Healthcare Provider Efforts vs Fees: Striking the Right Balance

Alena Podaneva*

August 2024

Abstract

This study explores how parties within the healthcare sector can achieve an equilibrium by utilizing internal enforcement alongside external enforcement mechanisms. It particularly investigates how the optimal balance between external and internal enforcement varies with changes in the sensitivity of healthcare output to the efforts of healthcare providers and service suppliers. The analysis is conducted within the framework of a repeated game with imperfect public monitoring under double moral hazard. The study examines an optimal relational contract by solving the game in a stationary environment. The main result suggests that with an increase in healthcare output sensitivity to parties' efforts or parties' patience, in the equilibrium external enforcement should increase along with internal enforcement.

Keywords: game theory, imperfect public monitoring, enforcement.
JEL codes: C7

*University of Luxembourg. Corresponding author. E-mail: podaneva.alenaphd@gmail.com, Tel.: +352 621 961 471. 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg.

Contents

1 Introduction

Several factors contribute to the decline in service quality when public services are outsourced. These factors include a lack of understanding of private service providers' cost structures and operations, contract incompleteness, challenges in monitoring due to high costs and employee turnover, as well as the pressure to reduce expenses. However, one factor that has not been thoroughly researched is lenient penalties for poor service quality. When mild penalties are in place, outsourcing firms may be incentivized to deliver lower-quality services without facing significant consequences. This can result in quality deterioration and potentially hinder both service providers and the government from achieving maximum payoffs. On the contrary, when penalties are low, they can act as a mechanism that encourages firms to enforce their compliance with contracts. Therefore, this study aims to determine the bonuses that the healthcare provider should pay to the outsourced firm when it surpasses or falls short of an agreed healthcare output. Their difference serves as a penalty, internally binding the parties' relationship. Consequently, the paper seeks to find an optimal balance between this internal enforcement and the government's efforts, namely the external enforcement level, to maximize the parties' joint output.

In this study, I analyze a repeated game with imperfect public monitoring. The parties involved can observe the stochastic output that arises from their interactions but can't verify it.[1] Both the principal's external enforcement effort and the agent's efforts are unverifiable and unobservable, leading to a double moral hazard (DMH) problem. I focus on the scenario where the parties' efforts are continuous. By virtue of Levin (2003), I search for an optimal relational contract. I solve the game in a stationary environment, where the principal offers a time-invariant base payment and discretionary bonus every period to the agent.

Indeed, under the DMH the first best is not possible to achieve, while to achieve second-best in static game is possible by any linear contract (Kim & Wang, 1998), or "share or noting and bonus" contract (Zhao, 2007). I explore what an optimal relational contract would be in a dynamic environment under DMH.

In static environment Cong and Zhou (2021) proves that under specific conditions hold, there always exists a "share or nothing and bonus" contract to achieve the second-best outcome. I prove in my paper that under a certain inequality, there always exists a share-or-nothing contract to achieve the second-best outcome in a dynamic environment simplified

[1]Seminal contributions include Bull (1987), Gibbons and Roberts (2013), Klein and Leffler (1981), and Levin (2003). See also Gibbons and Roberts (2013) for a review.

to a stationary environment, following Levin (2003) approach. The dynamic model was chosen to determine the optimal maximum and minimum contracts, $\bar{W}$ and $\underline{W}$, which bind the parties' relationship. The objective of the paper is to find a balance between the difference in these contracts (internal enforcement) and the efforts of the healthcare provider (external enforcement). The existence of at least one share-or-nothing contract under the second-best scenario enables me to achieve this dynamic trade-off.

The main contribution of this study is twofold. First, I extend the Levin (2003)'s model to the DMH scenario. Thus, I incorporate an external enforcement variable into my model by considering the principal's unobservable efforts. Second, I conduct a simulation of the model using Spaeter (1998)'s distribution. The simulation allows me to examine how variations in the distribution impact the optimal level of the principal's effort, which then exogenously influences the variation in contingent payments. This study aims to establish how parties can achieve more efficient outcomes by making use of internal enforcement and external enforcement in the optimal relational contract.[2]

The main result of this study is that with the increase in healthcare output sensitivity to the government's external enforcement and supplier's efforts (from delivery of facility management services to surgery equipment), external enforcement should increase along with the internal enforcement in order to encourage parties to sustain the relationship under optimal relational contract.

The remainder of this paper is organized as follows. Section 2 provides an extensive literature review. Section 3 discusses the model's relevance to real-world scenarios. The model is described in Section 4. Section 5 provides simulations. Section 6 concludes.

2 Literature review

This study is related to two strands of contract theory literature: relational contract modelling within a stationary environment and with persistent types of players (Levin, 2003; MacLeod, 2003; Malcomson, 2016; Pearce & Stacchetti, 1998), and the complementary nature of external and internal enforcement from a theoretical perspective (Dumav et al., 2022; Watson, 2021; Watson et al., 2020).

[2]To identify violations of contract terms, the authority engages in contract enforcement - a process aimed at compelling the involved parties to fulfil the actions outlined in the agreed-upon contract. The mechanisms for contract enforcement differ according to their applicability, type of contract and formality. For instance, contract enforcement can be formal and informal, public or private, internal (self-enforceable) or external (attracting a third party, public institution, regulatory authority, or law court).

Relational contracts have emerged as a result of developments in repeated game theory applied to principal-agent modelling.[3] This concept is derived from the dynamic enforcement constraint for the equilibria in repeated games. Rubinstein (1979) is the first to examine a repeated-game model of a principal-agent relationship with binary choices for both parties. Radner (1985) expands on this by considering the model in discrete time and studying equilibria that involve "review strategies". Spear and Srivastava (1987) further explore these equilibria within a larger context, analysing them as part of a dynamic program that incorporated continuation values.

