Performance Measures and Intra-Firm Spillovers: Theory and Evidence

Jan Bouwens
Cambridge Judge Business School
University of Cambridge
Trumpington Street
Cambridge CB2 1AG, UK
j.bouwens@jbs.cam.ac.uk

Christian Hofmann
LMU Munich
Chair for Managerial Accounting
80539 Munich, Germany
hofmann@bwl.lmu.de

Laurence van Lent
Frankfurt School of Finance and Management
Sonnemannstr. 9-11
60314 Frankfurt am Main, Germany
l.vanlent@fs.de

Current version: July 2017

We appreciate helpful comments from Brian Cadman (editor), two anonymous reviewers, Vedran Capkun, Yan Du, Raffi Indjejikian, Francois Larmande, DJ Nanda, Richard Sansing, Jack Stecher, Ahmed Tahoun, Andreas Weiler and seminar participants at HEC Paris, the University of Leuven, the University of Mannheim, the AAA Management Accounting Section midyear meeting, the workshop on Accounting and Economics, and GEABA.

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. 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 unit's effect on other business units or the other units' effect on the focal unit. Our theory highlights complementarity between above- and below-level measures and the existence of an interaction between the two directions of spillovers. Based on a survey of 122 business unit managers, we report evidence consistent with an interaction effect and with complementarity between above- and below-level measures. In particular, we show that firms increase the weighting on both of above- and below-levels measures when they are coping simultaneously with high levels of spillovers on other units and spillovers from other units.

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

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 spillovers that exist between them (Roberts 2004). While transfer pricing schemes could potentially align the incentives of business unit managers to deal effectively with intra-firm spillovers, it is well-understood that these schemes are often used for strategic tax purposes instead (Hiemann and Reichelstein 2012).[1] A solid insight from prior work is that performance measurement systems can also be designed to encourage cooperation. In particular, unlike "own-level" performance measures, "above-level" performance measures which summarize the performance of multiple business units (Bushman et al. 1995; Keating 1997; Abernethy et al. 2004) and "below-level" measures of a specific subset of a business unit manager's activities (Baiman and Baldenius 2009; Bouwens and van Lent 2007) can be useful to deal with intra-firm spillovers.[2] However, what is unclear from prior work is whether the use of below-level measures substitutes for or complements the use of above-level measures when dealing with spillovers.

Specifically, using above-level measures to evaluate managerial performance lets managers "internalize" their externalities (i.e., activities with effects on other units), but exposes managers to high compensation risk (Holmstrom and Tirole 1991). In turn, using below-level measures for performance evaluation supports balanced managerial activities (including externality activities); however, by nature,

[1] Under a perfect transfer-pricing scheme, business-unit performance would reflect any externalities arising from intra-firm spillovers (Bushman et al. 1995). We address in our paper those spillovers 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). In addition, when spillovers arise due to information (processing) frictions between units (Puranam et al. 2012), the difficulty in measuring spillovers impedes the use of transfer-pricing schemes.

[2] While own-level measures are summary performance metrics associated with a manager's actions captured at the manager's level within the firm's hierarchy, above-level measures are summary performance metrics captured at a higher hierarchical level. From the perspective of a business unit manager, examples include business unit profits (own-level metric), and firm-wide profits and stock return (above-level metrics). In contrast, below-level measures are non-summary performance metrics associated with a specific subset of a manager's activities. Examples include project milestones, cost control measures, productivity measures, safety and attendance measures, and quality measures (see, Ittner and Larcker 2002; Ittner et al. 1997).

such measures lack congruence (Feltham and Xie 1994; Datar et al. 2001).[3] It is a priori unclear, how the relation of the two trade-offs is affected by variations in spillover.

We examine theoretically and empirically how spillovers between business units of a firm affect the design of the performance measurement system of business unit managers, comprising own-level measures, above-level measures, and below-level measures. Considering the full spectrum of performance measures, the question of how each of these measures should be used optimally in incentive contracts is far from trivial. Moreover, no empirical evidence exists to date that evaluates the use of above-level versus below-level measures in the presence of varying degrees of spillovers. In the theoretical part of our study, we show that below-level measures complement above-level measures when dealing with intra-firm spillovers (i.e., in settings with large spillovers, the optimal performance measurement system comprises many performance measures), and in the empirical part of our study, we document evidence that is consistent with our predictions.

In our theoretical analysis, we extend the agency model of a multi-unit firm by Bushman et al. (1995) in several dimensions. First, reflecting the typical diversity of business-unit managers' activities, we assume that agents are responsible for multiple tasks. Second, we combine prior literature by proposing that spillovers can have an effort effect and a noise effect. Specifically, the effort effect reflects spillovers that vary the marginal product of an agent's effort in the output of other business units (Bushman et al. 1995), whereas the noise effect reflects spillovers that vary the noisiness of business-unit output (Keating 1997; Anctil and Dutta 1999; Baldenius and Michaeli 2017). Third, consistent with many empirical settings, we allow for the possibility that performance measurement comprises business-unit output, firm-wide output, and specific performance as own-, above-, and below-level performance measures.

By considering the full range of performance measures in our agency setting, we can tailor our empirical analysis more closely to the theoretically-relevant concept of the incentive ratio (i.e., the relative weightings of the performance measures); what's more, when in practice senior managers decide to place

[3] In addition, below-level measures can signal whether managers (i) are acting in a cooperative fashion (Baiman and Baldenius 2009) and (ii) are achieving the goals of the company (Hirsch 1994; Bouwens and van Lent 2007).

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. In the theoretical part of our study, we provide comparative statics regarding the effects of intra-firm spillovers on the weightings of performance measures relative to one another.

The theoretical analysis yields three highlights: First, there exists a non-trivial relation between the direction of spillover and the optimal performance measurement system. Specifically, spillovers to and from other business units differentially affect performance measure properties and, thus, the optimal weighting of performance measures.[4] Second, 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 a consequence of the multi-task setting and the modeling of effort and noise effects from spillovers. The omission of such an interaction effect in empirical tests likely yields misspecified regressions.

Third, below-level measures complement rather than substitute for above-level measures when dealing with intra-firm spillovers. This suggests a positive relation between the complexity of the performance measurement system and the level of spillovers (i.e., in settings with large spillovers, the optimal performance measurement system comprises many measures). In this sense, our study speaks to the controversy in the strategic management accounting literature about whether the optimal performance measurement system comprises a single summary measure of performance (such as, Economic Value Added or Total Shareholder Return) or, alternatively, a Balanced Scorecard of (potentially many) performance measures that reflect specific managerial activities.

Turning to the empirical analysis, we test for the relations between the direction of spillovers and the use of own-level, above-level, and below-level performance measures, while controlling for the degree

[4] Conceptually, the origin of spillovers is manifold, including joint production technologies, interrelations in factor, labor, or product markets (Burton and Obel 1984; Holmstrom and Milgrom 1990), and the exchange of goods, services, or information between business units (Thompson 1967). In turn, the direction of spillovers is related to the sequence of business unit decision-making and the flow of goods, services, or information between units (Puranam et al. 2012).

of decentralization and various other factors known to influence the use of performance measures (Ittner et al. 1997; Keating 1997; Abernethy et al. 2004). 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 (Luft and Shields 2003; Lanen 1995; Ittner and Larcker 2001). Our approach yields yet another benefit: While transfer pricing schemes are an alternative tool to deal with spillovers, these schemes are designed to deal with internal deliveries of goods or services and are less suitable to address spillovers that arise from other frictions between units, for example those due to information processing (Puranam et al. 2012). Our survey instrument for spillovers asks respondents to indicate the extent to which actions of other managers in the firm impact on their performance (as well the extent to which their own actions impact on the performance of other managers). As such, our instrument captures those spillovers that remain present after transfer pricing schemes have "done their work".

We implement several empirical strategies aimed at enhancing the robustness of our findings, 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 predictions that (i) spillovers can reinforce one another and thus affect the relative weightings of the performance measures and (ii) above-level and below-level measures complement each other when dealing with spillovers are consistent with our data. Our estimates of the economic effects are sizable especially when considering those cases in which business unit managers are facing both spillovers from other units and their unit has important spillovers to other units. For instance, holding the spillover from other units constant at its sample maximum, an increase in the spillover to other units from its minimum to its maximum value, goes together with a gain in incentive weight on above level measures of about 50 percent points. The relative weighting on own measures is reduced by an even larger extent, as business units place also a higher weight of about 17 percent point on below-level measures.

In conclusion, we note that business unit managers who have to deal with incoming and outgoing spillovers face 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. And thus, our findings support the wisdom of recent advice in the strategic management accounting literature to combine both above-level performance measures such as EVA with more detailed below-level measures as advocated by proponents of the balanced scorecard (Kaplan 2001). Our findings also highlight something not fully recognized in this literature, namely the role of spillovers in determining the use of these measures in control systems.

Section 2 considers the theoretical framework and derives predictions regarding the impact of intra-firm spillovers on the relative weighting of own-level, above-level, and below-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. Section 5 concludes.

2. Theoretical Framework

In this section, we analyze an agency model of a multi-unit firm where the performance of business unit managers is assessed based on own-level, above-level, and below-level measures. After presenting the model, we characterize the optimal incentive parameters and discuss comparative statics in terms of intra-firm spillovers.

We consider a multi-unit firm with given spillovers between business units.[5] A risk-neutral principal (representing firm owners) hires a risk-averse agent (i.e., manager) for each of the *N* business

[5] A long-held view in organization science proposes that the division of labor, which allows organizational units to specialize in distinct stages of the firm's production process, simultaneously creates a need for coordination between these units and makes

units.[6] After accepting the contract, each agent exerts unobservable effort in two tasks, where one task is devoted to own operations (i.e., operating activity) and the other task affects own operations and the operations of other units (i.e., externality activity).[7] Finally, performance measures are realized, the agents receive their compensation, and the principal obtains the net payoff.

Consistent with the organization science literature (Puranam et al. 2012; Kretschmer and Puranam 2008), externalities on other units can be affected by managerial effort and/or create noise in these units' output.[8] Specifically, the own-level performance or output of unit *i* (e.g., business-unit profit), y_i, is characterized by

$$y_i = b_i a_i + \tilde{\varepsilon}_i + \mu_{ii} e_i + \sum_{j \in F_i} (\mu_{ij} e_j + \tilde{\eta}_{ij}), \qquad i = 1, ..., N, \tag{1}$$

where $a_i \in \mathbb{R}$ represents agent *i*'s effort for the operating activity, $e_i \in \mathbb{R}$ represents agent *i*'s effort for the externality activity, b_i represents the marginal product of a unit of agent *i*'s operating effort, $\mu_{ij} \geq 0$ represents the marginal product of a unit of manager *j*'s externality effort in business unit *i*,[9] $\tilde{\varepsilon}_i$ is a normally distributed random variable capturing operating events beyond the agents' control, $\tilde{\varepsilon}_i \sim \mathrm{N}(0, \sigma_i^2)$, $\tilde{\eta}_{ij}$ is a normally distributed random variable capturing uncontrollable events from the externality of unit *j* on unit *i*, $\tilde{\eta}_{ij} \sim \mathrm{N}(0, \sigma_{\eta ij}^2)$, and F_i is the set of business units where spillovers *from* these units affect the

collaboration non-trivial (Lawrence and Lorch 1967; Kretschmer and Puranam 2008). By defining business units, choosing the business units' technology, and allocating tasks and responsibilities, the firm's owners predetermine the interactions among business units (Milgrom and Roberts 1992; Acemoglu et al. 2007). We thus focus on the firm's management of these interactions that are the outcome of preceding choices with respect to the organizational structure. In turn, the choice of the organizational structure can reflect the anticipated design of the firm's control system (Antle and Demski 1988; Maskin et al. 2000).

