GOVERNANCE AND THE EFFICIENCY OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 410

Merger Efficiency and Managerial Incentives

Matthias Kräkel *
Daniel Müller **

* University of Bonn
** University of Bonn

June 2013

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Sonderforschungsbereich/Transregio 15 · www.sfbtr15.de
Universität Mannheim · Freie Universität Berlin · Humboldt-Universität zu Berlin · Ludwig-Maximilians-Universität München
Rheinische Friedrich-Wilhelms-Universität Bonn · Zentrum für Europäische Wirtschaftsforschung Mannheim

Speaker: Prof. Dr. Klaus M. Schmidt · Department of Economics · University of Munich · D-80539 Munich,
Phone: +49(89)2180 2250 · Fax: +49(89)2180 3510

Merger Efficiency and Managerial Incentives*

Matthias Kräkel† and Daniel Müller‡

June 2, 2013

We consider a two-stage principal-agent model with limited liability in which a CEO is employed as agent to gather information about suitable merger targets and to manage the merged corporation in case of an acquisition. Our results show that the CEO systematically recommends targets with low synergies—even when targets with high synergies are available—to obtain high-powered incentives and, hence, a high personal income at the merger-management stage. We derive conditions under which shareholders prefer a self-commitment policy or a rent-reduction policy to deter the CEO from opportunistic recommendations.

JEL classification: D82; D86; G34

Keywords: acquisition; merger; moral hazard

---

*We have benefited from comments made by Martin Peitz, Andreas Roider, Patrick Schmitz, Monika Schnitzer, Jan Schymik, the seminar audience at the University of Regensburg, and the participants of the SFB/TR 15 meeting in Berlin (2012). All errors are of course our own.

†University of Bonn, Department of Economics, Institute for Applied Microeconomics, Adenauerallee 24-42, D-53113 Bonn, Germany, E-mail address: m.kraekel@uni-bonn.de,; Tel: +49-228-739211; Fax: +49-228-739210

‡University of Bonn, Department of Economics, Institute for Applied Microeconomics, Adenauerallee 24-42, D-53113 Bonn, Germany, E-mail address: daniel.mueller@uni-bonn.de, Tel: +49-228-739212, Fax: +49-228-739210.

1. Introduction

On May 07, 1998, the Daimler-Benz AG and the Chrysler Corporation merged into the DaimlerChrysler AG, one of the world's biggest car manufacturers with 442,000 employees and a market value of about $100 billion. The former Daimler Chief Executive Officer (CEO), Jürgen Schrempp, promised huge synergy savings in distribution, product design, and research & development. Leading newspapers were less optimistic. On the day following the merger, the New York Times stated that "at a news conference held here to proclaim the biggest industrial marriage in history, neither company could explain in detail where billions of dollars in savings from reduced expenses would come from" (Andrews 1998). In 2001, these fears were confirmed by the actual course of events—the market value of DaimlerChrysler shrank to $44 billion, which was nearly the pre-merger market value of the Daimler-Benz AG alone. Thus, synergies either remained unexploited or did not exist.[1]

Nevertheless, the merger had one clear winner—the 1998 Daimler CEO and later DaimlerChrysler CEO Jürgen Schrempp. Before merging, his estimated yearly income amounted to $2.9 million. After merging, the pay system for top executives at Daimler-Benz changed dramatically: at least 70 percent of top executive compensation became performance bonuses and other incentive payments (Bryant 1999). As a consequence, the new estimated income of Jürgen Schrempp (at least) doubled. There does not only exist anecdotal evidence for the observation that the income of an acquiring firm's CEO rises considerably—even after a merger that leads to low or no synergies. The empirical results of Bliss and Rosen (2001), Anderson et al. (2004), Grinstein and Hribar (2004), Bebchuk and Grinstein (2005), Girma et al. (2006), Harford and Li (2007), and Guest (2009) show that this observation can be considered as a stylized fact.[2]

With acquisitions leading to higher CEO compensation, an immediately related question is how the anticipation of this positive income effect affects the quality of acquisition decisions. In the following, we offer a rationale for why CEOs prefer low-synergy mergers over high-synergy mergers, and how they benefit from poor merger quality. We consider a two-stage principal-agent relationship between a CEO, on the one hand, and the board of directors or the shareholders—henceforth summarized as the "principal"—on the other. The CEO is protected by limited liability. In line with the observation in Anderson et al. (2004) that changes in CEO compensation following a merger are likely to reflect a restructuring of incentives, we assume this principal-agent relationship to be governed by a series of short-term contracts. In the first stage, the CEO gathers information on possible merger targets and recommends a target to the principal. At the end of the first stage, the principal decides on whether to acquire the target firm or not. In the second stage, in case of merging, the CEO is employed to manage the merged firm. At this stage, the principal can optimally fine-tune CEO incentives by using bonuses

[1] As the article "Dark Days at Daimler" published in BusinessWeek on August 15, 2005, put it: "Chrysler proved to be a massive rescue job that sucked up billions and absorbed German management for years [...]. Synergies have been few and far between."

[2] See Williams et al. (2008) for a literature survey.

that depend on the CEO's performance. Our analysis shows that if a CEO identifies both low- and high-synergy targets, he will tend to recommend a low-synergy one to make the principal choose high-powered incentives at the merger-management stage, yielding a large rent to the CEO. This result, providing one possible explanation for the low synergies from the Daimler-Chrysler merger, sheds light on how CEOs can manipulate their post-merger remuneration by making opportunistic merger recommendations. Besides the case of Daimler-Chrysler, there exists broad empirical evidence that merging often leads to poor or disastrous outcomes (e.g., Jensen and Ruback 1983, Jarrell et al. 1988, Bradley et al. 1988, Morck et al. 1990, Bruner 2005). This empirical literature is in line with our theoretical findings.

Regarding the CEO's recommendation of a merger target, we focus on decision-based incentives throughout the paper: while the synergies of the recommended target firm are verifiable for the principal, CEO pay in the first stage can only condition on the fact whether an acquisition takes place or not.[3] We find that the principal may benefit from offering the CEO a sufficiently high wage premium in case of an acquisition, although the quality of the CEO's recommendation of a merger target is not contractible. Offering a large acquisition premium acts as a commitment device for the principal not to approve low-synergy recommendations because low-synergy targets will not justify the high CEO pay. Consequently, the CEO is kept from opportunistically recommending a low-synergy merger target while identifying high-synergy targets at the same time.

In practice, CEOs often bear personal costs from merging (e.g., traveling between the headquarter and the newly acquired firm). The principal may be able to influence these costs (e.g., the frequency of traveling) and include them in the contractual terms offered to the CEO. In this case, it may be optimal to impose sufficiently high costs on the CEO to reduce his rents from a low-synergy merger and, thereby, influence his recommendation of merger targets. This rent-reduction strategy, however, leads to lower expected profits for the principal in case of a high-synergy merger because of a binding participation constraint of the CEO. We summarize conditions under which influencing the CEO's recommendation via endogenous personal costs is more profitable for the principal than the commitment solution described in the previous paragraph.

In a final step, we assume that information gathering by the CEO is endogenous. If the CEO exerts costly effort in the first stage of the game, he increases the number of target firms whose synergies he may then learn about. In this situation of repeated moral hazard, prospective rents from merger management in the second stage create implicit incentives for the CEO to gather information in the first stage. Information gathering can further be motivated by a first-period wage premium in case of acquisition of a target firm. We show that if the probability of detecting a low-synergy target and the principal's relative profit loss from opportunistic recommendation

[3]The incomplete contracting assumption of decision-based rewards was introduced by Dewatripont and Tirole (1999). According to this approach, incentive schemes condition on actual decisions but not on the content or quality of the information underlying these decisions.

are sufficiently large, the principal will benefit from *dis*incentivizing the CEO by offering a premium for *not* recommending a merger target in the first stage.

The rest of the paper is organized as follows. We start with a review of the related literature in Section 2. In Section 3, we introduce our basic model, which is analyzed in Section 4. Section 5 discusses several modifications of the basic model to check the robustness of our main finding. We conclude in Section 6. All proofs are deferred to the Appendix.

2. Related Literature

Besides the empirical work on post-merger CEO pay cited above, our paper is related to part of the merger literature that explains why CEOs sometimes choose merger targets with low synergies.[4] First, CEOs may receive a utility from empire building (e.g., higher prestige) and ignore synergies (Baumol 1959, Marris 1963, Williamson 1963, Jensen 1986). Second, overconfident CEOs may imagine to measure the true value of a target firm more precisely than the whole capital market, leading to the well-known hubris effect (Roll 1986). Third, CEOs may prefer to invest in those industries in which they are experts in order to entrench themselves (Shleifer and Vishny 1989). Fourth, a raider may decide to acquire a target firm to benefit from a breach of implicit contracts with the workforce and other stakeholders (Shleifer and Summers 1988, Schnitzer 1995, Brusco 1996). Finally, a risk averse CEO of an acquiring firm may benefit from the diversification of personal risk (Amihud and Lev 1981, Morck et al. 1990). These theories do not exclude the possibility that CEOs occasionally acquire merger targets with low synergies. However, our approach points out that CEOs may *systematically* prefer inefficient mergers to efficient ones and deliberately choose a poor merger target even when they have the opportunity to aquire a more profitable firm. This finding fits quite well to the conclusion of Williams et al. (2008) that managers seem to benefit from mergers that are not in their shareholders' best interest. Moreover, contrary to our paper, the aforementioned theories cannot explain why the incomes of the acquiring firms' CEOs increase and why mergers are often accompanied by a restructuring of a firm's compensation system.

Furthermore, our paper contributes to the literature on real authority and project recommendation.[5] While we share the basic information structure, our assumptions on authority, verifiability, and contractual form are complementary to those in Dow and Raposo (2005), who explore the determinants of a CEO's choice of corporate strategy. With no aspect of corporate strategy being verifiable, in Dow and Raposo (2005) only the long-term success of the firm is contractible. The principal-agent relationship is governed by a long-term contract, which is renegotiated once before the CEO decides which of the strategies to implement. Consequently,

[4]For an overview see DeBondt and Thompson (1992).

[5]The seminal papers by Aghion and Tirole (1997) and Baker et al. (1999) do not discuss the interplay of project recommendation and subsequent optimal incentive provision. Moreover, in our paper, the second-stage moral hazard problem endogenously implies the conflict of interests between principal and agent, which is exogenously given in Aghion and Tirole (1997) and Baker et al. (1999).

renegotiation takes place irrespective of whether a change in strategy occurs or the status quo is maintained. With our focus on M&A activity, we assume that whether an acquisition occurs is verifiable by a third party. Moreover, merger management is governed by a separate contract, which is stipulated only if the principal decided to acquire the merger target proposed by the CEO. Our sequence of events thus resonates well with the observation that "increases in compensation following a merger are likely to represent a restructuring of incentives to encourage managers to respond to the challenges of leading a more complex organization" (Williams et al., 2008, p. 333). Also our suggestion how to manage the arising conflict of interests is novel. In Dow and Raposo (2005) the conflict of interests is overcome by stipulating an "excessively high" bonus payment for long-term firm success in the initial contract, which imposes a floor on the wage the principal can offer in renegotiation. In the optimum, the initial wage is set sufficiently high such that under renegotiation the CEO is indifferent between a moderate and a drastic corporate strategy, which prevents withholding of information by the CEO at the cost of higher bonus payments if no conflict of interests prevails. Our commitment-based resolution via a sufficiently high acquisition premium is not feasible in Dow and Raposo (2005) where only long-term firm success is verifiable.

Our result that the principal may find it optimal to pay a high acquisition premium in order to commit herself not to approve low-synergy mergers is reminiscent of the recent finding by Berkovitch et al. (2010) regarding organizational design: if project recommendation is subject to managerial moral hazard, then implementation of the less efficient unitary functional structure (U-form) may favorably affect the manager's recommendation behavior by making projects preferred by the manager too costly to implement, thereby outperforming the more efficient multidivisional structure (M-form). In Berkovitch et al. (2010), however, the choice of organizational structure is the only way to influence the manager's recommendation behavior—monetary incentive schemes are assumed to be ineffective. Complementary to this approach, our paper, which endogenizes the manager's preferences over merger projects, explores information management in incentive problems via traditional monetary reward schedules. Furthermore, we go beyond the analyses in Dow and Raposo (2005) and Berkovitch et al. (2010) by addressing the incentivization of information acquisition in the shadow of the conflict of interests and how non-monetary means such as working conditions can help to overcome this conflict.

Our paper is also related to the literature on information management in principal-agent relationships. Early papers in the literature (e.g., Lambert 1986, Demski and Sappington 1987) consider moral hazard only with regard to information gathering but do not allow for this to be followed by a moral hazard situation.[6] In Lewis and Sappington (1997), the agent first decides on gathering information about the initially unknown state of the world and, thereafter, chooses cost-reducing effort. The principal-agent relationship is governed by a single, non-renegotiable

[6]Moreover, in these papers the agent cannot communicate his information and makes any productive decision himself.

long-term contract that covers both information gathering and effort provision. Lewis and Sappington (1997) find that very high-powered incentives are needed to induce effective information gathering by the agent.[7] In contrast, in our paper the principal-agent relationship is governed by a series of short-term contracts such that the actual outcome of the gathering of information affects the form of the subsequent incentive contract. As a consequence, in our model the principal may prefer to dampen incentives for the gathering of information if the agent's second-period rent is large and much information is rather detrimental to the principal (cf. Section 5.4).

Finally, with the gathering of information requiring costly effort (cf. Section 5.4), our paper adds to the principal-agent literature on sequential moral hazard with a risk-neutral and wealth-constrained agent. If an agent exerts effort in two subsequent periods, second-period rents can be utilized by the principal to optimally design first-period incentives. This effect was first emphasized by Schmitz (2005a) and further elaborated by Schmitz (2005b, 2012), Ohlendorf and Schmitz (2012), and Kräkel and Schöttner (2010). In our setting, contrary to these contributions, the principal may actually find it optimal to dampen first-period incentives that stem from prospective second-period rents.

3. THE MODEL

Consider a relationship between a principal (she) and an agent (he)—both risk neutral—that lasts for two periods, $t = 1, 2$. The agent is protected by limited liability, i.e., wage payments to the agent have to be nonnegative in each period. The principal wants to engage in merger-and-acquisition (M&A) activities, but lacks the expertise and/or the time to run these activities herself. The agent's task in the first period therefore is to identify a potential merger target, i.e. a firm that the principal might acquire. If acquisition takes place and the principal wants to proceed with the merger, the agent's task in the second period then is to manage the merged firm. The probability of the merged firm succeeding in the market depends on both the agent's effort and the synergies created by the merger. The agent is offered a new contract in each period. For simplicity, the agent's outside opportunity in each period is zero. Moreover, under the absence of any merger activity, the principal's business in each period generates a stand-alone profit of zero.[8]

M&A information gathering.—At the beginning of their relationship, both principal and agent know that there are $n \geq 2$ potential M&A targets, but they are uninformed about the targets' specific M&A synergies. At this point, both parties share the same prior probability distribution regarding the synergies of the potential target firms. Ex ante, all n target firms are identi-

[7] In the same vein Khalil et al. (2005), who consider the task design problem of Lewis and Sappington (1997) when implementation costs are not observable, as well as Taylor (1995), who considers a repeated game setting, assume that information gathering and incentive provision are governed by a single contract.

[8] By stand-alone profit we refer to the profit a firm generates if it is run according to the current modus operandi and conducts business as usual.

cal with the synergies δ^j of target firm j ($j = 1, \ldots, n$) being drawn from the set $\{-\infty, \delta_L, \delta_H\}$, where $0 < \delta_L < \delta_H$, according to some distribution F with probability $P(\delta^j = \delta) > 0$ for all $\delta \in \{-\infty, \delta_L, \delta_H\}$. In the first period, with probability $1 - i$, the agent remains completely uninformed. With probability $i \in (0, 1)$, on the other hand, the agent learns about the synergies of all n merger targets, where the profile of actual synergies is denoted by $\Delta = \{\delta^1, \cdots, \delta^n\}$. The question of whether information gathering was successful as well as the profile Δ of synergies in case of successful information gathering are private information of the agent. However, the agent can send a report r to the principal that points to a specific merger target. The report either recommends a particular target, $r = T \in \{1, \ldots, n\}$, or nothing, $r = \varnothing$. The agent can offer information on δ^T to justify his recommendation. While this information makes δ^T verifiable for the principal, communication r as well as the information on δ^T are unverifiable by a third party. Thus, the agent cannot be forced by the principal to fully reveal Δ in case of successful information gathering.[9] If the agent is indifferent between several targets, we assume that he will recommend the target with the highest synergies. Moreover, if the agent is indifferent between making a merger recommendation or not, he will make a recommendation. After the agent has made his report, the principal decides whether to proceed with the acquisition and (if so) which target firm to acquire.

