MACIE PAPER SERIES

Marburg Centre for Institutional Economics

MACIE

Nr. 2021/02

Strategic Delay in R&D Projects -An Agency Perspective

Elisabeth Schulte
MACIE, Philipps-Universität Marburg

Matthias Verbeck
MACIE, Philipps-Universität Marburg

Marburg Centre for Institutional Economics • Coordination: Prof. Dr. Elisabeth Schulte
c/o Research Group Institutional Economics • Barfuessertor 2 • D-35037 Marburg

Phone: +49 (0) 6421-28-23196 • Fax: +49 (0) 6421-28-24858 •
www.uni-marburg.de/fb02/MACIE • macie@wiwi.uni-marburg.de

Philipps Universität Marburg

Strategic Delay in R&D Projects - An Agency Perspective *

Elisabeth Schulte
University of Marburg

Matthias Verbeck
University of Marburg†

November 1, 2021

Abstract

We show that strategic delay can pose a problem in delegated R&D projects. In our model, a principal delegates a research project to an agent. Depending on the agent's effort provision in two time periods, the research project can be completed either early, late or never. Our central assumption is that the agent is able to opportunistically withhold possible early completion from the principal (strategic delay). We derive the conditions under which strategic delay poses a problem. There are two options for the contract's optimal adjustment that both fall short of the first-best solution. (1) The contract prevents strategic delay by separating between successful and unsuccessful agents after period 1, but thereby distorts the agent's working incentives in both periods. (2) The principal strategically delays the start of the research project until the second period. We discuss several model extensions and possible institutional remedies to mitigate the problem.

JEL codes: D82, D83, D86, M52

Keywords: Principal-agent problem, Moral hazard, Adverse selection, Information disclosure, Project completion, Strategic delay, Incentives in research and development.

*Part of the research was conducted at University of Texas at Dallas. Their hospitality is gratefully acknowledged.
†Corresponding author: matthias.verbeck@gmail.com

1 Introduction

The standard moral hazard model in the fashion of Holmström (1979) has contributed enormously to understanding efficiency losses in agency relationships, i.e. when tasks are delegated rather than performed by the bearer of the payoff consequences. One of the pivotal premises of the model is that the agent's actions are hidden, whereas the output, which is assumed to be stochastically related to the agent's actions, is common knowledge. The notion that the observability of output is unproblematic seems especially uncontroversial in the case of a one-shot interaction between the two parties. Even if an output is technically hidden to the principal (an assumption that is valid for many agency relationships), the reported output can still be verified by the principal. An agent in an optimal one-shot contract will see his wage increase with performance. Hence, the agent would never underreport his level of output, thus rationalizing the common-knowledge assumption.

However, when the interaction between principal and agent lasts more than one period, this simple reasoning becomes more questionable. We consider a two-period timescale in which an agent is supposed to work on a project and the agent's success (i.e. project completion) can occur early (period 1), late (period 2) or never. One characteristic of this scenario is that the desired effort level in the second period critically hinges on the first period's outcome, because only the case of a first period failure would arouse the principal's interest in a positive effort level for the second period. A rational agent who seeks to maximize his income might therefore be reluctant to immediately disclose an early success, anticipating that his or her second period effort (possibly resulting in an extra payment) will only be required in the case of a (reported) failure.

The unobservability of the timing of success may give rise to a conflict between creating working incentives on the one hand, and ensuring truthful reporting on the other. Where it is advantageous from the agent's perspective, s/he may then *strategically delay* the project's completion by leading the principal to believe that a project already finalized in period 1 was not actually completed before the end of period 2.

Indeed, projects finalized later than originally planned are the rule rather than the exception. Examples of projects that were eventually completed with delays of up to several years exist in abundance. According to Parkinson's law, "work expands so as to fill the time available for its completion"(Parkinson (1957)), an observation that will be familiar to anybody with experience in time-critical undertakings. Missed deadlines and unforeseen delays are a notorious problem that plagues project managers, home-builders and doctoral advisors alike. Besides being a permanent cause of nuisance, they are also of high economic relevance.

Frequently cited explanations for this phenomenon stem from concepts rooted in behavioral economics (e.g. ODonoghue and Rabin (1999)), such as limited self-control, time-inconsistent preferences and systematic overestimation of one's own abilities. Other rationales come from managerial economics, e.g. restrictive deadlines as a means of creating working incentives (e.g. Green and Taylor (2016)) or the strategic slowdown of projects due to the manager's intrinsic utility derived from holding the position of project leader (Katolnik and Schöndube (2019)). This paper aims at adding a novel perspective on the problem. The question we pose is whether the agent's informational advantage about a project's true completion status might be a (further) explanation for the omnipresence of belated project completions and violated deadlines in delegated research projects.

Our analysis intends to focus on *R&D projects*, which typically feature several characteristics that make them relatively unique and therefore worth investigating in their own right. We present a parsimonious economic model that captures these characteristics, and analyze the problem of delegated research from the perspective of principal-agent theory.

For our analysis, we follow the model of Holmström (1979) with an agent who chooses a continuous level of unobservable effort and generates a binary output, stochastically linked to the agent's input. These assumptions are a good approximation of conditions found in many research settings, where a high degree of effort and dedication is a necessary, but not a sufficient condition for the success of the project, which is typically characterized by the ever-present possibility of failure. Thus, in research undertakings, effort is typically less deterministically related to the output, compared to other kinds of projects (e.g. building projects). Therefore, it seems natural to model the project's outcome as binary (success or failure), e.g. researchers in the pharmaceutical industry either succeed in making a new drug market-ready or not. What is more, if a research endeavor remained fruitless in a given period, the effort invested in that period has no more than a very limited influence, if any, on the probability of completing the undertaking at a later stage. Often, research starts from scratch after a particular approach turns out to be a dead end. This characteristic distinguishes research projects from other kinds of projects where effort accumulates over time and early investments of effort do have an impact on the prospects of completing the project at a later stage (e.g. Toxvaerd (2006)).

R&D projects are furthermore often embedded in larger contexts which impose deadlines for their completion. Reasons for deadlines can be seen in the (likely) date by which a competitor will have solved a similar problem, or in the limited availability of resources (laboratories, scientific equipment). In our stylized model, we restrict the timescale to two periods and the deadline is at the end of period 2.

In addition, researchers typically exhibit a high degree of specific knowledge, making their actions and results hard to evaluate for any less knowledgeable party. Therefore, a (less informed) principal must to some extent rely on the agent's reports on project status. An asymmetry in observability is plausible: While a researcher will generally be unable to pretend that an unfinished project has been completed, s/he may very well be able to conceal its completion. This limited observability of the completion status gives rise to the principal's problem of discriminating between different types of agents, successful and unsuccessful ones after period 1, and a rational principal would already have to consider the agent's potentially untruthful reporting at the contracting stage. This is a typical problem of hidden information. Unlike the canonical model, however, in our model contracting takes place before the information asymmetry comes about, and the type distribution is *endogenously* determined, viz. by the effort choice for period 1, which in turn depends on the working incentives that are induced by the contract. The problems of hidden actions and hidden information are thus closely linked in our model.

For certain parameter constellations, specifically if early completion of the project is particularly desirable, strategic delay is neither a problem nor an observable phenomenon. The optimal incentive-compatible contract ensures the truthful revelation of the project status as a by-product. However, if obtaining a solution late rather than never is the driving interest, then the truth-telling constraint binds and the first best is not obtainable. The principal constructs the contract so as to deter the agent's strategic delay of a success report. In doing so, the agent's effort in the two periods can no longer be separately incentivized. As a consequence, either the agent's efforts deviate from his or her first-best levels in both periods or the principal strategically delays the start of the project, making early success impossible. Thus, strategic delay on the part of the agent is a problem, but not a phenomenon on the equilibrium path. In fact, if the principal does not strategically delay the start of the contract, early success will be more likely than in the first-best, due to the distortion of the agent's incentives. If this distortion is too severe, the principal will prefer to strategically delay the start of the project and the optimal contract induces first-best effort in the second period. In our model setting, if a strategic delay occurs, it is due to not allowing work to fill the time that would in principle be available.

The remainder of the paper is structured as follows: Section 2 relates our research to the literature. In Section 3, we present our model, derive the circumstances under which the first best is (not) attainable and characterize the optimal contract. We also discuss further contractual frictions that may be relevant in the context of our model, in particular limited liability and limited commitment power on the part of the principal. In Section 4, we extend the scope of our model. We discuss limits to the principal's commitment power, and possible institutional remedies to the problem of strategic delay, in particular the possibility of contracting with more than one agent. We conclude in Section 5. For proofs and mathematical details of our arguments, we refer to the Appendices.

2 Related Literature

Our work is related to various strands of the existing literature. Most broadly, our work contributes to the literature on (dynamic) agency, in which a principal wishes to set proper working incentives such that the agent's (intertemporal) performance maximizes the principal's benefit, e.g. Holmström (1979), Lambert (1983), Holmström and Milgrom (1987) and Malcomson and Spinnewyn (1988). More recent works in that strand of literature that - unlike the present paper - use continuous time frameworks comprise, among others, Sannikov (2008), Biais et al. (2007, 2010), Hoffmann and Pfeil (2010) and Williams (2015). None of these papers explicitly address the problem of deadlines; rather, the focus of these papers is on analyzing the agent's incentives to divert cash flows for private benefit and what contractual countermeasures can be taken by the principal.