[6] In line with Roberts (2004), we assume that incentive systems are more "malleable" than slow moving organization design elements such as the assignment of decision rights to agents. Similarly, Christie, Joye and Watts (2003) view production and information externalities among a firm's operating units as part of a firm's investment opportunity set, which is typically considered to be a determinant of the firm's compensation policy (Smith and Watts 1992).

[7] Examples of operating activities are efficiency improvements for products that are independent from other units' products, whereas examples of externality activities are design choices that affect downstream production cost or quality choices that affect the demand for complementary products offered by other units. The multi-task setting is consistent with surveyed business-unit managers, who are empowered with decision rights over diverse issues such as HRM, operations, marketing, or investments.

[8] For instance, when an agent's (externality) effort improves the quality of an interim product and, thus, reduces expected cost of downstream production, downstream output is essentially a summary performance metric of multiple agents. Consistent with Liang, Rajan and Ray (2008), problems with coordination and communication make it harder to measure the performance of multiple agents. Related, Anctil and Dutta (1999) and Baldenius and Michaeli (2017) assume that intra-firm trade introduces a variance effect in business-unit output, that can actually increase in the volume of intra-firm trade (Baldenius and Michaeli 2017).

[9] Assuming $\mu_{ij} \geq 0$ allows us to simplify the analysis. We will discuss later how negative spillovers, $\mu_{ij} < 0$, affect our results.

output of unit *i*; technically, $F_i = \{ j \mid \mu_{ij} \neq 0 \text{ or } \sigma^2_{\eta ij} \neq 0, \text{ for all } j \neq i \}$. For emphasis, $\text{Cov}[\tilde{\varepsilon}_i, \tilde{\varepsilon}_j] =$ $\text{Cov}[\tilde{\varepsilon}_i, \tilde{\eta}_{ij}] = \text{Cov}[\tilde{\varepsilon}_i, \tilde{\eta}_{ji}] = \text{Cov}[\tilde{\eta}_{ij}, \tilde{\eta}_{ji}] = 0$ for all $j \neq i$. The highlight of expression (1) is that spillovers from other units $j \in F_i$ either relate to the marginal product of agent *j*'s externality effort (μ_{ij}, see Bushman et al., 1995) and/or to performance measure noisiness ($\sigma^2_{\eta ij}$; see Keating 1997).

The firm's accounting system generates an above-level performance measure by aggregating business unit output to firm-wide or aggregate output (e.g., firm profits). Specifically, above-level performance, *x*, is characterized by $x = \sum_{i=1}^{N} y_i$. While the above-level measure captures the total marginal product of a unit of agent *i*'s externality effort, $\mu_{ii} + \mu_{iT}$ with $\mu_{iT} = \sum_{j \in T_i} \mu_{ji}$, aggregating own-level measures also yields a noisier above-level measure, $\sigma^2_x > \sigma^2_{yi}$ for all i, where

$$\sigma^2_x := \text{Var}[x] = \text{Var}[\sum_{i=1}^{N} y_i] = \sum_{i=1}^{N} \left(\sigma^2_i + \sum_{j \in F_i} \sigma^2_{\eta ij} \right) = \sum_{i=1}^{N} \left(\sigma^2_i + \sum_{j \in T_i} \sigma^2_{\eta ji} \right) \text{ and} \tag{2a}$$

$$\sigma^2_{yi} := \text{Var}[y_i] = \sigma^2_i + \sum_{j \in F_i} \sigma^2_{\eta ij}, \tag{2b}$$

and T_i is the set of business units where unit *i* has a spillover *to* the output of these other units; technically, $T_i = \{ j \mid \mu_{ji} \neq 0 \text{ or } \sigma^2_{\eta ji} \neq 0, \text{ for all } j \neq i \}$.

Finally, the accounting system generates below-level measures by reporting on the performance of specific activities (e.g., project milestones, cost control, or quality). Specifically, z_i denotes the below-level performance of agent *i*'s operating activity, $z_i = b_i a_i + \theta_i$, where θ_i is normally distributed noise, $\theta_i \sim N(0, \sigma^2_{zi})$, with $\text{Cov}[\theta_i, \theta_j] = \text{Cov}[\theta_i, \varepsilon_k] = \text{Cov}[\theta_i, \eta_{kl}] = 0$, $i, j, k, l = 1, \ldots, N$, $i \neq j$, and $k \neq l$.[10] While

[10] Our subsequent results are largely unaffected by the detailed assumptions about the below-level measure. For example, we get qualitatively similar results with measure $z = \sum_i a_i + \theta$.

the own-level and the above-level measure are summary performance metrics of multiple units, the below-level measure is a non-summary performance metric associated with a subset of managerial activities.[11]

The principal offers agent *i* a contract that depends on three performance measures: above-level performance, *x*, agent *i*'s own-level performance, y_i, and agent *i*'s below-level performance, z_i. Agent *i*'s compensation, w_i, is a linear function of the measures, i.e., $w_i = f_i + \alpha_i x + \beta_i y_i + \gamma_i z_i$, where f_i is the fixed payment, and α_i, β_i, and γ_i are incentive weights for above-level performance, own-level performance, and below-level performance, respectively.[12]

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, $\kappa_i = ½(a_i^2 + e_i^2)$. Normally distributed noise and negative exponential utility yield a simple representation of the agent's certainty equivalent (Holmstrom and Milgrom 1987), which is given by $CE_i = \mathrm{E}[w_i] - \kappa_i - ½ r \mathrm{Var}[w_i]$, representing expected compensation net of effort cost and risk premium.

The principal selects incentive weights that maximize expected firm profit net of the agents' compensation, subject to contract acceptance and effort choices by all agents. Specifically, the principal maximizes $\mathrm{E}[x] - \Sigma_i \mathrm{E}[w_i]$, subject to, for all agents, individual rationality constraints, $CE_i \geq 0$, and incentive compatibility constraints, $(a_i, e_i) \in \operatorname{argmax} CE_i$, where we scale without loss of generality each agent's reservation certainty equivalent to zero.[13]

Solving for the optimal incentive weights yields Lemma 1. Lemma 1 describes the relative weighting of above-level and own-level performance, $IRA_i^* = \alpha_i^* / \beta_i^*$, and the relative weighting of below-

---

[11] We assume that spillovers do not affect z_i. Specific measures are less susceptible to spillovers for at least two reasons: First, specific metrics such as cost control, productivity, or quality capture the outcome of processes and are rather unaffected by whether the produced goods are processed further in the same unit or shipped to another unit. Second, different to summary metrics such as business unit or firm-wide output, specific metrics can be designed to ignore spillovers.

[12] Limiting the contract to (x, y_i, z_i) is restrictive when $N > 2$. In this case, including the output of all business units separately would be optimal. However, none of the firms in our empirical sample includes the performance of another unit separately in the performance evaluation of a specific manager.

[13] 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.

level and own-level performance, $IRB_i^* = \gamma_i^* / \beta_i^*$. We will use the relative weights to characterize how changes in intra-firm spillovers affect the weighting of performance measures.

Lemma 1: *With optimal linear contracts, the relative incentive weights on x and* y_i *for agent i are*

$$IRA_i^* = \frac{\left(\mu_{iT}\left(\mu_{ii}+\mu_{iT}\right)\left(1+\frac{b_i^2}{r\,\sigma_{zi}^2}\right)+\frac{b_i^2\mu_{iT}^2}{r\,\sigma_{yi}^2}\right)\Big/\sigma_x^2}{\left(\left(\mu_{ii}-\frac{\sigma_{yi}^2}{\sigma_x^2}\left(\mu_{ii}+\mu_{iT}\right)\right)\left(\mu_{ii}+\mu_{iT}\right)\left(1+\frac{b_i^2}{r\,\sigma_{zi}^2}\right)+b_i^2\left(1-\frac{\sigma_{yi}^2}{\sigma_x^2}\right)\right)\Big/\sigma_{yi}^2} \qquad (3a)$$

and the relative incentive weights on z_i *and* y_i *are*

$$IRB_i^* = \frac{b_i^2\left(\left(1-\frac{\sigma_{yi}^2}{\sigma_x^2}\right)-\left(\mu_{ii}-\frac{\sigma_{yi}^2}{\sigma_x^2}\left(\mu_{ii}+\mu_{iT}\right)\right)\frac{\mu_{iT}}{r\,\sigma_{yi}^2}\right)\Big/\sigma_{zi}^2}{\left(\left(\mu_{ii}-\frac{\sigma_{yi}^2}{\sigma_x^2}\left(\mu_{ii}+\mu_{iT}\right)\right)\left(\mu_{ii}+\mu_{iT}\right)\left(1+\frac{b_i^2}{r\,\sigma_{zi}^2}\right)+b_i^2\left(1-\frac{\sigma_{yi}^2}{\sigma_x^2}\right)\right)\Big/\sigma_{yi}^2}, \qquad (3b)$$

where $\mu_{iT} = \sum_{j\in T_i}\mu_{ji}$ *and* σ_x^2 *and* σ_{yi}^2 *are defined in (2) and* $i = 1,\ldots,N$.

Proof: See Appendix A.

Various forces affect the relative incentive weights in the multi-task, three performance measure setting: First, following Banker and Datar {,1989}, incentive ratios reflect the adjusted sensitivity to noise of the associated performance measures (e.g., in case of above-level performance, the marginal product of externality effort relative to noise, μ_{iT} / σ_x^2). Second, in the multi-task setting, each tasks' adjusted sensitivity is weighted with its total marginal product (e.g., $\mu_{ii} + \mu_{iT}$ for externality effort). Third, consistent with (Feltham and Xie 1994), effort externalities imply that own-level performance can be non-congruent with above-level performance (i.e., when $\mu_{iT} \neq 0$). While the principal increases the relative weight on above-level performance to improve the agent's effort allocation across tasks, this effect is mitigated when own-level performance is noisy (this is the term $b_i^2\mu_{iT}^2 / (r\,\sigma_{yi}^2)$). Fourth, the measure of a specific activity enables the principal to influence the agent's effort allocation (Feltham and Xie 1994; Bushman et al. 1996). Because the operating activity can also be motivated via z_i, the principal heightens the relevance of the externality activity in (3a) and (3b) by increasing the adjusted sensitivity for the externality activity (this is the term $1+b_i^2/(r\sigma_{zi}^2)>1$). Intuitively, given z_i, it is more important to motivate externality effort via y_i and x than to motivate operating effort with these measures.

