{it}\left(\mu_{it} - \frac{\sigma_i^2}{\sigma_j^2} m_{it}\right)\right)} \frac{\sigma_i^2}{\sigma_j^2} - \frac{\left(\mu_{ik} - \frac{\sigma_i^2}{\sigma_j^2}(b_{ik} + m_{ik})\right)\left(\frac{\lambda_i^2}{\sigma_i^2} + r\,\Sigma_t\, b_{it} m_{it}\right)}{r\left(\Sigma_t\, b_{it}\left(\mu_{it} - \frac{\sigma_i^2}{\sigma_j^2} m_{it}\right)\right)^2} \frac{\sigma_i^2}{\sigma_j^2},$$ with n, t = 1,2 and $n \neq k$. The sign of this partial derivative is generally ambiguous. However, with $\sigma_j^2 = \sigma_i^2$, the partial derivative simplifies to

$$\frac{\partial IR_i^*}{\partial m_{ik}} = \frac{\mu_{ik}\left(-2\frac{m_{in}}{\sigma_i^2}\mu_{in} + r\right)}{r(\mu_{ik}^2 + \mu_{in}^2 - m_{in}^2)} \text{ at } m_{ik} = 0,$$

implying that with $m_{in}^2 < \mu_{ik}^2 + \mu_{in}^2$, $\partial IR_i^*/\partial m_{ik} > (<)\ 0$ if $r\sigma_i^2 > (<)\ 2m_{in}\mu_{in}$.

(ii) $\frac{\partial IR_i^*}{\partial m_{jk}} = 0;$

(iii) $\frac{\partial IR_i^*}{\partial \sigma_j^2} = -\frac{(\sum_k b_{ik}\mu_{ik})\left(\frac{\lambda_i^2}{\sigma_i^2} + r\sum_k b_{ik}m_{ik}\right)}{r\left(\sum_k b_{ik}\left(\mu_{ik} - \frac{\sigma_i^2}{\sigma_j^2}m_{ik}\right)\right)^2}\frac{\sigma_i^2}{\sigma_j^4} < 0;$

(iv) $\frac{\partial IR_i^*}{\partial \sigma_i^2} = -\frac{(\sum_k b_{ik}m_{ik})\left(\frac{\lambda_i^2}{\sigma_j^2} + r\sum_k b_{ik}\mu_{ik}\right)}{r\left(\sum_k b_{ik}\left(\mu_{ik} - \frac{\sigma_i^2}{\sigma_j^2}m_{ik}\right)\right)^2}\frac{1}{\sigma_j^2} > 0.$

APPENDIX 2

Appendix 2 reproduces the questionnaire items we used to construct our variables.

Weightings of above-level, own-level, and below-level measures
We are interested in the performance measures your superior uses to evaluate your annual performance. For each of the measures below, indicate your superior's weighting of each measure when he or she formally evaluates your annual performance. Your answers must sum to 100%.

1. Measures related to stock-price
2. Measures that summarize the performance of the whole company (e.g., firm-wide net income, firm-wide return-on-assets).
3. Measures that summarize the joint performance of the group in the firm of which your unit is part (e.g., group profit, divisional return-on-investment).
4. Measures that summarize the performance of your business unit (e.g., business unit profit, business unit return-on-investment).
5. Measures that summarize the performance of *specific* activities within your unit (e.g., sales of the marketing department, average costs of manufacturing, R&D expenses).
6. Other measures, please specify:

Above-level = (1) + (2) + (3)
Own-level = (4)
Below-level = (5)

Interdependencies
Spillovers to other managers
To what extent do your actions affect the performance of other units in the firm?

Spillovers from other managers
To what extent do the actions of other managers in the firm affect the performance of your unit?

Scale: 1 = not at all
4 = to some extent
7 = a great deal

Supply from others
What percentage of your unit's total production (services) uses products (services) supplied by other units in the firm?

Supply to others
What percentage of your unit's manufactured products (services) is supplied to other units in the firm?

%Time-Meet
What percentage of your total time available in the most recent month did you spend on meeting with managers from other units in the firm?

Independent business
To what extent could your unit operate as an independent company (i.e., detached from your current firm) in the marketplace?

Scale: 1 = not at all

4 = for about 50% of our activities
7 = for all of our activities

Decentralization
Please compare your influence in making decisions with the influence of your superior. If you or your subordinates in your unit make decisions without prior consent of your superior, you are considered to have complete influence.