M&A synergies.—If the principal acquires a target firm j with negative synergies, $\delta^j = -\infty$, she will go bankrupt after the first period and suffer an extreme loss of $-\infty$ in the second period—e.g., from losing everything she owns due to insolvency. In this sense, the mere acquisition of a target firm with negative synergies severely harms the principal's core business and forces her out of business.[10] If, on the other hand, the principal acquires a target firm j with strictly positive synergies, i.e., $\delta^j \in \{\delta_L, \delta_H\}$, she can then, at the beginning of the second period, choose between running two independent businesses or merging her two businesses. In the former case, each business generates its stand-alone profit. For simplicity, the stand-alone profit of a target firm with $\delta^j > 0$ is set to zero, such that the principal can acquire any such firm at the end of the first period at no cost, reflecting its market value. If the principal decides to conduct a merger of her two businesses in order to capitalize on the synergies, then she has to employ the agent to manage the merged firm—tasks such as identification and realization of potential cost savings, restructuring of assets, and reconfiguration of the organization all require managerial effort. The principal cannot replace the current agent by another one, because the current agent has acquired valuable target-specific knowledge that is not transferable to a new agent. The merged firm's success, π, depends on both the synergies created by the merger and

[9] The assumption of communication not being verifiable is made for the ease of exposition. Alternatively, we can think of a setting in which n (i.e., the number of identified merger targets) is stochastic and only the agent observes the realization of n. In such a setting, the agent can always claim that $n = 1$ when recommending target T.

[10] For example, imagine the case where the acquired firm realizes a huge loss ex post and the principal, as new owner, is liable for the loss. Alternatively, it is conceivable that the principal is harmed considerably when it turns out that the acquired firm is involved in criminal activities.

the second-period managing effort exerted by the agent. The agent's effort choice, $e \geq 0$, is unobservable for the principal and comes at cost $c(e)$ for the agent, where $c'(e) > 0$ for $e > 0$, $c''(e) > 0$, and $c(0) = c'(0) = 0$. If the agent exerts effort e and the acquired firm exhibits synergies $\delta > 0$, the merged firm's profit is high, $\pi = \pi_H$, with probability $p(e + \delta) \in (0, 1)$, and low, $\pi = \pi_L < \pi_H$, otherwise. The success probability is monotonically increasing and concave, $p' > 0$ and $p'' < 0$. In case of a merger, the agent bears an additional personal cost $\kappa \geq 0$, i.e., management of the merged firm leads to an additional disutility for the agent.[11] If the principal is indifferent between acquiring and not acquiring the target firm, she will not pursue the acquisition. If the principal is indifferent between a merger or running two independent businesses, she will pursue the merger.

Contracting.—With the information gathered and communicated by the agent in the first period being unverifiable, the first-period contract specifies a wage payment contingent on whether an acquisition has occurred or not.[12] The agent receives w_{1H} in case of an acquisition and w_{1L} otherwise.[13] Regarding the second-period contract, we assume that the merged firm's success π is not verifiable because it is realized in the distant future and thus cannot be used for current contracting purposes. Instead, there is a contractible binary performance measure on the agent's managerial effort, $\sigma \in \{\sigma_L, \sigma_H\}$. The higher the agent's managerial effort, the larger the probability of realization σ_H of the performance measure: $q(e) = P(\sigma = \sigma_H | e) \in [0, 1)$, with $q'(e) > 0$, $q''(e) \leq 0$, and $q(0) = 0$. The performance measure directly refers to the agent's activity level but is not affected by merger synergies. For example, if the CEO's compensation is equity based, signal σ might reflect short-term changes in the firm's stock value. These changes reflect how determined the CEO pursues the merger management, e.g., by renegotiating supply conditions or thinning out the work force, but do not yet reflect the actual merger's effect on long-term firm performance.[14] The second-period contract offered by the principal thus specifies wage payments contingent on the agent's performance: w_{2H} in case of good performance σ_H, and w_{2L} in case of bad performance σ_L. Due to the agent's limited liability, we have $w_{tL} \geq 0$ and $w_{tH} \geq 0$ for $t = 1, 2$.

Sequence of events.—The sequence of events is summarized in Figure 1. At the beginning of the first period, the principal (P) offers the agent (A) a contract $\boldsymbol{w_1} = (w_{1H}, w_{1L})$. If the agent rejects this contract, the game ends and both parties receive their zero reservation payoff for each period. If the agent accepts the contract, nature (N) determines whether he obtains

[11]This disutility might arise from the agent having to travel frequently between the firm's headquarters and the newly acquired firm, which keeps him away from his family or from having to cope with new employees who doubt his competence and question his authority.

[12]This assumption is in the spirit of decision-based incentives à la Dewatripont and Tirole (1999). We thus implicitly assume that courts are not willing to enforce message games according to Moore and Repullo (1990).

[13]Implicitly we assume that it is not contractible immediately after the acquisition of a target firm whether an actual merger of the two businesses took place. This seems plausible if one thinks of the merger as a long-term ongoing process of standardizing production and harmonizing work flows over the two businesses.

[14]It is conceivable, however, that also short-term firm success is already affected by merger synergies. We discuss such a setting in Section 5.2 below.

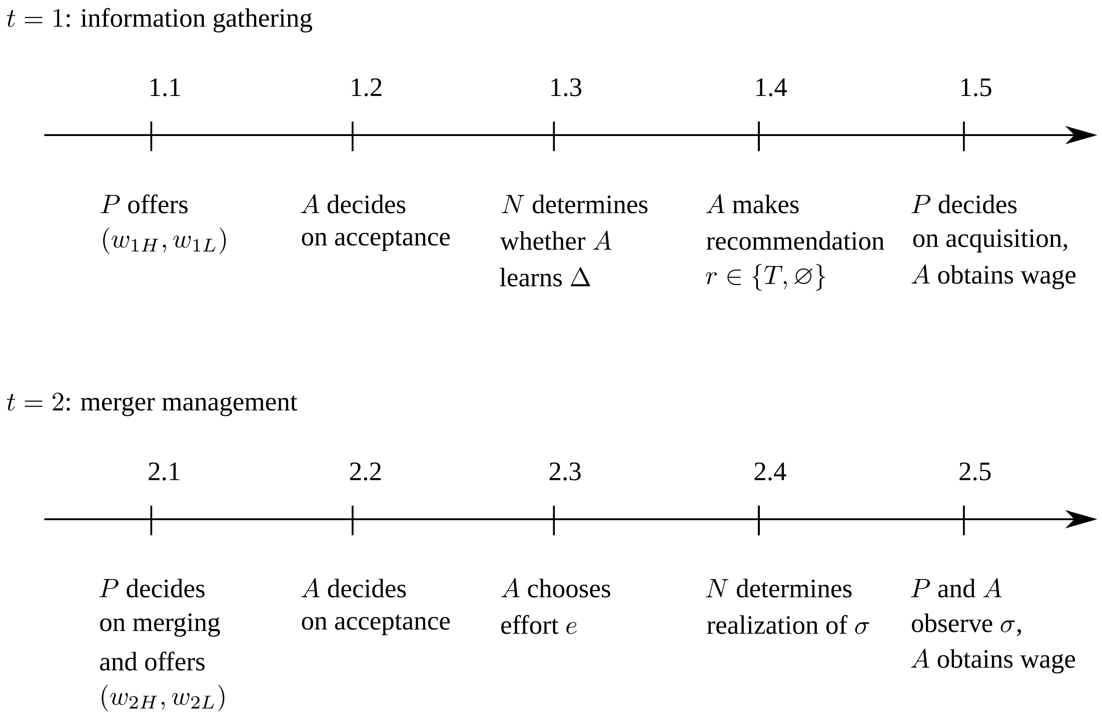

Figure 1: Timing of events

information about potential merger targets or not.[15] Subsequently, the agent hands over a report to the principal. The principal then decides whether to acquire a target firm and, if so, which one, and first-period wage payments are made according to contract $\boldsymbol{w_1}$. If the principal does not acquire a target firm or goes bankrupt after acquiring a target firm with negative synergies, the interaction of principal and agent concerning the M&A activity is terminated after the first period.[16] In this case, in the second period the agent obtains his zero reservation utility and the principal either earns zero profits from running only her core business or suffers an extreme loss from bankruptcy. If the principal has acquired a target firm with strictly positive synergies, the game continues in $t = 2$. At the beginning of the second period, the principal decides whether to conduct a merger or run her two businesses independently. In the latter case, the principal does not need the agent to manage the merged firm, the interaction of principal and agent concerning M&A activity is terminated, the agent obtains his zero second-period reservation utility, and the principal earns zero profits. In the former case, the principal has to employ the agent to manage the merged firm and offers him a contract $\boldsymbol{w_2} = (w_{2H}, w_{2L})$. If the agent rejects this contract, again the interaction of principal and agent ends, and each party obtains a zero payoff in the second period. If the agent accepts, he decides how much effort e to exert in managing the merged firm. After nature has determined the realization of the performance measure σ,

[15] A variant of the model, in which the agent can exert costly effort to improve information gathering, is discussed in Section 5.4.

[16] Note that this assumption does not rule out that agent and principal still collaborate on further tasks not considered in our paper. For example, it is conceivable that (unless in case of bankruptcy) a CEO continues to work for a corporation, although shareholders and the board have voted against merging.

second-period wage payments are made according to contract $\boldsymbol{w_2}$.

4. The Analysis

To facilitate the exposition of the following analysis, we first introduce some further notation. For a given set Δ of identified merger synergies, let

$$\Delta_+ := \{\delta \in \Delta | \delta > 0\} \tag{1}$$

refer to the subset of identified merger targets generating strictly positive synergies. Within this subset, let

$$\overline{\delta}(\Delta_+) := \max_{\delta \in \Delta_+} \delta \qquad \text{and} \qquad \underline{\delta}(\Delta_+) := \min_{\delta \in \Delta_+} \delta \tag{2}$$

denote the highest and lowest possible synergies, respectively, that can be realized given the available information.

4.1. First-Best Solution

As a benchmark solution, we can solve for the first-best second-period effort level which maximizes expected net surplus. Under the absence of contractual frictions like asymmetric information and limited liability, the principal would implement this effort level. Given that at the beginning of the second stage a merger occurred with a target firm generating synergies $\delta > 0$, first-best effort in the second stage, e^{FB}, maximizes expected second-period surplus,

$$S(e, \delta) - \kappa \tag{3}$$

with

$$S(e, \delta) := \pi_L + (\pi_H - \pi_L)\, p\,(e + \delta) - c\,(e)\,. \tag{4}$$

From the first-order condition, we obtain

$$\pi_H - \pi_L = \frac{c'\left(e^{FB}\right)}{p'\left(e^{FB} + \delta\right)} \tag{5}$$

as implicit description of first-best effort as a function of given synergies, $e^{FB}\,(\delta)$.

Suppose that merger synergies have been revealed in the first stage. Given that $\Delta_+ \neq \varnothing$, efficient merging requires $\kappa \leq S(e^{FB}(\overline{\delta}(\Delta_+)), \overline{\delta}(\Delta_+))$. If this condition is not met, or if $\Delta_+ = \varnothing$, merging is not efficient.

4.2. Merger Management

Suppose the principal acquired a firm endowed with merger potential $\delta > 0$ and hires the agent in the second period to manage the merged firm. Given the agent accepted the second-period contract $\boldsymbol{w_2}$, the agent chooses effort to maximize his expected second-period utility

$$EU_2(e) = q(e) \cdot w_{2H} + (1 - q(e)) \cdot w_{2L} - c(e) - \kappa. \tag{6}$$

The agent's effort choice is then implicitly characterized by the corresponding first-order condition,

$$w_{2H} - w_{2L} = \frac{c'(e^*)}{q'(e^*)}. \tag{7}$$

The principal chooses $\boldsymbol{w_2}$ to maximize her expected profit,

$$\Pi(\boldsymbol{w_2}) = \pi_L + (\pi_H - \pi_L)p(e^* + \delta) - w_{2L} - q(e^*)\left(w_{2H} - w_{2L}\right), \tag{8}$$

subject to the incentive constraint (7), the participation constraint $EU_2(e^*) \geq 0$, and the limited-liability constraint $w_{2H}, w_{2L} \geq 0$. The function

$$\Psi(e) := \frac{c'(e)}{q'(e)}q(e) - c(e) \tag{9}$$

combines $EU_2(e)$ with the incentive constraint (7) and describes the agent's second-period expected utility under a binding limited-liability constraint (i.e., $w_{2L} = 0$) and $\kappa = 0$. Note that Ψ is strictly increasing so that its inverse, Ψ^{-1}, exists. To guarantee strict concavity of the principal's objective function $\Pi(\boldsymbol{w_2})$, in what follows the function $\Psi(e)$ is assumed to be convex.[17] Letting $e_I^*(\delta)$ being implicitly characterized by

$$\frac{\partial S(e_I^*(\delta), \delta)}{\partial e} = \Psi'(e_I^*(\delta)), \tag{10}$$

the following proposition describes the optimal second-period contract and the associated effort level:

Proposition 1. *If*

(i) $\kappa < \Psi(e_I^*(\delta))$*, then* $e_I^*(\delta)$ *is implemented by* $w_{2H}^* = c'(e_I^*(\delta))/q'(e_I^*(\delta))$ *and* $w_{2L}^* = 0$*;*

(ii) $\kappa \in [\Psi(e_I^*(\delta)), \Psi(e^{FB}(\delta)))$*, then* $e_{II}^* := \Psi^{-1}(\kappa)$ *is implemented by* $w_{2H}^* = c'(e_{II}^*)/q'(e_{II}^*)$ *and* $w_{2L}^* = 0$*;*

[17] If only the participation constraint is binding, then $\Pi(\boldsymbol{w_2}) = \pi_L + (\pi_H - \pi_L)p(e^* + \delta) - c(e^*) - \kappa$, which is always well-behaved. However, if the limited-liability constraint is binding (i.e., $w_{2L} = 0$), then $\Pi(\boldsymbol{w_2}) = \pi_L + (\pi_H - \pi_L)p(e^* + \delta) - \Psi(e^*) - c(e^*)$, so that convexity of Ψ is sufficient to guarantee strict concavity of $\Pi(\boldsymbol{w_2})$. Note that for the family of power functions $c(e) = e^\alpha$ and $q(e) = e^\beta$ with $\alpha > 1$ and $\beta \in (0, 1]$, Ψ is always convex.

(iii) $\kappa \geq \Psi(e^{FB}(\delta))$*, then* $e^{FB}(\delta)$ *is implemented by* $w^*_{2H} = [c'(e^{FB}(\delta))/\, q'(e^{FB}(\delta))] + \kappa - \Psi(e^{FB}(\delta))$ *and* $w^*_{2L} = \kappa - \Psi(e^{FB}(\delta))$.

Moreover, $e^*_I(\delta) < e^*_{II} < e^{FB}(\delta)$.

Proposition 1 shows that the higher the agent's disutility from merging, κ, the more his limited-liability constraint is relaxed and the higher will be the effort level implemented by the principal. If the agent's disutility from merging is sufficiently small, he will earn a strictly positive rent and exert only moderate effort (case (i)). If his disutility κ exceeds the threshold $\Psi(e^*_I(\delta))$, implemented effort will monotonically increase in κ until a second threshold is reached, $\Psi(e^{FB}(\delta))$ (case (ii)). For this and higher levels of κ the principal induces the agent to choose first-best effort (case (iii)). Note that the two threshold levels $\Psi(e^*_I(\delta))$ and $\Psi(e^{FB}(\delta))$ depend on the magnitude of the merger synergies.

As an immediate corollary of Proposition 1 we obtain that, if only the limited-liability constraint is binding, a decrease in merger synergies strictly increases the agent's second-period wage for good performance.

Corollary 1. *If* $\kappa < \Psi(e^*_I(\delta))$*, then* $dw^*_{2H}/d\delta < 0$.