We next discuss how variations in intra-firm spillovers affect the relative weighting of the performance measures. Generally, variations in spillovers can change performance measure properties such as adjusted sensitivity, precision, and congruence, which affect the optimal incentive weights. Specifically, following expression (1), variations in spillovers *from* other business units can change the marginal product of other agents' externality efforts (i.e., from the perspective of business unit *i*, μ_{ij}) and the uncontrollable events from the externalities of other units (i.e., $\sigma_{\eta ij}^2$). By symmetry, variations in spillovers *to* other business units can change the marginal product of externality effort and the uncontrollable events imposed on other units (i.e., from the perspective of business unit *i*, μ_{ji} and $\sigma_{\eta ji}^2$). We label the change of the marginal product of externality effort as the *effort effect of spillovers* and the change of the uncontrollable events from externalities as the *noise effect of spillovers*.[14]

In practical settings, effort and noise effects are likely difficult to separate. Thus, we report on the joint effect of variations in spillovers, allowing for simultaneous consequences of spillovers for the marginal product of the externality effort and the uncontrollable events from externalities. We assume that the total marginal product of the externality activity and the total noise from externalities increase in the spillovers to and from other business units (i.e., $\partial(\Sigma_{j\in T_i}\mu_{ji})/\partial To>0$, $\partial(\Sigma_{j\in T_i}\sigma_{\eta ji}^2)/\partial To>0$, $\partial(\Sigma_{j\in F_i}\mu_{ij})/\partial From>0$, and $\partial(\Sigma_{j\in F_i}\sigma_{\eta ij}^2)/\partial From>0$, where, from the perspective of unit *i*, To and

[14] We verify whether in our sample noise and spillovers are associated in the manner described in the *noise effect*. Since we do not have questions related to noise in the survey, we construct a proxy based on publicly available data corresponding to noise in firm-level performance measures (i.e., σ_x^2 in our model). Specifically, we follow Ittner et al. (1997) and define SDROA (SDROE, SDROS) as the standard deviation of median industry return on assets (return on equity, return on sales) over the five year period before the survey data was collected. Industries are classified based on 3-digit SIC codes. We compute Spearman correlations between these proxies for noise (σ_x^2) and empirical measures of intra-firm spillovers (described in detail below). We find positive correlations between noise and *spillover to other business units* as well as between noise and *spillover from other business units*, albeit that the latter association is only weakly significant. Details are available on request from the authors. These findings are consistent with earlier empirical results; see, e.g., Table 2, Panel B in Abernethy et al. (2004).

$From$ represent spillovers to and from other units).[15] Observation 1 summarizes the implications of Lemma 1 regarding the effects of variations in spillover on the relative weighting of the performance measures (see Appendix A for a proof and a detailed analysis of the separate effort and noise effects).

Observation 1: *Variations in intra-firm spillovers have the following effects on performance measure weightings:*

(i) Spillovers to and from other business units differentially affect performance measure weightings.

(ii) Spillovers to and from other business units interact in their effects on performance measure weightings.

(iii) The weightings of above-level and below-level performance, relative to own-level performance, change in a similar fashion with variations in spillover to and from other business units.

Part (i) of the observation states that there exists a non-trivial relation between the direction of spillover and the optimal weighting of performance measures. Consistent with the well-established relation between sensitivity and precision (Banker and Datar 1989), the joint effect of spillovers depends on the relative importance of the effort and noise effects. For example, for large spillovers to other units, the above-level metric becomes relatively more important if the increase of the total noise from externalities relative to the increase of the marginal product of the externality activity is sufficiently small (i.e., $\partial IRA_i^* / \partial To > 0$ if $[\partial(\Sigma_{j\in T_i}\sigma^2_{\eta ji}) / \partial To] / [\partial(\Sigma_{j\in T_i}\mu_{ji}) / \partial To] < \hat{g}_A$, where $\hat{g}_A$ is defined in the appendix). Intuitively, for a high marginal product of externality effort, a relatively large weight on above-level performance more strongly motivates the agent to consider his externality effort, whereas a noisy above-level metric makes it more costly to choose a large weight on this measure. In contrast, for large spillovers from other units, the above-level metric becomes unambiguously more important (i.e., $\partial IRA_i^* / \partial From > 0$

---

[15] The results in Observation 1 hold even if there are some negative externalities such as $\mu_{ji} < 0$. When spillovers decrease $\Sigma_{j\in T_i}\mu_{ji}$ (i.e., more negative externalities), the sign of the comparative statics flips. However, when there are large negative externalities, it is plausible that firm owners will restructure the organizational units to let the manager fully internalize the total effects of his actions. We also note that key insights of our analysis (i.e., the types of spillovers (i) differentially affect and (ii) interact in their effect on performance measurement choices) are unaffected by the presence of positive or negative spillovers μ_{ji}.

). Variations of other agents' marginal product do not affect the metrics' informativeness about the focal agent's activities, whereas noisy own-level and above-level metrics make it less costly to choose a relatively large weight on above-level performance.

Part (ii) of the observation predicts an interaction between spillovers to other units and spillovers from other units, which would affect the relative weightings of the performance measures. Specifically, the cutoffs for the joint effect of spillovers to other units depend on the spillovers from other units (e.g., $\hat{g}_A$ varies with $From$). In other words, spillovers *from* other units affect the noise in the unit's own-level performance which, in turn, affects the relation between the optimal weighting of performance measures and the spillovers *to* other units. For instance, with small spillovers from other units, the unit's own-level performance and, hence, above-level performance are sufficiently precise (noting the summing up nature of above-level performance); then, increasing the spillover to other units substantially increases the noise in above-level performance, implying that the noise effect dominates and the relative weight placed on above-level performance is reduced. With large spillovers from other units, own-level and above-level performance are quite noisy; then, increasing the spillover to other units has a minor influence on the noise in above-level performance, implying that the effort effect dominates and the relative weight placed on above-level performance is increased.

Finally, Part (iii) of the observation states that the weightings of the above-level and the below-level measure, relative to the own-level measure, change in a similar fashion with variations in spillover. For example, for large spillovers to other units, the above-level measure is noisy. Then, relative to the weight on own-level performance, the principal decreases the weight on above-level performance to reduce the agent's risk premium and the weight on below-level performance to more strongly motivate the agent to consider his externality activity rather than his operating activity. For large spillovers from other units the own-level measure is noisy, and strong motivation of the agent's externality effort via this measure is costly. Thus, the principal reduces the weight on own-level performance, relative to the weights on above-level and below-level performance.

In summary, the analysis of incentive problems in a multi-unit firm yields the following results: First, there exists a non-trivial relation between the direction of spillover and performance measurement, i.e., spillovers to and from other units differentially affect the optimal weighting of performance measures. Second, the model predicts an interaction between spillovers to other units and spillovers from other units, which would affect the relative weightings of the performance measures. Third, the weightings of the above-level and below-level measure, relative to the own-level measure, change in a similar fashion with variations in spillover. This suggests the complexity of performance measurement (in terms of the number of performance measures) to increase in spillovers; in this sense, below-level measures complement rather than substitute for the use of above-level measures when dealing with intra-firm spillovers. The interaction of the two directions of spillover in determining the optimal incentive weights is a subtle consequence of (i), a multi-task setting, (ii), the distinction between the two directions of spillover, and (iii), the differentiation between an effort and a noise effect caused by intra-firm spillovers. 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 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 in the context of a broader research program on the relation between organizational design and performance measures. However, the data in this study have not been used in any 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.[16] 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 this bias (following Podsakoff et al. 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).

[16] Appendix B reproduces the questionnaire items we used to construct our variables.

Table 1, 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 1 about here. -----------------

As Table 1, 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 constructs IRA^* and IRB^* 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.

We combined the separate weightings of stock-price related,[17] 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 and below-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,

[17] 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 spillovers in very much the same way as other firm-wide measures (Keating 1997; Bouwens and van Lent 2007).

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 et al. 1996).

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).[18] Details are reported in Table 2, Panel A.

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

The total 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 2 Panel A must sum to 100 percent, *above* + *own* + *below* = 100%. The ratio of our variables yield proxies for our theoretical constructs; *above/own* (*below/own*) proxies for our theoretical construct of the relative incentive weights IRA^* (IRB^*) for the weighting on above-level (below-level) and own-level performance.

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 (Aitchison 1986; Abernethy et al. 2013).We use *Log(above/own)*, the natural logarithm of the weighting of above-level measures divided by the weighting of own-level measures, and *Log(below/own)*, the natural logarithm of the weighting of below-level measures divided by the weighting of own-level measures, as the dependent

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

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, To and $From$, 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 et al. 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 BUs* (Spillovers to other business units) 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 To of the overall spillovers from the focal business unit to other business units (likely capturing the effect of spillovers on both, the marginal product of the externality effort and the uncontrollable events from externalities). Similarly, *Spillovers from other BUs* (Spillovers from other business units) 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 *From* of the overall spillovers from other business units. 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 2, 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 2, 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 BUs*. 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 BUs* and find a correlation of 0.48 (*p*-value<0.01). Finally, we correlate both *Spillovers to other BUs* and *Spillovers from other BUs* 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 BUs* is 0.48, *p*-value<0.01; corr. with *Spillovers from other BUs* 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 BUs* is −0.47, *p*-value<0.01; corr. with *Spillovers from other BUs* is –0.36, *p*-value<0.01). Together these tests suggest that our spillover metrics are valid.[19]

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

[19] The percentage of a unit's total production delivered to (sourced from) other units is a directional variable and could therefore be used as an alternative proxy for *To* (*From*). Our inferences are unchanged when we replace the *Spillovers to (from) other BUs* with this alternative proxy. Details are available from the authors upon request.

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 BUs* and *Spillovers from other BUs* (Table 2, 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 (Ittner et al. 1997; Ittner and Larcker 2001; Bushman et al. 1995; Keating 1997; Abernethy et al. 2004; Bouwens and van Lent 2007). Table 3 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.[20] 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.

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

[20] 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.

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 3 about here. -----------------

Size is the natural logarithm of the number of full time employees of a given business unit (see Table 1). 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 (Hair et al. 1998; Chenhall 2005). 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.[21]

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).[22] 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).

[21] 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).

[22] The estimated correlation between the residuals of the two equations is 0.49.

4. Empirical findings

We first discuss the summary statistics for our main variables as reported in Table 2 as well as their Pearson correlations, which are reported in Table 4. Table 5 presents the estimation results of the seemingly unrelated regressions that examine the association between spillovers and the use of above-level and below-level measures relative to own-level measures, and Figure 1 presents evidence from a simulation of first differences (i.e., the difference between the expected value of the weight placed on own-, above-, 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 own-, above-, and below-level performance measures in Table 2, 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 spillover variables *Spillovers to other BUs* and *Spillovers from other BUs*, as well as for the variables we use to establish convergent validity (*Supply to other BUs*, *Supply from other BUs*, *%Time-Meet*, and *Independent business*). The test variables span the full theoretical range of one to seven and show on average substantial intra-firm spillovers (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 spillover 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 4, which presents the correlations among the variables in our study, the weightings of the three performance-measure categories (i.e., own-level, above-level, and below-level) are, as we would expect given that they are constrained to sum to unity, negatively associated. Note that these simple correlations between the three performance measures cannot be used to test our prediction that above-level and below-level measures are complements. For this, we first need to control for the (exogenous) determinants of the performance measures as well as deal with the statistical consequences of the summing up to unity constraint. We discuss this further in Section 4.2.