There is also a rich body of literature that explicitly analyzes R&D settings from an agency perspective (e.g. Manso (2011)). In many of these contributions, the parties involved learn about the project value over time and make the continuation of the project dependent on intermediate project outputs (e.g. Bergemann and Hege (1998, 2005), Hörner and Samuelson (2013)). Typically, the informational friction between principal and agent combined with the project's unknown returns results in the funding of research projects being stopped too early. Unlike these works, we do not consider the case of an uncertain or steadily updated project value in our own model.

Instead, we focus on the information asymmetry that arises between principal and agent because of the unobservability of the project's completion status. There are a few contributions that have chosen a similar approach. Most notably, Green and Taylor (2016) analyze a two-stage project setting where the project's progress level is visible only to the agent and the agent's self-reported project status is directly contractible. The optimal incentives scheme uses the potential termination of the research project as an incentive device to prevent both shirking and making false statements about the project's actual progress level. In a similar fashion, Lewis and Ottaviani (2008), Lewis (2012) and Ulbricht (2016) analyze models of delegated search where the agent is able to underreport the obtained results. All three contributions substantially differ from ours in numerous ways. In Lewis and Ottaviani (2008), the search revenues are taken from a continuous distribution and are decreasing over time, rendering the optimal speed of search the central question of the analysis. In the paper of Lewis (2012), a search for the best alternative is analyzed. Once more, deadlines appear endogenously from the principal's quest to not fund an agent who has already made a valuable discovery. Ulbricht (2016) focusses on a delegated search where the distribution of search revenues is unknown, but can be disclosed by the agent. In contrast to the works cited above, our paper uses a discrete two-time-period setting, we do not impose wealth constraints on the agent, and, most importantly, the possibilities to choose the contract duration are limited due to the presence of an exogenous deadline in our model. Campbell et al. (2014) present an insightful model in which multiple agents can complete a joint project and might withhold output from each other (but not from a principal, as in our work).

Moreover, our paper is related to further contract-theory contributions that deal with incentives for project completion or potential delays. Mason and Välimäki (2015) study optimal contracts for project completion in a continuous time setting and analyze the resulting optimal contracts with and without the principal's ability to commit to a long-term contract. Toxvaerd (2006) models the execution of a multistage project under agency that requires multiple milestones to be achieved for its completion. Due to the agent's risk aversion the per-period effort level is generally lower than in a first-best setting, resulting in longer completion times as compared to a non-delegated project. Katolnik and Schöndube (2019) present a model in which the agent derives a private benefit from holding the position of project manager and therefore has an incentive to delay the project's completion. Models that analyze the problem of delayed output as a result of an agent's time-inconsistent preferences are provided, for example, by O'Donoghue and Rabin (1999) and Herweg and Müller (2011).

3 The Model

A risk-neutral (female) principal seeks to complete a research project before its deadline, which is two periods ahead. She can delegate the research to a risk-neutral (male) agent, whose research efforts are not observable. The principal's preferred effort choices are our benchmark for the first best. Thus, we abstract from the possibility that the completion of the research project has a social value beyond the principal's benefits. The unobservable effort $e_t \in [0, 1]$ in period $t = 1, 2$ determines the probability $\rho(e_t)$ of successfully completing the project in period t, $\rho(e_t) = e_t$. The project can either be successfully completed early ($t = 1$), late ($t = 2$) or never (ultimate project failure). Research effort comes at a strictly convex cost $C(e_t)$, with $C(0) = C'(0) = 0$.

Project success is observable only to the researcher (i.e. to the agent), and it is verifiable. A failure is not verifiable, but payments can condition on a failure to report a success. The principal's payoff from the project depends on the project status and the timing of its revelation, as depicted in Table 1.

The principal obtains a payoff $Y_E > 0$ if an early success (i.e. in period 1) is reported, and $Y_L > 0$ if a success is achieved and reported in period 2. If the project ultimately ends in failure or a success is never reported, the principal obtains $Y_N < Y_E, Y_L$. If the agent withholds an early success and reports a successful completion of the project only in period 2, the principal obtains Y_D. While the principal realizes the payoff consequence of this strategic delay, it is not verifiable whether the success has been achieved early or late.

<table>
<tr><td>Outcome in $t = 1$</td><td colspan="2">Success</td><td colspan="3">Failure</td></tr>
<tr><td>Report in $t = 1$</td><td>Success</td><td colspan="4">Failure</td></tr>
<tr><td>Payoff in $t = 1$</td><td>Y_E</td><td colspan="4">0</td></tr>
</table>

<table>
<tr><td>Outcome in $t = 2$</td><td colspan="2">-</td><td colspan="2">Success</td><td>Ultimate Failure</td></tr>
<tr><td>Report in $t = 2$</td><td>-</td><td>Success</td><td>Success</td><td colspan="2">Ultimate Failure</td></tr>
<tr><td>Payoff in $t = 2$</td><td>-</td><td>Y_D</td><td>Y_L</td><td colspan="2">Y_N</td></tr>
</table>

Table 1: The principal's payoffs as a function of outcomes and reports in $t = 1, 2$

The principal can specify three payments for the verifiable events, w_E, w_L, w_N, i.e. the agent earns a wage w_E for reporting a success in period 1, w_L for reporting a success in period 2, and w_N in case he fails to report a success in either period.

We assume that the principal and agent are rational, and that they maximize their expected discounted payoff (net of the cost of effort). We assume a common discount factor δ and that our parameter constellation satisfies:

A1 $Y_E - \delta \cdot Y_N < C'(1), Y_L - Y_N < C'(1)$.

A2 $Y_E > \delta \cdot Y_D$.

We impose these assumptions in order to avoid case distinctions in our statements. Assumption A1 rules out that maximal effort provision (and hence certain completion of the project) is optimal in any period. We choose a parsimonious model representation based on this parameter constellation, remaining agnostic about its underlying reasons, as there may be many factors that impact on it. For instance, a successful project completion may induce a constant stream of payoffs starting at the moment of completion, in which case $Y_E > Y_L$, or it may give rise to a payoff (e.g. attention as the main source of the principal's benefit) only in the period of its publication, in which case Y_L could be greater than Y_E. Ultimate project failure may be more severe than a (preliminary) failure to achieve a success in the first period ($Y_N < 0$), or it may be equally bad. We assume that a strategically delayed success report harms the principal in order to rule out a (trivial) preference-based explanation for strategic delay (Assumption A2). If instead $\delta \cdot Y_D \geq Y_E$, the principal actually prefers the agent to strategically delay the report of an early success (maybe because she herself only receives research funds for as long as the project remains incomplete). In this case, strategic delay would be a phenomenon that occurs due to the principal's preferences, but it would not be a problem. Several reasons come to mind why a strategic delay can be harmful. The research result may leak, the exploitability of the discovery may be limited, or the principal may have to bear costs for continuing the research in the second period. Moreover, the priority principle in science contributes to the desire to avoid a delay in publishing a research result.

3.1 First Best

For now, we will neglect the agency problems and consider the principal's optimization problem as though she were able to control the effort levels directly (e.g., performing the research herself). We refer to the solution that maximizes the principal's payoff as first best. The principal's ex ante objective function can then be written as:

$$V_1 = e_1 \cdot Y_E - C(e_1) + (1 - e_1) \cdot \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - C(e_2)) . \tag{1}$$

If the project was completed in the first period, no further research is carried out. If the project was not yet completed in the first period, the optimal effort in the second period maximizes:

$$V_2 = Y_N + e_2 \cdot (Y_L - Y_N) - C(e_2), \tag{2}$$

where $C'(0) = 0$ makes sure that the project is worth continuing in $t = 2$ and rules out the corner solution $e_2 = 0$. The first order condition reads:

$$Y_L - Y_N = C'(e_2). \tag{3}$$

As the marginal benefit is constant and the marginal cost increases in e_2, problem (2) has a unique solution. The corner solution $e_2 = 1$ is ruled out by Assumption A1. Consequently,

(3) defines the optimal level e_2^*.[1] Denote with V_2^* the (optimized) expected value of the second period research, given a failure in the first period (evaluated in the second period):

$$V_2^* = Y_N + e_2^* \cdot (Y_L - Y_N) - C(e_2^*), \tag{4}$$

where e_2^* solves (3). Note that e_2^* is independent of the effort level chosen in the first period. The principal chooses e_1 so as to maximize:

$$V_1 = e_1 \cdot (Y_E - \delta \cdot V_2^*) - C(e_1), \tag{5}$$

where V_2^* is defined in (4) and e_2^* satisfies (3). The condition for an interior optimum reads:

$$Y_E - \delta \cdot V_2^* = C'(e_1). \tag{6}$$

If equation (6) has a solution, it is unique, as the marginal cost of effort is strictly increasing and the marginal benefit is constant. The corner solution $e_1^* = 1$ is ruled out by Assumption A1.

We find that:

Proposition 1 *The first-best combination of efforts e_1^*, e_2^* is unique. It is optimal to report an early success in the first period.*
(i) If $Y_E > \delta \cdot V_2^*$, where V_2^* is defined in (4), $e_1^*, e_2^* \in (0, 1)$, e_1^* is defined by (6) and e_2^* is defined by (3).
(ii) If $\delta \cdot V_2^* > Y_E, e_1^* = 0, e_2^* \in (0, 1)$, and e_2^* is defined by (3).