The intuition for this finding is rooted in the concavity of the probability function p. The smaller δ (i.e., the lower the synergies from the merger), the smaller will be the argument of the probability function, $e^* + \delta$. Low synergies thus make the agent choose his efforts at a high marginal productivity level p'. In this situation, the principal benefits much stronger from high-powered incentives than under high synergies, which are associated with lower values of p'. In other words, low synergies and, hence, exceedingly poor prospects of the merged firm induce the principal to create strong incentives to encourage the agent to save the merger project.[18] Note that this effect is not specific to the substitutability of managerial effort and merger synergies within the probability function p. In Section 5.1, we consider the case of e and δ being complements in p.

According to Proposition 1, the principal's second-period profit *under merging* is

$$\Pi(\delta,\kappa) = \begin{cases} S(e^*_I(\delta),\delta) - \Psi(e^*_I(\delta)) & \text{if } \kappa < \Psi(e^*_I(\delta)) \\ S(e^*_{II},\delta) - \kappa & \text{if } \kappa \in [\Psi(e^*_I(\delta)), \Psi(e^{FB}(\delta))) \\ S(e^{FB}(\delta),\delta) - \kappa & \text{if } \ \kappa \geq \Psi\left(e^{FB}(\delta)\right). \end{cases} \tag{11}$$

As depicted in Figure 2 and summarized in Lemma 1, the function $\Pi(\delta,\kappa)$ is nonincreasing and weakly concave in the agent's disutility from merging κ.

Lemma 1. *For* $\delta \in \{\delta_L, \delta_H\}$,

(i) $\partial\Pi(\delta,\kappa)/\partial\kappa = 0$ *for* $\kappa \leq \Psi(e^*_I(\delta))$*, and* $\partial\Pi(\delta,\kappa)/\partial\kappa < 0$ *otherwise;*

[18] Similar forces can drive rational self-sabotage in teams, see Kräkel and Müller (2012).

(ii) $\partial^2\Pi(\delta,\kappa)/\partial\kappa^2 < 0$ *for* $\kappa \in (\Psi(e_I^*(\delta)), \Psi(e^{FB}(\delta)))$*, and* $\partial^2\Pi(\delta,\kappa)/\partial\kappa^2 = 0$ *otherwise.*

Moreover, $\Pi(\delta_L,\kappa) < \Pi(\delta_H,\kappa)$ *for all* $\kappa \geq 0$*.*

Intuitively, for $\kappa < \Psi(e_I^*(\delta))$ the agent obtains a strictly positive second-period rent such that an increase in κ only reduces this rent but leaves the principal's second-period profit under merging unchanged. If κ becomes so high that the agent's second-period participation constraint is binding, the principal has to compensate the agent for any increase in κ in order to ensure his participation such that the principal's profit decreases in κ. Finally, note that in terms of second-period profits under merging the principal benefits from higher merger synergies.

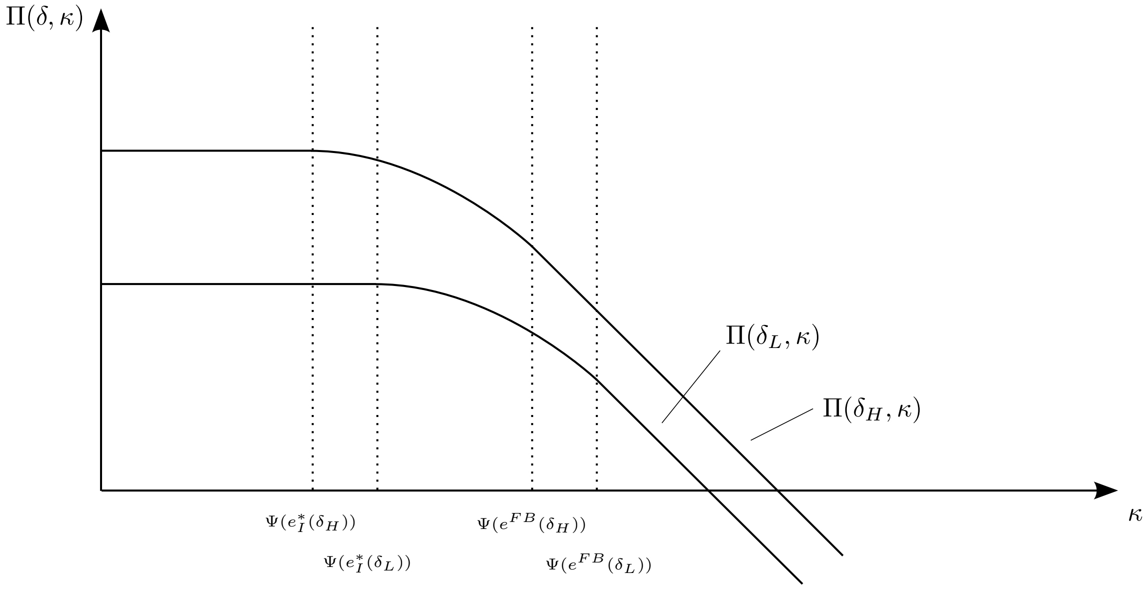

Figure 2: Principal's second-period profit

In order to focus on the conflict of interest between the principal and the agent, we assume that $\Pi(\delta_L, 0) > 0$ for the rest of the paper. Note that this assumption does not preclude post-merger losses in the form of $\pi_L < 0$.[19] After acquisition of a target firm with positive merger potential $\delta > 0$, the principal can still opt for running two independent businesses, each of which generates zero stand-alone profits. Therefore, her effective second-period profits after acquisition of a target firm with $\delta > 0$ is

$$\Pi_2(\delta,\kappa) := \max\{0, \Pi(\delta,\kappa)\}. \tag{12}$$

4.3. Merger Recommendation and Acquisition Decision

At the end of the first period, at date 1.5, for a given first-period contract $\boldsymbol{w_1} = (w_{1H}, w_{1L})$ the principal has to decide whether to make an acquisition or not. If the agent does not make

[19]To see this, note that $p(\delta_L)\pi_H + (1 - p(\delta_L))\pi_L > 0$ or, equivalently, $\pi_L > -\pi_H p(\delta_L)/(1 - p(\delta_L))$, is a sufficient condition for $\Pi(\delta_L, 0) > 0$ (where we made use of $\Psi(0) = c(0) = 0$).

a recommendation ($r = \varnothing$) or recommends a merger with negative synergies ($r = T$ with $\delta^T = -\infty$), the principal will refrain from making an acquisition in order to avoid the risk of bankruptcy. If, on the other hand, the principal faces a recommendation $r = T$ pointing to a merger target with strictly positive synergies $\delta^T > 0$, she will then acquire the merger target in question if

$$\Pi_2(\delta^T, \kappa) > w_{1H} - w_{1L}. \tag{13}$$

This implies that the principal never acquires the target firm if $\Pi(\delta^T, 0) \leq w_{1H} - w_{1L}$ because the increase in first-period wage cost in case of an acquisition exceeds the increase in second-period productivity. For $w_{1H} - w_{1L} < 0$, in contrast, the principal will always acquire the target firm because even running two independent businesses is more profitable than paying the high first-period wage w_{1L} if no acquisition takes place. For $0 \leq w_{1H} - w_{1L} < \Pi(\delta^T, 0)$ a necessary condition to acquire the target firm is that the principal prefers merging over running two independent businesses. As illustrated in Figure 3, the principal will acquire the target firm (and subsequently merge the two businesses) if and only if $\kappa < \tilde{\kappa}(\delta^T, w_{1L}, w_{1H})$, where

$$\Pi(\delta^T, \tilde{\kappa}(\delta^T, w_{1L}, w_{1H})) \equiv w_{1H} - w_{1L}. \tag{14}$$

If the agent's disutility from merging equals or exceeds this threshold, the principal will forgo acquisition of the target firm because, with the agent's second-period participation constraint being binding, from the principal's point of view the synergies δ^T do not justify compensating the agent for his disutility in case of a merger.

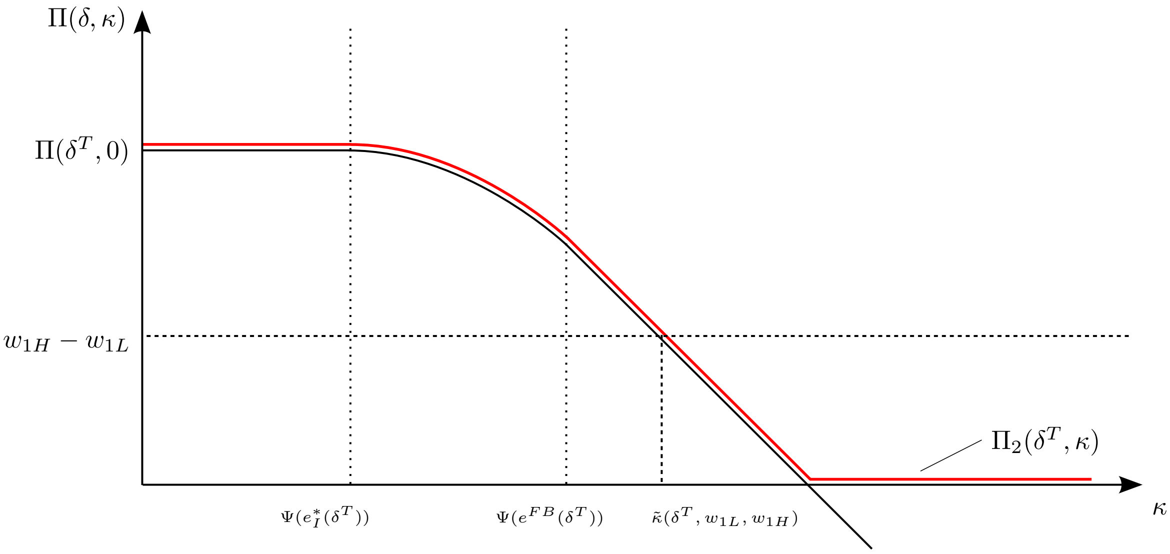

Figure 3: Acquisition decision for $0 \leq w_{1H} - w_{1L} < \Pi(\delta^T, 0)$

At date 1.4, if the agent succeeded in gathering information ($\Delta_+ \neq \varnothing$), he has to decide whether to make a recommendation and, if so, what merger target to recommend.[20] Restricting attention (with some foresight) to contracts with a nonnegative first-period wage spread, we obtain the following result regarding the agent's reporting decision:[21]

Proposition 2. *Let $w_{1L} \leq w_{1H}$ and suppose that the agent has identified merger synergies with $\Delta_+ \neq \varnothing$. Then $r = T$ with*

(i) $\delta^T = \underline{\delta}(\Delta_+)$ *if* $\kappa < \Psi(e_I^*(\delta_L))$ *and* $0 \leq w_{1H} - w_{1L} < \Pi(\delta_L, 0)$*;*

(ii) $\delta^T = \overline{\delta}(\Delta_+)$ *otherwise.*

According to part (i) of Proposition 2, if the agent's disutility from merger management is sufficiently small and the first-period acquisition premium is not too high, he will propose the least productive merger, $\delta^T = \underline{\delta}(\Delta_+)$, and the principal will be willing to follow this recommendation. In particular, this means that the agent will go against the principal's interest whenever he identifies both low-synergy and high-synergy target firms and recommend a low-synergy merger. The agent's incentive to propose the least productive merger is twofold—ensuring a positive rent and maximizing it. First and foremost, recommendation of the least productive merger avoids that production becomes too profitable (from the agent's perspective) and that the principal implements a high effort level, thereby extracting all rents. In addition, according to Corollary 1, given the principal does not extract rents to the full, recommendation of the least productive merger yields a maximum wage for the agent and thus a maximum rent.

In the remaining cases, i.e., part (ii) of Proposition 2, the agent is willing to act in the principal's best interest and recommends the most productive merger target he has identified. On the one hand, this willingness may arise because the agent is indifferent between any recommendation he could make—with the principal extracting all rents or rejecting the recommendation irrespective of the agent's recommendation. On the other hand, and more interestingly, this willingness may also be rooted in the principal's unwillingness to acquire anything but a high-synergy target: if $\Pi(\delta_L, 0) \leq w_{1H} - w_{1L} < \Pi(\delta_H, 0)$ and $\kappa < \tilde{\kappa}(\delta_H, w_{1L}, w_{1H})$, then the only way for the agent to obtain the high first-period wage $w_{1H} > w_{1L}$ (and possibly a strictly positive second-period rent in addition) is to present the principal a high-synergy target. Note that Proposition 2 implies that the principal always decides to merge when the agent recommends the least productive target that just avoids bankruptcy, but that she may reject a target when the agent recommends the most productive one.

Proposition 2 sheds new light on the case of former Daimler CEO Jürgen Schrempp mentioned in the introduction. In the light of Proposition 2, Schrempp may not have opted for the

[20] If the agent did not succeed in gathering information about synergies, he cannot back up his recommendation with evidence. In this case, irrespective of whether the agent makes a recommendation, the principal will not make an acquisition to avoid the risk of bankruptcy. If the agent learned about synergies and $\Delta_+ = \varnothing$, he makes a useless recommendation with $\delta^T = -\infty$ and the principal refrains from acquiring the target.

[21] As we show in the proof of Proposition 3, it is never optimal for the principal to offer a first-period contract with $w_{1L} > w_{1H}$.

acquisition of Chrysler to realize benefits from empire building, entrenchment, or personal diversification. Instead, he aimed at manipulating his post-merger remuneration. By suggesting a low-synergy target, he made the board choose high-powered incentives, thereby maximizing his personal rent. This conjecture is in line with the general conclusion of Anderson et al. (2004, p. 8) that the rise of CEO pay following a merger results from a restructured compensation package meant to encourage the CEO to cope with the challenges of the new complex corporation.

4.4. First-Period Contracting

At date 1.1, anticipating the agent's recommendation decision and her own acquisition decision, the principal offers the first-period contract $\boldsymbol{w_1} = (w_{1H}, w_{1L}) \in \mathbb{R}^2_{\geq 0}$ in order to maximize her expected overall profits

$$\Pi_1 = P_{\text{acquisition}}\{\mathbb{E}[\Pi_2(\delta^T, \kappa)|\text{acquisition}] - w_{1H}\} + (1 - P_{\text{acquisition}})(-w_{1L}), \tag{15}$$

where $P_{\text{acquisition}}$ denotes the probability of an acquisition occurring.[22]

With our focus on first-period contracts with a non-negative wage spread, $w_{1H} \geq w_{1L}$, it follows immediately that the agent (weakly) prefers an acquisition to occur because he then obtains a (weakly) higher wage and possibly a second-period rent. In consequence, the agent will, whenever feasible, recommend a target firm with $\delta^T > 0$ instead of making no recommendation ($r = \varnothing$) or a useless recommendation with $\delta^T = -\infty$. The agent's decision whether to recommend a target firm with $\delta^T = \delta_L$ or $\delta^T = \delta_H$, however, does not directly depend on first-period wages but is governed by the principal's acquisition decision as well as prospective second-period rents. According to (13), when faced with a recommendation $\delta^T \in \{\delta_L, \delta_H\}$, the principal's acquisition decision is determined by the interplay of the agent's disutility from merging, κ, and the difference in first-period wages, $w_{1H} - w_{1L}$. With absolute levels of first-period wages playing no role regarding the agent's recommendation decision, the principal optimally sets $w^*_{1L} = 0$. Let $P(\delta_k \in \Delta)$ denote the probability that at least one identified merger target has synergies δ_k ($k = L, H$), $P(\delta_k \in \Delta, \delta_j \in \Delta)$ denote the probability that at least one target has synergies δ_k and at least one other target δ_j, and $P(\delta_k \in \Delta, \delta_j \notin \Delta)$ the probability that at least one target has synergies δ_k, but no other target has synergies δ_j ($k, j = L, H$; $k \neq j$). We obtain the following result for the principal's optimal first-period contract offer:

Proposition 3. *The optimal first-period contract specifies*

(i) $w^*_{1L} = 0$ *and* $w^*_{1H} = \Pi(\delta_L, 0)$ *if* $\kappa < \Psi(e^*_I(\delta_L))$ *and*

$$\frac{\Pi(\delta_H, \kappa)}{\Pi(\delta_L, 0)} > 2 + \frac{P(\delta_L \in \Delta, \delta_H \notin \Delta) + P(\delta_H \in \Delta, \delta_L \notin \Delta)}{P(\delta_H \in \Delta, \delta_L \in \Delta)} \ ;$$

[22] Note that participation of the agent is not an issue because of non-negativity of wages due to limited liability.