More importantly, we find that weightings of above-level measures are significantly positively associated with *Spillovers from other BUs*, but not with *Spillovers to other BUs*. Conversely, weightings of below-level measures are significantly positively correlated with *Spillovers to other BUs*, but not with *Spillovers from other BUs.* In line with these results, weightings of own-level measures are negatively associated with both types of spillovers (albeit only weakly with *Spillovers from other BUs*). Taken together, it appears that below-level and above-level measures are both used in response to increasing intra-firm spillovers, 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 ratios, however, these linear associations should be interpreted with some care.

----------------- Insert Table 4 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 of the above-level and below-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.

Table 5 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, from the perspective of a focal business unit, spillovers *to* others (To) and spillovers *from* others ($From$). 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).[23]

---------------- Insert Table 5 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.[24] We find qualitatively similar results for the control variables in the remaining models, described below, which we will not discuss again.

Next, we estimate a regression that includes the two signed spillover proxies (i.e., *Spillovers to other BUs* and *Spillovers from other BUs*) 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

[23] 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.

[24] Some 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 et al. 2009; Raith 2008). Information asymmetry tends to be highly correlated with decentralization (both in our study as in prior work). Including a control variable for information asymmetry based on the Dunk (1993) instrument does not affect any of our conclusions.

theoretical analysis, then this regression is misspecified and its findings should be interpreted with care. The results are presented in Table 5 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 BUs* 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 BUs* 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 5 ("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 BUs* (*Spillovers from other BUs*) constant at its sample average, *Spillovers from other BUs* (*Spillovers to other BUs)* is not associated with the weight placed on above-level measures relative to the weight on own-level measures.[25] While there is no significant relation between the spillover proxies and the log-ratio (above/own) when considering

[25] Specifically, suppose we have the following equation: $\mathrm{Y} = \beta_0 + \beta_1\mathrm{X} + \beta_2\mathrm{Z} + \beta_3\mathrm{X} \times \mathrm{Z} + \Sigma\beta_j\mathrm{Controls}_j + \epsilon$. The marginal effect of X on Y when Z is held at its sample minimum ($\underline{\mathrm{Z}}$) is defined as $\partial\mathrm{Y}/\partial\mathrm{X}|_{\mathrm{Z}=\underline{\mathrm{Z}}}$, 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.

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 BUs* (*Spillovers to other BUs*) 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.[26]

Overall, our results in Table 5 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. Our seemingly unrelated regression framework also allows us to directly test the complementarity between the weight placed on above-level and below-level measures. Following Brynjolfsson and Milgrom (2013), we compute the partial correlation of the weight on these performance measures conditional on those exogenous factors that affect these choices. In other words, we compute the correlation between the residuals of the two regression equations that model the log-ratios (above/own) and (below/own). A (negative) positive correlation suggests that the decision to place more weight on above-level measures relative to own-level measures is complementary to (a substitute for) placing more weight on below level measures relative to the own-level measures.

[26] In addition, we can interpret the findings reported in Table 5 through the lens of the comparative statics summarized in Appendix A. Consider the relative incentive weights for above-level versus own-level measures. We find no significant association between this ratio and either type of spillover empirically (when holding constant the other spillover at its mean). Table A.1 shows that for *Spillover to other business unit* the effort and noise effects can have different signs, which potentially cancel out against each other. In addition, the effort effect for *Spillover from other business unit* predicts no change in the relative weight placed on above-level measures when this type of spillover increases. Similarly, for the relative incentive weight on below-level measures versus own-level measures, we find empirically a negative association with *Spillover to other BUs*. This is consistent with either the noise effect dominating the effort effect or with both effects having a negative impact on the relative incentive weight. The empirical finding of a positive coefficient on *Spillover from other BUs* is consistent with the theory's prediction of a positive noise effect. Thus, the variations in Table 5 suggest that both, effort and noise effects are present in our sample of business units.

We find that the partial correlation (i.e., the correlation between the regression residuals) is 0.43 ($p < 0.001$), which suggest that the decisions to move more weight to above-level measures and to below-level measures relative to own-level measures are indeed complements.

4.3 Simulation results

Ultimately, we are interested in reconciling our empirical findings in Table 5 with how the percentage weighting of each performance measure changes as intra-firm spillovers vary. However, these relations cannot be simply inferred from the results in Table 5: 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 BUs* and *Spillovers from other BUs* to affect the relative use of each performance measure both directly and indirectly.[27] Third, our dependent variable is a log-ratio and not the percentage weights placed on the performance measures.

Hence, using our regression results to answer questions of practical relevance – such as, what happens to the weighting of above-level measures when the firm increases collaboration between units – is not straightforward. Indeed, consider the implications of increasing *Spillovers to other BUs* 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 BUs* would ultimately affect the weight placed on the above-level measures.

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 from other BUs* (*Spillovers to other BUs*) is changed from its sample minimum to its sample maximum, while holding *Spillovers to other BUs* (*Spillovers from other BUs*) constant at representative sample values (i.e., at the minimum, 10th, 25th, 50th, 75th, 90th, and

[27] The non-linear specification implies that the partial derivative of the log-ratio (above/own) with respect to *Spillovers to other BUs* (*Spillovers from other BUs*) depends on the level of *Spillovers from other BUs* (*Spillovers to other BUs*). Indeed, both sign and significance of the partial derivative may change with different levels of the interacting variable (Jaccard and Turrisi 2003).

maximum).[28] 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).

Figure 1 summarizes the findings from the analysis of first differences. Each marker symbol on the (extrapolated) lines represents the average (computed over 1,000 simulations) of the weight placed on the associated performance measure for minimum (dotted lines) and maximum values (solid lines) of *Spillovers from other BUs* (*Spillovers to other BUs*), respectively holding constant the other spillover variable at a given level of its sample distribution. For example, at point A, where *Spillovers to other BUs* is at its sample maximum the weight on own-level measures changes from approximately 0.83 to 0.22, when *Spillovers from other BUs* is increased from its sample minimum to the maximum. The first difference is simply the vertical distance between the solid and dotted lines for each performance measure. At point A, this difference is -0.61 and unreported tests show that the difference is significant at the 1 percent level. While Figure 1 plots first differences (as the vertical distance between the dotted and solid lines), it also helps to evaluate what happens to the weight on each performance measure as one type of spillover is increased (moving along the x-axis from minimum to maximum values), while the other type of spillover is held constant at its minimum (dotted line) or maximum values (solid line).

----------------- Insert Fig. 1 about here. -----------------

We find that spillovers have economically meaningful consequences for incentive weight choices. Recall that our theory yielded three observations, summarized in Observation 1. First, the direction of spillovers matters for the incentive weighting. We observe differential outcomes in Panels A and B, which

[28] Specifically, the simulation involves taking *M* draws from the multivariate normal distribution with mean $\hat{\beta}$ and variance matrix $V(\hat{\beta})$, the coefficient matrix and the estimated variance-covariance matrix for the coefficients, resp., from the seemingly unrelated regression model. We simulate first differences by setting the value for *Spillovers to other BUs* (*Spillovers from other BUs*) to its sample minimum, and *Spillovers from other BUs* (*Spillovers to other BUs*) to a given constant value. 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 BUs* (*Spillovers to other BUs*) to its sample maximum and continue to hold *Spillovers to other BUs* (*Spillovers from other BUs*) at a given sample value. 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. To ascertain significance, we could repeat this procedure 1,000 times to approximate the distribution of first differences (see, Zelner 2009; Abernethy et al. 2013; King et al. 2000), but in the current case, we simply plot the expected values.

reflect *Spillovers to other BUs* and *Spillovers from other BUs*, respectively. More importantly, however, and this segues into our second theoretical observation, these two spillovers interact in their effects on performance measure weightings. For example, in Panel A, the weight on above-level measures increases when *Spillovers from other units* is held constant at its sample *maximum* and *Spillovers to other BUs* increases from its minimum to maximum values, but in contrast decreases when *Spillovers from other BUs* is evaluated at the sample *minimum* while changing *Spillovers to other BUs* in the same way as before. More generally, consistent with Observation 1 (ii), we find that the slope of each line in the Panels A and B is not independent of the value of the other spillover type; hence, the two spillovers interact with each other in their effects on performance measure weightings. Finally, Figure 1 also confirms the prediction in Observation 1 (iii), namely that above-level and below-level performance measures (compared with own-level measures) respond in a similar way to changes in spillovers to and from other business units. This holds in particular in Panel A (depicting the first differences for *Spillovers from other BUs*), when considering increasingly higher levels of spillovers to others units, both performance measures display negative first differences at first and gradually move towards positive first differences. What's more, shifting attention not to the first difference but to the weighting on each performance measure, holding constant *Spillovers from other BUs* at its maximum (solid line) or minimum value (dotted line), we find that the weight on both below and above-level measures increases when changing the *Spillovers to other BUs* (at the maximum value of *Spillovers from other BUs*), whereas for minimum values of *Spillovers from other BUs*, we observe that the weight on both below and above-level performance measures decreases.

Note that Observation 1, part (ii) implies that with small spillovers from other units, the unit's own-level performance and above-level performance are sufficiently precise; then, increasing the spillovers to other units increases the noise in above-level measures and the weight on above-level measures is reduced. We find exactly this pattern in our simulation in Panel A. The dotted line with diamonds, representing the weight on above-level measures when *Spillovers from other BUs* is at its minimum, decreases as the *Spillovers to other BUs* increases. Our theoretical observation also implies that when the spillovers from other units are large (empirically represented by the solid line), the unit's own-level performance and above-level performance are noisy; then, increasing the spillovers to other units has relatively little influence on the noise in above-level performance and more weight is placed on the above-level measure. Again, the upward sloping solid line with diamonds, representing the weight on the above-level measure, is consistent with this prediction.

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 the business unit experiences high levels of one spillover and the other type of spillovers (i.e., *to* or *from* other BUs) becomes increasingly more prevalent. To a lesser extent, firms use below-level measures in the same way, especially when coping with increased *Spillovers to other BUs* and experiencing high levels of *Spillovers from other BUs*. While the implications of variations in spillover are identical, the underlying reasoning differs. Following the agency setting in Section 2, *Spillovers to other BUs* yield larger effort externalities, implying that above-level measures more fully reflect the consequences of managerial actions. The associated strengthened use of above-level measures in providing incentives can even exceed any discount caused by noisy above-level measures; below-level measures can then be used to fine tune incentives. In turn, from the perspective of a focal manager, *Spillovers from other BUs* merely create noise in own-level measures, implying that above-level and below-level measures are relatively precise metrics and, thus, are used more intensively to provide incentives.[29] Collectively, these results are consistent with the statement that below-level measures complement above-level measures when dealing with intra-firm spillovers.