Meanwhile, MacLeod and Malcomson (1989) extend the study of Radner (1985) by incorporating continuous effort. They demonstrate that when output measures are not verifiable and performance is commonly known, continuation values in long-term relationships can be utilised to incentivise individuals. Pearce and Stacchetti (1998) explore how both verifiable and non-verifiable performance measures could be combined in optimal contracts, providing a foundational framework for the recursive approach in contract theory. Levin (2003) develops a model on Pearce and Stacchetti (1998)'s work by introducing a nonpersistent and observable outcome that cannot be verified. Additionally, Levin (2003) includes the shock to the cost of effort, while the environment in MacLeod and Malcomson (1989) does not involve any uncertainty.[4]

Theories in microeconomics have been developed based on certain contract enforcement assumptions. Hereafter, incentive theory focuses on how agents attempt to resolve information asymmetry, allowing for a perfect enforcement agreement by an external mechanism. In contrast, transaction cost theory eliminates this binding assumption. It establishes how economic agents in the real world solve contracting problems under incomplete contracts, where contractual breach penalties and/or supervising authorities can punish the present. In the late 1980s, incomplete contract theory was developed, focusing on the information asymmetry between parties interacting within hierarchical relationships. This theory focuses on imperfections in institutional systems, aimed at making contracts enforceable and allowing for the unverifiability of key variables. Recently, game theorists have modeled relational contracts as long-term contractual relationships encompassing both formal (externally enforced) and informal (self-enforcing) contracts. Notably, a commonality across these theories is the mutual exclusion of either self-enforcing or externally enforced contracts, with a limited number of studies considering their combined application. Nevertheless, the

[3]See MacLeod (2007) and MacLeod and Malcomson (2023) for a comprehensive review of the relational contracts theoretical literature. For empirics, follow Macchiavello (2022).

[4]Similar to Levin (2003) but with a risk-averse agent, there is a study by MacLeod (2003).

relational contract literature has made a step further, and recently written papers prove that external enforcement complements internal enforcement.

The concept of self-enforcement and external enforcement in finite period models was first distinguished by Bull (1987). Recently, other studies built on the Levin (2003) model have considered the complementarity of external and internal enforcement in infinite-period models. Notably, Dumav et al. (2022) study a repeated principal-agent game coupling that limited external enforcement with persistent productivity shocks. Watson et al. (2020) demonstrate that in more realistic setting where players can successfully renegotiate every period both components of their long-term contract every following any history, the external-enforcement technology always complements self-enforcement in an optimal contract. This can be achieved even after deviation. Watson (2021) congregates relational incentive contracts on a class of repeated-game style models of ongoing relationships with moral hazard. He establishes external enforcement through an external contract that prescribes contractual provisions to the external enforcer on how to intervene in a relationship. This contract is introduced by compelling monetary transfers as a function of verifiable information. The internal part of the contract records how the contracting parties agree to act. In comparison to the aforementioned studies, I introduce external enforcement into the model through the principal's external enforcement effort and establish a link to the self-enforcement constraint through the tightness of the dynamic enforcement constraint.

This study also contributes to existing research on DMH. DMH has been applied in various applications to determine optimal contract conditions, including franchising (Bhattacharyya & Lafontaine, 1995), warranty contracts (Cooper & Ross, 1985), supply chains (Corbett & DeCroix, 2001; Corbett et al., 2005), collaborative services such as financial planning, consulting, and IT outsourcing (Roels et al., 2010), joint product improvement by client and customer support centers (Bhattacharya et al., 2014), repair and restoration services (Jain et al., 2013), and justice production (Roussey & Soubeyran, 2018). These studies share the common assumption that the parties' efforts contribute to a joint outcome. In the healthcare sector, any relationship should, in my view, be considered within the double-sided framework since all actors obtain a common output, joint production of health, and improvement of patient care.

Nevertheless, there is limited research on the interrelationships between different actors in healthcare services from the perspective of double-sided asymmetric information. Those that exist are mainly focused on adverse selection due to common problems with health insurance, where individuals have private health risk information, and doctors have private patient health condition information. Classical problems within this framework are: lack of coverage, overtreatment and undertreatment, inappropriate referral to a specialist.[5] This paper focuses on the less covered in the healthcare literature asymmetric information problem, DMH, leaving aside adverse selection.[6]

There have been trials by other authors to solve for an optimal contract within patient-physician and hospital-physician relationships under a DMH in a static environment. Specifically, Schneider (2004) covers the DMH problem in patient-physician relationship based on the Cooper and Ross (1985)'s paper, while Leonard and Zivin (2005) introduce a regulator as a third party to the problem, who observes physician effort and forces him to provide its particular level. Wang (2015) describes hospital and physician relationship under a double-sided model.[7] Finally, Bhattacharya et al. (2015) solve for an optimal contract within the research and development partnerships of a provider that conducts research and a client that provides activities in the pharmaceutical industry. Thus, they simultaneously resolve the holdup, DMH, and risk-aversion problems. All these studies assume that one of the actors is risk-averse, that is, the patient, physician, or service provider. In contrast, I focus on the interaction of parties in the healthcare industry who are indifferent to risk over time.

My research also adds to the existing knowledge on contract enforcement. In the following subsections, I provide a comprehensive review of recent literature on contract enforcement in economics, covering theoretical frameworks and empirical findings. To enhance clarity, I organize the discussion into separate subsections for public and private contract enforcement.

2.1 Public enforcement by law court

Enforcing a contract may require engaging a third party, such as a court, to serve as an overseer in situations in which the parties violate the agreement.

[5] For example, optimal contract establishment between the National Health Service (NHS) (public, or private insurer) and a healthcare provider (such as a hospital or GP) for a specific health service (De Fraja, 2000) or when there are multiple treatments available for a given diagnosis (Siciliani, 2006); between health planner and patients (Frank et al., 2000); between hospitals varying in terms of doctor abilities and unit mass of patients (Makris & Siciliani, 2013).

[6] Major (2019) widely covers various combinations of asymmetric information among different actors involved in providing healthcare services including patient-physician relationships, physicians working within medical institutions, government interactions with medical institutions, and patients dealing with governmental or state agencies.

[7] Follow McCullough and Snir (2010) and Culyer and Newhouse (2000) for moral hazard problem in between hospital and physician relationship.