(ii) $w_{1L}^* = w_{1H}^* = 0$ *otherwise.*

According to Proposition 3, if the agent's disutility from merging is high, $\kappa \geq \Psi(e_I^*(\delta_L))$, the principal optimally offers a zero first-period fixed wage, $w_{1L}^* = w_{1H}^* = 0$. Intuitively, since the agent never obtains a second-period rent, he is, according to Proposition 2, willing to act in the principal's best interest and recommend the most productive target firm he identified, i.e., $\delta^T = \overline{\delta}(\Delta_+)$. With no need arising to influence the agent's behavior, the principal economizes on wage cost as much as possible.

The situation is different if the agent's disutility from merging is low, $\kappa < \Psi(e_I^*(\delta_L))$. With the agent recommending the least productive target firm he identified, i.e., $\delta^T = \underline{\delta}(\Delta_+)$, the principal may actually benefit from offering the agent a sufficiently high wage premium in case of an acquisition, $w_{1H}^* = \Pi(\delta_L, 0)$, even though the content (or quality) of the agent's recommendation itself is not contractible. The reason is that the high acquisition premium acts as a commitment device for the principal not to acquire any target firm associated with positive synergies except target firms associated with high synergies. This, in turn, deters the agent from withholding a high-synergy recommendation and making a low-synergy recommendation instead because he cannot reap the higher second-period rent associated with lower synergies. In accordance with these observations, the decision whether the principal offers an acquisition premium is driven by two effects. First, offering such a premium will be profitable if the higher second-period profits from a high-synergy merger, $\Pi(\delta_H, \kappa)$, are large relative to second-period profits from a low-synergy merger, $\Pi(\delta_L, 0)$. Second, the principal will prefer offering such a premium if it is likely that the agent is tempted not to recommend the most productive acquisition, i.e., the higher $P(\delta_H \in \Delta, \delta_L \in \Delta)$.

5. Discussion

In the following, we address the robustness of our results by considering four natural extensions of our basic model. First, we consider the case of synergies and effort being complements. Second, we allow for merger synergies to affect not only the merged firm's prospect of success but also the realization of the agent's performance measure. Third, we assume that the principal can influence the agent's personal cost from merging. Finally, we allow for the agent to exert costly effort in order to improve the gathering of information.

5.1. Synergies and Efforts as Complements

In this section, let the probability of high firm profits be described by $p\,(e \cdot \delta)$. From Proposition 1 we know that the agent's second-period incentive constraint is given by $w_{2H} = c'\,(e^*)\,/q'\,(e^*)$. If we consider the agent's effort as the principal's choice variable, she implements e^*, being implicitly described by the first-order condition

$$(\pi_H - \pi_L)p'(e^*\delta)\delta - c'(e^*) - \Psi'(e^*) = 0$$

with $\Psi(e)$ being defined in (9). Let w_{2H}^* denote the corresponding wage. Thus,

$$\frac{dw_{2H}^*}{d\delta} = \frac{c''(e^*)\,q'(e^*) - c'(e^*)\,q''(e^*)}{[q'(e^*)]^2}\frac{de^*}{d\delta}$$

where

$$\frac{de^*}{d\delta} = -(\pi_H - \pi_L)\frac{p'(e^*\delta) + p''(e^*\delta)e^*\delta}{(\pi_H - \pi_L)p''(e^*\delta)\delta^2 - c''(e^*) - \Psi''(e^*)}.$$

Hence, $sign\,(dw_{2H}^*/d\delta) = sign\,(de^*/d\delta) < 0 \Leftrightarrow p'(e^*\delta) + p''(e^*\delta)e^*\delta < 0$. In words, the result of Corollary 1 still holds as long as $p'(e\delta) + p''(e\delta)\,e\delta$ is negative in the relevant range.[23]

5.2. Interaction between Synergies and Performance Measure

In the basic model we assumed that the agent's performance measure and, hence, the probability of a favorable realization of the measure, $q(\cdot)$, is purely effort based. It is also conceivable, however, that the performance measure (e.g., short-term firm success) may already have been influenced by the merger synergies. In that case, the probability of a favorable outcome of the performance measure should increase in the synergies created by the merger. This clearly creates an incentive for the agent to recommend a merger target with high synergies in the first period, thereby increasing his likelihood of good performance in the second period. Our main result, however, may also prevail under these circumstances, i.e., even with successful merger management being more likely for a more productive merger, the agent may nevertheless recommend the merger target with the lowest synergies.

For the sake of exposition, let $\kappa = 0$ such that the agent's limited-liability constraint imposes a binding restriction. Further, let $q(e + \delta)$ denote the probability of high second-period performance of the agent, which now depends on the sum of effort and merger synergies. Otherwise, the model is the same as before. Proceeding in analogy to our previous analysis of merger management (see Section 4.2), the agent chooses second-period effort according to the incentive constraint

$$w_{2H} - w_{2L} = \frac{c'(e^*)}{q'(e^* + \delta)}. \tag{16}$$

With the agent's participation not being an issue, the principal sets $w_{2L} = 0$ and $w_{2H} = c'(e^*)/q'(e^* + \delta)$. Considering the effort level as the principal's choice variable, she implements $e^*(\delta)$, which is implicitly characterized by the first-order condition

$$(\pi_H - \pi_L)p'(e^*(\delta) + \delta) - c'(e^*(\delta)) - \Psi_e(e^*(\delta), \delta) = 0 \tag{17}$$

[23] For an example see the Additional Material.

where as before $\Psi(e,\delta) := q(e+\delta)c'(e)/q'(e+\delta) - c(e)$ is assumed to be convex in effort e, $\Psi_{ee}(e,\delta) \geq 0$.

In our baseline model, with κ sufficiently low, the agent's incentive to recommend the least productive merger arose from the desire to boost his own second-period incentive pay (cf. Corollary 1). Suppressing the dependency of $q(\cdot)$ and $c(\cdot)$ on effort and/or merger synergies, we have

$$\frac{dw_{2H}^{*}}{d\delta} = \frac{c''q' - c'q''}{[q']^2} \cdot \frac{de^{*}}{d\delta} - \frac{c'q''}{[q']^2} \tag{18}$$

where

$$\frac{de^{*}}{d\delta} = \frac{\Psi_{e\delta}(e^{*},\delta) - p'' \cdot (\pi_H - \pi_L)}{p'' \cdot (\pi_H - \pi_L) - \Psi_{ee}(e^{*},\delta) - c''} \tag{19}$$

and

$$\Psi_{e\delta}(e^{*},\delta) = \frac{[c''q' - c'q''][(q')^2 - 2qq''] + qq'[c''q'' - c'q''']}{[q']^3}. \tag{20}$$

Inspection of (18) to (20) reveals that $q'' \approx 0$ and $q''' \approx 0$ (i.e., if $q(\cdot)$ is sufficiently flat in the relevant range) is a sufficient condition for $\Psi_{e\delta}(e^{*},\delta) > 0$, which, in turn, implies $de^{*}/d\delta < 0$ such that $dw_{2H}^{*}/d\delta < 0$.

Altogether, the agent's second-period expected utility (or rent) can be written as follows:

$$EU_2(e^{*}(\delta)) = q(e^{*}(\delta)+\delta) \cdot w_{2H}^{*}(\delta) - c(e^{*}(\delta)). \tag{21}$$

Applying the envelope theorem yields

$$\frac{dEU_2}{d\delta} = q'(e^{*}(\delta)+\delta) \cdot w_{2H}^{*}(\delta) + q(e^{*}(\delta)+\delta) \cdot \frac{dw_{2H}^{*}}{d\delta}. \tag{22}$$

Hence, there are two effects that work into opposite directions. The first expression in (22) is positive and measures the increase in the agent's success probability if he recommends a merger target with higher synergies. As discussed in the paragraph before, the second expression in (22) can be negative so that the agent benefits from lower synergies due to an increase in his wage payment in case of successful merger management. Note that the first effect is absent in the model of Section 3. If the second effect dominates the first effect, we will still have the result that an agent who has identified positive merger synergies prefers to recommend the least profitable one in order to increase his second-period rent.

To illustrate that second-period incentive pay decreasing in merger synergies may indeed dominate the first effect, consider the following example. Let $q(e+\delta) = \alpha \cdot (e+\delta)$ and $p(e+\delta) = \beta \cdot \ln(e+\delta)$ with $\alpha, \beta > 0$ being sufficiently small to guarantee $q, p \in (0,1)$ in the optimum. Second-period effort costs are described by $c(e) = \frac{\gamma}{2}e^2$ with $\gamma > 0$. For this specification the agent will focus on the wage-increasing effect of low merger synergies if γ is sufficiently large so that the optimal effort is very small anyway. In consequence, $dEU_2/d\delta < 0$.[24]

According to the above discussion, if merger synergies do not only affect the success of the merged firm but also the agent's performance measure, there are forces at work that dampen the agent's incentive to report a low-synergy target if he has also identified a high-synergy one. If the CEO's compensation is equity based and short-term firm success is affected by actual merger synergies, then this observation is in line with the idea stated in Bliss and Rosen (2001) that CEOs with a greater percentage of stock-based compensation make fewer wealth-reducing mergers than CEOs with a greater percentage of cash compensation.

5.3. Endogenous Costs of Merging

So far the agent's personal costs from merging, κ, were assumed to be exogenously given. In practice, however, the principal can often influence these costs. For example, the principal can decide how often the agent has to travel between headquarters and the newly acquired firm or how often and in what detail the agent has to report the progress of merger management. In this subsection, we allow for the principal to endogenously choose the agent's working conditions under merging, $\kappa \in [0, \infty)$, at date 1.1 such that the extended first-period contract takes the form $\boldsymbol{w_1} = (w_{1H}, w_{1L}, \kappa)$. We assume that if the principal is indifferent between different values of κ, she will prefer the one that is best for the agent. According to Proposition 3, with profits under merging (weakly) decreasing in κ (see Fig. 2), there are three candidates for an optimal first-period contract: (i) $\boldsymbol{w_1} = (\Pi(\delta_L, 0), 0, 0)$, (ii) $\boldsymbol{w_1} = (0, 0, \Psi(e_I^*(\delta_L)))$, and (iii) $\boldsymbol{w_1} = (0, 0, 0)$.

By stipulating a high acquisition premium, cf. case (i), the principal commits herself to merge only with high-synergy targets. As we know from Proposition 2, the agent willingly recommends the most productive merger target in this case. With profits under a high-synergy merger weakly decreasing in the agent's personal cost from merging, the principal prefers not to make the agent's life harder than necessary and sets $\kappa = 0$.

With wages $w_{1L} = w_{1H} = 0$, cf. cases (ii) and (iii), the principal is generally willing to acquire both low- and high-synergy target firms. While the principal does not prefer a positive κ for a *given* value of merger synergies (because second-period profits are decreasing in κ), she may nevertheless benefit from choosing a positive κ to influence the agent's recommendation. In particular, the principal may be interested in implementing a sufficiently large value of κ in order to reduce the agent's rent, thereby preventing him from recommending a low-synergy target in cases where he identified both low- and high-synergy target firms. According to Proposition 2, to do so the principal optimally chooses $\kappa = \Psi(e_I^*(\delta_L))$: while a smaller κ fails to induce the desired recommendation behavior, a larger κ achieves this goal but decreases profits in case of

[24]See the Additional Material or Kräkel and Müller (2012) on the specification used in the example.

a merger.[25] Alternatively, the principal may opt for not influencing the agent's recommendation behavior, in which case she minimizes his personal merger costs (i.e., $\kappa = 0$) as he has to be compensated for κ under a binding participation constraint.

Comparison of the principal's ex ante expected profits under these candidate solutions reveals the following observation regarding the optimal first-period contract, $\boldsymbol{w_1^*} = (w_{1H}^*, w_{1L}^*, \kappa^*)$.

Proposition 4. *There exist* Π^{min} *and* Π^{max} *such that:*

(i) if $\Pi(\delta_L, 0) < \Pi^{min}$*, then* $\boldsymbol{w_1^*} = (\Pi(\delta_L, 0), 0, 0)$*;*

(ii) if $\Pi(\delta_L, 0) > \Pi^{max}$*, then* $\boldsymbol{w_1^*} = (0, 0, 0)$.

According to Proposition 4, the principal will not make use of κ to influence the agent's recommendation decision for rather low or rather high values of $\Pi(\delta_L, 0)$. On the one hand, if profits from a low-synergy merger are low, the opportunity cost from adopting the self-commitment strategy are also low, which makes offering the commitment-based contract $\boldsymbol{w_1^*} = (\Pi(\delta_L, 0), 0, 0)$ optimal. On the other hand, if profits from a low-synergy merger are exceedingly high, the gains from preventing opportunistic recommendation behavior by the agent are too low to outweigh the opportunity cost associated with the contracts based on commitment or rent reduction. Consequently, the principal prefers to offer the contract $\boldsymbol{w_1^*} = (0, 0, 0)$, which is referred to as laissez-faire contract in the following. For intermediate profits of low-synergy mergers it is not as clear which contract the principal prefers to offer. A necessary condition for the principal to directly influence the agent's recommendation behavior by choosing $\kappa = \Psi(e_I^*(\delta_L))$ is that[26]

$$\frac{\Pi(\delta_H, \Psi(e_I^*(\delta_L)))}{\Pi(\delta_H, 0)} > 1 - \frac{P(\delta_H \in \Delta, \delta_L \in \Delta)}{P(\delta_H \in \Delta)} \left[1 - \frac{1}{\frac{P(\Delta_+ \neq \varnothing)}{P(\delta_L \in \Delta, \delta_H \in \Delta)} + 1} \right], \tag{23}$$

where $P(\Delta_+ \neq \varnothing) \equiv P(\delta_H \in \Delta, \delta_L \in \Delta) + P(\delta_H \in \Delta, \delta_L \notin \Delta) + P(\delta_L \in \Delta, \delta_H \notin \Delta)$ denotes the probability that the agent identifies at least one target firm with positive synergies, respectively. Thus, there seems to be scope for the principal to put her new contractual instrument to use, in particular when $P(\delta_L \in \Delta, \delta_H \in \Delta)/P(\delta_H \in \Delta)$ is large. To understand this intuitively, suppose that high-synergy targets (if identified at all) are rarely observed exclusively but mostly together with low synergy targets. Then, under a laissez-faire contract, the agent when identifying a high-synergy target, will almost always recommend a low-synergy merger instead, making this contract form rather unattractive. If, in addition, $P(\delta_L \in \Delta, \delta_H \in \Delta)/P(\Delta_+ \neq \varnothing)$ is small, then the likelihood of identifying low-synergy merger targets is relatively high because high-synergy targets are rarely identified alone, but given the agent identifies target firms with positive synergies at all, he will rarely identify both types of target firms at the same time. Since

[25] Remember that $\Pi(\delta, \Psi(e_I^*(\delta_L))) > \Pi(\delta, \kappa)$ for $\delta \in \{\delta_L, \delta_H\}$ and $\kappa > \Psi(e_I^*(\delta_L))$, see Figure 2.
[26] See the Additional Material.

we started from the hypothesis that profits from low-synergy mergers are not too low, a commitment contract also is not overly attractive because low-synergy mergers do not take place and the respective profits are not realized.

5.4. Endogenous Information Gathering

In the previous sections, we assumed that the agent has no influence on the outcome of information gathering. One might imagine, however, that the more effort the agent exerts in information gathering, the more likely he might identify a target firm that generates positive synergies. We address this issue by positing a positive relationship between the number of identified merger targets and the agent's effort exerted in information gathering. In particular, we will show that even though implicit incentives created by prospective second-period rents make first-period incentive provision comparatively easy, the principal may nevertheless prefer to disincentivize information gathering in order to reduce the scope for opportunistic recommendation behavior of the agent.