It is worthwhile to point out that own-level measures are given very significant weights if the spillover effect is only strong in one direction. For example, when *Spillovers from other BUs* are at their sample maximum value (and *Spillovers to other BUs* 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.

[29] Of course, these interpretations are ex post and specific to our sample; in particular, it is an empirical question whether the effort effect caused by *Spillovers to other BUs* dominates the noise effect.

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 spillover is present in the firm and, more importantly, on whether these spillovers 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 division's performance is affected by other units in the firm, suggesting that above-level measures complement rather than substitute for below-level measures.

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 spillovers. We use information gathered directly from business unit managers who are affected by these spillovers and/or are encouraged to work with other managers to overcome the problems such spillovers cause. These respondents also provide us with data about the performance measures their superiors 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. That said, in our empirical work, we use proxies for the signed spillovers that are based on a single questionnaire item. Our choice is in part driven by the desire to allow readers to make cross-study comparisons with other papers that have investigated the effects of spillovers on the use of performance measures (Abernethy et al. 2004; Keating 1997). Nevertheless, to mitigate concerns about measurement error, follow-up studies should consider developing new questionnaire instruments that are not just based on more than a single item, but also allow for a separate examination of the effort and noise effects of spillovers that our theoretical analysis has highlighted to be important.

Overall, our study provides evidence that above- and below-level measures are complements when dealing with intra-firm spillovers. Our study also highlights that spillovers are "directional"; that is, for the purpose of performance measurement, it matters a great deal whether spillovers affect a focal unit or whether it is the focal unit, which "impacts" on the performance of other units 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

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

———. 2013. The role of performance measures in the intertemporal decisions of business unit managers. *Contemporary Accounting Research* 30 (3):925-961.

Acemoglu, D., P. Aghion, C. Lelarge, J. Van Reenen, and F. Zilibotti. 2007. Technology, Information, and the Decentralization of the Firm. *The Quarterly Journal of Economics* 122 (4):1759-1799.

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

Anctil, R. M., and S. Dutta. 1999. Negotiated transfer pricing and divisional vs. firm-wide performance evaluation. *The Accounting Review* 74 (1):87-104.

Antle, R., and J. S. Demski. 1988. The controllability principle in responsibility accounting. *Accounting Review*:700-718.

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

Baldenius, T., and B. Michaeli. 2017. Investments and Risk Transfers. *The Accounting Review*.

Banker, R., and S. Datar. 1989. Sensitivity, precision and the linear aggregation of signals for performance evaluation. *Journal of Accounting Research* 27:21-39.

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

Brynjolfsson, E., and P. Milgrom. 2013. Complementarity in organizations. In *The handbook of organizational economics*, edited by R. Gibbons and J. Roberts. Princeton and Oxford: Princeton University Press.

Burton, R. M., and B. Obel. 1984. *Designing Efficient Organizations: Modelling and Experimentation*: Nort Holland.

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

Bushman, R. M., R. J. Indjejikian, and A. Smith. 1996. CEO compensation: The role of individual performance evaluation. *Journal of Accounting & Economics* 21 (2):161-193.

Chenhall, R. H. 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.

Chin, W. W., and P. R. Newsted. 1999. Structural equation modeling analysis with small samples using partial least squares. In *Statistical strategies for small sample research*, edited by R. H. Hoyle. Thousand Oaks: Sage, 307-341.

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

Datar, S., S. L. Kulp, and R. A. Lambert. 2001. Balancing performance measures. *Journal of Accounting Research* 39 (1):75-92.

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

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

Feltham, G., and J. Xie. 1994. Performance measure congruity and diversity in multi-task principal agent relations. *The Accounting Review* 69 (3):429-453.

Fowler Jr., F. J. 1995. *Improving survey questions*. Vol. 38. Thousand Oaks: Sage.

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

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

Hair, J. F., R. E. Anderson, R. L. Tatham, and W. C. Black. 1998. *Multivariate data analysis*. 5 ed. Upper Saddle River, NJ: Prentice Hall.

Harman, H. H. 1967. *Modern factor analysis*. 2nd ed. Chicago, IL: University of Chicago Press.

Hiemann, M., and S. Reichelstein. 2012. Transfer pricing in multinational corporations: An integrated management-and tax perspective. In *Fundamentals of International Transfer Pricing in Law and Economics*: Springer, 3-18.

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

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

———. 1990. Regulating trade among agents. *Journal of Institutional and Theoretical Economics* 146:85-105.

Holmstrom, B., and J. Tirole. 1991. Transfer Pricing and Organizational Form. *Journal of Law Economics & Organization* 7 (2):201-228.

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

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

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

———. 2002. Determinants of performance measure choices in worker incentive plans. *Journal of Labor Economics* 20 (2):S58-S90.

Ittner, C. D., D. F. Larcker, and M. V. Rajan. 1997. The choice of performance measures in annual bonus contracts. *Accounting Review* 72 (2):231-255.

Jaccard, J., and R. Turrisi. 2003. *Interaction effects in multiple regression*. Edited by M. S. Lewis-Beck. 2 ed. Vol. 72. Thousand Oaks: Sage Publications.

Kaplan, R. S. 2001. Integrating shareholder value and activity-based costing with the balanced scorecard. *Balanced Scorecard Report* 15.

Keating, A. S. 1997. Determinants of divisional performance evaluation practices. *Journal of Accounting and Economics* 24:243-274.

Khandwalla, P. 1972. The effect of different types of competition on the use of management controls. *Journal of Accounting Research* 10 (2):275-285.

King, G., M. Tomz, and J. Wittenberg. 2000. Making the most of statistical analysis: Improving interpretation and presentation. *American Journal of Political Science* 44 (2):341-355.

Kretschmer, T., and P. Puranam. 2008. Integration Through Incentives Within Differentiated Organizations. *ORGANIZATION SCIENCE* 19 (6):860-875.

Lanen, W. N. 1995. Discussion of aggregate performance measures in business unit manager compensation: The role of intrafirm dependencies. *Journal of Accounting Research* 33 (Supplement):129-134.

Lawrence, P., and J. Lorch. 1967. *Organizations and environment*. Homewood: Irwin.

Liang, P. J., M. V. Rajan, and K. Ray. 2008. Optimal team size and monitoring in organizations. *The Accounting Review* 83 (3):789-822.

Luft, J., and M. Shields. 2003. Mapping management accounting research: graphics and guidelines for theory-consistent empirical research. *Accounting, Organizations and Society* 28 (2-3):169-250.

Maskin, E., Y. Qian, and C. Xu. 2000. Incentives, Information, and Organizational Form. *The Review of Economic Studies* 67 (2):359-378.

Milgrom, P., and J. Roberts. 1992. *Economics, organization and management*. International Edition ed. Englewood Cliffs, NJ: Prentice Hall.

Moon, H. R., and B. Perron. 2008. Seemingly unrelated regressions. In *The New Palgrave Dictionary of Economics*, edited by S. N. Durlauf and L. E. Blume: Palgrave Macmillan.

Nunnally, J. C., and I. H. Bernstein. 1994. *Psychometric theory*. New York: McGraw-Hill.

Podsakoff, P. M., S. B. MacKenzie, J. Y. Lee, and N. P. Podsakoff. 2003. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology* 88 (5):879-903.

Puranam, P., M. Raveendran, and T. Knudsen. 2012. Organization Design: The Epistemic Interdependence Perspective. *Academy of Management Review* 37 (3):419-440.

Raith, M. 2008. Specific knowledge and performance measurement. *RAND Journal of Economics* 39 (4):1059-1079.

Roberts, J. 2004. *The modern firm: Organizational design for performance and growth*. Oxford: Oxford University Press.

Smith, C. W., and R. L. Watts. 1992. The investment opportunity set and corporate financing, dividend and compensation policies. *Journal of Financial Economics* 32:263-292.

Thompson, J. D. 1967. *Organizations in action*. New York: McGraw-Hill.

Wooldridge, J. M. 2000. *Introductory econometrics: a modern approach*: South-Western College Publishing.

Zelner, B. A. 2009. Using simulation to interpret results from logit, probit, and other nonlinear models. *Strategic Management Journal* 30 (12):1335-1348.

APPENDIX A

Proof of Lemma 1:

The principal selects incentive weights that maximize expected firm profit net of the agents' compensation, subject to contract acceptance and effort choices by the agents. Specifically, the principal solves the following problem (Holmstrom and Milgrom 1990):

$$\max_{\{f_i,\alpha_i,\beta_i,\gamma_i,a_i,e_i\}_{i=1,\dots,N}} \mathrm{E}[x\,|\,\boldsymbol{a},\boldsymbol{e}]-\sum_{i=1}^{N}\mathrm{E}[w_i\,|\,f_i,\alpha_i,\beta_i,\gamma_i,\boldsymbol{a},\boldsymbol{e}] \tag{A.1a}$$

subject to, for all $i=1,\dots,N$,

$$CE_i(f_i,\alpha_i,\beta_i,\gamma_i,\boldsymbol{a},\boldsymbol{e})\geq 0 \text{, and} \tag{A.1b}$$

$$(a_i,e_i) = \text{argmax } CE_i(f_i,\alpha_i,\beta_i,\gamma_i,\boldsymbol{a},\boldsymbol{e}), \tag{A.1c}$$

with $\boldsymbol{a}=(a_1,\dots,a_N)$, $\boldsymbol{e}=(e_1,\dots,e_N)$,[30] and

$$\begin{aligned}CE_i(f_i,\alpha_i,\beta_i,\gamma_i,\boldsymbol{a},\boldsymbol{e})= f_i+\alpha_i\mathrm{E}[x\,|\,\boldsymbol{a},\boldsymbol{e}]+\beta_i\mathrm{E}[y_i\,|\,a_i,\boldsymbol{e}]+\gamma_i\mathrm{E}[z_i\,|\,a_i]-½(a_i^2+e_i^2)\\ -½r(\alpha_i^2\sigma_x^2+\beta_i^2\sigma_{yi}^2+\gamma_i^2\sigma_{zi}^2+2\alpha_i\beta_i\sigma_{yi}^2)\end{aligned}. \tag{A.2}$$

From the incentive compatibility constraints, (A.1c), agent *i*'s effort choice is given by:

$$a_i^*=b_i(\alpha_i+\beta_i+\gamma_i) \quad \text{and} \quad e_i^*=(\mu_{ii}+\textstyle\sum_{j\in T_i}\mu_{ji})\alpha_i+\mu_{ii}\beta_i. \tag{A.3}$$