2.1.1 Theoretical approach

According to the literature on principal-agent contracts under unverifiable performance measures, contractual terms may not be enforced by a third party or could involve excessively high enforcement costs (Baker et al., 1994; Fuchs, 2007; Levin, 2003; MacLeod, 2003; MacLeod & Malcomson, 1989). Doornik (2010) explores an intermediate scenario where parties need to bear some cost for contract enforcement, but the costs are not prohibitive that external enforcement becomes impossible, e.g. liquidated damages clauses. He demonstrates that the structure of principal-agent contracts can impact the likelihood of incurring enforcement costs, which has implications for determining the optimal contract choice. In this context, the two parties are able to draft contracts detailing payments based on output. However, unlike a standard principal-agent model, verifying output and enforcing payments comes with associated expenses.

Under incomplete contract theory, enforcement can be viewed from an ex-ante versus ex-post perspective. In a principal-agent model with adverse selection, costly enforcement technology (Guasch et al., 2008) and lack of enforcement (Chong et al., 2006; Estache, 2006; Guasch, 2004; Saussier et al., 2009) lead to renegotiation occurrences. However, an efficient judicial system may attract more renegotiation because of the court's ability to force a solution (Domingues & Sarmento, 2016; Guasch et al., 2008). Al-Najjar et al. (2002) focus on incomplete contracts leading judges to void clauses and initiate renegotiation. The procurement and regulation models proposed by Laffont (2003) and Laffont and Meleu (2000) integrate the concept of adverse selection, in which enforcement of penalties may be influenced by enforcement expenditures. These models align with the perspectives of the Chicago school. Although modern contract theory recognizes the importance of law enforcement, it has not yet fully addressed this specific aspect.

2.1.2 Empirical approach

There is a growing body of empirical research on enforcement of public contracts. For example, studies by Girth (2014) and Coviello et al. (2018) examine the impact of external factors on the enforceability of public contracts, while Coviello et al. (2018) also discusses a strategy that is not only within the discretion of public managers but also aligns with the theoretical literature on internal verifiability (Kvaløy & Olsen, 2009, 2010). Using an extensive dataset on Italian public procurement, Coviello et al. (2018) empirically analyze the effects of court inefficiency on the performance of public works. They observe that in situations where courts are inefficient, there are longer delays in delivering public works, and these delays increase for more valuable contracts. In contrast to this focus on accountability, private agreements emphasize self-enforcement through mutually binding pacts, where reputation plays a crucial role. Giacomelli and Menon (2017) demonstrate how inadequate contract enforcement can significantly influence firms' incentives for expansion - their findings suggest that reducing judicial proceedings duration by 10% leads to a 2% growth in local firms' average size.

2.2 Public enforcement through institutions

Contracts can also be enforced by institutions. This subsection explores how legal frameworks and institutions adhere to contractual agreements.

Theoretical approach

The study by Laffont and Meleu (2000) examines the concept of incentive regulation in developing nations, where inadequate institutions lead to contracts that cannot be fully enforced. In this framework, the agent must decide whether to accept a regulatory contract based on the menu before knowing its cost implications. The effort made by the agent is not visible, which reduces the apparent costs. Therefore, contracts that reward agents based on observed costs are more likely to be enforced when the principal's expenditure level increases. The effective enforcement of regulatory contracts means that high-cost agents encounter reduced benefits.

Laffont (2003) analyses the structure of incentives and defines the most effective regulatory agreements and enforcement expenses. By contrast, Garcia et al. (2005) proposes that an agent's legal costs impact the likelihood of successfully enforcing the original contract. The effort put forth by an agent in carrying out a project is influenced by its ability to pursue litigation over contract terms to recoup any cost overruns. Garcia et al. (2005) focuses on how an agent's motivation for reducing project costs is impacted by their option to engage in expensive litigation. Their model illustrates that, when large-scale public projects are procured through contracts with strong incentives for private firms, excessive litigation may occur in weak institutional settings. They demonstrate that committing to a predetermined level of government-led litigation alongside weaker incentive contracts serves as a more efficient procurement method. Guccio et al. (2017) discover that the quality of local conditions can influence public officials' motivations to act efficiently.

Empirical approach

The role of institutional weaknesses in infrastructure procurement has been discussed widely. The general result of the studies mentioned below is that the quality of the institutional environment matters in infrastructure procurement. For example, Estache et al. (2015) demonstrate that "weak institutions" do not uniformly advocate for or oppose private finance. Instead, different weaknesses push in different directions. If a regulator is weak and faces significant information asymmetry, there could be potential cost reductions on which the government may fail to capitalize through pricing. This situation leads to private finance as a means of increasing the likelihood of cost reduction and taking advantage of lower expected costs.

Cavalieri et al. (2020) empirically demonstrate that certain dimensions of institutional quality have a more significant impact on performance in contract execution within the transport infrastructure. Other studies (Baldi et al., 2016; Coviello & Gagliarducci, 2017; Finocchiaro Castro et al., 2014, 2018) focusing on corruption in the institutional environment find an association between the characteristics of the local area and outcomes in public procurement, impacting measures such as price differentials, cost overruns, and execution time (Bandiera et al., 2009; Coviello & Gagliarducci, 2017; Finocchiaro Castro et al., 2014).

2.3 Private enforcement: self-enforcing agreements

Self-enforcement is often overlooked in the literature as a possible mechanism for maintaining future partnerships and preventing breach of contracts. Telser (1980) is one of the first to develop the theory of self-enforcing agreements. One crucial assumption in his paper is that parties only consider uncertain future outcomes, without taking into account their past relationship history. In other words, an agreement is self-enforcing, as long as both parties expect greater utility from continuing the relationship than from breaking the present contract. Noorderhaven (1992) argues that under this assumption, as in agency theory, only transactions with an immediate and simultaneous performance exchange can occur. Nevertheless, it is impossible to examine agency theory for these exchanges. Further theoretical models have been developed to address contract self-enforcement when cooperation is optimal. In supply contracts, the long-term returns from the current relationship must be equivalent to the present value of returns from the spot market for the product involved in order for both parties to continue trading with each other.