Formally, as before, the probability of the agent becoming informed at all is exogenously given by $i \in [0,1]$. However, the number of the target firms the agent identifies in case of successful information gathering now depends on the effort exerted by him in information gathering, $I \in \{0,1\}$, which is chosen at date 1.3:[27] if the agent exerts little effort, $I = 0$, he identifies only $n(0) \geq 1$ target firms, whereas if he exerts high effort, $I = 1$, he identifies $n(1) > n(0)$ target firms. His effort choice, whether information gathering was successful, and the number of identified target firms are private information of the agent.[28] Exerting effort I leads to costs $C(I) = C \cdot I$ for the agent, where $C > 0$. If the agent is indifferent between high and low effort, he chooses the effort level the principal wants him to choose. As in the previous subsections, synergies can take one of three possible values: $-\infty$, δ_L, or δ_H. At the beginning of the game, both principal and agent know that the synergies of any identified merger target are stochastically independent, where synergies $-\infty$ are realized with probability $p_0 \in (0,1)$ and synergies δ_k with probability $p_k \in (0,1)$ ($k = L, H$). Given effort $I \in \{0,1\}$, the ex-ante probabilities of no target generating positive synergies, at least one target generating synergies $\delta_k \in \{\delta_L, \delta_H\}$, and at least one target generating high synergies without another target generating low synergies are given by $P(\Delta_+ = \varnothing | I) = p_0^{n(I)}$, $P(\delta_k \in \Delta | I) = 1 - (1-p_k)^{n(I)}$, and $P(\delta_H \in \Delta, \delta_L \notin \Delta | I) = (1-p_L)^{n(I)} - p_0^{n(I)}$, respectively. Finally, to condense the analysis, we adopt the simplifying assumption of Subsection 5.2 that $\kappa = 0$.[29] All other assumptions of Section 3 are still valid. Let $\Psi_k := \Psi(e_I^*(\delta_k))$ with $k \in \{L, H\}$ denote the agent's second-period rent.

[27] The assumption of a binary effort choice is made to ease exposition.

[28] Here, we could set $n(0) = 1$ and $n(1) = n \geq 2$. This would be in accordance with footnote 9 in Section 3 that the agent could always claim to have identified only one target firm.

[29] This assumption rules out cases where the agent never obtains a positive rent in $t = 2$ such that there would be no conflict of interests between principal and agent when the latter recommends a merger target.

Following Proposition 2, we have to distinguish two types of contracts—the commitment-based contract and the laissez-faire contract. Under a laissez-faire contract with $w_{1H} - w_{1L} < \Pi(\delta_L, 0)$ the principal acquires both low-synergy and high-synergy targets. As we will demonstrate next, it may actually be optimal for the principal in this case to disincentivize information gathering. Thus, even though prospective second-period rents make incentive provision for information gathering cheap (maybe even costless), the principal may prefer to deter high effort by paying the agent a strictly positive wage w_{1L}. To see this formally, suppose the principal offers $w_{1H} - w_{1L} \geq -\Psi_H$ such that the agent prefers a productive merger to not making a recommendation and recommends the low-synergy target whenever he can. The agent's expected utility from exerting effort I is

$$\begin{aligned} EU_1(w_{1H}, w_{1L}, I) = i\{&[1 - (1 - p_L)^{n(I)}][\Psi_L + w_{1H}] \\ &+ [(1 - p_L)^{n(I)} - p_0^{n(I)}][\Psi_H + w_{1H}]\} + [(1 - i) + ip_0^{n(I)}]w_{1L} - C \cdot I. \end{aligned} \tag{24}$$

The agent is not willing to exert high effort if $EU_1(w_{1H}, w_{1L}, 1) < EU_1(w_{1H}, w_{1L}, 0)$, or equivalently,

$$w_{1L} - w_{1H} > \Psi_H + [\Psi_L - \Psi_H] \frac{(1 - p_L)^{n(0)} + (1 - p_L)^{n(1)}}{p_0^{n(0)} - p_0^{n(1)}} - \frac{C}{i[p_0^{n(0)} - p_0^{n(1)}]} =: \eta. \tag{25}$$

According to (25), the agent is more inclined to exert high effort if, ceteris paribus, the difference in first-period wages ($w_{1H} - w_{1L}$), the minimum second-period rent (Ψ_H), or the difference in second-period rents ($\Psi_L - \Psi_H$) is large. Intuitively, exerting high effort in information gathering benefits the agent for two reasons: First, identifying a larger number of merger targets reduces the probability of identifying only useless targets, thereby making the occurrence of a productive merger more likely in which case he obtains a second-period rent of at least Ψ_H and, if $w_{1H} - w_{1L} > 0$, even a larger wage payment. Second, a larger number of observations increases the probability of identifying at least one low-synergy target firm, which benefits the agent because he then obtains the large second-period rent Ψ_L instead of only Ψ_H.

Note that for $\eta > 0$, the agent prefers to exert high effort even with no direct incentives in place, i.e., for $w_{1L} = w_{1H} = 0$. The principal's objective then is to maximize her expected profits,

$$\begin{aligned} \Pi_1(w_{1H}, w_{1L}, I) = i\{&[1 - (1 - p_L)^{n(I)}][\Pi(\delta_L, 0) - w_{1H}] \\ &+ [(1 - p_L)^{n(I)} - p_0^{n(I)}][\Pi(\delta_H, 0) - w_{1H}]\} - [(1 - i) + ip_0^{n(I)}]w_{1L}, \end{aligned} \tag{26}$$

subject to the incentive constraint (25), the limited-liability constraint, and the additional constraint that $-\Psi_H \leq w_{1H} - w_{1L} < \Pi(\delta_L, 0)$. Suppose $\eta > 0$. With wage payments reducing the principal's profits, the optimal way to deter high effort is to offer $w_{1H} = 0$ and $w_{1L} = \eta$. The principal prefers to deter high effort if $\Pi_1(0,\eta,0) > \Pi_1(0,0,1)$, or equivalently,

$$\eta < \frac{i(p_0^{n(0)} - p_0^{n(1)})\Pi(\delta_H,0)}{(1-i)+ip_0^{n(0)}}\left\{\frac{(1-p_L)^{n(0)}-(1-p_L)^{n(1)}}{p_0^{n(0)}-p_0^{n(1)}}[1-\lambda]-1\right\} =: \tilde{\Omega}(\lambda), \quad (27)$$

where $\lambda := \Pi(\delta_L,0)/\Pi(\delta_H,0)$. A necessary condition for the principal to prefer disincentivizing information gathering is that $\tilde{\Omega}(\lambda)$ is strictly positive, which in turn requires that $(1-p_L)^{n(0)} - p_0^{n(0)} > (1-p_L)^{n(1)} - p_0^{n(1)}$. Condition (27) thus captures the principal's primary rationale to deter high effort: if a low-synergy merger is rather unprofitable compared to a high-synergy merger (λ small), the principal may prefer the agent to exert low effort if the probability of the agent recommending a high-synergy target decreases as the number of observations increases.

It remains to analyze whether deterring provision of high effort is optimal not only in the class of laissez-faire contracts, but also in comparison to commitment-based contracts. In the appendix we show that the profit under a commitment-based contract is bounded above by

$$\Pi_1(\Pi(\delta_L,0),0,1) = i[1-(1-p_H)^{n(1)}][\Pi(\delta_H,0)-\Pi(\delta_L,0)]. \quad (28)$$

Given $0 < \eta < \min\{\tilde{\Omega}(\lambda), \Psi_H\}$, it follows from (26) and (28) that the principal indeed prefers the laissez-faire contract with effort deterrence if $\Pi_1(0,\eta,0) > \Pi_1(\Pi(\delta_L,0),0,1)$, or equivalently, if

$$\eta < \frac{i\Pi(\delta_H,0)}{(1-i)+ip_0^{n(0)}}\{\lambda[2-(1-p_L)^{n(0)}-(1-p_H)^{n(1)}] - [1-(1-p_L)^{n(0)}-(1-p_H)^{n(1)}+p_0^{n(0)}]\} =: \hat{\Omega}(\lambda). \quad (29)$$

Noting that $\eta > 0$ can be made arbitrarily close to zero by the appropriate choice of C, we compile the above sufficient conditions for effort deterrence to be optimal in the following

Proposition 5. *If $0 < \eta < \min\{\tilde{\Omega}(\lambda), \hat{\Omega}(\lambda), \Psi_H\}$, then the optimal contract stipulates $w_{1H} = 0$ and $w_{1L} = \eta$.*

For the conditions in Proposition 5 to be possibly met, we must have that $\tilde{\Omega}(\lambda)$ and $\hat{\Omega}(\lambda)$ are both strictly positive. Whether this holds or not depends on the parameter values and thus is unclear in general. For rather extreme values of λ, deterrence of high effort will not be optimal. If $\lambda \approx 0$, i.e., profits from a low-synergy merger are very low, then $\hat{\Omega}(\lambda) < 0$ because low opportunity costs make the commitment-based contract too attractive.[30] If $\lambda \approx 1$, on the other hand, then $\tilde{\Omega}(\lambda) < 0$ because, with low-synergy and high-synergy mergers resulting in almost equal profits, a laissez-faire contract with zero wage payments is the better choice for

[30] Note that $1-(1-p_L)^{n(0)}-(1-p_H)^{n(1)}+p_0^{n(0)} > 1-[(1-p_L)^{n(0)}-p_0^{n(0)}]-(1-p_H)^{n(0)} = 1-P(\delta_H \in \Delta, \delta_L \notin \Delta|0) - P(\delta_H \notin \Delta|0) > 0$.

the principal. As the following parameter specification illustrates, deterrence of high effort may nevertheless be optimal for intermediate values of λ: for $n(0) = 1$, $n(1) = 5$, $p_L = 0.8$, and $p_0 = p_H = 0.1$ it is readily verified that $\tilde{\Omega}(\lambda) > 0$ and $\hat{\Omega}(\lambda) > 0$ as long as $\lambda \in (0.256, 0.499)$. This numerical example also points to the main intuition of why disincentivizing the agent may be rational for the principal: if the probability of detecting a low-synergy merger is rather high (here, $p_L = 0.8$), the threat of opportunistic recommendation is considerable. In this situation, the principal may prefer to mitigate this problem by reducing the number of merger targets identified by the agent.

6. Conclusion

In this paper, we offer a rationale why CEOs systematically prefer to recommend low-synergy merger targets instead of high-synergy ones when identifying both kinds of targets at the same time. Since the CEO is protected by limited liability, he may earn a positive rent under the optimal contract. By recommending a low-synergy target, the CEO increases both his chances of obtaining a positive rent and, if so, its magnitude. We identify two possible solutions for shareholders to influence the CEO's recommendation behavior. First, offering a large acquisition premium to the CEO can serve as a commitment device for shareholders to accept only sufficiently productive targets. Second, if the CEO's personal merger costs can be endogenously influenced via the CEO contract, shareholders can benefit from sufficiently large costs so that the CEO no longer receives a positive rent. As a consequence, the CEO is not interested in recommending poor merger targets to manipulate his post-merger remuneration and, hence, his expected rent.

In our setting, low post-merger profits were allowed to become negative. Together with the finding that CEOs prefer mergers which ex ante are less likely to succeed, this fits well to empirical cases (e.g., DaimlerChrysler) where merging is indeed value reducing. If we reinterpret the synergy parameter δ as the CEO's target-specific ability of running the merged corporation, a CEO will prefer a merger target for which he is poorly suited at the merger-management stage, i.e., merging with a business in which he is not an expert, to maximize his post-merger remuneration. This prediction differs from the traditional entrenchment hypothesis mentioned in Section 2, according to which CEOs have an incentive to expand in those industries in which they are experts in order to protect their jobs.

A. Proofs

Proof of Proposition 1. We can procede similarly to the proof of Proposition 1 in Kräkel and Schöttner (2010). Since the incentive constraint $w_{2H} = [c'(e^*)/q'(e^*)] + w_{2L}$ together with

$w_{2L} \geq 0$ already implies that $w_{2H} \geq 0$, the Lagrangian can be written as

$$L(w_{2L}, w_{2H}) = \pi_L + (\pi_H - \pi_L)p(e^* + \delta) - w_{2L} - q(e^*)(w_{2H} - w_{2L}) \\ + \lambda_1 [w_{2L} + (w_{2H} - w_{2L}) q(e^*) - \kappa - c(e^*)] + \lambda_2 w_{2L}, \quad \text{(A.1)}$$

with e^* being a monotonically increasing function of $w_{2H} - w_{2L}$, implicitly defined by (7). Computing the partial derivatives with respect to w_{2L} and w_{2H} yields

$$\frac{\partial L}{\partial w_{2L}} = (\pi_H - \pi_L)p'(e^* + \delta)\frac{\partial e^*}{\partial w_{2L}} - 1 - q'(e^*)\frac{\partial e^*}{\partial w_{2L}}(w_{2H} - w_{2L}) + q(e^*) \\ + \lambda_1 + \lambda_1 (w_{2H} - w_{2L}) q'(e^*)\frac{\partial e^*}{\partial w_{2L}} - \lambda_1 q(e^*) - \lambda_1 c'(e^*)\frac{\partial e^*}{\partial w_{2L}} + \lambda_2 = 0 \quad \text{(A.2)}$$

and

$$\frac{\partial L}{\partial w_{2H}} = (\pi_H - \pi_L)p'(e^* + \delta)\frac{\partial e^*}{\partial w_{2H}} - q'(e^*)\frac{\partial e^*}{\partial w_{2H}}(w_{2H} - w_{2L}) - q(e^*) \\ + \lambda_1 q(e^*) + \lambda_1 (w_{2H} - w_{2L}) q'(e^*)\frac{\partial e^*}{\partial w_{2H}} - \lambda_1 c'(e^*)\frac{\partial e^*}{\partial w_{2H}} = 0. \quad \text{(A.3)}$$

As $\partial e^*/\partial w_{2L} = -\partial e^*/\partial w_{2H}$, we have that $\lambda_1 + \lambda_2 = 1$, implying that either (i) only the limited-liability constraint is binding, or (ii) both the limited-liability and the participation constraints are binding, or (iii) only the participation constraint is binding.

In case (i), $\lambda_2 = 1$, $\lambda_1 = 0$, and $w_{2L} = 0$. Inserting in (A.3) and using incentive constraint (7) yields

$$(\pi_H - \pi_L)p'(e^* + \delta) = c'(e^*) + \frac{q(e^*)}{\partial e^*/\partial w_{2H}}. \quad \text{(A.4)}$$

The comparison with (5) shows that $e^* < e^{FB}$, since the second-period surplus function (3) is strictly concave. Note that, in this situation, the agent earns a strictly positive rent: $EU_2(e^*) > 0 \Leftrightarrow \Psi(e^*) > \kappa$ with $\Psi(e^*)$ being defined in (9). By using

$$\Psi'(e^*) = \frac{c''(e^*)q'(e^*) - c'(e^*)q''(e^*)}{q'(e^*)^2} q(e^*) \quad \text{(A.5)}$$

and the fact that $\partial e^*/\partial w_{2H} = q(e^*)/\Psi'(e^*)$ we can rewrite (A.4) as

$$(\pi_H - \pi_L)p'(e^* + \delta) - c'(e^*) - \Psi'(e^*) = 0. \quad \text{(A.6)}$$

In case (ii), we have $\lambda_1, \lambda_2 > 0$ as well as $w_{2L} = 0$ and $EU_2(e^*) = 0 \Leftrightarrow \Psi(e^*) = \kappa$. Using again $\partial e^*/\partial w_{2H} = q(e^*)/\Psi'(e^*)$, we can rewrite (A.3) as

$$\lambda_1 = 1 - \frac{(\pi_H - \pi_L)p'(e^* + \delta) - c'(e^*)}{\Psi'(e^*)}. \quad \text{(A.7)}$$

Since (due to $\lambda_1 + \lambda_2 = 1$ and $\lambda_1, \lambda_2 > 0$) the multiplier λ_1 is smaller than one, we must have that $(\pi_H - \pi_L)p'(e^* + \delta) - c'(e^*) > 0$. Strict concavity of the second-period surplus function (3) implies that $e^* < e^{FB}$. Combining $\lambda_1 > 0$ with (A.7) yields

$$(\pi_H - \pi_L)p'(e^* + \delta) - c'(e^*) - \Psi'\left(e^*\right) < 0. \tag{A.8}$$

Note that $(\pi_H - \pi_L)p(e^* + \delta) - c(e^*) - \Psi\left(e^*\right)$ describes a strictly concave function of e^* since Ψ is convex. Hence, the optimal effort in (A.8) is strictly larger than the optimal effort implicitly described by (A.6).

In case (iii), $\lambda_1 = 1$ and $\lambda_2 = 0$. Inserting in (A.3) immediately leads to equation (5). Hence, $e^* = e^{FB}$. From the binding participation constraint $EU_2\left(e^*\right) = 0$ and the non-binding limited-liability constraint $w_{2L} > 0$ we obtain $\Psi\left(e^*\right) < \kappa$.