1. Strategic decisions
2. Investment decisions
3. Marketing decisions
4. Decisions on internal processes
5. Human resource decisions

Scale: 1 = I have complete influence
4 = my superior and I share influence almost equally
7 = my superior has complete influence

Scale is reverse coded

Competition
Please indicate the degree of competition your unit faces with regard to the following:
1. Prices
2. Marketing and distribution
3. Quality of products
4. Product mix

Scale: 1 = almost no competition
4 = moderate competition
7 = strong competition

Growth opportunities
Please indicate your expectations about the following:
1. The growth opportunities that exists within the industry in which you compete.
2. The growth opportunities your unit faces.

Scale: 1 = significant decline
4 = no growth
7= significant increase

Size
How many people work in your unit (in full time equivalents)?

Varcomp

Your total compensation may vary with your performance. Please indicate the maximum amount (as a percentage of your fixed salary) available as performance-dependent pay (either as cash bonus or as stocks or options).

Information asymmetry
Please compare the amount of information you have relative to your superior.

1. Of you and your superior, who is in possession of better information regarding the activities undertaken in your unit?
2. Of you and your superior, who is more familiar technically with the work of your unit?
3. Of you and your superior, who is more certain of the performance potential of your unit?
4. Of you and your superior, who is better able to assess the potential effect factors external to your unit may have on your activities?
5. Of you and your superior, who has a better understanding of what can be achieved in your unit?

Scale: 1 = my superior does
4 = my superior and I do, almost equally
7 = I do

T A B L E 1

Comparative Statics of Optimal Relative Incentive Weights, IR_i^*, for Agent i =A,B

$\partial IR_i^* / \partial o$	**Congruity effect**	**Noise effect**
Spillover to other business unit ($o = \delta_i$)	$\gtreqless 0$	< 0
Spillover from other business unit ($o = \delta_j$)	0	> 0

TABLE 2

Summary Statistics on Business Units in the Sample and on Survey Respondents

The sample consists of 122 observations, and information is collected via a questionnaire. The survey respondents are business unit managers.

Panel A: Characteristics of the business unit

Variable	**Mean**	**Std. Dev.**	**Min.**	**Median**	**Max.**
Service Industry (1 = service, 0 = manufacturing)	0.27	0.45	0.00	0.00	1.00
Longevity of business unit in firm (in years)	17.50	9.88	0.00	20.00	39.00
Size of the business unit (measured in number of full-time employees	1,005.03	1,479.19	9.00	290.00	6,500.00
Relative size of business unit in firm (as a % of total sales)	12.62	17.55	0.00	5.00	100.00

Panel B: Respondent characteristics

Variable	**Mean**	**Std. Dev.**	**Min.**	**Median**	**Max.**
Age	43.85	8.13	25.00	44.00	63.00
Tenure in current job	3.16	3.55	0.00	2.00	30.00
Longevity in business unit	5.93	6.50	0.00	4.00	43.00
Experience in industry compared with that of superior(*)	–3.73	11.97	–30.00	–3.00	30.00
Experience in firm compared with that of superior(*)	–4.75	11.09	–27.00	–5.00	42.00

(*) Negative numbers indicate that the respondent has less experience than his/her superior does.

T A B L E 3

Summary Statistics on Weightings of Different Types of Performance Measures and Interdependency variables

Panel A: Summary statistics for weightings of above-level, own-level, and below-level performance measures. N = 122

	Mean	Std. Dev.	Min.	Median	Max.
Weightings of performance measures					
Above	0.30	0.28	0.00	0.25	1.00
Own	0.58	0.30	0.00	0.50	1.00
Below	0.13	0.17	0.00	0.00	0.80
Weightings of performance measures for those companies using that measure (weight > 0)					
Above (N = 87)	0.41	0.24	0.06	0.35	1.00
Own (N = 113)	0.62	0.27	0.05	0.50	1.00
Below (N = 56)	0.28	0.15	0.04	0.30	0.80

Panel B: Summary statistics for Spillovers to other managers, Spillovers from other managers, and corresponding convergent validity test variables. N = 122. We report statistics before mean-centering the variables.

Interdependency variables	**Mean**	**Std. Dev.**	**Min.**	**Median**	**Max.**
Spillovers to other managers	4.02	1.52	1.00	4.00	7.00
Spillovers from other managers	4.12	1.55	1.00	4.00	7.00
Supply to other managers (in %)	15.89	27.20	0.00	5.00	100.00
Supply from other managers (in %)	26.45	31.95	0.00	10.00	100.00
%Time-Meet	9.97	8.24	0.00	8.50	40.00
Independent business	4.86	2.05	1.00	6.00	7.00
Unsigned spillovers	8.15	2.73	2.00	8.00	14.00