Watson (2021) explains that self-enforcement involves coordinated actions by the parties involved, aligned with their individual motivations. Self-enforcing agreements allow one party (the principal) to end a contract with the other party (the agent) if undesirable actions are identified. The threat of contract termination serves as a deterrent for the agent's misbehavior, particularly when they derive greater benefits from the relationship compared to outside it. Despite enforceable legal contracts, various aspects of parties' conduct and performance remain uncontracted due to multiple uncertainties and information imbalances. Consequently, self-enforcing agreements are widespread in business relationships (Gil & Zanarone, 2014; Gil & Marion, 2013) whether through informal arrangements alone (Levin, 2003) or in conjunction with formal contracts (Zanarone, 2013). There is also extensive literature on purely self-enforcing contracts that can be found in studies such as those by Bull (1987), Levin (2003), and MacLeod and Malcomson (1989).

A significant breach in the incentive contract literature comes from the impossibility of considering both public and private enforcement mechanisms. However, these are typically excluded. One way to deal with contract complexity and incomplete contracts and eliminate possible hold-up problems is to rely on informal commitments, such as relational contracting (Macaulay, 2018). Reliance on relational contracts in the public domain is unfeasible, owing to their uniqueness and contract longevity. Therefore, using relational contracts, I investigate internal and external contract enforcement within public services procurement in the healthcare industry.

3 Real-world scenario relevance of the model

The model analysed in this study can be used to describe relationships between a pharmaceutical company and a healthcare provider, as well as between a medical equipment maintaining firm and a hospital, and other actors providing healthcare services.[8] The choice to focus on the healthcare sector was based on the feasibility of applying risk-neutral parties' interaction to numerous different relations within this industry.[9] In general, the model is suitable to describe various relationships in other industries, like franchiser-franchisee interactions, that satisfy particular criteria:

[8] New technology development, like telemedicine devices, may not fit this model because the developer could be a monopolist. This might lead to the government being pressured to extend the contract even if the service is unsatisfactory. In such cases, even though there may be penalties for the supplier, the contract will not be terminated, which does not align with the grim-trigger strategy examined in this study.

[9] In the healthcare sector, any firm that focuses on preventive measures, optimizing healthcare delivery, or enhancing patient outcomes instead of prioritizing revenue generation exhibits the behaviour of a risk-neutral firm.

- risk-neutral parties;
- presence of double-sided moral hazard problem in the parties relationships;
- contract is of the type: fixed fee and contingent on the output bonus or fee;
- parties interact every period with the possibility to continue relationships endlessly;
- one's party prevarication ends the relationship forever;
- introduction of external contract enforcement in the relationship through the principal's effort level complements or substitutes for the agent's efforts without directly affecting their efficiency.

Below, there is a detailed example of a scenario within the healthcare industry that fulfills these criteria. Specifically, medical equipment maintaining firm and a hospital.

Parties risk-neutrality and external contract enforcement efforts: a healthcare organisation offers a contract to a provider for maintenance services on medical equipment to perform preventive maintenance and repairs on medical devices. The risk-neutrality of the provider is influenced by their reputation and long-term relationships with hospitals, which are more important than short-term risk-taking behaviour. Both parties' actions impact the chance of failure, leading to a DMH problem where both the hospital and service provider may have incentives to neglect proper care for the equipment. To address this issue, the healthcare organization employs various forms of external enforcement, such as performance benchmarks, contractual penalties, and quality assurance measures.

Contract: a hospital offers a contract to a maintenance firm that includes fixed payment and possible fees or bonuses. The firm can be fined or have the contract terminated if it does not respond promptly to equipment breakdowns, leading to potential issues with critical medical devices. In contrast, a hospital or equipment manufacturer covers the costs associated with replacements or major repairs, and the maintenance firm earns an additional bonus for coordinating and facilitating the repair or replacement process.

Parties interaction: multiple firms can maintain healthcare medical equipment because the market is not monopolistic. Since service failure directly impacts patient health, there is a high probability of contract termination for unsatisfactory maintenance. A possibility of renewing or extending the contract between parties due to the new equipment maintenance can be equalized to a newly offered contract by the principal.

Principal's contract enforcement effort level

Efforts of the principal in the principal-agent model, particularly in the context of external contract enforcement, can take various forms beyond direct monitoring. These efforts are designed to incentivize and ensure compliance from the agent without necessarily incurring additional monitoring costs. Here are some examples:

- Clear contractual terms and incentive structures: the principal can invest effort in designing clear and comprehensive contractual terms, including performance metrics, standards, and incentive structures. By clearly defining expectations and rewards for achieving desired outcomes, the principal creates a framework that aligns the agent's incentives with organizational objectives, reducing the need for extensive monitoring.

- Regular communication and feedback mechanisms: the principal can establish channels for regular communication and feedback between the principal and the agent. By providing timely feedback on performance and addressing any concerns or issues promptly, the principal fosters transparency, trust, and accountability, which can motivate the agent to adhere to contractual obligations without the need for constant monitoring.

- Training and capacity building: the principal can invest in training programs and capacity-building initiatives to enhance the agent's skills, knowledge, and capabilities. By providing the agent with the necessary resources and support to perform their duties effectively, the principal empowers the agent to fulfill their contractual obligations autonomously, reducing the likelihood of non-compliance and the need for continuous monitoring.

- Performance reviews and recognition: the principal can conduct periodic performance reviews and provide recognition or rewards for exemplary performance. By acknowledging and rewarding the agent's achievements, the principal reinforces desired behaviors and outcomes, motivating the agent to maintain high levels of performance and compliance with contractual obligations.

- Escalation mechanisms and dispute resolution processes: the principal can establish escalation mechanisms and dispute resolution processes to address any disagreements or disputes that may arise during the contract period. By providing a structured framework for resolving conflicts and addressing grievances, the principal ensures that issues are addressed promptly and fairly, mitigating the risk of non-compliance and minimizing the need for external monitoring.

Overall, these efforts by the principal contribute to effective contract enforcement and compliance without imposing significant additional monitoring costs. By investing in clear communication, training, feedback mechanisms, and performance incentives, the principal can incentivize the agent to fulfill their contractual obligations autonomously, fostering a mutually beneficial relationship that maximizes value creation and minimizes the need for external oversight.