The optimal wages directly follow from the respective incentive, participation and limited-liability constraints. □

Proof of Corollary 1. For $\kappa < \Psi(e_I^*(\delta))$, we have $w_{2H}^* = c'(e_I^*(\delta))/q'(e_I^*(\delta))$, such that

$$\frac{dw_{2H}^*}{d\delta} = \frac{c''(e_I^*(\delta))q'(e_I^*(\delta)) - c'(e_I^*(\delta))q''(e_I^*(\delta))}{[q'(e_I^*(\delta))]^2} \cdot \frac{de_I^*(\delta)}{d\delta}. \tag{A.9}$$

Differentiation of (10) with respect to δ reveals that

$$\frac{de_I^*(\delta)}{d\delta} = -\frac{(\pi_H - \pi_L)p''(e_I^*(\delta) + \delta)}{(\pi_H - \pi_L)p''(e_I^*(\delta) + \delta) - c''(e_I^*(\delta)) - \Psi''(e_I^*(\delta))} < 0, \tag{A.10}$$

which establishes $dw_{2H}^*/d\delta < 0$. □

Proof of Lemma 1. We first prove parts (i) and (ii). Suppose the principal has merged with a firm associated with synergies $\delta > 0$. For $\kappa \leq \Psi(e_I^*(\delta))$, the principal implements effort level $e_I^*(\delta)$, as defined in (10), which is independent of κ. Hence, $\Pi(\delta, \kappa)$ is a constant function of κ. For $\kappa \geq \Psi(e^{FB}(\delta))$, the principal implements $e^{FB}(\delta)$, as defined in (5), which is independent of κ, and $\Pi(\delta, \kappa)$ is linearly decreasing in κ. It remains to show that $\Pi(\delta, \kappa)$ is strictly decreasing and strictly concave in κ for $\kappa \in (\Psi(e_I^*(\delta)), \Psi(e^{FB}(\delta)))$. The principal implements effort level e_{II}^* characterized by $\Psi(e_{II}^*) = \kappa$. With $de_{II}^*/d\kappa = 1/\Psi'(e_{II}^*) > 0$,

$$\begin{aligned} \frac{\partial \Pi(\delta, \kappa)}{\partial \kappa} &= [(\pi_H - \pi_L)p'(e_{II}^* + \delta) - c'(e_{II}^*)]\frac{de_{II}^*}{d\kappa} - 1 \\ &= [(\pi_H - \pi_L)p'(e_{II}^* + \delta) - c'(e_{II}^*) - \Psi'(e_{II}^*)]\,\frac{de_{II}^*}{d\kappa} \overset{\text{(A.8)}}{<} 0. \end{aligned} \tag{A.11}$$

From the proof of Proposition 1, we know that $(\pi_H - \pi_L)p'(e_{II}^* + \delta) - c'(e_{II}^*) - \Psi'(e_{II}^*)$ is zero for $e_{II}^* = e_I^*(\delta)$ and negative for $e_{II}^* \in (e_I^*(\delta), e^{FB}(\delta)]$, which establishes that $\Pi(\delta, \kappa)$ is strictly decreasing in κ. Strict concavity of $\Pi(\delta,\kappa)$ follows from

$$\frac{\partial^2\Pi(\delta,\kappa)}{\partial\kappa^2} = [(\pi_H-\pi_L)p'(e^*_{II}+\delta)-c'(e^*_{II})]\,\frac{d^2e^*_{II}}{d\kappa^2} + [(\pi_H-\pi_L)p''(e^*_{II}+\delta)-c''(e^*_{II})]\left(\frac{de^*_{II}}{d\kappa}\right)^2 \quad \text{(A.12)}$$

together with $d^2e^*_{II}/d\kappa^2 = -\Psi''(e^*_{II})/[\Psi'(e^*_{II})]^3 < 0$ (because $\Psi''(e^*_{II}) \geq 0$ by assumption) and $[(\pi_H-\pi_L)p'(e^*_{II}+\delta)-c'(e^*_{II})] > 0$ (because $e^*_{II} < e^{FB}(\delta)$, see Proposition 1).

It remains to establish that $\Pi(\delta_L,\kappa) < \Pi(\delta_H,\kappa)$. For $\kappa < \Psi(e^*_I(\delta))$,

$$\frac{\partial\Pi(\delta,\kappa)}{\partial\delta} = \left[\frac{\partial S(e^*_I(\delta),\delta)}{\partial e} - \Psi'(e^*_I(\delta))\right]\frac{de^*_I(\delta)}{d\delta} + \frac{\partial S(e^*_I(\delta),\delta)}{\partial\delta} = \frac{\partial S(e^*_I(\delta),\delta)}{\partial\delta} = (\pi_H-\pi_L)p'(e^*_I(\delta)+\delta) > 0, \quad \text{(A.13)}$$

where the second equality follows from the definition of $e^*_I(\delta)$ in (10). Likewise, for $\kappa \geq \Psi(e^*_I(\delta))$ and $\tilde{e} \in \{e^*_{II}, e^{FB}(\delta)\}$, we have

$$\frac{\partial\Pi(\delta,\kappa)}{\partial\delta} = \frac{\partial S(\tilde{e},\delta)}{\partial e}\cdot\frac{d\tilde{e}}{d\delta} + \frac{\partial S(\tilde{e},\delta)}{\partial\delta} = \frac{\partial S(\tilde{e},\delta)}{\partial\delta} = (\pi_H-\pi_L)p'(e+\delta) > 0 \quad \text{(A.14)}$$

where $\partial e/\partial\delta = 0$ for $\tilde{e} = e^*_{II}$, and $\partial S(\tilde{e},\delta)/\partial e = 0$ for $e = e^{FB}(\delta)$. By parts (i) and (ii) established above, the functions $\Pi(\delta_L,\kappa)$ and $\Pi(\delta_H,\kappa)$ have the same qualitative shape. Therefore we must have that $\Pi(\delta_L,\kappa) < \Pi(\delta_H,\kappa)$ for all κ, even though both thresholds $\Psi(e^*_I(\delta))$ and $\Psi(e^{FB}(\delta))$ are shifted to the left if synergies δ increase, i.e.

$$\frac{\partial\Psi(e^*_I(\delta))}{\partial\delta} = \Psi'(e^*_I(\delta))\cdot\frac{\partial e^*_I(\delta)}{\partial\delta} \overset{\text{(A.10)}}{<} 0 \quad \text{(A.15)}$$

and

$$\frac{\partial\Psi(e^{FB}(\delta))}{\partial\delta} = \Psi'(e^{FB}(\delta))\frac{(\pi_H-\pi_L)p''(e^{FB}+\delta)}{-(\pi_H-\pi_L)p''(e^{FB}+\delta)+c''(e^{FB})} < 0. \quad \text{(A.16)}$$

□

Proof of Proposition 2. From Proposition 1, we know that if the agent recommends $r = T$ with $\delta^T > 0$ and $\kappa < \Psi(e^*_I(\delta^T))$, then under the optimal second-period contract the principal implements effort $e^*_I(\delta^T)$, the agent's participation constraint is slack and he obtains a strictly positive rent, i.e., $EU_2(e^*_I(\delta^T)) = \Psi(e^*_I(\delta^T)) - \kappa$. From the proof of Lemma 1, we know that $\Psi(e^*_I(\delta_H)) < \Psi(e^*_I(\delta_L))$.

Note that for $w_{1H} \geq w_{1L}$ the agent always (weakly) prefers an acquisition to occur because he obtains a (weakly) higher wage and possibly a second-period rent. Therefore, whenever feasible, the agent prefers recommending $r = T$ with $\delta^T \in \{\delta_L,\delta_H\}$ over not making a recommendation ($r = \varnothing$) or making a useless recommendation $r = T$ with $\delta^T = -\infty$.

Anticipating the principal's acquisition decision, the agent chooses whether to make a recommendation and (if so) what recommendation to make in order to maximize his expected utility. Given $\Delta_+ \neq \varnothing$ and $0 \leq w_{1L} \leq w_{1H}$, we have to distinguish three cases:

Case 1: $\Pi(\delta_H, 0) \leq w_{1H} - w_{1L}$

Even if the agent makes a merger recommendation $r \neq \varnothing$, the principal never acquires the target firm and the agent always obtains w_{1L}. In consequence, the agent is indifferent between any recommendation he can make and therefore recommends $r = T$ with $\delta^T = \overline{\delta}(\Delta_+)$.

Case 2: $\Pi(\delta_L, 0) \leq w_{1H} - w_{1L} < \Pi(\delta_H, 0)$

If $\kappa \geq \tilde{\kappa}(\delta_H, w_{1L}, w_{1H})$, then the principal never acquires the target firm and the agent always obtains w_{1L}. Consequently, the agent recommends $r = T$ with $\delta^T = \overline{\delta}(\Delta_+)$.

If $\kappa < \tilde{\kappa}(\delta_H, w_{1L}, w_{1H})$, then the principal acquires the target firm if $\delta^T = \delta_H$ and does not acquire the target firm otherwise. Therefore, if $\delta_H \in \Delta_+$, the agent recommends $r = T$ with $\delta^T = \delta_H = \overline{\delta}(\Delta_+)$, thereby obtaining w_{1H} (and possibly a second-period rent) whereas any other recommendation would only yield $w_{1L} \leq w_{1H}$. If $\delta_H \notin \Delta_+$, then no recommendation the agent can make leads to acquisition of the target firm and he always obtains w_{1L}. Therefore, the agent recommends $r = T$ with $\delta^T = \delta_L = \overline{\delta}(\Delta_+)$.

Case 3: $0 \leq w_{1H} - w_{1L} < \Pi(\delta_L, 0)$

For $\delta^T \in \{\delta_L, \delta_H\}$, the principal acquires the target firm for $\kappa < \tilde{\kappa}(\delta^T, w_{1L}, w_{1H})$, where

$$\Psi(e_I^*(\delta_H)) < \Psi(e_I^*(\delta_L)) < \tilde{\kappa}(\delta_L, w_{1L}, w_{1H}) < \tilde{\kappa}(\delta_H, w_{1L}, w_{1H}). \tag{A.17}$$

If $\kappa < \Psi(e_I^*(\delta_H))$, then for $\delta^T \in \{\delta_L, \delta_H\}$ the principal acquires the target firm and the agent obtains a strictly positive second-period rent equal to $\Psi(e_I^*(\delta^T)) - \kappa$. In both cases the agent obtains $w_{1H} \geq w_{1L}$. Since $\Psi(e_I^*(\delta_H)) < \Psi(e_I^*(\delta_L))$—cf. the proof of Lemma 1—the agent strictly prefers to recommend $r = T$ with $\delta^T = \delta_L = \underline{\delta}(\Delta_+)$ whenever $\delta_L \in \Delta_+$. The agent recommends $\delta^T = \delta_H = \underline{\delta}(\Delta_+)$ whenever $\delta_L \notin \Delta_+$.

If $\Psi(e_I^*(\delta_H)) \leq \kappa < \Psi(e_I^*(\delta_L))$, then the principal acquires the target firm for $\delta^T \in \{\delta_L, \delta_H\}$. The agent obtains a strictly positive second-period rent for $\delta^T = \delta_L$ whereas for $\delta^T = \delta_H$ the agent obtains no second-period rent. Since in both cases the agent earns $w_{1H} \geq w_{1L}$, he strictly prefers to recommend $r = T$ with $\delta^T = \delta_L = \underline{\delta}(\Delta_+)$ whenever $\delta_L \in \Delta_+$. If $\delta_L \notin \Delta_+$, then the agent recommends $\delta^T = \delta_H = \underline{\delta}(\Delta_+)$.

If $\Psi(e_I^*(\delta_L)) \leq \kappa < \tilde{\kappa}(\delta_L, w_{1L}, w_{1H})$, then for $\delta^T \in \{\delta_L, \delta_H\}$ the principal acquires the target firm and the agent obtains $w_{1H} \geq w_{1L}$. The agent does not obtain a second-period rent in either case. Since the agent is indifferent between $\delta^T = \delta_L$ and $\delta^T = \delta_H$, he recommends $\delta^T = \overline{\delta}(\Delta_+)$.

If $\tilde{\kappa}(\delta_L, w_{1L}, w_{1H}) \leq \kappa < \tilde{\kappa}(\delta_H, w_{1L}, w_{1H})$, then the principal acquires the target firm for $\delta^T = \delta_H$ and the agent obtains $w_{1H} \geq w_{1L}$, whereas the principal does not acquire the target firm for $\delta^T = \delta_L$ and the agent obtains w_{1L}. The agent does not obtain a strictly positive second-period rent in either case. Therefore, the agent recommends $r = T$ with $\delta^T = \delta_H$ whenever

$\delta_H \in \Delta_+$. If $\delta_H \notin \Delta_+$, then the agent obtains w_{1L} irrespective of his recommendation and he recommends $r = T$ with $\delta^T = \delta_L = \overline{\delta}(\Delta_+)$.

If $\tilde{\kappa}(\delta_H, w_{1L}, w_{1H}) \leq \kappa$, then the principal never acquires the target firm and the agent always obtains w_{1L}. In consequence, the agent is indifferent between any recommendation he can make and therefore recommends $r = T$ with $\delta^T = \overline{\delta}(\Delta_+)$. □

Proof of Proposition 3. First, we consider first-period contracts with a non-negative wage spread $w^*_{1H} - w^*_{1L} \geq 0$. As was argued in the text, the principal optimally sets $w^*_{1L} = 0$. If the principal offers $w_{1H} \geq \Pi(\delta_H, 0)$, then an acquisition never occurs and $\Pi_1 = 0$.

If the principal sets $w_{1H} \in [\Pi(\delta_L, 0), \Pi(\delta_H, 0))$, then for $\kappa \geq \tilde{\kappa}(\delta_H, 0, w_{1H})$ an acquisition never occurs and $\Pi_1 = 0$. For $\kappa < \tilde{\kappa}(\delta_H, 0, w_{1H})$ an acquisition occurs if the agent recommends $r = T$ with $\delta^T = \delta_H$. According to Proposition 2, if $\Delta_+ \neq \varnothing$, then the agent recommends $r = T$ with $\delta^T = \overline{\delta}(\Delta_+)$, such that $\Pi_1 = i \cdot P(\delta_H \in \Delta) \cdot [\Pi(\delta_H, \kappa) - w_{1H}]$. The optimal wage w_{1H} to choose for the principal from the range $[\Pi(\delta_L, 0), \Pi(\delta_H, 0))$ is $w_{1H} = \Pi(\delta_L, 0)$: this not only minimizes the wage cost in case of an acquisition, but also makes it most likely that the principal realizes strictly positive profits from M&A because $\tilde{\kappa}(\delta_H, 0, \Pi(\delta_L, 0)) > \tilde{\kappa}(\delta_H, 0, w_{1H})$ for all $w_{1H} \in (\Pi(\delta_L, 0), \Pi(\delta_H, 0))$. In summary, for $w_{1L} = 0$ and $w_{1H} = \Pi(\delta_L, 0)$,

$$\Pi_1 = \begin{cases} i \cdot P(\delta_H \in \Delta) \cdot [\Pi(\delta_H, \kappa) - \Pi(\delta_L, 0)] & \text{if } \kappa < \tilde{\kappa}(\delta_H, 0, \Pi(\delta_L, 0)) \\ 0 & \text{if } \kappa \geq \tilde{\kappa}(\delta_H, 0, \Pi(\delta_L, 0)). \end{cases} \tag{A.18}$$