Panel C: Spearman correlations among the interdependency variables. N=122

	(1)	(2)	(3)	(4)	(5)	(6)
1. Spillovers to other managers	1					
2. Spillovers from other managers	0.58	1				
3. Supply to other managers (in %)	0.34	0.22	1			
4. Supply from other managers (in %)	0.09	0.48	0.03	1		
5. %Time-Meet	0.48	0.54	0.56	0.36	1	
6. Independent business	-0.47	-0.36	-0.36	-0.35	-0.41	1
7. Unsigned spillovers	0.87	0.88	0.31	0.36	0.57	-0.49

T A B L E 4

Summary Statistics for Questionnaire Items Used to Construct the Control Variables

Table 4 presents the distribution of each questionnaire item used to construct the control variables in this study. The sample consists of 122 observations, and information is collected via a questionnaire. The theoretical range for the manifest variables associated with *Decentralization, Growth opportunities,* and *Competition* is 1–7. Full details about the survey instruments are provided in Appendix 2.

Survey items	**Mean**	**Std. Dev.**	**Min.**	**Median**	**Max.**
Decentralization: Please compare your influence on decision-making with your superior's influence on decision-making. If you or your subordinates in your unit make decisions without your superior's prior consent, you are considered to have complete influence.					
Strategic decisions	4.64	1.34	1.00	5.00	7.00
Investment decisions	4.49	1.33	1.00	4.00	7.00
Marketing decisions	3.09	1.54	1.00	3.00	7.00
Internal processes decisions	2.60	1.36	1.00	2.00	6.00
Human resource decisions	3.29	1.39	1.00	3.00	7.00
Growth opportunities: Please indicate your expectations about the following:					
The growth opportunities that exist within the industry in which you compete.	5.40	1.02	1.00	5.00	7.00
The growth opportunities your unit faces.	5.84	0.88	1.00	6.00	7.00
Competition: Please indicate the degree of competition with regard to the following.					
Prices	5.44	1.32	2.00	6.00	7.00
Marketing and distribution	4.54	1.62	1.00	5.00	7.00
Quality of products	5.00	1.22	0.00	5.00	7.00
Product mix	4.50	1.51	0.00	5.00	7.00
Varcomp: Your total compensation may vary with your performance. Please indicate the maximum amount (as a percentage of your fixed salary) available as performance-dependent pay (either as cash bonus, stocks, or options).	37.91	25.09	0.00	30.00	100.00

T A B L E 5

Pearson Correlations among All Variables Included in the Regression Model, and Cronbach's Alpha for All Latent Variables

Table 5 presents Pearson correlations among all variables used in this study. Variable definitions appear in Appendix 2. Also presented are Cronbach's alpha statistics for composite measures (Nunnally and Bernstein, 1994). The sample consists of 122 observations. Correlations (in absolute value) larger than 0.15 are significant at the 10% level.

	Cronbach's alpha	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1. Above		1.00										
2. Own		–0.84	1.00									
3. Below		–0.14	–0.42	1.00								
4. Spillovers from other managers		0.18	–0.13	–0.06	1.00							
5. Spillovers to other managers		0.11	–0.19	0.16	0.58	1.00						
6. Unsigned spillovers		0.17	–0.18	0.04	0.89	0.90	1.00					
7. Decentralization	0.71	–0.19	0.03	0.26	–0.08	0.13	0.05	1.00				
9. Growth opportunities	0.85	0.09	–0.09	–0.01	–0.08	–0.05	–0.08	0.01	1.00			
8. Competition	0.54	–0.03	0.13	–0.19	0.08	0.08	0.10	–0.07	–0.08	1.00		
6. Varcomp		–0.04	0.19	–0.28	–0.08	–0.19	–0.15	–0.14	0.16	0.15	1.00	
10. Size		0.03	0.11	–0.25	0.06	0.00	0.03	–0.13	–0.23	0.24	0.36	1.00
11. Service		–0.01	0.04	–0.05	0.01	0.05	0.01	–0.02	–0.07	0.05	0.01	–0.10

TABLE 6

Seemingly Unrelated Regressions Examining the Association between the Weight on Above-, Own-, and Below-level Performance Measures and Interdependencies

Table 6 presents seemingly unrelated regressions based on 122 observations. Dependent variables are log-ratios of weightings of above-level performance measures versus own-level performance measures and the weightings of below-level performance measures versus own-level performance measures. We analyze log-ratios to account for the compositional nature of our weightings of performance-measure variables (see, Abernethy et al., 2013). Log(above/own) measures the percentage weighting of above-level measures relative to the percentage weighting of own-level measures. Log(below/own) measures the percentage weighting of below-level measures relative to the percentage weighting of own-level measures. Independent variables are *Spillovers from others* and *Spillovers to others*, which have been mean-centered, and their interaction term *To×From*, or *Unsigned spillovers* which measures the interdependencies between business units within a firm. We control for *Varcomp*, *Decentralization*, *Competition*, *Growth opportunities*, *Size*, and *Service*. These variables are defined in Appendix 2. Standard errors are based on a bootstrapping procedure and are reported in parentheses. †, *, **, *** denotes significance at the 15%, 10%, 5%, and 1%-levels, respectively, and are two-tailed.