4 The Model

This section describes an applied model of this study. Initially, I provide the timing of the game. Then, I discuss how agents interact in the stage game. Finally, I extend the stage game to the infinitely-repeated version, where I solve for the optimal relational contract that can be achieved by the parties in equilibrium under certain constraints.

I examine the dynamic relationship between two long-lived parties ($i = P, A$): a risk-neutral principal, referred to as "she", and a risk-neutral agent, "he", responsible for providing services on behalf of the principal. To illustrate this relationship, I consider the context of contracting healthcare-oriented services (the detailed example provided in the previous section 3). Interactions occur at regular intervals over periods $t = 0, 1, 2, ...$. In the vein of MacLeod and Malcomson (1989) and Levin (2003), I assume that the parties' commitment to the relationship is observable, ensuring that identical promised payments in every period are honoured.[10] The physical environment remains unchanged at each date.

4.1 Timing

In each period t, the parties play the following stage game, as depicted in Fig. 1.

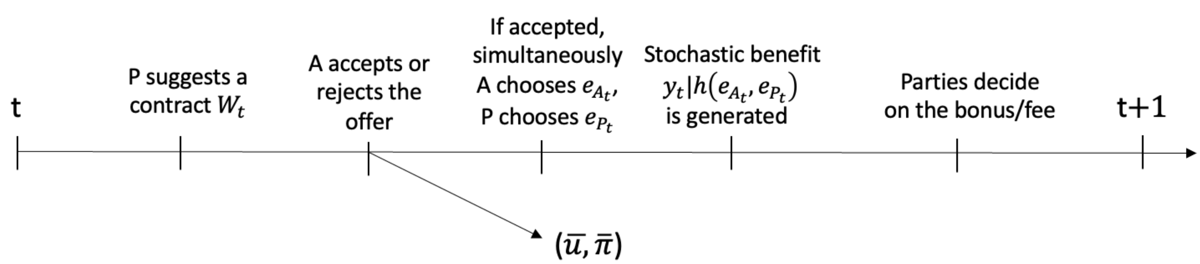

Figure 1: Timing

[10]See Halac (2012) for relaxing the assumption that commitment to the relationship is certain.

The principal makes a take-or-leave-it offer to the agent. Upon receiving the offer, he decides whether to accept or reject it. If declined, the agent gains $\overline{u} \in \mathbb{R}$ (by signing a contract with another principal) in the current period and for all subsequent periods, while the principal earns $\overline{\pi} \in \mathbb{R}$ (by providing services in-house or outsourcing to an alternative supplier).[11] If accepted, both the agent and the principal simultaneously choose their respective effort levels, e_{A_t} and e_{P_t}.

The agent exerts effort $e_{A_t} \in E_A \subseteq [0, \overline{e}_A]$, incurring cost $c(e_{A_t})$, where $c_e, c_{ee} > 0, c(0) = 0$, and $\lim_{e \to \overline{e}_A} c_e = \infty$.[12] His nonpersistent level of effort is unobservable and non-contractible.[13] Simultaneously, the principal chooses efforts to incentivize and ensure compliance from the agent without necessarily incurring additional monitoring costs, $e_{P_t} \in E_P \subseteq [0, \overline{e}_P]$. She bears a cost of effort $\phi(e_{P_t})$, where $\phi_e, \phi_{ee} > 0, \phi(0) = 0$, and $\lim_{e \to \overline{e}_P} \phi_e = \infty$. The choice of a cost function with increasing marginal cost for the principal aligns with risk-neutral decision-making and reflects the realistic cost structure of her activities. In particular, the principal's external enforcement effort incurs increasing costs as it becomes more intensive. Similarly, her nonpersistent effort is unobservable and nonverifiable for the agent. Thus, I face a DMH problem.

The parties' actions yield a stochastic output y_t, which is a random variable conditional on e_{A_t} and e_{P_t} through a composite effort function $h(e_{A_t}, e_{P_t}) : \mathbb{R}^2_+ \to \mathbb{R}$, where $h_{e_A}(e_{A_t}, e_{P_t}), h_{e_P}(e_{A_t}, e_{P_t}) > 0$ and $h_{e_A e_A}(e_{A_t}, e_{P_t}), h_{e_P e_P}(e_{A_t}, e_{P_t}) \leq 0$.[14] The output is distributed as $F(y_t | h(e_{A_t}, e_{P_t}))$ with the corresponding twice continuously differentiable density function $f(y_t | h(e_{A_t}, e_{P_t}))$ on the support $Y = [\underline{y}, \overline{y}]$.[15] This is a repeated game with imperfect public monitoring, where the output y is publicly observed by both the agent and the principal, but is non-verifiable by a third party. To allow for the validity of the first-order approach (Cong & Zhou, 2021; Jewitt, 1988; Milgrom, 1981; Rogerson, 1985), I assume

[11] This reflects a traditional grim-trigger strategy.

[12] The limit $\lim_{e \to \overline{e}_A} \phi_e = \infty$ indicates that the model acknowledges the infeasibility of extremely high effort levels. In particular, the marginal cost of effort becomes prohibitively high or approaches infinity when the agent attempts to exert effort beyond the maximum level.

[13] The assumption that the agent's effort, and the principal's later, is not persistent means that their choice of effort in one period does not provide any information about their effort level in the next period.

[14] I employ a composite effort function to simplify the analysis in the vein of Bhattacharyya and Lafontaine (1995), Cong and Zhou (2021), and Kim and Wang (1998). The conditions for the first and second order derivatives of the joint production function justify complementary efforts by parties that correspond to reality.

[15] A restriction on the finite continuum of outputs is essential for the existence of an optimal contract. Without such a limitation, continuous efforts allow the principal's payoff decrease in output, potentially incentivizing her to sabotage strong performance. For more details on this issue, follow Innes (1990).

that $\forall h \in \mathbb{R}$, $\frac{d}{dy}\left(\frac{f_h(y_t|h(e_{A_t},e_{P_t}))}{f(y_t|h(e_{A_t},e_{P_t}))}\right) > 0$ and $F(y_t|h(e_{A_t}, e_{P_t}))$ is convex in h for any $y \in [\underline{y}, \overline{y}]$.[16]

Finally, upon the realization of the output y_t, the principal compensates the agent with an agreed fixed payment $w_t \geq 0$. Additionally, adjusting this payment with a promised, but not guaranteed, output-contingent bonus $b_t : Y \to \mathbb{R}$. Consequently, I examine an asymmetric information scenario, where both parties' efforts are unobservable, and they have equal levels of moral hazard. The scenario where the parties' hidden actions vary is beyond the scope of this study.