If the principal sets $w_{1H} \in [0, \Pi(\delta_L, 0))$, then for $\kappa \geq \tilde{\kappa}(\delta_H, 0, w_{1H})$ an acquisition never occurs and $\Pi_1 = 0$. For $\kappa \in [\tilde{\kappa}(\delta_L, 0, w_{1H}), \tilde{\kappa}(\delta_H, 0, w_{1H})$ an acquisition occurs if the agent recommends $r = T$ with $\delta^T = \delta_H$. According to Proposition 2, if $\Delta_+ \neq \varnothing$, then the agent recommends $r = T$ with $\delta^T = \overline{\delta}(\Delta_+)$, such that $\Pi_1 = i \cdot P(\delta_H \in \Delta) \cdot [\Pi(\delta_H, \kappa) - w_{1H}]$. For $\kappa \in [\Psi(e^*_I(\delta_L)), \tilde{\kappa}(\delta_L, 0, w_{1H}))$, an acquisition occurs if the agent recommends $r = T$ with $\delta^T \in \{\delta_L, \delta_H\}$. According to Proposition 2, if $\Delta_+ \neq \varnothing$, then the agent recommends $r = T$ with $\delta^T = \overline{\delta}(\Delta_+)$, such that $\Pi_1 = i \cdot \{P(\delta_H \in \Delta) \cdot \Pi(\delta_H, \kappa) + P(\delta_L \in \Delta, \delta_H \notin \Delta) \cdot \Pi(\delta_L, \kappa) - P(\Delta_+ \neq \varnothing) w_{1H}\}$. For $\kappa < \Psi(e^*_I(\delta_L))$ an acquisition occurs if the agent recommends $r = T$ with $\delta^T \in \{\delta_L, \delta_H\}$. According to Proposition 2, if $\Delta_+ \neq \varnothing$, then the agent recommends $r = T$ with $\delta^T = \underline{\delta}(\Delta_+)$, such that $\Pi_1 = i \cdot \{P(\delta_H \in \Delta, \delta_L \notin \Delta) \cdot \Pi(\delta_H, \kappa) + P(\delta_L \in \Delta) \cdot \Pi(\delta_L, \kappa) - P(\Delta_+ \neq \varnothing) w_{1H}\}$. The optimal wage w_{1H} to choose for the principal from the range $[0, \Pi(\delta_L, 0))$ is $w_{1H} = 0$: this not only minimizes the wage cost in case of an acquisition, but for $\kappa \geq \Psi(e^*_I(\delta_L))$, where the agent reports $\delta^T = \overline{\delta}(\Delta_+)$ whenever $\Delta_+ \neq \varnothing$, also makes it most likely that the principal realizes strictly positive profits from M&A because $\tilde{\kappa}(\delta^T, 0, 0) > \tilde{\kappa}(\delta^T, 0, w_{1H})$ for all $w_{1H} \in (0, \Pi(\delta_L, 0)$ and $\delta^T \in \{\delta_L, \delta_H\}$. In summary, for $w_{1L} = w_{1H} = 0$,

$$\Pi_1 = \begin{cases} i \cdot \{P(\delta_H \in \Delta, \delta_L \notin \Delta) \cdot \Pi(\delta_H, \kappa) \\ \quad +P(\delta_L \in \Delta) \cdot \Pi(\delta_L, 0)\} & \text{if } \kappa < \Psi(e_I^*(\delta_L)) \\ i \cdot \{P(\delta_H \in \Delta) \cdot \Pi(\delta_H, \kappa) \\ \quad +P(\delta_L \in \Delta, \delta_H \notin \Delta) \cdot \Pi(\delta_L, \kappa)\} & \text{if } \kappa \in [\Psi(e_I^*(\delta_L)), \tilde{\kappa}(\delta_L, 0, 0)) \\ i \cdot P(\delta_H \in \Delta) \cdot \Pi(\delta_H, \kappa) & \text{if } \kappa \in [\tilde{\kappa}(\delta_L, 0, 0), \tilde{\kappa}(\delta_H, 0, 0)) \\ 0 & \text{if } \kappa \geq \tilde{\kappa}(\delta_H, 0, 0). \end{cases} \tag{A.19}$$

Comparison of (A.18) and (A.19) reveals that the principal optimally offers $w_{1L}^* = w_{1H}^* = 0$ for $\kappa \geq \Psi(e_I^*(\delta_L))$ (where for $\kappa \geq \tilde{\kappa}(\delta_H, 0, 0)$ this statement is without loss of generality because the principal is indifferent). For $\kappa < \Psi(e_I^*(\delta_L))$, on the other hand, she optimally offers $w_{1L}^* = w_{1H}^* = 0$ if

$$i \cdot \{P(\delta_H \in \Delta, \delta_L \notin \Delta) \cdot \Pi(\delta_H, \kappa) + P(\delta_L \in \Delta) \cdot \Pi(\delta_L, \kappa)\} \\ \geq i \cdot P(\delta_H \in \Delta) \cdot [\Pi(\delta_H, \kappa) - \Pi(\delta_L, 0)], \tag{A.20}$$

or equivalently (making use of the fact that $\Pi(\delta_L, \kappa) = \Pi(\delta_L, 0)$ for $\kappa < \Psi(e_I^*(\delta_L))$)

$$\frac{\Pi(\delta_H, \kappa)}{\Pi(\delta_L, 0)} \leq \frac{P(\delta_L \in \Delta) + P(\delta_H \in \Delta)}{P(\delta_H \in \Delta) - P(\delta_H \in \Delta, \delta_L \notin \Delta)}, \tag{A.21}$$

and $w_{1L}^* = 0$ and $w_{1H}^* = \Pi(\delta_L, 0)$ otherwise. With regard to (A.21), note that $P(\delta_k \in \Delta) = P(\delta_k \in \Delta, \delta_j \in \Delta) + P(\delta_k \in \Delta, \delta_j \notin \Delta)$ ($k, j = L, H$; $k \neq j$).

To finally establish the desired result, it remains to show that it is not optimal for the principal to offer $w_{1L} > w_{1H} \geq 0$. In this case, if $\Delta_+ \neq \varnothing$ and the agent recommends a target firm with $\delta^T > 0$, then the principal will always acquire the target because $\Pi_2(\delta^T, \kappa) \geq 0$. The agent, however, will recommend $r = T$ with $\delta^T > 0$ only if he obtains a second-period rent and this rent outweighs obtaining the high first-period wage w_{1L}. Formally, the agent prefers recommending $r = T$ with $\delta^T \in \{\delta_L, \delta_H\}$ over not making a recommendation ($r = \varnothing$) or making a useless recommendation (with $\delta^T = -\infty$) if

$$\kappa \leq \Psi(e_I^*(\delta^T)) - (w_{1L} - w_{1H}). \tag{A.22}$$

Now, suppose $\Delta_+ \neq \varnothing$. With $\Psi(e_I^*(\delta_L)) > \Psi(e_I^*(\delta_H))$, if $\kappa \leq \Psi(e_I^*(\delta_H)) - (w_{1L} - w_{1H})$, then the agent recommends $r = T$ with $\delta^T = \delta_L$ whenever $\delta_L \in \Delta_+$, and $r = T$ with $\delta^T = \delta_H$ otherwise. If $\kappa \in (\Psi(e_I^*(\delta_H)) - (w_{1L} - w_{1H}), \Psi(e_I^*(\delta_L)) - (w_{1L} - w_{1H})]$, then the agent recommends $r = T$ with $\delta^T = \delta_L$ whenever $\delta_L \in \Delta_+$, and makes no recommendation or a useless recommendation otherwise. Thus, the principal's expected overall profit for a first-period contract with $w_{1L} > w_{1H} \geq 0$ is

$$\Pi_1 = \begin{cases} [(1-i) + i \cdot P(\Delta_+ = \varnothing)](-w_{1L}) & \\ + i \cdot \{P(\delta_H \in \Delta, \delta_L \notin \Delta) \cdot \Pi(\delta_H, \kappa) & \\ \quad +P(\delta_L \in \Delta) \cdot \Pi(\delta_L, 0) & \\ \quad -P(\Delta \neq \varnothing) w_{1H}\} & \text{if } \kappa \leq \Psi(e_I^*(\delta_H)) - (w_{1L} - w_{1H}) \\ [(1-i) + i \cdot P(\delta_L \notin \Delta)]\,(-w_{1L}) & \\ + i \cdot P(\delta_L \in \Delta) \cdot [\Pi(\delta_L, 0) - w_{1H}] & \text{if } \kappa \in (\Psi(e_I^*(\delta_H)) - (w_{1L} - w_{1H}), \\ & \quad \Psi(e_I^*(\delta_L)) - (w_{1L} - w_{1H})] \\ -w_{1L} & \text{if } \kappa > \Psi(e_I^*(\delta_L)) - (w_{1L} - w_{1H}), \end{cases} \tag{A.23}$$

where we made use of the fact that $\Pi(\delta_L, \kappa) = \Pi(\delta_L, 0)$ for $\kappa < \Psi(e_I^*(\delta_L))$. Comparison of (A.19) and (A.23) reveals that for all $\kappa \geq 0$ the principal is better off offering $w_{1L} = w_{1H} = 0$ instead of $w_{1L} > w_{1H} \geq 0$ □

Proof of Proposition 4. From Proposition 3, we know that the principal prefers the commitment-based contract $\boldsymbol{w_1} = (\Pi(\delta_L, 0), 0, 0)$ over the laissez-faire contract $\boldsymbol{w_1} = (0, 0, 0)$ if and only if

$$\Pi(\delta_L, 0) < \frac{P(\delta_H \in \Delta, \delta_L \in \Delta)}{P(\delta_L \in \Delta) + P(\delta_H \in \Delta)} \Pi(\delta_H, 0) =: \hat{\Pi}_L.$$

Furthermore, the principal prefers the contract $\boldsymbol{w_1} = (0, 0, \Psi(e_I^*(\delta_L)))$ based on rent reduction over the laissez-faire contract $\boldsymbol{w_1} = (0, 0, 0)$ if and only if

$$\begin{aligned} i \cdot [P(\delta_L \in \Delta, \delta_H \notin \Delta) \cdot \Pi(\delta_L, \Psi(e_I^*(\delta_L))) + P(\delta_H \in \Delta) \cdot \Pi(\delta_H, \Psi(e_I^*(\delta_L)))] \\ > i \cdot [P(\delta_L \in \Delta) \cdot \Pi(\delta_L, 0) + P(\delta_H \in \Delta, \delta_L \notin \Delta) \cdot \Pi(\delta_H, 0)], \end{aligned} \tag{A.24}$$

or equivalently, making use of the fact that $\Pi(\delta_L, \Psi(e_I^*(\delta_L))) = \Pi(\delta_L, 0)$, $P(\delta_H \in \Delta) = P(\delta_H \in \Delta, \delta_L \in \Delta) + P(\delta_H \in \Delta, \delta_L \notin \Delta)$, and $P(\delta_L \in \Delta) = P(\delta_L \in \Delta, \delta_H \in \Delta) + P(\delta_L \in \Delta, \delta_H \notin \Delta)$,

$$\Pi(\delta_L, 0) < \Pi(\delta_H, \Psi(e_I^*(\delta_L))) - \frac{P(\delta_H \in \Delta, \delta_L \notin \Delta)[\Pi(\delta_H, 0) - \Pi(\delta_H, \Psi(e_I^*(\delta_L)))]}{P(\delta_H \in \Delta, \delta_L \in \Delta)} =: \bar{\Pi}_L.$$

Finally, contract $\boldsymbol{w_1} = (\Pi(\delta_L, 0), 0, 0)$, based on self-commitment, is better for the principal than contract $\boldsymbol{w_1} = (0, 0, \Psi(e_I^*(\delta_L)))$, based on rent reduction, if and only if

$$\begin{aligned} i \cdot P(\delta_H \in \Delta) \cdot [\Pi(\delta_H, 0) - \Pi(\delta_L, 0)] \\ > i \cdot [P(\delta_L \in \Delta, \delta_H \notin \Delta) \cdot \Pi(\delta_L, 0) + P(\delta_H \in \Delta) \cdot \Pi(\delta_H, \Psi(e_I^*(\delta_L)))] \end{aligned} \tag{A.25}$$

or equivalently,

$$\Pi(\delta_L, 0) < \frac{P(\delta_H \in \Delta) \cdot [\Pi(\delta_H, 0) - \Pi(\delta_H, \Psi(e_I^*(\delta_L)))]}{(\delta_L \in \Delta, \delta_H \notin \Delta) + P(\delta_H \in \Delta)} =: \tilde{\Pi}_L.$$

The desired result then immediately follows by defining $\Pi^{min} \equiv \min\{\hat{\Pi}_L, \tilde{\Pi}_L\}$ and $\Pi^{max} \equiv \max\{\bar{\Pi}_L, \hat{\Pi}_L\}$. □

Proof of Proposition 5. To establish the proposition, it remains to derive the upper bound on the principal's profit under a commitment-based contract. If the principal offers a commitment-based contract with $w_{1H} - w_{1L} \geq \Pi(\delta_L, 0)$, then the agent recommends the most productive merger he has identified and the principal acquires only high-synergy targets. The agent's expected utility from exerting effort I in the first period is

$$EU_1(w_{1H}, w_{1L}, I) = i[1 - (1 - p_H)^{n(I)}][\Psi_H + w_{1H}] + [(1 - i) + i(1 - p_H)^{n(I)}]w_{1L} - C \cdot I. \quad \text{(A.26)}$$

The agent is willing to exert high effort if $EU_1(w_{1H}, w_{1L}, 1) \geq EU_1(w_{1H}, w_{1L}, 0)$, or equivalently,

$$w_{1H} - w_{1L} \geq \frac{C}{i[(1 - p_H)^{n(0)} - (1 - p_H)^{n(1)}]} - \Psi_H. \quad \text{(A.27)}$$

The incentive constraint (A.27) reflects the usual result obtained in models of repeated moral hazard with a risk-neutral, wealth-constrained agent: prospective second-period rents act as a reward and punishment for the first period and can therefore be used partially to circumvent the limited-liability constraint. In our case, the higher rent Ψ_H, the more motivated the agent to gather information without being incentivized via w_{1H}. The principal chooses first-period wages to maximize her expected profits,

$$\Pi_1(w_{1H}, w_{1L}, I) = i[1 - (1 - p_H)^{n(I)}][\Pi(\delta_H, 0) - w_{1H}] - [(1 - i) + i(1 - p_H)^{n(I)}]w_{1L},$$

subject to the above incentive constraint (A.27), the limited-liability constraint, and the additional constraint that $w_{1H} - w_{1L} \geq \Pi(\delta_L, 0)$. Note that incentive and limited-liability constraints together imply that the participation constraint is satisfied. Clearly, the principal optimally sets $w_{1L} = 0$. Moreover, note that, ceteris paribus, the principal prefers the agent to exert high effort $I = 1$, because this increases the likelihood of strictly positive merger profits to be realized. With w_{1H} being bounded below, the best the principal can thus hope for is the agent exerting high effort at the minimum wage such that the maximum profit under a commitment-based contract is $\Pi_1(\Pi(\delta_L, 0), 0, 1) = i[1 - (1 - p_H)^{n(1)}][\Pi(\delta_H, 0) - \Pi(\delta_L, 0)]$. This establishes the desired result. □

References

Aghion, P. and J. Tirole (1997): "Formal and Real Authority in Organizations." *Journal of Political Economy*, Vol. 105, 1-29.

Amihud, Y. and B. Lev (1981): "Risk Reduction as a Managerial Motive for Conglomerate Mergers." *Bell Journal of Economics*, Vol. 12, 605–617.

Anderson, C.W., Becher, D.A. and T.L. Campbell II (2004): "Bank Mergers, the Market for Bank CEOs, and Managerial Incentives." *Journal of Financial Intermediation*, Vol. 13, 6-27.

Andrews, E.L. (1998): "Shaping a Global Giant: The Deal; Daimler and Chrysler Trade Vows Before Complex Union." *New York Times*, May 08.

Baker, G., Gibbons, R. and K.J. Murphy (1999): "Informal Authority in Organizations." *Journal of Law, Economics, and Organizations*, Vol. 15, 56-73.

Baumol, W.J. (1959): *Business Behavior, Value and Growth*. New York.

Berkovitch, E., R. Israel and Y. Spiegel (2010): "A Double Moral Hazard Model of Organization Design." *Journal of Economics and Management Strategy*, Vol. 19, 55-85.

Bebchuk, L. and Y. Grinstein (2005): "Firm Expansion and CEO Pay." *NBER Working Paper Series*, Working Paper 11886, http://www.nber.org/papers/w11886.

Bliss, R.T. and R.J. Rosen (2001): "CEO Compensation and Bank Mergers." *Journal of Financial Economics*, Vol. 61, 107-138.

Bradley, M., Desai, A. and E.H. Kim (1988): "Synergistic Gains from Corporate Acquisitions and their Division between the Stockholders of Target and Acquiring Firms." *Journal of Financial Economics*, Vol. 21, 3-40.

Bruner, R.F. (2005): *Deals from Hell*. Hoboken.

Brusco, S. (1996): "Bankruptcy, Takeovers, and Wage Contracts." *Journal of Economics and Management Strategy*, Vol. 5, 515-534.

Bryant, A. (1999): "The World: Raising the Stakes; American Pay Rattles Foreign Partners." *New York Times*, January 17.

De Bondt, W.F.M. and H.E. Thompson (1992): "Is Economic Efficiency the Driving Force behind Mergers?." *Managerial and Decision Economics*, Vol. 13, 31-44.

Demski, J.S. and D.E.M. Sappington (1987): "Delegated Expertise." *Journal of Accounting Research*, Vol. 25, 68-89.