	Model (1): **[Unsigned spillovers]**		Model (2) **[Signed spillovers; no interaction]**		Model (3) **[Signed spillovers; with interaction]**	
Dependent variable	$Log\left(\frac{above}{own}\right)$	$Log\left(\frac{below}{own}\right)$	$Log\left(\frac{above}{own}\right)$	$Log\left(\frac{below}{own}\right)$	$Log\left(\frac{above}{own}\right)$	$Log\left(\frac{below}{own}\right)$
<u>*Test variables:*</u>						
Unsigned spillovers	0.214 (0.201)	0.102 (0.165)				
Spillovers to other managers			0.137 (0.398)	-0.443† (0.284)	0.140 (0.393)	-0.441† (0.287)
Spillovers from other managers			0.293 (0.420)	0.661** (0.314)	0.255 (0.415)	0.624** (0.306)
To × From					0.474** (0.200)	0.467*** (0.159)

Control variables:						
Varcomp	-0.033	-0.048***	-0.033	-0.044***	-0.031	-0.042***
	(0.021)	(0.017)	(0.022)	(0.017)	(0.021)	(0.016)
Decentralization	-0.841*	0.102	-0.868*	-0.0888	-0.522	0.252
	(0.482)	(0.414)	(0.505)	(0.416)	(0.492)	(0.401)
Competition	-0.722	-0.449	-0.727	-0.483	-0.741	-0.497
	(0.556)	(0.380)	(0.570)	(0.385)	(0.570)	(0.396)
Growth opportunities	0.722	0.144	0.715	0.0946	0.332	-0.283
	(0.600)	(0.479)	(0.623)	(0.485)	(0.657)	(0.510)
Size	0.237	-0.131	0.236	-0.137	0.194	-0.178
	(0.320)	(0.255)	(0.322)	(0.252)	(0.323)	(0.248)
Service	0.621	-0.799	0.609	-0.883	0.0345	-1.450*
	(0.789)	(0.737)	(0.797)	(0.733)	(0.858)	(0.754)
Chi-squared	11.61	19.36	11.68	25.03	20.56	39.93
p-value	0.11	0.01	0.17	0.00	0.01	0.00
adj. R^2	9.17%	14.4%	9.22%	17.9%	15.2%	25.8%

T A B L E 7

Analysis of First-Differences in Expected Value of Weight Placed on Performance Measures for Varying Degrees of Interdependencies

The table presents simulations of first differences, that is the difference in the expected value of the weight placed on a performance measure (above-level, own-level, below-level) when *Spillover to others* (*Spillover from others*) is changed from its sample minimum to its sample maximum value, while holding *Spillover from others* (*Spillover to others*) constant at its sample maximum and all other variables constant at their mean. Specifically, the simulation involves taking M draws from the multivariate normal distribution with mean $\widehat{\beta}$, the coefficient matrix from the seemingly unrelated regression model, and variance matrix $V(\hat{\beta})$, the estimated variance-covariance matrix for the coefficients in the model. We simulate first differences by setting the value for each type of spillover effect to its sample minimum, while holding the value of the other type of spillover at its sample maximum and holding all other variables constant at their means. We then generate the expected value of the outcome variable (the log-ratio of weight on performance measures) conditional on these starting values for the explanatory variables by taking one draw from the normal distribution. Next, we set the value for each type of spillover effect to its sample maximum, while continuing to condition the value of the other spillover to its sample maximum and holding all other variables at their means. We generate the expected value of the log-ratio conditional on these ending values for the explanatory variables. The first difference is simply the difference between these two expected values of the log-ratio. We then repeat this procedure 1,000 times to approximate the distribution of first differences (see, King et al., 2000, Zelner, 2009). Cell entries represent the average (over 1,000 simulations) weight placed on each type of performance measure (above-level, own-level, and below-level) Note that we account for the interaction between the spillover effects when computing the expected values. Rows marked "Diff." present the simulated first difference. Significance is evaluated from the empirical distribution of first differences. *, **, *** denotes significance at the 10 percent, 5 percent, and 1 percent level, respectively.

		Weight on Above-level measures	Weight on Own-level measures	Weight on Below-level measures
Spillovers from other managers	Minimum	0.09	0.91	0.00
	Maximum	0.60	0.22	0.18
	First diff.	0.52***	−0.69***	0.17***
Spillovers to other managers	Minimum	0.11	0.81	0.08
	Maximum	0.60	0.22	0.18
	First diff.	0.49***	−0.59***	0.10