When $b_t(y_t) \geq 0$, the principal decides whether to fulfill or withdraw the bonus payment at the end of period t. Conversely, when $b_t(y_t) < 0$, the decision belongs to the agent. Allowing for a negative bonus payment removes limited liability from the model. Let W_t denote the total compensation, where $W_t = w_t + b_t(y_t)$ if the contingent payment is honored, and $W_t = w_t$ otherwise.

To streamline notations within a dynamic environment, henceforth in this paper, I define: $h = h(e_{A_t}, e_{P_t})$, $h_{e_P} = h_{e_P}(e_{A_t}, e_{P_t})$, $h_{e_A} = h_{e_A}(e_{A_t}, e_{P_t})$, $h_{e_P e_P} = h_{e_P e_P}(e_{A_t}, e_{P_t})$ and $h_{e_A e_A} = h_{e_A e_A}(e_{A_t}, e_{P_t})$. Hence, $f(y_t|h) = f(y_t|h(e_{A_t}, e_{P_t}))$, $f_h(y_t|h) = f_h(y_t|h(e_{A_t}, e_{P_t}))$, $F(y_t|h) = F(y_t|h(e_{A_t}, e_{P_t}))$ and $F_h(y_t|h) = F_h(y_t|h(e_{A_t}, e_{P_t}))$. In a static environment, the time subscripts will be omitted.

4.2 Stage game

A discussion of the stage game in this section serves as a preliminary step towards addressing the DMH problem within a dynamic framework, which constitutes the central focus of this paper.

I denote a stage game as G. The stage game G determines the optimal contract $W(\cdot)$, when parties interact only during one period. Following Kim and Wang (1998) and Zhu and Wang (2005), under the binding first order approach (FOA) and combination of the parties' incentive constraints, the relaxed problem that solves for the optimal linear contract is as follows:

$$\max_{e_P, e_A} \int_{\underline{y}}^{\overline{y}} y f(y|h) dy - \phi(e_P) - c(e_A) - \overline{u}, \tag{1}$$

[16]These assumptions make the Mirrlees-Rogerson conditions, the monotone likelihood ratio property (MLRP) and the convexity of the distribution function condition (CDFC), valid and allow for the first-order conditions to bind. Follow the Appendix A1 for more details.

$$\int_{\underline{y}}^{\overline{y}} yf(y|h)dy = \frac{c'(e_A)}{h_{e_A}} + \frac{\phi'(e_P)}{h_{e_P}}. \tag{2}$$

Defining the solution of the problem by (e_A^*, e_P^*) that is independent of any specific contract and slightly rearranging (2), it is clear that $\int_{\underline{y}}^{\overline{y}} yf_h(y|h(e_A^*, e_P^*))h_{e_A}dy > c'(e_A^*)$. More precisely, given e_P^*, the expected marginal payoff from e_A^* is strictly larger than its marginal cost, and vice versa. This inequality explains why (e_A^*, e_P^*) is the second-best effort choice. An impossibility of achieving first-best outcome within this problem is related to the balancing-budget problem, i.e. despite the output, the sum of parties payment is always equal to the whole output (Holmstrom, 1982).

Kim and Wang (1998) and Zhu and Wang (2005) have proved that under the DMH problem without limited liability the linear sharing contract that can always achieve second-best outcome is of the form $W^*(y) = \frac{c'(e_A^*)}{R_{e_A}(e_A^*, e_P^*)}[y - R(e_A^*, e_P^*)] + \overline{u} + c(e_A^*)$, where $R(e_A^*, e_P^*) = \int_{\underline{y}}^{\overline{y}} yf(y|h)dy$. Cong and Zhou (2021) simplify the optimal linear sharing contract to the form $W^*(y) = \gamma^*(y - \hat{y}^*)$, where $\hat{y}^*$ and γ^* are defined below:

$$W^*(y) = \underbrace{\frac{c'(e_A^*)}{\int_{\underline{y}}^{\overline{y}} yf_h(y|h)h_{e_A}dy}}_{\gamma^*} \left(y - \underbrace{\int_{\underline{y}}^{\overline{y}} yf(y|h)dy + \frac{c(e_A^*) + \overline{u}}{\gamma^*}}_{\hat{y}^*} \right).$$

4.3 Multiple-periods contracting problem

Let consider $G^\infty(\delta)$ as an infinitely repeated version of the two players stage game G, where parties maximize their discounted payoff streams. Specifically, parties care about the future of their interactions, so from date t onward, their expected payoffs are discounted by the same factor $\delta \in (0, 1)$. δ corresponds to the probability that the interaction will continue until the next date, i.e. after each stage, there is a probability $(1 - \delta)$ that the game will end. In essence, the game will end in finite time but just randomly. Expected lifetime payoffs are normalized by $(1 - \delta)$ to show them as per-period averages.

Principal's expected payoff:

$$\pi_t \equiv (1 - \delta)E \sum_{\tau=t}^{\infty} \delta^{\tau - t}[\Lambda_\tau[y_\tau - W_\tau - \phi(e_{P_\tau})] + (1 - \Lambda_\tau)\overline{\pi}]. \tag{3}$$

Agent's expected payoff:

$$u_t \equiv (1-\delta)E\sum_{\tau=t}^{\infty}\delta^{\tau-t}[\Lambda_\tau[W_\tau - c(e_{A_\tau})] + (1-\Lambda_\tau)\overline{u}]. \tag{4}$$

In both equations, (3) and (4), $\Lambda_\tau \in [0,1]$ is a probability that determines the agent's response to the take-or-leave-it offer. When $\Lambda_\tau = 1$, he accepts the offer, whereas a value of $\Lambda_\tau = 0$ means rejection. Once Λ_τ is zero, it remains zero forever, indicating a strict punishment rule as in Abreu (1988). Then, the expected surplus is $s_t = u_t + \pi_t$. In a dynamic environment, the principal's offer in period t depends on the information she obtains when she makes an offer. The information available at the beginning of period t is denoted as $h^t = (w_0, \Lambda_0, b_0, y_0, e_{A_0}, e_{P_0} \cdots w_{t-1}, \Lambda_{t-1}, b_{t-1}, y_{t-1}, e_{A_{t-1}}, e_{P_{t-1}})$ as a public history up to period t, and H^t is a set of all possible period-t public histories.