Proof: See Appendix A.

The optimal reporting behavior follows from Assumption A2. If the LHS of (6) is negative (case (ii) in the above Proposition), the principal prefers not to provide any effort at all in the first period at all, i.e. $e_1^* = 0$. In the following, we assume that the parameter constellation is such that the LHS of (6) is strictly positive, such that it is optimal to induce some effort in period 1. Otherwise, if no effort is optimally provided in period 1, we obtain another preference-based explanation for a delay of the project, and the strategic delay of a success report is not an issue.

The benefit of a marginal unit of effort is $Y_E - \delta \cdot V_2^*$ in the first period, and $Y_L - Y_N$ in the second period. The component $\delta \cdot V_2^*$ represents the expected payoff from continuing the research project in the second period, which is forgone if the project is completed early. When $\delta \cdot V_2^*$ exceeds $Y_E - (Y_L - Y_N)$, the benefit from providing a marginal unit of effort early rather than late, it follows from the first order conditions that $e_2^* > e_1^*$. In such parameter constellations, it is ex ante more likely that the project will be completed close to the deadline than ahead of the deadline due to the optimal choice of efforts. A failure in $t = 1$ tends to be less harmful than a failure in $t = 2$, since a successful completion of the research project is still possible after the former, but not after the latter event.

3.2 Delegated Research

We now turn to the analysis of our agency problems. We intend to identify the circumstances under which the incentives for efficient effort provision are (in)compatible with the incentives for truthfully revealing an early success. Remember that a success can only be reported if the project has indeed been successfully completed. If the agent reports an early success, he is rewarded with w_E. If the agent delays the report of a success until the second period, his reward

[1] The second order conditions are presented in the proof of Proposition 1 in Appendix A.

is w_L in the second period, the same as if he reports a success achieved in period 2. If the agent fails to report a success in both periods, he obtains w_N in the second period. The agent's ex ante participation constraint for $t = 1$ is therefore:

$$R_1 = e_1 \cdot (w_E - \delta \cdot w_N) - C(e_1) + (1 - e_1) \cdot \delta \cdot (e_2 \cdot (w_L - w_N) - C(e_2)) + \delta \cdot w_N \geq 0. \quad (7)$$

The agent acts rationally and anticipates his choices in the second period. If the project has not been completed in the first period, he chooses the effort in the second period so as to maximize his expected payoff in that period. In order to induce the efficient choice of effort in the second period, the principal needs to align the agent's interests with her own. Consequently, by using the first order approach and making use of (3), the incentive constraint for efficient effort provision in the second period reads:

$$w_L - w_N = Y_L - Y_N. \quad (8)$$

Denote with R_2 the expected rent (possibly) granted to the agent in period 2:

$$R_2 = e_2 \cdot (w_L - w_N) - C(e_2) + w_N, \quad (9)$$

with $e_2 = e_2^*$ if (8) holds.

In order to align interests in the first period, making use of (6), (8) and (9), the principal has to set the reward for a success reported in the first period as follows:

$$w_E = Y_E - \delta \cdot (V_2^* - R_2), \quad (10)$$

where $V_2^* - R_2 = Y_N - w_N$ if (8) holds, such that:

$$w_E = Y_E - \delta \cdot (Y_N - w_N). \quad (11)$$

In the absence of the strategic delay problem, the agent's choice of efforts coincides with the first-best solution if and only if (8) and (11) both hold. The principal can then satisfy the agent's ex ante participation constraint by choosing w_N to satisfy:

$$e_1^* \cdot w_E - C(e_1^*) + (1 - e_1^*) \cdot \delta \cdot (e_2^* \cdot w_L + (1 - e_2^*) \cdot w_N - C(e_2^*)) \geq 0. \quad (12)$$

No loss of efficiency is triggered by the delegation of the choice of efforts alone, as (8), (11) and (12) can be satisfied simultaneously.

However, the agent's discretion regarding the choice of whether and when to report a success imposes additional constraints. In the second period, truthful reporting is a by-product of incentive provision, as $w_L > w_N$. A problem may arise in the first period, as the agent may profit from a strategic delay of his report of a success. Truth-telling requires:

$$w_E \geq \delta \cdot w_L. \quad (13)$$

The conditions for efficient effort provision are compatible with truth-telling if and only if:

$$Y_E - \delta \cdot (Y_N - w_N) \geq \delta \cdot (Y_L - Y_N + w_N), \quad (14)$$

that is, if and only if:

$$Y_E \geq \delta \cdot Y_L. \quad (15)$$

If (15) is violated, the agent strategically withholds the report of a success in the first period. Anticipating the strategic delay, it is no longer (11) which guides the agent's effort choice in the first period, because he does not plan to cash in w_E, but $\delta \cdot w_L$ for an early success. We conclude:

Proposition 2 *If the principal delegates research to the agent, the effects of an unobservable project completion status are as follows:*

(i) If $Y_E \geq \delta \cdot Y_L$, the first best is implementable.

(ii) If $Y_E < \delta \cdot Y_L$, the first best is not implementable. If the principal naïvely offers the contract that is optimal in the absence of the strategic delay problem, the agent's effort choice is inefficiently high in the first period and he strategically delays the report of an early success.

If $Y_E \geq \delta \cdot Y_L$, the incentive to report truthfully is implied by the incentives that induce efficient effort provision. Strategic delay is neither a phenomenon nor a problem (in our otherwise agency-friendly model). The agent's working incentives can be set such that the agent effectively internalizes the principal's payoff consequences when choosing his effort levels. The principal can choose the agent's payoffs so as to maximize the surplus by satisfying (8) and (11), and she can extract the entire surplus by satisfying (12) with equality:

$$w_N = -\left(\frac{1}{\delta} \cdot (e_1^* \cdot Y_E - C(e_1^*)) + (1 - e_1^*) \cdot (e_2^* \cdot (Y_L - Y_N) - C(e_2^*))\right). \tag{16}$$

Strategic delay (possibly) occurs only if $\delta \cdot Y_L > Y_E$. In this case, the incentives for efficient effort provision are not compatible with the incentives to immediately report an early success. Then, strategic delay is a problem (because we assumed that $\delta \cdot Y_D < Y_E$), and, if unaccounted for, it is certainly also a phenomenon. If strategic delay on the part of the agent is deterred in the optimal contract (and is hence not a relevant phenomenon on the equilibrium path), this deterrence comes at the cost of distorting the incentives for effort provision.

The good news from our analysis so far is that for a range of plausible parameter constellations, strategic delay does not appear to be a problem. However, this finding should be interpreted cautiously given our choice of a particularly agency-friendly model framework (no risk aversion, no limits to liability). In our model, the principal can extract the entire surplus while aligning the agent's payoff consequences of his actions with her own, except for those of a delayed report. In a less agency-friendly setting, the problem of strategic delay will certainly interact with other sources of distortion.

In the next section, we show how the principal addresses the problem of strategic delay in her optimal contract with the agent.

3.3 Optimal Contracting Under Strategic Delay

In this section, we focus on the parameter constellation $Y_E < \delta \cdot Y_L$ in order to study optimal contracting when strategic delay is a problem. For the complementary case, the optimal contract is defined by the incentive-compatibility constraints (8) and (11) and the participation constraint (12).

The principal has four options: She could allow the agent to delay the report of an early success, deliberately violating the truth-telling constraint (13) by offering a *pooling contract*. With such a contract, the project's completion, if achieved, would be revealed in $t = 2$ in any case. She could ensure the agent truthfully reports a success in the first period by designing the contract as a *separating contract*. She could strategically *delay the start of the project* by offering a one-shot contract for $t = 2$. Lastly, she could *terminate* the contractual relationship *early* at the end of the first period. We will analyze these four options in turn.

We start by showing that a pooling contract can never be optimal.

Proposition 3 *A pooling contract is never optimal.*

Proof: Suppose the principal offers some pooling contract w_E, w_L, w_N, where (13) is violated such that w_E is too low to be relevant on the equilibrium path. The contract induces certain effort levels on the part of the agent, and the late reporting of an early success, in which case the principal would earn $\delta \cdot Y_D$. If the principal offered a wage $w_E = \delta \cdot w_L$ instead, the agent would be willing to report an early success in the first period, in which case the principal earned $Y_E > \delta \cdot Y_D$. As $\delta \cdot w_L$ is in any case the relevant incentive to provide effort in the first period, the agent's effort choices would not be affected by such a modification of the contract, nor would his incentive to participate in the contract. Hence, the proposed modification of the contract would leave the principal better off and the agent as well off as under the original (pooling) contract. It follows that the pooling contract is not optimal.□

Next, we turn to the optimal separating contract. We use the superscript S for the endogenous variables and characterize the optimal separating contract as follows:

Proposition 4 *Suppose $Y_E < \delta \cdot Y_L$. Compared to the first-best, the optimal separating contract induces effort levels such that $e_1^S > e_1^*$ and $e_2^S < e_2^*$.*

In Appendix A, we provide a complete derivation of the optimal separating contract. We demonstrate that it is not optimal for the principal to leave either the participation constraint or the truth-telling constraint slack. The binding truth-telling constraint $w_E^S = \delta \cdot w_L^S$ effectively leaves two choice variables for the principal, a payment in the case of a successful completion of the research project (appropriately discounted if it is reported early), and a (negative) payment in the case of an ultimate failure. The principal chooses the spread to guide the agent's incentives, and she uses the failure payment in order to satisfy the participation constraint. She therefore cannot target the agent's effort choices in both periods separately, which leads to the compromise in between the optimal levels, as stated in the proposition. It is interesting to note that our model predicts a *higher* probability of observing an early success in delegated research over the two periods than if the principal carries out the research on her own.