To obtain the optimal incentive contract, for each agent *i*, we set (A.2) equal to zero (based on (A.1b)) and solve for f_i^*. Substituting f_i^*, a_i^*, and e_i^* into (A.1a), differentiating with respect to the incentive weights, and solving the associated first-order conditions yields the optimal incentive weights,

$$\alpha_i^*=D^{-1}\left(\left(\Sigma_{j\in T_i}\mu_{ji}\right)\left(\mu_{ii}+\Sigma_{j\in T_i}\mu_{ji}\right)\left(1+\tfrac{b_i^2}{r\,\sigma_{zi}^2}\right)+\frac{b_i^2\left(\Sigma_{j\in T_i}\mu_{ji}\right)^2}{r\,\sigma_{yi}^2}\right), \tag{A.4a}$$

$$\beta_i^*=D^{-1}\left(\left(\mu_{ii}-\tfrac{\sigma_{yi}^2}{\sigma_x^2}\,(\mu_{ii}+\Sigma_{j\in T_i}\mu_{ji})\right)\left(\mu_{ii}+\Sigma_{j\in T_i}\mu_{ji}\right)\left(1+\tfrac{b_i^2}{r\,\sigma_{zi}^2}\right)+b_i^2\left(1-\tfrac{\sigma_{yi}^2}{\sigma_x^2}\right)\right)\frac{\sigma_x^2}{\sigma_{yi}^2}\text{, and} \tag{A.4b}$$

[30] Agent *i* conjectures agent *j*'s actions. For expositional brevity, we do not explicitly highlight these conjectures. In equilibrium, the conjectures are correct.

$$\gamma_i^* = D^{-1} b_i^2 \left(\left(1 - \tfrac{\sigma_{yi}^2}{\sigma_x^2}\right) - \left(\mu_{ii} - \tfrac{\sigma_{yi}^2}{\sigma_x^2} \left(\mu_{ii} + \Sigma_{j \in T_i} \mu_{ji}\right)\right) \frac{\left(\Sigma_{j \in T_i} \mu_{ji}\right)}{r\, \sigma_{yi}^2} \right) \frac{\sigma_x^2}{\sigma_{zi}^2}, \tag{A.4c}$$

where $D = b_i^2 \left(\tfrac{\sigma_x^2}{\sigma_{yi}^2} - 1\right) + \left(\left(\Sigma_{j \in T_i} \mu_{ji}\right)^2 + \mu_{ii}^2 \left(\tfrac{\sigma_x^2}{\sigma_{yi}^2} - 1\right) + r(\sigma_x^2 - \sigma_{yi}^2) \right) \left(1 + \tfrac{b_i^2}{r\, \sigma_{zi}^2}\right) + \frac{b_i^2 \left(\Sigma_{j \in T_i} \mu_{ji}\right)^2}{r\, \sigma_{yi}^2} > 0$. The optimal relative incentive weights (3a) and (3b) follow from substituting (A.4a), (A.4b), and (A.4c) into $IRA_i^* = \alpha_i^* / \beta_i^*$ and $IRB_i^* = \gamma_i^* / \beta_i^*$.

Proof of Observation 1:

To provide a benchmark for the relative incentive weights in the multi-task, three performance measure setting of Observation 1, we first consider the special case of a single-task setting where the principal only uses own-level and above-level performance to provide incentives. Consistent with Bushman et al. (1995), each agent only provides externality effort (i.e., $b_i = 0$), implying $\gamma_i^* = 0$ and $IRB_i^* = 0$. Then, expression (3a) simplifies to an expression which is similar to Bushman et al. (1995), Proposition 1, i.e.,

$$IRA_i^* = \frac{\mu_{iT} \big/ \sigma_x^2}{\left(\mu_{ii} - \tfrac{\sigma_{yi}^2}{\sigma_x^2} \left(\mu_{ii} + \mu_{iT}\right)\right) \Big/ \sigma_{yi}^2}. \tag{A.5}$$

To rule out negative incentive weights, we assume $0 \le \mu_{iT} / \mu_{ii} \le \sigma_x^2 / \sigma_{yi}^2 - 1$.

For our main setting with two tasks and three performance measures, the relative incentive weights are given by (3a) and (3b). To rule out negative incentive weights, we assume

$$0 \le \frac{\mu_{iT}}{\mu_{ii}} \le \begin{cases} \dfrac{1}{2}\left(\dfrac{\sigma_x^2}{\sigma_{yi}^2} - 1\right)\left(1 - \sqrt{1 - \dfrac{4r(\sigma_{yi}^2 / \mu_{ii})^2}{\sigma_x^2 - \sigma_{yi}^2}}\right) & \text{if } r \le \dfrac{(\sigma_x^2 - \sigma_{yi}^2)}{4(\sigma_{yi}^2 / \mu_{ii})^2} \\ \left(\dfrac{\sigma_x^2}{\sigma_{yi}^2} - 1\right) & \text{otherwise} \end{cases}.$$

Table A.1 summarizes the comparative statics of variations in spillovers to and from other business units. Panel A reports the comparative statics for the single-task, two performance measure setting and

Panel B reports the comparative statics for the multi-task, three performance measure setting. In each panel, we differentiate between (i) the effort effect of spillovers; (ii) the noise effect of spillovers; and (iii) the joint effect of spillovers. The effort effect of spillovers describes the consequences from varying the marginal product of externality effort either of the focal agent (spillovers to other units) or the agent of another unit (spillovers from other units). The noise effect of spillovers describes the consequences from varying the noise in own-level performance either of another unit (spillovers to other units) or the focal unit (spillovers from other units). Finally, the joint effect of spillovers describes the consequences of variations in spillover in terms of both, marginal product and uncontrollable events, allowing for effects in multiple units.

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

Part (i) of the observation is apparent from Table A.1. Part (ii) follows from investigating the cutoffs derived for the joint effect of spillovers in Panel B (see below). A sufficient condition for Part (iii) is when the cutoffs $\hat{g}_i$ are replaced by either $\min\{\hat{g}_A, \hat{g}_B\}$ (such that $\partial IRA_i^* / \partial To > 0$ and $\partial IRB_i^* / \partial To > 0$) or by $\max\{\hat{g}_A, \hat{g}_B\}$ (such that $\partial IRA_i^* / \partial To < 0$ and $\partial IRB_i^* / \partial To < 0$). Proofs of the comparative statics follow below.

Panel A. Single-task, two performance measure setting

(i) Effort effect of spillovers:

$$\frac{\partial IRA_i^*}{\partial \mu_{ji}} = \left(\frac{1}{\mu_{iT}} + \frac{\sigma_{yi}^2 / \sigma_x^2}{\mu_{ii} - \frac{\sigma_{yi}^2}{\sigma_x^2}(\mu_{ii} + \mu_{iT})} \right) IRA_i^* > 0 \text{ and} \tag{A.6a}$$

$$\frac{\partial IRA_i^*}{\partial \mu_{ij}} = 0\,. \tag{A.6b}$$

(ii) Noise effect of spillovers:

$$\frac{\partial IRA_i^*}{\partial \sigma_{\eta ji}^2} = -\frac{1/\sigma_x^2}{\left(\mu_{ii} - \frac{\sigma_{yi}^2}{\sigma_x^2}(\mu_{ii} + \mu_{iT})\right)} IRA_i^* < 0 \text{ and} \tag{A.7a}$$

$$\frac{\partial IRA_i^*}{\partial \sigma_{\eta ij}^2} = \frac{1}{\sigma_{yi}^2}\left(1 + \frac{\frac{\sigma_{yi}^2}{\sigma_x^2}\mu_{iT}}{\left(\mu_{ii} - \frac{\sigma_{yi}^2}{\sigma_x^2}(\mu_{ii}+\mu_{iT})\right)}\right) IRA_i^* > 0 . \qquad \text{(A.7b)}$$

(iii) Joint effect of spillovers:

The joint effect of spillovers allows for multiple consequences of variations in spillover in terms of marginal product and uncontrollable events. Let To and $From$ represent spillovers to and from other units. Then, the joint effect of spillovers to other units is given by

$$\frac{\partial IRA_i^*}{\partial To} = \left(\left(\sigma_x^2 - \sigma_{yi}^2\right) - \mu_{iT} \frac{\frac{\partial(\Sigma_{j \in T_i} \sigma_{\eta ji}^2)}{\partial To}}{\frac{\partial \mu_{iT}}{\partial To}} \right) \frac{\mu_{ii}\frac{\partial \mu_{iT}}{\partial To}}{\mu_{iT}\left(\mu_{ii} - \frac{\sigma_{yi}^2}{\sigma_x^2}(\mu_{ii}+\mu_{iT})\right)\sigma_x^2} IRA_i^* ;$$

assuming $\partial \mu_{iT} / \partial To > 0$, $\partial IRA_i^* / \partial To > 0$ if, and only if,

$$\frac{\frac{\partial(\Sigma_{j \in T_i} \sigma_{\eta ji}^2)}{\partial To}}{\frac{\partial \mu_{iT}}{\partial To}} < g = \frac{\sigma_x^2 - \sigma_{yi}^2}{\mu_{iT}} . \qquad \text{(A.8a)}$$

The joint effect of spillovers from other units, assuming $\partial(\Sigma_{j \in F_i} \sigma_{\eta ji}^2) / \partial From > 0$, is given by

$$\frac{\partial IRA_i^*}{\partial From} = \left(\frac{1}{\sigma_{yi}^2} + \frac{\mu_{iT}}{\left(\mu_{ii} - \frac{\sigma_{yi}^2}{\sigma_x^2}(\mu_{ii}+\mu_{iT})\right)\sigma_x^2} \right) \frac{\partial(\Sigma_{j \in F_i} \sigma_{\eta ij}^2)}{\partial From} IRA_i^* > 0 , \qquad \text{(A.8b)}$$

where we use $\partial \sigma_{yi}^2 / \partial From = \partial \sigma_x^2 / \partial From = \partial(\Sigma_{j \in F_i} \sigma_{\eta ij}^2) / \partial From$ because of $x = \sum_i y_i$.