The principal's strategy σ_P specifies a decision whether or not to offer a contract to the agent, a fixed payment $w_t(h^t)$, a contingent bonus $b_t(h^t, y_t)$, and an effort level $e_{P_t}(h^t, w_t)$. The agent's strategy σ_A specifies a decision whether or not to accept an offer from the principal, and an effort level $e_{A_t}(h^t, w_t)$. Let ζ_w is a flow payoff from a verifiable fixed compensation, while ζ_b is a flow payoff from non-verifiable contingent bonus payment.

Hence, $(\sigma_P, \sigma_A, \zeta_w, \zeta_b)$ is a *relational contract* that is a complete plan of the relationship. It identifies for each period t and every history $h^t \in H^t$ a fixed compensation the principal offers w_t, the agent's participation decision Λ_t, in the event of acceptance, the principal's effort level e_{P_t} along with the agent's effort level e_{A_t}, and the variable bonus payment $b_t(y_t)$ given the output observable realization. A relational contract is *self-enforcing* if the players' strategies constitute a perfect public equilibrium (PPE) of the repeated game that describes the behavior on and off the equilibrium path (Fudenberg & Levine, 1994). In the PPE, players' decisions are based solely on publicly available information in equilibrium. Specifically, they can not base their actions on their previous efforts, while both parties can consider past payments and outcomes. This analysis focuses on optimal relational contracts, which are defined as PPEs that achieve Pareto efficiency within the set of PPE payoffs.

Further, I first describe the constraints that must be satisfied for the payoff pair (π, u) to be within the PPE payoff set. Then, in Proposition 1, I characterize the parties' efforts that can be sustained in a PPE. Next, I define an optimal relational contract under a specific scenario in Proposition 2.

Constraints

I denote the set of PPE payoffs by Φ. Each payoff pair $(\pi, u) \in \Phi$ is associated with the profile of actions $(e_A, e_P, w, b(y))$ and continuation payoffs $(\pi(y), u(y))$ as functions of observable but unverifiable output y, where $\pi(y)$ is the principal's continuation payoff and $u(y)$ is the agent's continuation payoff.[17] The continuation payoffs $(\pi(y), u(y))$ are generated by the continuation contract $W(y) \equiv w + b(y)$ imposition.

Parties' expected payoffs under the continuation contract $W(y)$ that are equal to the weighted sum of current and future payoffs, i.e.,

$$\pi \equiv (1-\delta) \int_{\underline{y}}^{\overline{y}} [y - W(y) - \phi(e_P)]\, f(y|h)dy + \delta \int_{\underline{y}}^{\overline{y}} \pi(y) f(y|h)dy,$$

$$u \equiv (1-\delta) \int_{\underline{y}}^{\overline{y}} [W(y) - c(e_A)]\, f(y|h)dy + \delta \int_{\underline{y}}^{\overline{y}} u(y) f(y|h)dy,$$

with the followed expected contract surplus $s \equiv u + \pi$, i.e.,

$$s \equiv (1-\delta) \int_{\underline{y}}^{\overline{y}} [y - \phi(e_P) - c(e_A)] f(y|h)dy + \delta \int_{\underline{y}}^{\overline{y}} s(y) f(y|h)dy,$$

where $s(y) \equiv \pi(y) + u(y)$ is the continuation surplus.

For this paper, there are five constraints, four as in Levin (2003) and one more, principal's incentive constraint, that identify whether the continuation contract $W(y) \equiv w + b(y)$ is self-enforcing and ensure the viability of the PPE payoff pair (π, u) through pure actions. In particular,

(i) individual rationality constraints that ensure the wiliness of parties to initiate the contract: $u \geq \bar{u}$ and $\pi \geq \bar{\pi}$ for the agent and the principal, respectively. These constraints are defined as (IR_A) and (IR_P);

(ii) agent's incentive constraint that guarantees a choice by the agent of a certain level of effort e_A:

$$e_A \in \arg\max_{e_A \in E_A} \left\{ \int_{\underline{y}}^{\overline{y}} \left[W(y) + \frac{\delta}{1-\delta} u(y) \right] f(y|h)dy - c(e_A) \right\};$$

[17] A continuation payoff refers to the expected payoff that a player anticipates receiving in future periods of a repeated game, based on the strategies and actions chosen by all players in preceding periods.

(iii) principal's incentive constraint that guarantees a choice by the principal of a certain level of effort e_P:

$$e_P \in \arg\max_{e_P \in E_P} \left\{ \int_{\underline{y}}^{\overline{y}} \left[y - W(y) - \frac{\delta}{1-\delta}\pi(y) \right] f(y|h)dy - \phi(e_P) \right\};$$

(iv) self-enforcement constraints (called by Li and Matouschek (2013) as truth-telling) that ensure wiliness of parties to make the variable payment for all output $y \in Y$. I define them as the agent's dynamic-enforcement constraint (DEC_A) and the principal's dynamic-enforcement constraint (DEC_A). Specifically, for the agent, it is preferable to set $e_A = 0$ rather than commit to any agreement with $e_A > 0$ unless the following condition is met:

$$(DEC_A) \qquad \delta\{\text{future gain to the agent}\} \equiv \delta(u(y) - \overline{u}) \geq -(1-\delta)b(y),$$

which stipulates that the agent will adhere to the relationship with the principal if the agent's future gain, $u(y) - \overline{u}$, is equal to or exceeds the variable payment that he can receive. Similarly, the principal prefers not to exert any effort, $e_P = 0$, and maintains relationships with the agent only if the following condition holds:

$$(DEC_P) \qquad \delta\{\text{future gain to principal}\} \equiv \delta(\pi(y) - \overline{\pi}) \geq (1-\delta)b(y),$$

which stipulates that her future gains from interacting with him, $\pi(y) - \overline{\pi}$, must be greater than or equal to the highest possible variable payment that she should make to the agent. It is important to note that when (DEC_A) and (DEC_P) are satisfied, constraints (IR_A) and (IR_P) are implicitly met. More precisely, the imposition of (DEC_A) and (DEC_P) eliminates the necessity of applying (IR_A) and (IR_P).