Next, we address the principal's third option, i.e. strategically delaying the start of the project. Not being exposed to the possibility of doing research in the first period, the agent cannot hide an early success, making the truth-telling constraint irrelevant. Due to the convex cost of effort, the agent has an interest in smoothing his effort. Thus, if the agent anticipates that a contract will be offered to him in the second period, effort-smoothing can only be prevented if the principal can effectively exclude the agent from the research technology. She has an interest to do so if Y_E is sufficiently small:

Proposition 5 *Consider $Y_E < \delta \cdot Y_L$. There is a threshold $\overline{Y}_E > \delta \cdot Y_L$ such that for $Y_E < \overline{Y}_E$, the principal prefers to strategically delay the start of the project until period 2. For $Y_E > \overline{Y}_E$, she prefers to offer the optimal separating contract.*

Proof: Remember Proposition 1(ii): For $Y_E = \delta \cdot V_2^*$ we have $e_1^* = 0$, such that for this parameter range, the proposition is in fact a straightforward corollary to Proposition 1(ii). Consider the case $Y_E = \delta \cdot V_2^* + \epsilon$, with $\epsilon > 0$. For ϵ close to zero, e_1^* is close to zero and so is the first period's contribution to the principal's ex ante expected payoff. The spread in the optimal separating contract $w_L^S - w_N^S = \Delta$ induces an effort level in the first period that is discretely higher than e_1^* and induces an effort level in the second period that is strictly below e_2^*, such that the principal's expected payoffs are lower than first best in both periods. In fact, the deviation of Δ from $Y_L - Y_N$, (the level that induces $e_2 = e_2^*$), is mainly due to preventing an excessively high effort level in period 1. Thus, for ϵ close to zero, the principal prefers to enforce $e_1 = 0$ without distorting the second period incentives. This proves the first part of the proposition.

At the other end of the relevant parameter range, i.e. for $Y_E = \delta \cdot Y_L - \epsilon$, ϵ close to zero, the spreads that induce the first-best levels of effort in period 1 and 2, respectively, differ only marginally. Respecting the truth-telling constraint induces only small distortions of the agent's incentives from their first-best levels in the optimal separating contract. Moreover, the principal's expected payoffs in both periods deviate only marginally from their first-best levels, and they contribute almost equally to her overall expected payoff. Hence, the principal strictly prefers the separating contract over a delay of a project start. The existence of a (unique) threshold follows from the fact that the principal's payoff is strictly increasing in Y_E in the separating contract, and is not affected by it when delaying the start of the project. □

Finally, we consider the principal's last option, i.e. concluding a contract which incentivizes effort only in period 1, which would render the strategic delay problem obsolete.

Proposition 6 *Consider $Y_E < \delta \cdot Y_L$. A contract that deters research in period 2 is dominated by a separating contract.*

Proof: If the principal does not incentivize effort in the second period, her payoff in the second period (conditional on a failure in the first period) is Y_N, the lowest possible value for V_2. Denote the payments to the agent in the first period conditional on a reported success and failure, respectively, with $\tilde{w}_E$ and $\tilde{w}_N$. The optimal contract in this class of contracts satisfies the agent's participation constraint with equality, and the optimal spread $\tilde{w}_E - \tilde{w}_N$ equals $Y_E - \delta \cdot Y_N$.

Suppose the principal expands this contract to include payments to the agent $\tilde{w}_L, \tilde{w}'_N$ for a success and a failure in the second period, respectively, as follows: $\tilde{w}_L - \tilde{w}'_N = \epsilon > 0$, $\epsilon < \min\{Y_E, \frac{\tilde{w}_E}{\delta}\}$, $\tilde{w}_L \leq \frac{\tilde{w}_E}{\delta}$ (such that the truth-telling constraint is satisfied), and $C'(\epsilon) \cdot \epsilon + \tilde{w}'_N = C(\epsilon)$ (such that the agent's expected rent in the second period is zero). It is easy to verify that such payments exist: Choose any $\epsilon < \min\{Y_E, \frac{\tilde{w}_E}{\delta}\}$ and set $\tilde{w}'_N = C(\epsilon) - C'(\epsilon) \cdot \epsilon$, and $\tilde{w}_L = \epsilon + \tilde{w}'_N$. As $\tilde{w}'_N < 0$ (due to the convexity of $C(\cdot)$), $\tilde{w}_L < \frac{\tilde{w}_E}{\delta}$.

These modifications to the original contract mean that neither the agent's participation constraint nor his working incentives in period 1 are affected. The agent provides a positive level of effort in period 2, $\tilde{e}_2 = \epsilon < Y_E < \delta \cdot (Y_L - Y_N) < Y_L - Y_N$, which gives rise to an expected payoff for the principal strictly between Y_N and V_2^* in period 2 (conditional on a failure in period 1).

The principal's expected payoff in the first period is the same as in the original contract, whereas the expected payoff in the second period is strictly higher. Thus, she is better off under the modified contract than under the original contract. □

To summarize the results of our analysis: The principal deals with the strategic delay problem either by offering a separating contract, or by strategically delaying the start of the project. The separating contract distorts the agent's efforts in both periods, and leads to a strictly higher probability of solving the research problem in the first period. A delay of the start of the project allows the first-best effort to be implemented in the second period, but the opportunity to find a solution in the first period is completely forgone.

The latter case is somewhat reminiscent of the work of Bhaskar (2014), who presents a long-term agency model in which the principal learns about the agent's ability over time and an inefficient zero-effort in early periods can be optimal.

4 Model extensions

In this section, we intend to explore the strategic delay problem in a richer institutional setting. We comment on the limits to our model, and we augment our model in several directions.

4.1 Limits to Enforceability

So far, we have assumed that the principal is able to enforce the ex ante optimal contract. This assumption is crucial for the viability of both solutions to the strategic delay problem, the separating contract and the strategic delay of the project start. The optimal separating contract induces an inefficiently low effort level in period 2, which gives rise to scope for renegotiation and challenges the principal's commitment power.

Likewise, whether the principal is able to delay the project start and effectively prevent early research depends on whether the agent can anticipate the contract to be concluded in the second period. If he can do that, he is better off smoothing the cost of effort, and he would already start the research (provided he has access to the appropriate research technology) in period 1. In such a case, a strategic delay of the project start is de facto impossible and the principal would be better off offering the optimal separating contract.

If the principal cannot commit not to renegotiate a separating contract upon a failure in period 1, she will optimally propose a renegotiation-proof contract in the first place, which implements first-best incentives for the second period.[2] This fact, the binding truth-telling constraint and the binding participation constraint jointly pin down the optimal separating contract, which induces an even higher effort level in period 1 than the optimal separating contract in the case with commitment. This distortion further decreases the principal's ex ante expected payoff and makes the separating contract less attractive. As a consequence, the parameter range for which a strategic delay of the project start is optimal increases. An analogous reasoning to the proof of Proposition 5 implies that there remains a parameter range for which the optimal solution to the strategic delay problem is a separating contract.

4.2 Multiple Agents

In many real world settings it is plausible to assume that there will be more than one agent working to achieve a particular research goal. In this section, we analyze whether the principal can profit from inducing competition among multiple agents. Making any agent's payment contingent on that agent's individual report *and* also on the report of other agents seems like a promising way to mitigate, or sidestep altogether, the problem of strategic delay. In this section, we outline the consequences of introducing a second agent, such that in total there are two agents, named A and B. We refer to Appendix B for a detailed derivation of the results presented here.

In what follows, we stipulate that the parameter constellation is such that for both agents the respective effort levels yield interior solutions which are period-wise symmetric, i.e. $e_t^A = e_t^B$ for $t = 1, 2$. Let $P(Y_E) = (1 - (1 - e_1^A) \cdot (1 - e_1^B))$ denote the probability of an early success. Likewise, we define $P(Y_L) = (1 - (1 - e_2^A) \cdot (1 - e_2^B))$. The principal's maximization problem now reads as:

$$P(Y_E) \cdot Y_E - 2 \cdot C(e_1^A) + (1 - P(Y_E)) \cdot \delta \cdot (Y_N + P(Y_L) \cdot (Y_L - Y_N)) - 2 \cdot C(e_2^A)). \qquad (17)$$

[2]For an in-depth analysis of the principal's ability to commit to long-term contracts and its role in project completion, see Toxvaerd (2006).

Due to the symmetry, it suffices to express the solution to the principal's problem in terms of the effort levels demanded from agent A (analogous conditions with the superscripts reversed apply to agent B's effort levels).

Optimal effort levels are implicitly defined by:

$$C'(e_1^A) = (1 - e_1^B) \cdot (Y_E - \delta \cdot V_2^{**}), \tag{18}$$

$$C'(e_2^A) = (1 - e_2^B) \cdot (Y_L - Y_N), \tag{19}$$

where $V_2^{**} \geq V_2^*$, because the expected profit attainable with two agents is generally higher than with one agent. Due to the introduction of a second agent, the contracting space has become larger. In particular, we can condition any agent's compensation level on the reports obtained from both agents. Therefore, we analyze a scenario in which if only one agent reports a failure after period 1, that agent has to pay a penalty w_F. Apart from that, all other payments remain defined as in the single-agent case.