Dewatripont, M. and J. Tirole (1999): "Advocates." *Journal of Political Economy*, Vol. 107, 1-39.

Dow, J. and C. C. Raposo (2005): "CEO Compensation, Change, and Corporate Strategy." *Journal of Finance*, Vol. 60, 2701-2727.

Girma, S., Thompson, S. and P.W. Wright (2006): "The Impact of Merger Activity on Executive Pay in the United Kingdom." *Economica*, Vol. 73, 321-339.

Grinstein, Y. and P. Hribar (2004): "CEO Compensation and Incentives: Evidence from M&A Bonuses." *Journal of Financial Economics,* Vol. 73, 119-143.

Guest, P.M. (2009): "The Impact of Mergers and Acquisitions on Executive Pay in the United Kingdom." *Economica*, Vol. 76, 149-175.

Harford, J. and K. Li (2007): "Decoupling CEO Wealth and Firm Performance: The Case of Acquiring CEOs." *Journal of Finance*, Vol. 62, 917-949.

Jarrell, G.A., Brickley, J.A. and J.M. Netter (1988): "The Market for Corporate Control: The Empirical Evidence since 1980." *Journal of Economic Perspectives*, Vol. 2, No. 1, 49-68.

Jensen, M.C. (1986): "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers." *American Economic Review, Papers and Proceedings*, Vol. 76, 323-329.

Jensen, M.C. and R.S. Ruback (1983): "The Market for Corporate Control: The Scientific Evidence." *Journal of Financial Economics,* Vol. 11, 5-50.

Khalil, F., D. Kim and D. Shin (2006): "Optimal Task Design: To Integrate or to Separate Planning and Implementation?" *Journal of Economics and Management Strategy*, Vol. 15, 457-478.

Kräkel, M. and D. Müller (2012): "Sabotage in Teams." *Economics Letters*, Vol. 115, 289-292.

Kräkel, M. and A. Schöttner (2010): "Minimum Wages and Excessive Effort Supply." *Economics Letters,* Vol. 108, 341-344.

Lambert, R.A. (1986): "Executive Effort and Selection of Risky Projects." *RAND Journal of Economics*, Vol. 17, 77-88.

Lewis, T.R. and D.E.M. Sappington (1997): "Information Management in Incentive Problems." *Journal of Political Economy*, Vol. 105, 796-821.

Marris, R. (1963): "A Model of the ‚Managerial Enterprise'." *Quarterly Journal of Economics*, Vol. 77, 185-209.

Moore, J. and R. Repullo (1990): "Nash Implementation: A Full Characterization." *Econometrica*, Vol. 28, 1083-1099.

Morck, R., Shleifer, A. and R.W. Vishny (1990): "Do Managerial Objectives Drive Bad Acquisitions?." *Journal of Finance*, Vol. 45, 31-48.

Ohlendorf, S. and P.W. Schmitz (2012): “Repeated Moral Hazard and Contracts with Memory: The Case of Risk-Neutrality.” *International Economic Review*, Vol. 53, 433-452.

Roll, R. (1986): "The Hubris Hypothesis of Corporate Takeovers." *Journal of Business*, Vol. 59, 197-216.

Schmitz, P.W. (2005a): “Should Contractual Clauses that Forbid Renegotiation Always be Enforced?” *Journal of Law, Economics, and Organization*, Vol. 21, 315-329.

Schmitz, P.W. (2005b): “Allocating Control in Agency Problems with Limited Liability and Sequential Hidden Actions.” *RAND Journal of Economics*, Vol. 36, 318-336.

Schmitz, P.W. (2012): “Job Design with Conflicting Tasks Reconsidered.” *Working Paper*.

Schnitzer, M. (1995): ",Breach of Trust’ in Takeovers and the Optimal Corporate Charter." *Journal of Industrial Economics*, Vol. 43, 229-259.

Shleifer, A. and L.H. Summers (1988): "Breach of Trust in Hostile Takeovers." In: A.J. Auerbach (ed.), *Corporate Takeovers: Causes and Consequences*. Chicago, 33-56.

Shleifer, A. and R.W. Vishny (1989): "Management Entrenchment: The Case of Manager-Specific Investments." *Journal of Financial Economics*, Vol. 25, 123-139.

Taylor, C. (1995): "The Economics of Breakdowns, Checkups, and Cures." *Journal of Political Economy*, Vol. 103, 53-74.

Williams, M.A., Michael, T.B. and E.R. Waller (2008): "Managerial Incentives and Acquisitions: A Survey of the Literature." *Managerial Finance*, Vol. 34, 328-341.

Williamson, O.E. (1963): "Managerial Discretion and Business Behavior." *American Economic Review*, Vol. 53, 1030-1057.

Example for synergies and effort being complements:

Let $q(e) = \alpha \cdot e$ and $p(e\delta) = 1 - \exp(-\beta e\delta)$ with $\alpha, \beta > 0$ guaranteeing $q, p \in (0, 1)$ in the optimum. In addition let $c(e) = \frac{\gamma}{2}e^2$ with $\gamma > 0$ and $\kappa = 0$. In the second period, the agent maximizes

$$EU_2(e) = \alpha e w_{2H} - \frac{\gamma}{2}e^2,$$

leading to the incentive constraint $\alpha w_{2H}^* = \gamma e^*$, and the strictly positive second-period rent $EU_2^*(e) = \frac{1}{2}\frac{\alpha^2}{\gamma}(w_{2H}^*)^2$, which increases in the wage w_{2H}^*. The principal solves

$$\begin{aligned} &\max_{w_{2H}} \pi_L + (\pi_H - \pi_L)(1 - \exp(-\beta e^* \delta)) - \alpha e^* w_{2H}^* \\ &= \max_{w_{2H}} \pi_L + (\pi_H - \pi_L)\left(1 - \exp\left(-\beta \frac{\alpha w_{2H}^*}{\gamma}\delta\right)\right) - \frac{\alpha^2 (w_{2H}^*)^2}{\gamma}. \end{aligned}$$

The first-order condition leads to

$$(\pi_H - \pi_L)\beta\delta \exp\left(-\beta\frac{\alpha w_{2H}^*}{\gamma}\delta\right) = 2\alpha w_{2H}^* \iff w_{2H}^* = \frac{\gamma}{\alpha\beta\delta}W\left((\pi_H - \pi_L)\frac{\beta^2\delta^2}{2\gamma}\right)$$

with W denoting the Lambert W function (or omega function), which is defined as $W(x)$ with $x = W(x)\exp(W(x))$. Differentiating w_{2H}^* with respect to δ yields

$$\frac{\partial w_{2H}^*}{\partial \delta} = \frac{\gamma}{\alpha\beta\delta^2}\frac{1 - W\left((\pi_H - \pi_L)\frac{\beta^2\delta^2}{2\gamma}\right)}{1 + W\left((\pi_H - \pi_L)\frac{\beta^2\delta^2}{2\gamma}\right)}W\left((\pi_H - \pi_L)\frac{\beta^2\delta^2}{2\gamma}\right),$$

which is negative iff $W\left((\pi_H - \pi_L)\frac{\beta^2\delta^2}{2\gamma}\right) > 1$, that is, if $(\pi_H - \pi_L)\frac{\beta^2\delta^2}{2\gamma}$ is sufficiently large. (Note that we must have that $q(e^*) = \alpha e^* < 1 \Leftrightarrow \frac{\alpha^2}{\gamma}w_{2H}^* < 1 \Leftrightarrow W\left((\pi_H - \pi_L)\frac{\beta^2\delta^2}{2\gamma}\right) < \frac{\beta\delta}{\alpha}$, which is always in line with $W\left((\pi_H - \pi_L)\frac{\beta^2\delta^2}{2\gamma}\right) > 1$ for α being sufficiently small.)

On the Example in Section 5.2:

Let $q(e+\delta) = \alpha \cdot (e+\delta)$ and $p(e+\delta) = \beta \cdot \ln(e+\delta)$ with $\alpha, \beta > 0$ being sufficiently small to guarantee $q, p \in (0, 1)$ in the optimum. Second-period effort costs are described by $c(e) = \frac{\gamma}{2}e^2$ with $\gamma > 0$. Hence, the incentive constraint is given by $w_{2H} = \gamma e/\alpha$ and the principal solves

$$\max_{w_{2H}} \pi_L + (\pi_H - \pi_L)\beta \ln\left(\frac{\alpha}{\gamma}w_{2H} + \delta\right) - \alpha\left(\frac{\alpha}{\gamma}w_{2H} + \delta\right)w_{2H}. \tag{A.28}$$

Since the objective function is strictly concave, the optimal wage w_{2H}^* is described by the respective first-order condition, leading to

$$w_{2H}^{*} = \frac{\gamma\sqrt{8(\pi_H - \pi_L)\frac{\beta}{\gamma} + \delta^2} - 3\gamma\delta}{4\alpha}. \tag{A.29}$$

For a feasible solution, let $(\pi_H - \pi_L)\beta > \gamma\delta^2$. The agent's second-period rent reads as

$$\begin{aligned} EU_2(e) &= \alpha(e(w_{2H}^{*}) + \delta) \cdot w_{2H}^{*} - \frac{\gamma}{2} e(w_{2H}^{*})^2 \\ &= \frac{(\pi_H - \pi_L)\beta}{4} + \frac{\gamma\delta}{16}\left(\sqrt{\delta^2 + \frac{8\beta}{\gamma}(\pi_H - \pi_L)} - 7\delta\right). \end{aligned} \tag{A.30}$$

Differentiating with respect to δ yields

$$\frac{4\beta(\pi_H - \pi_L) + \gamma\delta^2 - 7\gamma\delta\sqrt{\delta^2 + \frac{8\beta}{\gamma}(\pi_H - \pi_L)}}{8\sqrt{\delta^2 + \frac{8\beta}{\gamma}(\pi_H - \pi_L)}}, \tag{A.31}$$

which is negative if $\beta^2(\pi_H - \pi_L)^2 - 24\beta\gamma\delta^2(\pi_H - \pi_L) - \gamma^2\delta^4 < 0$. In particular, if γ is sufficiently large so that the optimal effort is very small anyway, then the agent will focus on the wage-increasing effect of low merger synergies.

Derivation of Condition (30):

A necessary condition for the principal to directly influence the agent's recommendation behavior by choosing $\kappa = \Psi(e_I^{*}(\delta_L))$ is that $\tilde{\Pi}_L < \bar{\Pi}_L$.

In the following, $P(\Delta_+ \neq \varnothing) \equiv P(\delta_H \in \Delta, \delta_L \in \Delta) + P(\delta_H \in \Delta, \delta_L \notin \Delta) + P(\delta_L \in \Delta, \delta_H \notin \Delta)$ denotes the probability that the agent identifies at least one target firm with positive synergies. Furthermore, we use the shorter notation $P_i := P(\delta_i \in \Delta)$ (with $i = H, L$), $P_{\neg i} := P(\delta_i \notin \Delta)$ (with $i = H, L$), $P_{ij} \equiv P_{ji} := P(\delta_i \in \Delta, \delta_j \in \Delta)$ (with $i, j = H, L$; $i \neq j$), $P_{i\neg j} := P(\delta_i \in \Delta, \delta_j \notin \Delta)$ (with $i, j = H, L$; $i \neq j$) and so on. Moreover, let $\Pi_i(0) := \Pi(\delta_i, 0)$ (with $i = H, L$) and $\Pi_H(\Psi) := \Pi(\delta_H, \Psi(e_I^{*}(\delta_L)))$.

We have

$$\bar{\Pi}_L > \tilde{\Pi}_L \Leftrightarrow \frac{(P_{HL} + P_{H\neg L})\Pi_H(\Psi) - P_{H\neg L}\Pi_H(0)}{P_{HL}} > \frac{P_H\left[\Pi_H(0) - \Pi_H(\Psi)\right]}{P_{L\neg H} + P_H}$$

$$\Leftrightarrow (P_{L\neg H} + P_H)\left[P_{HL}\Pi_H(\Psi) - P_{H\neg L}\left(\Pi_H(0) - \Pi_H(\Psi)\right)\right] > P_{HL}P_H\left[\Pi_H(0) - \Pi_H(\Psi)\right]$$

$$\begin{aligned} \Leftrightarrow (P_{L\neg H} + P_H)\, P_{HL}\Pi_H(\Psi) &> \left[\Pi_H(0) - \Pi_H(\Psi)\right]\left[P_{HL}P_H + P_{H\neg L}\left(P_{L\neg H} + P_H\right)\right] \\ &= \left[\Pi_H(0) - \Pi_H(\Psi)\right]\left(P_H^2 + P_{L\neg H}P_{H\neg L}\right), \end{aligned}$$

where the last equality follows from

$$
\begin{aligned}
P_{HL}P_H + P_{H\neg L}\left(P_{L\neg H} + P_H\right) &= P_{HL}P_H + \left(P_H - P_{HL}\right)\left(P_{L\neg H} + P_H\right) \\
&= P_{HL}P_H + P_H\left(P_{L\neg H} + P_H\right) - P_{HL}P_{L\neg H} - P_{HL}P_H \\
&= P_H^2 + P_{L\neg H}\left(P_H - P_{HL}\right) \\
&= P_H^2 + P_{L\neg H}P_{H\neg L}.
\end{aligned}
$$

The inequality

$$
\left(P_{L\neg H} + P_H\right)P_{HL}\Pi_H(\Psi) > \left[\Pi_H(0) - \Pi_H(\Psi)\right]\left(P_H^2 + P_{L\neg H}P_{H\neg L}\right)
$$

can be rewritten as follows:

$$
\begin{aligned}
&\left(P_{L\neg H} + P_H\right)P_{HL}\Pi_H(\Psi) > \left[\Pi_H(0) - \Pi_H(\Psi)\right]\left(P_H^2 + P_{L\neg H}P_{H\neg L}\right) \\
\Leftrightarrow\ &\Pi_H(0)\left(P_H^2 + P_{L\neg H}P_{H\neg L}\right) < \Pi_H(\Psi)\left[\left(P_{L\neg H} + P_H\right)P_{HL} + P_H^2 + P_{L\neg H}P_{H\neg L}\right] \\
&\qquad = \Pi_H(\Psi)\left[P_{L\neg H}\left(P_{HL} + P_{H\neg L}\right) + P_H P_{HL} + P_H^2\right] \\
&\qquad = \Pi_H(\Psi)\left[P_H\left(P_{L\neg H} + P_{HL}\right) + P_H^2\right] \\
&\qquad = \Pi_H(\Psi)P_H\left(P_L + P_H\right)
\end{aligned}
$$

or equivalntly,

$$
\begin{aligned}
\frac{\Pi_H(\Psi)}{\Pi_H(0)} &> \frac{P_H^2 + P_{L\neg H}P_{H\neg L}}{P_H\left(P_L + P_H\right)} \\
&= \frac{P_H^2 + \left(P_L - P_{LH}\right)\left(P_H - P_{HL}\right)}{P_H\left(P_L + P_H\right)} \\
&= \frac{P_H^2 + P_H P_L - \left(P_L + P_H\right)P_{HL} + P_{HL}^2}{P_H\left(P_L + P_H\right)} \\
&= 1 - \frac{\left(P_L + P_H\right)P_{HL} - P_{HL}^2}{P_H\left(P_L + P_H\right)} \\
&= 1 - \frac{P_{HL}}{P_H}\frac{P_L + P_H - P_{HL}}{P_L + P_H} \\
&= 1 - \frac{P_{HL}}{P_H}\left[1 - \frac{P_{HL}}{P_L + P_H}\right].
\end{aligned}
$$

Since

$$
\begin{aligned}
P_L + P_H &= \left(P_{L\neg H} + P_{LH}\right) + \left(P_{H\neg L} + P_{HL}\right) \\
&= \left(P_{L\neg H} + P_{H\neg L} + P_{LH}\right) + P_{LH} = P(\Delta_+ \neq \varnothing) + P_{LH},
\end{aligned}
$$

we obtain

$$\frac{\Pi_H(\Psi)}{\Pi_H(0)} > 1 - \frac{P_{HL}}{P_H}\left[1 - \frac{P_{HL}}{P_L + P_H}\right] \Leftrightarrow \frac{\Pi_H(\Psi)}{\Pi_H(0)} > 1 - \frac{P_{HL}}{P_H}\left[1 - \frac{1}{\frac{P(\Delta_+ \neq \varnothing)}{P_{HL}} + 1}\right].$$