Panel B. Multi-task, three performance measure setting

(i) Effort effect of spillovers:

$$\frac{\partial IRA_i^*}{\partial \mu_{ji}} \propto \left\{ \left[\left(\mu_{ii}+\mu_{iT}\right)\left(1+\frac{b_i^2}{r\sigma_{zi}^2}\right) + \frac{b_i^2 \mu_{iT}}{r\sigma_{yi}^2} \right] \left[\mu_{ii}\left(\mu_{ii}+\mu_{iT}\right)\left(1+\frac{b_i^2}{r\sigma_{zi}^2}\right) + 2b_i^2 \right] - \mu_{ii} b_i^2 \left(1+\frac{b_i^2}{r\sigma_{zi}^2}\right) \right\} \left(1-\frac{\sigma_{yi}^2}{\sigma_x^2}\right) + \frac{b_i^2 \mu_{iT}}{r\sigma_{yi}^2} \mu_{ii} \left(\mu_{ii} - \frac{\sigma_{yi}^2}{\sigma_x^2}(\mu_{ii}+\mu_{iT})\right)\left(1+\frac{b_i^2}{r\sigma_{zi}^2}\right) > 0 , \qquad \text{(A.9a)}$$

because of the non-negative incentive weights, and

$$\frac{\partial IRA_i^*}{\partial \mu_{ij}} = 0 . \tag{A.9b}$$

Further,

$$\begin{aligned}\frac{\partial IRB_i^*}{\partial \mu_{ji}} \propto & \left[\left(2\left(\mu_{ii} + \mu_{iT}\right) - \mu_{ii}\right)\frac{\sigma_{yi}^2}{\sigma_x^2} - \mu_{ii}\right]\left[\left(1 + \frac{b_i^2}{r\,\sigma_{zi}^2} + \frac{b_i^2}{r\,\sigma_{yi}^2}\right)\left(1 - \frac{\sigma_{yi}^2}{\sigma_x^2}\right)\right. \\ & \left. + \frac{\mu_{ii}}{r\,\sigma_{yi}^2}\left(\mu_{ii} - \frac{\sigma_{yi}^2}{\sigma_x^2}(\mu_{ii} + \mu_{iT})\right)\left(1 + \frac{b_i^2}{r\,\sigma_{zi}^2}\right)\right] ; \\ & + \mu_{ii}\frac{\sigma_{yi}^2}{\sigma_x^2}\left(\left(1 - \frac{\sigma_{yi}^2}{\sigma_x^2}\right) - \left(\mu_{ii} - \frac{\sigma_{yi}^2}{\sigma_x^2}(\mu_{ii} + \mu_{iT})\right)\frac{\mu_{iT}}{r\,\sigma_{yi}^2}\right)\left(1 + \frac{b_i^2}{r\,\sigma_{zi}^2}\right)\end{aligned} \tag{A.10a}$$

a sufficient condition for $\partial IRB_i^* / \partial \mu_{ji} > 0$ is $\left(2\left(\mu_{ii} + \mu_{iT}\right) - \mu_{ii}\right)\frac{\sigma_{yi}^2}{\sigma_x^2} - \mu_{ii} > 0$, or if

$$\frac{\mu_{iT}}{\mu_{ii}} > g_B = \frac{1}{2}\frac{\sigma_x^2 - \sigma_{yi}^2}{\sigma_{yi}^2} > 0 ;$$

$$\frac{\partial IRB_i^*}{\partial \mu_{ij}} = 0 . \tag{A.10b}$$

(ii) Noise effect of spillovers:

$$\frac{\partial IRA_i^*}{\partial \sigma_{\eta ji}^2} = -\frac{\left(\mu_{ii}\left(\mu_{ii} + \mu_{iT}\right)\left(1 + \frac{b_i^2}{r\,\sigma_{zi}^2}\right) + b_i^2\right) IRA_i^* \Big/ \sigma_x^2}{\left(\left(\mu_{ii} - \frac{\sigma_{yi}^2}{\sigma_x^2}(\mu_{ii} + \mu_{iT})\right)\left(\mu_{ii} + \mu_{iT}\right)\left(1 + \frac{b_i^2}{r\,\sigma_{zi}^2}\right) + b_i^2\left(1 - \frac{\sigma_{yi}^2}{\sigma_x^2}\right)\right)} < 0 , \tag{A.11a}$$

$$\frac{\partial IRA_i^*}{\partial \sigma_{\eta ij}^2} = \frac{\mu_{iT}\left(\mu_{ii} + \mu_{iT}\right)\left(1 + \frac{b_i^2}{r\,\sigma_{zi}^2}\right)\left(1 + IRA_i^*\right) \Big/ \sigma_x^2}{\left(\mu_{ii} - \frac{\sigma_{yi}^2}{\sigma_x^2}(\mu_{ii} + \mu_{iT})\right)\left(\mu_{ii} + \mu_{iT}\right)\left(1 + \frac{b_i^2}{r\,\sigma_{zi}^2}\right) + b_i^2\left(1 - \frac{\sigma_{yi}^2}{\sigma_x^2}\right)} > 0 , \tag{A.11b}$$

$$\frac{\partial IRB_i^*}{\partial \sigma_{\eta ji}^2} \propto -\mu_{iT}\sigma_{yi}^2\left(\left(\mu_{ii} + \mu_{iT}\right)\left(1 + \frac{b_i^2}{r\,\sigma_{zi}^2}\right) + b_i^2\frac{\mu_{iT}}{r\,\sigma_{yi}^2}\right) \Big/ \sigma_x^4 < 0 \text{, and} \tag{A.11c}$$

$$\frac{\partial IRB_i^*}{\partial \sigma_{\eta ij}^2} \propto \left(\mu_{ii}\left(\mu_{ii} + \mu_{iT}\right)\left(1 + \frac{b_i^2}{r\,\sigma_{zi}^2}\right) + b_i^2\right)\left(\frac{\sigma_x^2}{\sigma_{yi}^2} - 1\right) + \frac{b_i^2\mu_{iT}^2}{r\,\sigma_{yi}^2} > 0 . \tag{A.11d}$$

(iii) Joint effect of spillovers:

The joint effect of spillovers to other units is given by

$$\frac{\partial IRA_i^*}{\partial To} \propto c_0^A - c_1^A \frac{\frac{\partial(\Sigma_{j\in T_i}\sigma_{\eta ji}^2)}{\partial To}}{\frac{\partial \mu_{iT}}{\partial To}}, \tag{A.12a}$$

where $c_0^A = \left(2\left(\mu_{ii}+\mu_{iT}\right)\left(1+IRA_i^*\right) - \mu_{ii}\left(1+\frac{\sigma_x^2}{\sigma_{yi}^2}IRA_i^*\right)\right)\left(1+\frac{b_i^2}{r\,\sigma_{zi}^2}\right) + 2\frac{b_i^2\mu_{iT}}{r\,\sigma_{yi}^2} \geq 0$ and

$c_1^A = \left(b_i^2 + \mu_{ii}\left(\mu_{ii}+\mu_{iT}\right)\left(1+\frac{b_i^2}{r\,\sigma_{zi}^2}\right)\right) IRA_i^* \big/ \sigma_{yi}^2 \geq 0$. Thus, $\partial IRA_i^* / \partial To > 0$ if, and only if,

$$\frac{\frac{\partial(\Sigma_{j\in T_i}\sigma_{\eta ji}^2)}{\partial To}}{\frac{\partial \mu_{iT}}{\partial To}} < \hat{g}_A = \frac{c_0^A}{c_1^A} > 0, \tag{A.12b}$$

and

$$\frac{\partial IRB_i^*}{\partial To} \propto c_0^B - c_1^B \frac{\frac{\partial(\Sigma_{j\in T_i}\sigma_{\eta ji}^2)}{\partial To}}{\frac{\partial \mu_{iT}}{\partial To}}, \tag{A.12c}$$

where $c_0^B = \frac{b_i^2}{r\,\sigma_{zi}^2}\left(2\mu_{iT} - \mu_{ii}\left(\frac{\sigma_x^2}{\sigma_{yi}^2}-1\right)\right) + IRB_i^*\left(2\mu_{iT} - \mu_{ii}\left(\frac{\sigma_x^2}{\sigma_{yi}^2}-2\right)\left(1+\frac{b_i^2}{r\,\sigma_{zi}^2}\right)\right)$ and

$c_1^B = \frac{b_i^2}{r\,\sigma_{zi}^2}\left(\mu_{ii}\frac{\mu_{iT}}{r\,\sigma_{yi}^2}-1\right) + \frac{IRB_i^*}{\sigma_{yi}^2}\left(b_i^2 + \mu_{ii}\left(\mu_{ii}+\mu_{iT}\right)\left(1+\frac{b_i^2}{r\,\sigma_{zi}^2}\right)\right) \geq 0$. Thus, $\partial IRB_i^* / \partial To > 0$ if, and only if,

$$\frac{\frac{\partial(\Sigma_{j\in T_i}\sigma_{\eta ji}^2)}{\partial To}}{\frac{\partial \mu_{iT}}{\partial To}} < \hat{g}_B = \frac{c_0^B}{c_1^B}, \tag{A.12d}$$

where $\mu_{iT} / \mu_{ii} > g_B$ is a sufficient condition for $c_0^B > 0$. We note that $\hat{g}_A$ and $\hat{g}_B$ depend on σ_x^2 and σ_{yi}^2, implying that $\hat{g}_A$ and $\hat{g}_B$ vary with spillovers from other units, $From$.

The joint effect of spillovers from other units, assuming $\partial(\Sigma_{j\in F_i}\sigma_{\eta ji}^2) \big/ \partial From > 0$, is given by

$$\frac{\partial IRA_i^*}{\partial From} = \frac{\mu_{iT}\left(\mu_{ii}+\mu_{iT}\right)\left(1+\frac{b_i^2}{r\,\sigma_{zi}^2}\right)\left(1+IRA_i^*\right)\Big/\sigma_x^2}{\left(\mu_{ii}-\frac{\sigma_{yi}^2}{\sigma_x^2}\left(\mu_{ii}+\mu_{iT}\right)\right)\left(\mu_{ii}+\mu_{iT}\right)\left(1+\frac{b_i^2}{r\,\sigma_{zi}^2}\right)+b_i^2\left(1-\frac{\sigma_{yi}^2}{\sigma_x^2}\right)}\frac{\partial(\Sigma_{j\in F_i}\sigma_{\eta ij}^2)}{\partial From} > 0\,, \tag{A.13a}$$

$$\frac{\partial IRB_i^*}{\partial From} = \left(\frac{\left(\frac{\sigma_x^2-\sigma_{yi}^2}{\sigma_{yi}^2}+\frac{\mu_{iT}^2}{r\,\sigma_{yi}^2}\right)\Big/\sigma_x^2}{\left(1-\frac{\sigma_{yi}^2}{\sigma_x^2}\right)-\left(\mu_{ii}-\frac{\sigma_{yi}^2}{\sigma_x^2}\left(\mu_{ii}+\mu_{iT}\right)\right)\frac{\mu_{iT}}{r\,\sigma_{yi}^2}} + \frac{\mu_{iT}\left(\mu_{ii}+\mu_{iT}\right)\left(1+\frac{b_i^2}{r\,\sigma_{zi}^2}\right)\Big/\sigma_x^2}{\left(\mu_{ii}-\frac{\sigma_{yi}^2}{\sigma_x^2}\left(\mu_{ii}+\mu_{iT}\right)\right)\left(\mu_{ii}+\mu_{iT}\right)\left(1+\frac{b_i^2}{r\,\sigma_{zi}^2}\right)+b_i^2\left(1-\frac{\sigma_{yi}^2}{\sigma_x^2}\right)}\right) IRB_i^*\,\frac{\partial(\Sigma_{j\in F_i}\sigma_{\eta ij}^2)}{\partial From} > 0\,. \tag{A.13b}$$

TABLE A.1

Change of Optimal Relative Incentive Weights, IRA_i^* and IRB_i^*, for Given Variations in Spillover

Column (i) reports the effort effect of spillovers (i.e., a variation of the marginal product of externality effort as a consequence of spillovers), Column (ii) reports the noise effect of spillovers (i.e., a variation of performance measure noise as a consequence of spillovers), and Column (iii) reports the joint effect of spillovers. Spillovers *to* other units influence the set of affected units (T_i), the marginal product of externality effort (μ_{ji}), and/or the externality noise ($\sigma^2_{\eta ji}$), whereas spillovers *from* other units influence the set of affecting units (F_i), the marginal product of externality effort (μ_{ij}), and/or the externality noise ($\sigma^2_{\eta ij}$). The cutoffs g, g_B, $\hat{g}_A(From)$, and $\hat{g}_B(From)$ are defined in the proofs.