Then, agent A chooses e_1^A so as to maximize:

$$e_1^A \cdot w_E + (1 - e_1^A) \cdot (e_1^B \cdot w_F + (1 - e_1^B) \cdot \delta \cdot R_2) - C(e_1^A). \tag{20}$$

The adjusted truth-telling constraint reads as:

$$w_E \geq e_1^B \cdot w_F + (1 - e_1^B) \cdot \delta \cdot w_L. \tag{21}$$

In Appendix B we show that condition (21) can be rearranged to give:

$$Y_E \geq \delta \cdot (Y_L - C(e_2^A)). \tag{22}$$

Comparing (22) with (15), we observe that the parameter range for which truth-telling is incentive-compatible with first-best incentives has increased. Interestingly, it is not the punishment but the presence of the second agent alone that causes this parameter shift.[3] In fact, w_F can be chosen arbitrarily. In order to guide first period incentives for effort provision effectively, w_E has to be correspondingly increased if we impose a punishment. As a consequence, the punishment is canceled out when it comes to truth-telling.

The fact that strategic delay does indeed help reduce the parameter range (as compared to the single-agent case) for which strategic delay is relevant follows from the interplay of both agents' incentive compatibility constraints that make the agent on the margin the residual claimant to the social consequences of his actions. If an agent reports truthfully in the first period, the principal gets Y_E instead of entering the next period. If instead she enters the next period, the other agent will be present as well. Taking that as a given, the prospective earning is only $Y_L - C(e_2^A)$ (instead of Y_L in the single-agent case), which causes the truth-telling condition to be adjusted.

4.3 Replacement of Agents

If strategic delay poses a problem, i.e. it reduces the principal's expected gain compared to the first-best benchmark, a further obvious solution to the problem would be to conclude two separate short-term contracts with two distinct agents. The first-best effort levels would be implementable for the principal in two separate one-shot contracts, one contract for agent 1 in

[3] It is beyond the scope of this work to fully analyze the complete set of possible payments as a function of the agents' reports. Still, our result is indicative that punishments are not conducive to mitigating the problem of strategic delay in settings with multiple agents.

$t = 1$ and a separate one for agent 2 in $t = 2$, which is offered to a second agent only in case of a failure in the first period.[4] Truthful reporting of a project success would then not pose a problem in either period, since in a one-shot contract, truth-telling is a by-product of efficient effort provision. However, one obvious objection that could be raised here is that a series of short-term contracts might cause extra costs compared to one single long-term contract, e.g. due to substantial transaction costs associated with searching for and hiring agents.[5] Secondly, it can be assumed that non-negligible costs are also associated with the (initial) training of the new agent. Formally, this means that $C(0) > 0$ in $t = 1$ and also in $t = 2$, provided that the agent is replaced after the first period. These additional costs might outweigh the loss caused by the information asymmetry which only accrues in the single-agent case.

4.4 Monitoring

A further coping strategy from the principal's perspective is to actively reduce the information asymmetry between herself and the agent. Assume that the principal can invest some amount k in costly monitoring, such that the actual project completion status after period 1 becomes observable to her with certainty (and verifiable to any third party).[6] The contract could then include a punishment, i.e. a payment $P < 0$, if the agent's strategic delay is discovered (a less formal means of punishment would be to expose his misconduct to the scientific community). If the principal can commit to engage in costly verification of a reported failure in period 1 with some probability q, the first-best solution would (almost) be achievable for the principal, as long as there is no upper limit to the amount of the punishment. This is the case because the principal will choose q at just high enough a level to satisfy the (modified) truth-telling constraint $w_E \geq \delta \cdot w_L + q \cdot P$. The harsher the punishment becomes, the closer to zero q can be set. Hence, the principal has (virtually) no monitoring costs, while the agent is still willing to accept the contract. On the equilibrium path, he will never choose not to report an early success and is therefore never punished. Potential upper limits to the amount of the punishment, the principal's inability to commit and the difficulty to prove an unreported success in court constrain this idea's viability in practice. If the principal cannot commit to an ex ante probability of monitoring, the possibility of monitoring and punishing gives rise to an inspection game with an equilibrium in mixed strategies (see Verbeck (2018)).

4.5 Plurality of Research Methods

In many research settings a specific research goal (the development of a new vaccine, say) can be achieved using one of a number of methods or technologies. The choice of technology, then, could also be harnessed to prevent strategic delay if we assume that the technology used to find a solution to the research problem can be inferred from the solution, and that this information is contractible and observable to parties other than the agent. Then, the choice of research technology could be made an explicit part of the contract. Suppose, for example, that in $t = 1$ research is to be carried out with technology a and, upon failure, with research technology b in $t = 2$. Hence, early completion with technology a cannot be incorrectly presented as late completion if the contract specifies that technology b should have been used in the second period and that the incorrect choice of technology can be deterred by inflicting sufficiently severe punishments on the agents. If technology b is potentially less promising or powerful than

[4] We need to exclude the possibility of side-contracting between agent 1 and 2.

[5] This argument is related to Coase's rationale for the superiority of hierarchies over markets (Coase (1937)).

[6] Related ideas can be found in the literature on costly state verification, e.g. Townsend (1979). There is also a section on monitored search in Lewis and Ottaviani (2008).

technology a, the principal faces a trade-off between the advantage of not being confronted with the problem of strategic delay and the disadvantage of having the research performed with an inferior technology (see also Verbeck and Schulte (2016) where we present an in-depth study of the problem of technology choice).

5 Discussion and Conclusion

We have provided an (admittedly stylized and parsimonious) agency model that analyzes the tension between providing working incentives on the one hand and incentives to truthfully reveal information on the other in the context of R&D projects. We have identified conditions where this tension actually exists. If the contract offered does not take into account the possibility that the agent might withhold information on an early success despite being incentivized to do so are induced by a naïvely concluded contract, the agent will "overinvest but underreport" in $t = 1$, i.e. he will put in more effort than the principal actually desires but delay the disclosure of an early success. The principal's rational adjustment of the contract hinges on her possibility to effectively prevent the agent from conducting research in the first period. If this is not feasible, a separating contract that disincentivizes strategic delay but comes at the cost of distorted effort levels is the principal's best option. If a strategic delay of the project start is feasible, the principal will prefer this option if her payoff from an early success is sufficiently low.

The good news from our analysis is that parameter regions exist where truthful reporting is a by-product of incentive provision, in which case strategic delay is neither a problem nor a phenomenon. If the principal does not suffer from strategic delay (i.e. if our Assumption A2 does not apply), strategic delay will be a phenomenon if and only if it is not a problem. These results should however be interpreted with caution, as we study a setting with otherwise ideal contracting conditions (e.g. no risk aversion, no constraints on payments). If the principal cannot fully extract the surplus due to other contractual frictions, the problem of strategic delay is likely to interact with the distortions from such frictions.

Possible extensions and avenues for further research are an analysis of the agent's risk preferences or consideration of more than only two periods. Furthermore, specifications of the researcher's yield and cost functions could be analyzed. In our current setting, we do not consider learning effects, such that the effort in $t = 1$ has no effect at all on the probability of generating a success in $t = 2$ (or on the cost of conducting research in that period). While research efforts that lead to an initial failure can often be considered sunk costs, a more flexible representation has the potential to enrich our analysis.

References

Bergemann, D. and U. Hege (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance 22*(6), 703–735.

Bergemann, D. and U. Hege (2005). The financing of innovation: Learning and stopping. *The RAND Journal of Economics 36*(4), 719–752.

Bhaskar, V. (2014). The ratchet effect: A learning perspective. *Working Paper*.

Biais, B., T. Mariotti, G. Plantin, and J.-C. Rochet (2007). Dynamic security design: Convergence to continuous time and asset pricing implications. *The Review of Economic Studies 74*(2), 345–390.

Biais, B., T. Mariotti, J.-C. Rochet, and S. Villeneuve (2010). Large risks, limited liability, and dynamic moral hazard. *Econometrica 78*(1), 73–118.

Campbell, A., F. Ederer, and J. Spinnewijn (2014). Delay and deadlines: Freeriding and information revelation in partnerships. *American Economic Journal: Microeconomics 6*(2), 163–204.

Coase, R. (1937). The nature of the firm. *Economica 4*(16), 386–405.

Green, B. and C. Taylor (2016). Breakthroughs, deadlines, and self-reported progress: Contracting for multistage projects. *American Economic Review 106*(12), 3660–3699.

Herweg, F. and D. Müller (2006). Performance of procrastinators: On the value of deadlines. *Theory and Decision 70*(3), 329–366.

Hoffmann, F. and S. Pfeil (2010). Reward for luck in a dynamic agency model. *Review of Financial Studies 23*(9), 3329–3345.

Holmström, B. (1979). Moral hazard and observability. *Bell Journal of Economics 10*(2), 324–340.

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

Hörner, J. and L. Samuelson (2013). Incentives for experimenting agents. *The RAND Journal of Economics 44*(4), 632–663.

Katolnik, S. and J. Schöndube (2019). Don't kill the goose that lays the golden eggs: Strategic delay in project completion. *Working Paper*.

Lambert, R. (1983). Long-term contracts and moral hazard. *Bell Journal of Economics 14*(2), 441–452.

Lewis, T. (2012). A theory of delegated search for the best alternative. *The RAND Journal of Economics 43*(3), 391–416.

Lewis, T. and M. Ottaviani (2008). Search agency. *Working Paper*.

Malcomson, J. and F. Spinnewyn (1988). The multiperiod principal-agent problem. *The Review of Economic Studies 55*(3), 391–407.

Manso, G. (2011). Motivating innovation. *Journal of Finance 66*(5), 18231860.

Mason, R. and J. Välimäki (2015). Getting it done: Dynamic incentives to complete a project. *Journal of the European Economic Association 13*(1), 62–97.

O'Donoghue, T. and M. Rabin (1999). Incentives for procrastinators. *The Quarterly Journal of Economics 114*(3), 769–816.

Parkinson, C. (1957). *Parkinson's Law, and Other Studies in Administration.* Random House.

Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *The Review of Economic Studies 75*(3), 957–984.

Townsend, R. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory 21*(2), 265–293.

Toxvaerd, F. (2006). Time of the essence. *Journal of Economic Theory 129*(1), 252–272.

Ulbricht, R. (2016). Optimal delegated search with adverse selection and moral hazard. *Theoretical Economics 11*(1), 253–278.

Verbeck, M. (2018). The inspection game in science. *MACIE Working Paper Series*.

Verbeck, M. and E. Schulte (2016). Contracting with researchers. *MAGKS Working Paper Series*.

Williams, N. (2010). A solvable continuous time dynamic principal-agent model. *Journal of Economic Theory 159*(B), 989–1015.

Appendices

A Proofs

Proof of Proposition 1

The candidates for the optimum have been derived in the text. It remains to be shown that the second order conditions are satisfied. We have:

$$\begin{aligned}\frac{\partial^2 V_1}{\partial e_1^2} &= -C''(e_1)\\ \frac{\partial^2 V_1}{\partial e_2^2} &= -(1-e_1)\cdot\delta\cdot C''(e_2)\\ \frac{\partial^2 V_1}{\partial e_1 \partial e_2} &= -\delta\cdot(Y_L - C'(e_2))\end{aligned}$$

$\frac{\partial^2 V_1}{\partial e_1^2}$ and $\frac{\partial^2 V_1}{\partial e_2^2}$ are strictly negative, as $C(\cdot)$ is convex. $\frac{\partial^2 V_1}{\partial e_1 \partial e_2}$ is zero when e_2 satisfies the first order condition, such that $Y_L = C'(e_2)$. Thus, $\frac{\partial^2 V_1}{\partial e_1^2}\cdot\frac{\partial^2 V_1}{\partial e_2^2} - \left(\frac{\partial^2 V_1}{\partial e_1 \partial e_2}\right)^2 > 0$, and the conditions referred to in Proposition 1 do indeed characterize the maximum of the principal's objective function. □

Proof of Proposition 4

We assume that $w_L^S > w_N^S$ such that $e_2^S > 0$ in any separating contract. The complementary case is discussed (and ruled out) in Proposition 6. For the sake of readability, we omit the superscript "S" in the following, as there is no scope for ambiguity.

If the agent accepts a contract (w_E, w_L, w_N) that satisfies the truth-telling constraint, he chooses his effort levels e_1, e_2 to maximize:

$$e_1 \cdot w_E - C(e_1) + (1-e_1)\cdot\delta\cdot(e_2\cdot(w_L - w_N) - C(e_2) + w_N). \tag{23}$$

The agent's effort in the second period is defined by:

$$e_2 = C'^{-1}(w_L - w_N), \tag{24}$$

with $\frac{\partial e_2}{\partial w_L} = -\frac{\partial e_2}{\partial w_N} = \frac{1}{C''(w_L - w_N)} > 0$.

Anticipating a truthful report of an early success, effort in the first period satisfies:

$$e_1 = C'^{-1}\left(w_E - \delta \cdot (w_N + e_2 \cdot (w_L - w_N) - C(e_2))\right), \tag{25}$$

if the argument of $C'^{-1}(\cdot)$ in the above equation is positive. As $w_E \geq \delta \cdot w_L$, this is the case for all $w_L \geq w_N$. Using (24):

$$e_1 = C'^{-1}\left(w_E - \delta \cdot (w_N + (w_L - w_N) \cdot C'^{-1}(w_L - w_N) - C(C'^{-1}(w_L - w_N)))\right), \tag{26}$$

with $\frac{\partial e_1}{\partial w_E} = \frac{1}{C''(w_E - \delta \cdot (w_N + (w_L - w_N) \cdot C'^{-1}(w_L - w_N) - C(C'^{-1}(w_L - w_N))))} > 0$, $\frac{\partial e_1}{\partial w_L} = -\delta \cdot C'^{-1}(w_L - w_N) \cdot \frac{\partial e_1}{\partial w_E} = -\delta \cdot e_2 \cdot \frac{\partial e_1}{\partial w_E} < 0$ and $\frac{\partial e_1}{\partial w_N} = -\delta \cdot (1 - C'^{-1}(w_L - w_N)) \cdot \frac{\partial e_1}{\partial w_E} = -\delta \cdot (1 - e_2) \cdot \frac{\partial e_1}{\partial w_E} < 0$.

The principal chooses (w_E, w_L, w_N) so as to maximize her expected payoff:

$$e_1 \cdot (Y_E - w_E) + (1 - e_1) \cdot \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N))) \tag{27}$$

subject to the incentive-compatibility constraints (24), (26), the agent's participation constraint (28) and the truth-telling constraint (29):

$$e_1 \cdot w_E - C(e_1) + \delta \cdot (1 - e_1) \cdot (w_N + e_2 \cdot (w_L - w_N) - C(e_2)) \geq 0, \tag{28}$$

$$w_E \geq \delta \cdot w_L. \tag{29}$$

Thus, we have two binding constraints and two weak inequalities to satisfy. We use the Karush-Kuhn-Tucker conditions in order to characterize the principal's optimal choice. We seek to minimize:

$$-e_1 \cdot (Y_E - w_E) - (1 - e_1) \cdot \delta \cdot (Y_N + e_2 \cdot Y_L - (w_N + e_2 \cdot (w_L - w_N))) \tag{30}$$

satisfying:

$$\begin{aligned}
e_2 - C'^{-1}(w_L - w_N) &= 0 && (31)\\
e_1 - C'^{-1}\left(w_E - \delta \cdot (w_N + e_2 \cdot (w_L - w_N) - C(e_2))\right) &= 0 && (32)\\
-e_1 \cdot w_E + C(e_1) - (1 - e_1) \cdot \delta \cdot (w_N + e_2 \cdot (w_L - w_N) - C(e_2)) &\leq 0 && (33)\\
-w_E + \delta \cdot w_L &\leq 0 && (34)\\
\mu_1, \mu_2 &\geq 0 && (35)\\
\mu_1 \cdot (e_1 \cdot w_E + C(e_1) - (1 - e_1) \cdot \delta \cdot (w_N + e_2 \cdot (w_L - w_N) - C(e_2))) &= 0 && (36)\\
\mu_2 \cdot (-w_E + \delta \cdot w_L) &= 0 && (37)
\end{aligned}$$

Conditions (31)-(34) are the primary feasibility constraints, (35) is needed for dual feasibility and (36), (37) are the conditions for complementary slackness. We use the multipliers λ_1, λ_2 for constraints (31), (32) and μ_1, μ_2 for (33), (34), respectively.

(36) and (37) allow for four cases:

1. $\mu_1 = \mu_2 = 0$,

2. $\mu_1 = 0, \mu_2 \neq 0$, (34) is binding,

3. $\mu_1 \neq 0, \mu_2 = 0$, (33) is binding,

4. (33), (34) are both binding.

Case 1 Suppose the solution satisfies the conditions for Case 1 (both inequalities are slack). Then, it needs to satisfy the following Karush-Kuhn-Tucker optimality conditions:

$$\begin{aligned} e_1 + \frac{\partial e_1}{\partial w_E} \cdot (-Y_E + w_E + \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) & \\ + \lambda_2 \cdot \underbrace{\left(\frac{\partial e_1}{\partial w_E} - \frac{1}{C''\left(w_E - \delta \cdot (w_N + e_2 \cdot (w_L - w_N) - C(e_2))\right)} \right)}_{=0} & = 0, \quad (38) \end{aligned}$$

$$\begin{aligned} \underbrace{\delta \cdot e_2 \cdot \frac{\partial e_1}{\partial w_E}}_{=-\frac{\partial e_1}{\partial w_L}} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) & \\ -(1-e_1) \cdot \delta \left(-e_2 + \frac{\partial e_2}{\partial w_L} \cdot (Y_L - Y_N - (w_L - w_N)) \right) & \\ + \lambda_1 \cdot \underbrace{\left(\frac{\partial e_2}{\partial w_L} - \frac{1}{C''(w_L - w_N)} \right)}_{=0} & \\ + \lambda_2 \cdot \underbrace{\left(-\delta \cdot e_2 \cdot \frac{\partial e_1}{\partial w_E} + \frac{\delta \cdot \left(e_2 + \frac{\partial e_2}{\partial w_L} \cdot \underbrace{((w_L - w_N) - C'(e_2))}_{=0} \right)}{C''\left(w_E - \delta \cdot (w_N + e_2 \cdot (w_L - w_N) - C(e_2))\right)} \right)}_{=0} & = 0, \quad (39) \end{aligned}$$