	(i) Effort effect of spillovers	**(ii) Noise effect of spillovers**	**(iii) Joint effect of spillovers**
Panel A. Single-task, two performance measure setting ($b_i = 0$, where $IRA_i^* > 0$ and $IRB_i^* = 0$)			
Spillovers *to* other units	$\frac{\partial IRA_i^*}{\partial \mu_{ji}} > 0$	$\frac{\partial IRA_i^*}{\partial \sigma^2_{\eta ji}} < 0$	$\frac{\partial IRA_i^*}{\partial To} > 0$ iff $\frac{\frac{\partial(\Sigma_{j\in T_i}\sigma^2_{\eta ji})}{\partial To}}{\frac{\partial(\Sigma_{j\in T_i}\mu_{ji})}{\partial To}} < g$
Spillovers *from* other units	$\frac{\partial IRA_i^*}{\partial \mu_{ij}} = 0$	$\frac{\partial IRA_i^*}{\partial \sigma^2_{\eta ij}} > 0$	$\frac{\partial IRA_i^*}{\partial From} > 0$
Panel B. Multi-task, three performance measure setting ($b_i > 0$, where $IRA_i^* > 0$ and $IRB_i^* > 0$)			
Spillovers *to* other units	$\frac{\partial IRA_i^*}{\partial \mu_{ji}} > 0$ $\frac{\partial IRB_i^*}{\partial \mu_{ji}} > 0$ if $\frac{\sum_{j\in T_i}\mu_{ji}}{\mu_{ii}} > g_B$	$\frac{\partial IRA_i^*}{\partial \sigma^2_{\eta ji}} < 0$ $\frac{\partial IRB_i^*}{\partial \sigma^2_{\eta ji}} < 0$	$\frac{\partial IRA_i^*}{\partial To} > 0$ iff $\frac{\frac{\partial(\Sigma_{j\in T_i}\sigma^2_{\eta ji})}{\partial To}}{\frac{\partial(\Sigma_{j\in T_i}\mu_{ji})}{\partial To}} < \hat{g}_A(From)$ $\frac{\partial IRB_i^*}{\partial To} > 0$ iff $\frac{\frac{\partial(\Sigma_{j\in T_i}\sigma^2_{\eta ji})}{\partial To}}{\frac{\partial(\Sigma_{j\in T_i}\mu_{ji})}{\partial To}} < \hat{g}_B(From)$
Spillovers *from* other units	$\frac{\partial IRA_i^*}{\partial \mu_{ij}} = \frac{\partial IRB_i^*}{\partial \mu_{ij}} = 0$	$\frac{\partial IRA_i^*}{\partial \sigma^2_{\eta ij}}, \frac{\partial IRB_i^*}{\partial \sigma^2_{\eta ij}} > 0$	$\frac{\partial IRA_i^*}{\partial From}, \frac{\partial IRB_i^*}{\partial From} > 0$

APPENDIX B

Appendix B 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 BUs
To what extent do your actions affect the performance of other units in the firm?

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

Supply to other BUs
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?

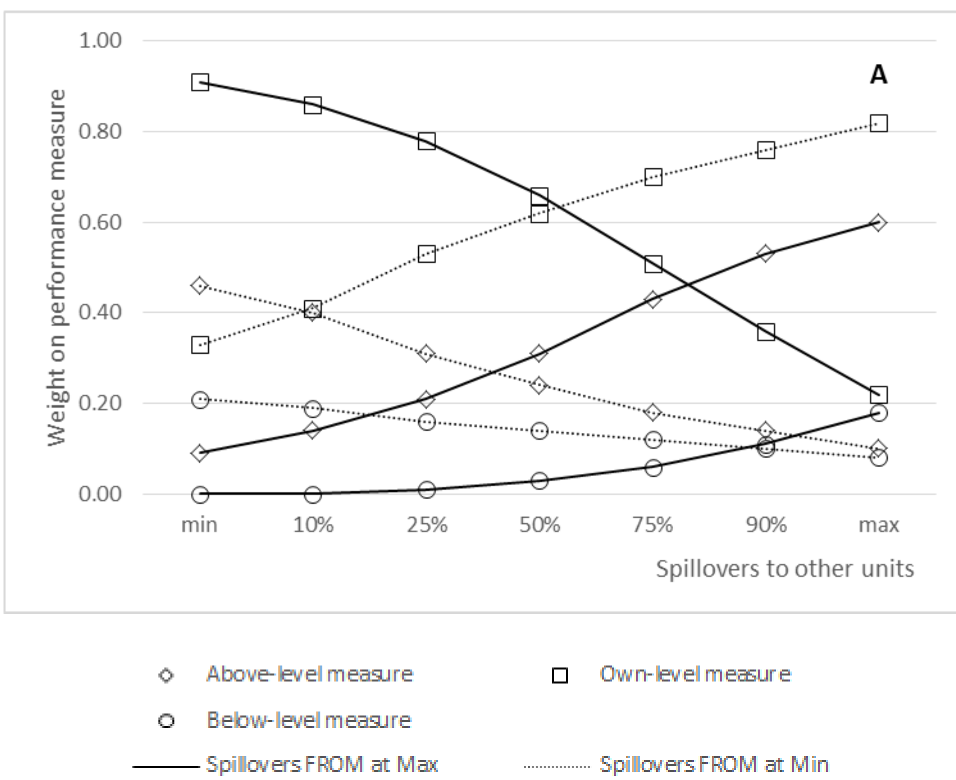
Scale: 1 = my superior does
4 = my superior and I do, almost equally
7 = I do

F I G U R E 1

Analysis of the Expected Percentage Weight Placed on Performance Measures for Varying Degrees of Spillovers

The figure presents simulations of the expected percentage weight placed on a performance measure (above-level, own-level, below-level) when *Spillovers to other BUs* (*Spillovers from other BUs*) is changed from its sample minimum to its sample maximum value, while holding *Spillovers from other BUs* (*Spillovers to other BUs*) constant at the sample minimum or the 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 percentage weights 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 a given percentile 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 the given percentile 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. Note that we account for the interaction between the spillover effects when computing the expected values. The figure in Panel A (B) plots the expected percentage weights for different values of Spillovers to (from) other units.

Panel A: Simulation of percentage weight on performance measures for different values of *Spillovers to other BUs*



Panel B: Simulation of percentage weight on performance measures for different values of *Spillovers from other BUs*

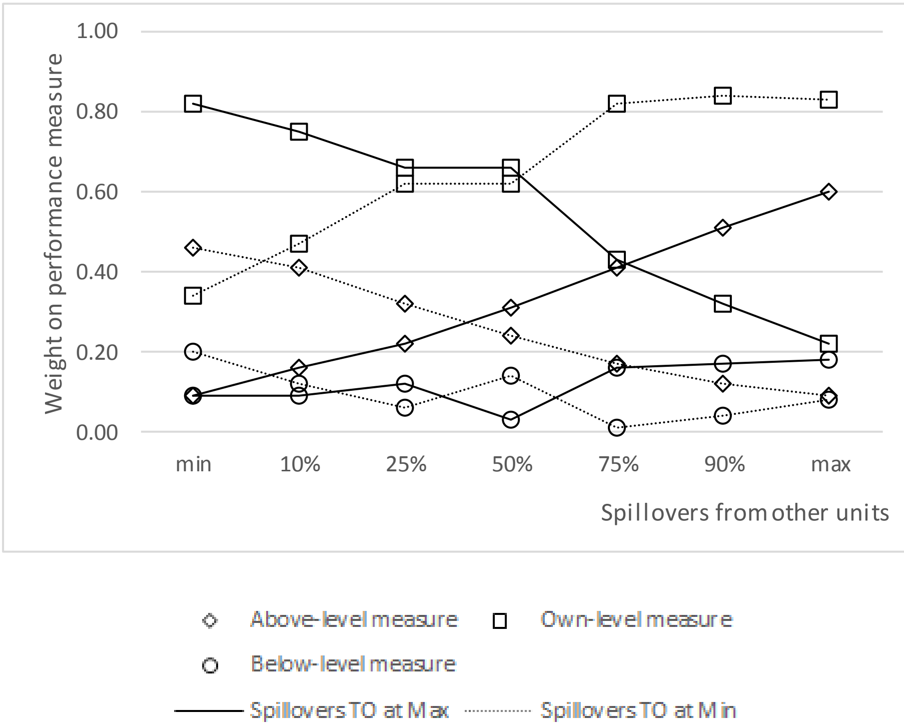


TABLE 1

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.

TABLE 2

Summary Statistics on Weightings of Different Types of Performance Measures and Spillover 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 BUs, Spillovers from other BUs, and corresponding convergent validity test variables. N = 122. We report statistics before mean-centering the variables.

Spillover variables	Mean	Std. Dev.	Min.	Median	Max.
Spillovers to other BUs	4.02	1.52	1.00	4.00	7.00
Spillovers from other BUs	4.12	1.55	1.00	4.00	7.00
Supply to other BUs (in %)	15.89	27.20	0.00	5.00	100.00
Supply from other BUs (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 spillover variables. N=122

	(1)	(2)	(3)	(4)	(5)	(6)
1. Spillovers to other BUs	1					
2. Spillovers from other BUs	0.58	1				
3. Supply to other BUs (in %)	0.34	0.22	1			
4. Supply from other BUs (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 3

Summary Statistics for Questionnaire Items Used to Construct the Control Variables

Table 3 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 B.

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

TABLE 4

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

Table 4 presents Pearson correlations among all variables used in this study. Variable definitions appear in Appendix B. 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 to other BUs		0.11	–0.19	0.16	1.00							
5. Spillovers from other BUs		0.18	–0.13	–0.06	0.58	1.00						
6. Unsigned spillovers		0.17	–0.18	0.04	0.90	0.89	1.00					
7. Decentralization	0.71	–0.19	0.03	0.26	0.13	–0.08	0.05	1.00				
8. Growth opportunities	0.85	0.09	–0.09	–0.01	–0.05	–0.08	–0.08	0.01	1.00			
9. Competition	0.54	–0.03	0.13	–0.19	0.08	0.08	0.10	–0.07	–0.08	1.00		
10. Varcomp		–0.04	0.19	–0.28	–0.19	–0.08	–0.15	–0.14	0.16	0.15	1.00	
11. Size		0.03	0.11	–0.25	0.00	0.06	0.03	–0.13	–0.23	0.24	0.36	1.00
12. Service		–0.01	0.04	–0.05	0.05	0.01	0.01	–0.02	–0.07	0.05	0.01	–0.10

TABLE 5

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

Table 5 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 other BUs* and *Spillovers to other BUs*, which have been mean-centered, and their interaction term *To×From*, or *Unsigned spillovers* which measures the spillovers between business units within a firm. We control for *Varcomp*, *Decentralization*, *Competition*, *Growth opportunities*, *Size*, and *Service*. These variables are defined in Appendix B. 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.

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