$$\begin{aligned} \underbrace{\delta \cdot (1-e_2) \cdot \frac{\partial e_1}{\partial w_E}}_{=-\frac{\partial e_1}{\partial w_N}} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) & \\ -(1-e_1) \cdot \delta \cdot \left(-(1-e_2) - \underbrace{\frac{\partial e_2}{\partial w_L}}_{=-\frac{\partial e_2}{\partial w_N}} \cdot (Y_L - Y_N - (w_L - w_N)) \right) & \\ + \lambda_1 \cdot \underbrace{\left(-\frac{\partial e_2}{\partial w_L} + \frac{1}{C''(w_L - w_N)} \right)}_{=0} & \\ + \lambda_2 \cdot \left(-\delta \cdot (1-e_2) \cdot \frac{\partial e_1}{\partial w_E} + \frac{\delta \cdot \left(1 - e_2 + \frac{\partial e_2}{\partial w_N} \cdot \underbrace{((w_L - w_N) - C'(e_2))}_{=0} \right)}{C''\left(w_E - \delta \cdot (w_N + e_2 \cdot (w_L - w_N) - C(e_2))\right)} \right) & \end{aligned}$$

$$= 0. \quad (40)$$

The program can be simplified to:

$$e_1 - \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) = 0, \quad (41)$$

$$\delta \cdot e_2 \cdot \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) \\ -(1 - e_1) \cdot \delta \left(-e_2 + \frac{\partial e_2}{\partial w_L} \cdot (Y_L - Y_N - (w_L - w_N)) \right) = 0, \quad (42)$$

$$\delta \cdot (1 - e_2) \cdot \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) \\ -(1 - e_1) \cdot \delta \cdot \left(-(1 - e_2) - \frac{\partial e_2}{\partial w_L} \cdot (Y_L - Y_N - (w_L - w_N)) \right) = 0. \quad (43)$$

Summing (42) and (43) yields:

$$\delta \cdot \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) \\ -(w_N + e_2 \cdot (w_L - w_N)))) + (1 - e_1) \cdot \delta = 0. \quad (44)$$

Using (41):

$$\delta \cdot e_1 + \delta \cdot (1 - e_1) = 0, \quad (45)$$

which can only be satisfied for the (uninteresting) case that $\delta = 0$. We can thus rule out Case 1.

Case 2 Suppose the solution satisfies the conditions for Case 2. Then it needs to satisfy the following Karush-Kuhn-Tucker optimality conditions:

$$e_1 - \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) - \mu_2 = 0, \quad (46)$$

$$\delta \cdot e_2 \cdot \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) \\ -\delta \cdot (1 - e_1) \cdot \left(-e_2 + \frac{\partial e_2}{\partial w_L} \cdot (Y_L - Y_N - (w_L - w_N)) \right) + \delta \cdot \mu_2 = 0, \quad (47)$$

$$\delta \cdot (1 - e_2) \cdot \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (e_2 \cdot Y_L - Y_N - (w_N + e_2 \cdot (w_L - w_N)))) \\ -(1 - e_1) \cdot \delta \left(-(1 - e_2) - \frac{\partial e_2}{\partial w_L} \cdot (Y_L - Y_N - (w_L - w_N)) \right) = 0. \quad (48)$$

We multiply (46) by δ, sum all three equations and once more obtain:

$$\delta \cdot e_1 + \delta \cdot (1 - e_1) = 0, \quad (49)$$

which means we can rule out Case 2.

Case 3 Suppose the solution satisfies the conditions for Case 3. This is the case when the participation constraint binds: $-e_1 \cdot w_E + C(e_1) - \delta \cdot (1-e_1) \cdot (w_N + e_2 \cdot (w_L - w_N) - C(e_2)) = 0$, and the truth-telling constraint is slack.

Then, it needs to satisfy the following Karush-Kuhn-Tucker optimality conditions:

$$\begin{aligned} & e_1 - \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) \\ & + \mu_1 \cdot \left(-e_1 + \frac{\partial e_1}{\partial w_E} \cdot \underbrace{(-w_E + C'(e_1) + \delta \cdot (w_N + e_2 \cdot (w_L - w_N) - C(e_2)))}_{=0} \right) = 0, \end{aligned} \tag{50}$$

$$\begin{aligned} & \delta \cdot e_2 \cdot \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) \\ & \qquad - (1-e_1) \cdot \delta \cdot \left(-e_2 + \frac{\partial e_2}{\partial w_L} \cdot (Y_L - Y_N - (w_L - w_N)) \right) \\ & \qquad + \mu_1 \cdot (-(1-e_1) \cdot \delta \cdot e_2) \\ & + \underbrace{\mu_1 \cdot \left(-\delta \cdot e_2 \cdot \frac{\partial e_1}{\partial w_E} \cdot (-w_E + C'(e_1) + \delta \cdot (w_N + e_2 \cdot (w_L - w_N) - C(e_2))) \right)}_{=0} \\ & \qquad + \underbrace{\mu_1 \cdot \left(\frac{\partial e_2}{\partial w_L} \cdot (-(1-e_1) \cdot \delta) \cdot \underbrace{(w_L - w_N - C'(e_2))}_{=0} \right)}_{=0} = 0, \end{aligned} \tag{51}$$

$$\begin{aligned} & \delta \cdot (1-e_2) \cdot \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) \\ & \qquad - (1-e_1) \cdot \delta \cdot \left(-(1-e_2) - \frac{\partial e_2}{\partial w_L} \cdot (Y_L - Y_N - (w_L - w_N)) \right) \\ & \qquad + \mu_1 \cdot (-(1-e_1) \cdot \delta \cdot (1-e_2)) \\ & \qquad + \underbrace{\mu_1 \cdot \left(-\delta \cdot (1-e_2) \cdot \frac{\partial e_1}{\partial w_E} \cdot (\ldots) \right)}_{=0} \\ & \qquad + \underbrace{\mu_1 \cdot \left(\frac{\partial e_2}{\partial w_N} \cdot (-(1-e_1) \cdot \delta) \cdot \underbrace{(w_L - w_N - C'(e_2))}_{=0} \right)}_{=0} \\ & \qquad = 0. \end{aligned} \tag{52}$$

The program can be simplified to:

$$\begin{aligned} & e_1 - \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N)))) \\ & \qquad - e_1 \cdot \mu_1 = 0, \end{aligned} \tag{53}$$

$$\delta \cdot e_2 \cdot \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N))))$$
$$-(1-e_1) \cdot \delta \cdot \left(-e_2 + \frac{\partial e_2}{\partial w_L} \cdot (Y_L - Y_N - (w_L - w_N)) + e_2 \cdot \mu_1 \right) = 0, \quad (54)$$
$$\delta \cdot (1-e_2) \cdot \frac{\partial e_1}{\partial w_E} \cdot (Y_E - w_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N))))$$
$$-(1-e_1) \cdot \delta \cdot \left(-(1-e_2) - \frac{\partial e_2}{\partial w_L} \cdot (Y_L - Y_N - (w_L - w_N)) + (1-e_2) \cdot \mu_1 \right) = 0. \quad (55)$$

Again, we multiply (53) by δ and sum all three equations:

$$\begin{aligned} \delta \cdot e_1 \cdot (1-\mu_1) - \delta \cdot (1-e_1) \cdot (-1+\mu_1) &= 0 \\ \Leftrightarrow \delta \cdot (1-\mu_1) &= 0 \\ \Leftrightarrow \mu_1 &= 1. \end{aligned} \quad (56)$$

Plugging $\mu_1 = 1$ into (53) yields:

$$w_E = Y_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N))), \quad (57)$$

which can be used to derive from (54) or (55):

$$w_L - w_N = Y_L - Y_N, \quad (58)$$

such that:

$$w_E = Y_E + \delta \cdot w_N. \quad (59)$$

The truth-telling constraint requires that:

$$\begin{aligned} & w_E \geq \delta \cdot w_L \\ \Leftrightarrow \quad & Y_E - \delta \cdot (Y_N - w_N) \geq \delta \cdot (Y_L - Y_N) + \delta \cdot w_N, \end{aligned} \quad (60)$$

which is violated in the parameter constellation under consideration.

Case 4 We could rule out all cases but Case 4, where the participation constraint and the truth-telling constraints are binding. When the truth-telling constraint binds, $w_E = \delta \cdot w_L$, we effectively have only two variables to choose and we write the objective function to be minimized as:

$$-e_1 \cdot (Y_E - \delta \cdot w_L) - (1-e_1) \cdot \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - (w_N + e_2 \cdot (w_L - w_N))). \quad (61)$$

We can write the (binding) participation constraint as:

$$-e_1 \cdot \delta \cdot (w_L - w_N) + C(e_1) - \delta \cdot w_N - (1-e_1) \cdot \delta \cdot (e_2 \cdot (w_L - w_N) - C(e_2)) = 0. \quad (62)$$

The binding participation constraint allows us to restate (61) as follows:

$$-e_1 \cdot Y_E - (1-e_1) \cdot \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N)) + C(e_1) + (1-e_1) \cdot \delta \cdot C(e_2). \quad (63)$$

We denote $\Delta = w_L - w_N$. The agent's optimization problem can be expressed as follows:

$$e_1 \cdot \delta \cdot \Delta - C(e_1) + (1-e_1) \cdot \delta \cdot (e_2 \cdot \Delta - C(e_2)) + \delta \cdot w_N. \quad (64)$$

such that the optimal effort levels are defined by:

$$e_2 = C'^{-1}(\Delta), \quad (65)$$
$$e_1 = C'^{-1}(\delta \cdot (\Delta \cdot (1-e_2) + C(e_2))). \quad (66)$$

We have $e_1 \leq e_2$ if and only if:

$$\Delta \geq \delta \cdot (\Delta \cdot (1-e_2) + C(e_2)) \quad (67)$$
$$\Leftrightarrow \delta \leq \frac{\Delta}{\Delta \cdot (1-e_2) + C(e_2)}. \quad (68)$$

As $\delta \leq 1$, the condition above has no bite if

$$\Delta \geq \Delta \cdot (1-e_2) + C(e_2) \quad (69)$$
$$\Leftrightarrow \Delta \geq \frac{C(e_2)}{e_2} \quad (70)$$
$$\Leftrightarrow C'(e_2) \geq \frac{C(e_2)}{e_2}. \quad (71)$$

The above inequality is strictly satisfied due to the convexity of $C(\cdot)$. We conclude that $e_1 < e_2$ for all $\Delta > 0$.

Moreover,

$$\frac{\partial e_2}{\partial \Delta} = \frac{1}{C''(\Delta)} > 0, \quad (72)$$
$$\frac{\partial e_1}{\partial \Delta} = \frac{\delta \cdot (1-e_2)}{C''(\delta \cdot (\Delta \cdot (1-e_2) + C(e_2)))} > 0. \quad (73)$$

The principal's optimal choice of Δ is characterized by:

$$\frac{\partial e_1}{\partial \Delta} \cdot \big(-Y_E + C'(e_1) + \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - C(e_2))\big) + \frac{\partial e_2}{\partial \Delta} \cdot \delta \cdot (1-e_1) \cdot \big(-(Y_L - Y_N) + C'(e_2)\big) = 0. \quad (74)$$

The first summand is zero if $C'(e_1) = Y_E - \delta \cdot (Y_N + e_2 \cdot (Y_L - Y_N) - C(e_2))$, which requires $\Delta = \Delta_1^* = \frac{Y_E}{\delta \cdot (1-e_2)} - Y_N - \frac{Y_L - e_2}{1-e_2}$. The second summand is zero if $C'(e_2) = Y_L - Y_N$, which requires $\Delta = \Delta_2^* = Y_L - Y_N$. As she has only one instrument, Δ, at her disposal that affects the agent's effort choices in both periods simultaneously, these effort levels can in general not be implemented. Δ_1^* is smaller than Δ_2^* as $Y_E < \delta \cdot Y_L$.

Using (65) and (66), (74) can be expressed as follows:

$$\frac{\partial e_1}{\partial \Delta} \big(-Y_E + \delta \cdot Y_N + \delta \cdot \Delta + \delta \cdot e_2^S \cdot (Y_L - Y_N - \Delta)\big) + \frac{\partial e_2^S}{\partial \Delta} \cdot \delta \cdot (1-e_1^S) \cdot (-(Y_L - Y_N) + \Delta) = 0. \quad (75)$$

The LHS of (75) strictly increases in Δ. Suppose the principal sets $\Delta = \Delta_2^* = Y_L - Y_N$, the spread that induces first best effort in the second period. If $Y_E < \delta \cdot Y_L, \Delta^S < Y_L - Y_N$. This implies that $\Delta^S < \Delta_2^* = Y_L - Y_N$ in the optimum, such that $e_2^S < e_2^*$ and $e_1^S > e_1^*$ and deviations from the first-best are inevitable. □

B Multiple Agents - Full Exposition

In the following, we provide a detailed exposition of the results presented in Section 4.2. The principal's maximization problem equals:

$$(1-(1-e_1^A)\cdot(1-e_1^B))\cdot Y_E - C(e_1^A) - C(e_1^B) \quad (76)$$
$$+(1-e_1^A)\cdot(1-e_1^B)\cdot\delta\cdot(Y_N+(1-(1-e_2^A)\cdot(1-e_2^B))\cdot(Y_L-Y_N)-C(e_2^A)-C(e_2^B))$$

which can be rewritten as (17). First, we will show by means of an example that parameter constellations exist that satisfy our assumptions and yield a symmetric solution to the principal's problem. Therefore, we suppose that our cost function is defined by $C(e_t) = {e_t}^2$ and show that the desired properties are fulfilled for this particular case.

Given its structure, it is permissible to treat the problem as two separate bivariate optimization problems with decision variables e_t^A and e_t^B for any t. We start with $t = 2$. From the first-order condition, it follows that the optimal effort levels for agent A and B, respectively, are defined by

$$(1-e_2^B)\cdot(Y_L-Y_N) = 2\cdot e_2^A. \quad (77)$$

and:

$$(1-e_2^A)\cdot(Y_L-Y_N) = 2\cdot e_2^B. \quad (78)$$

Because $C'(1) > Y_L - Y_N$, the optimal e_2^A (conditional on e_2^B) is below 1 for $e_2^B = 0$ and equal to 0 for $e_2^B = 1$ and vice versa. Hence, there must be a unique intersection at $e_2^A = e_2^B = \frac{(Y_L-Y_N)}{(Y_L-Y_N+2)}$ which is the unique and symmetric candidate for an optimum.

Given the symmetry of efforts, the problem becomes in fact an univariate optimization problem and the sufficient conditions for an optimum to exist are

$$-C''(e_2^A) < 0 \quad (79)$$

and

$$C''(e_2^A)^2 - (Y_L-Y_N)^2 > 0. \quad (80)$$

The fulfillment of both conditions can be readily verified.

For $t = 1$, the rationale for interior, unique and symmetric solutions is exactly the same as for $t = 2$, with the sole exception that "$Y_L - Y_N$" has to be replaced by "$Y_E - \delta\cdot V_2^{**}$", where V_2^{**} refers to the maximized surplus in the second period when employing two agents.

Like in the case with only a single agent, we assume that $Y_E - \delta\cdot V_2^{**} > 0$ must hold, a condition that is more restrictive than in the single-agent case, because $V_2^{**} > V_2^*$.

Having verified that the case in which the principal wants to employ two agents with identical effort levels is relevant in our setting, it suffices to look at agent A's incentives.

When the second period is reached, the agent chooses e_2^A so as to maximize:

$$e_2^A\cdot w_L + (1-e_2^A)\cdot w_N - C(e_2^A), \quad (81)$$

with its solution characterized by:

$$C'(e_2^A) = w_L - w_N. \quad (82)$$

Bringing incentives into line to induce the socially desired effort requires:

$$w_L - w_N = (1-e_2^B)\cdot(Y_L-Y_N). \quad (83)$$

Supposing we implement that spread, and making use of the period-wise symmetry of efforts, the agent's expected rent in $t = 2$ (if reached) is:

$$R_2 = w_N + e_2^A \cdot (1 - e_2^A) \cdot (Y_L - Y_N) - C(e_2^A), \tag{84}$$

and the maximized surplus in the second period is:

$$V_2^{**} = Y_N + (1 - (1 - e_2^A)^2) \cdot (Y_L - Y_N) - 2 \cdot C(e_2^A), \tag{85}$$

so that:

$$\begin{aligned} V_2^{**} - R_2 = Y_N + (1 - (1 - e_2^A)^2 - e_2^A \cdot (1 - e_2^A)) \cdot (Y_L - Y_N) - C(e_2^A) - w_N \\ = Y_N + e_2^A \cdot (Y_L - Y_N) - C(e_2^A) - w_N. \end{aligned} \tag{86}$$

The agent chooses e_1^A so as to maximize (20) with its solution characterized by:

$$C'(e_1^A) = w_E - (e_1^B \cdot w_F + (1 - e_1^B) \cdot \delta \cdot R_2). \tag{87}$$

Bringing incentives into line to induce the socially desired effort requires:

$$w_E - (e_1^B \cdot w_F + (1 - e_1^B) \cdot \delta \cdot R_2) = (1 - e_1^B) \cdot (Y_E - \delta \cdot V_2^{**}), \tag{88}$$

so that:

$$w_E = e_1^B \cdot w_F + (1 - e_1^B) \cdot (Y_E - \delta \cdot (V_2^{**} - R_2)). \tag{89}$$

Truth-telling is ensured by (21), that is:

$$\begin{aligned} & e_1^B \cdot w_F + (1 - e_1^B) \cdot (Y_E - \delta \cdot (V_2^{**} - R_2)) \geq e_1^B \cdot w_F + (1 - e_1^B) \cdot \delta \cdot w_L \\ \Leftrightarrow\; & Y_E - \delta \cdot (Y_N + e_2^A \cdot (Y_L - Y_N) - C(e_2^A) - w_N) \geq \delta \cdot w_L \\ \Leftrightarrow\; & Y_E - \delta \cdot (Y_N + e_2^A \cdot (Y_L - Y_N) - C(e_2^A) - w_N) \geq \delta \cdot ((1 - e_2^A) \cdot (Y_L - Y_N) + w_N) \\ \Leftrightarrow\; & Y_E \geq \delta \cdot (Y_L - C(e_2^A)) \end{aligned} \tag{90}$$

where (90) is referred to as (22) in the main text.