(v) feasibility constraints that are met by the stage game G settings. Namely, the sequence of nonnegativity constraints: $e_A, e_P, w, b(y) \geq 0$. Furthermore, each continuation contract should be self-enforcing (i.e. self-generating). In particular, for each y, the pair of continuation payoffs $(u(y), \pi(y))$ should correspond to a self-enforcing contract that will be initiated in the next period.

It follows from Theorem 1 of Levin (2003) that an existence of self-enforcing contract that generates an expected surplus s that is strictly larger than an outside option $\bar{s}$ guarantees an existence of a pair (u, π) that gives the same expected payoff, where $u \geq \bar{u}$, $\pi \geq \bar{\pi}$ and $s + \pi$.

Hence, I focus on contracts that maximize the parties joint surplus subject to the constraint that each continuation contract is self-enforcing. If there is a self-enforcing contract (or PPE) that achieves a joint surplus s, there are also self-enforcing contracts that achieve any individually rational split of this surplus.

Given the mentioned above constraints, I formulate the problem:

$$s = \max_{e_A, e_P} \left\{ (1-\delta) \int_{\underline{y}}^{\overline{y}} [y - \phi(e_P) - c(e_A)] f(y|h) dy + \delta \int_{\underline{y}}^{\overline{y}} s(y) f(y|h) dy \right\}$$

s.t.

$$(IC_A) \qquad e_A \in \arg\max_{e_A \in E_A} \left\{ \int_{\underline{y}}^{\overline{y}} \left[W(y) + \frac{\delta}{1-\delta} u(y) \right] f(y|h) dy - c(e_A) \right\},$$

$$(IC_P) \qquad e_P \in \arg\max_{e_P \in E_P} \left\{ \int_{\underline{y}}^{\overline{y}} \left[y - W(y) - \frac{\delta}{1-\delta} \pi(y) \right] f(y|h) dy - \phi(e_P) \right\},$$

$$(DEC_A) \qquad W(y) + \frac{\delta}{1-\delta} \int_{\underline{y}}^{\overline{y}} [W(y) - c(e_A)] f(y|h) dy \geq 0 + \frac{\delta}{1-\delta} \bar{u},$$

$$(DEC_P) \qquad -W(y) + \frac{\delta}{1-\delta} \int_{\underline{y}}^{\overline{y}} [y - W(y) - \phi(e_P)] f(y|h) dy \geq 0 + \frac{\delta}{1-\delta} \overline{\pi},$$

where incentive constraints, (IC_A) and (IC_P), ensure that parties choose the efforts that maximize their utility. (DEC_A) and (DEC_P) constraints guarantee that the parties' expected payoff from the interaction in the contractual relationship is higher than what they would receive from pursuing their individual reservation utility or outside options.

To characterize a PPE, I employ a factorization technique, as called by Abreu et al. (1986, 1990), or a decomposition to stationary contracts, as termed by Levin (2003). This approach involves a characterization of a PPE in terms of payoffs rather than strategies. The idea is that in any Perfect Bayesian Equilibrium, the rewards can be delineated into current and future payoffs. Within a PPE, all continuation payoffs must align with the PPE scenarios. These payoffs can then be further decomposed, creating a recursive structure. The concept is akin to a principal-agent problem, where the principal motivates the agent by offering certain incentives or penalties linked to future rewards. The challenge lies in ensuring that these promised incentives and penalties correspond to payoffs within a PPE of the continuation game, rather than being monetary payoffs specified in an enforceable court contract. Thus, I rely on Theorem 2 from Levin (2003), which asserts that "if an optimal contract exists, there is a stationary contract that is optimal." This theorem allows to simplify the problem and find an optimal relational contract by focusing on stationary contracts.

The optimal relational contract

It is well known from Thomas and Worrall (1988) that once the participation constraints are hit in all states, the contract is determined only by the current state and no longer by the history. To decompose the relational contract into a stationary one I need to impose additional constraints. These constraints ensure that the parties' efforts sustain in a PPE. Therefore, I develop Proposition 1 based on Levin (2003)'s Theorem 3. Proposition 1 delineates these conditions. It adds (IC_P) as an additional constraint to Levin (2003)'s Theorem 3 due to the unobservability of principal's efforts.

The proof follows a similar approach to that of Theorem 3 in Levin (2003), as the additional principal's effort does not affect the necessity and sufficiency of the dynamic enforcement constraint (DEC) that binds the parties' variation in contingent payments by the future gains from the relationship. Thus, I confirm one of Levin (2003)'s key findings, that there exists a stationary contract with identical payoffs for any set of non-stationary actions and transfers. Moreover, I demonstrate that this finding holds true when both parties' efforts are unobservable.

Proposition 1. *Parties efforts, e_P and e_A, that generate expected surplus s can be implemented with a stationary contract if and only if there is a payment schedule $W : Y \to \mathbb{R}$ satisfying:*

$$(IC_A) \qquad e_A \in \arg\max_{e_A \in E_A} \mathbb{E}_y \left[W(y)|h\right] - c(e_A),$$

$$(IC_P) \qquad e_P \in \arg\max_{e_P \in E_P} \mathbb{E}_y \left[y - W(y)|h\right] - \phi(e_P),$$

$$(DEC) \qquad \frac{\delta}{1-\delta}(s - \bar{s}) \geq \sup_{y \in y} W(y) - \inf_{y \in y} W(y).$$

Proof. Consider a self-enforcing contract with efforts, e_P and e_A, payments $W(y) = w + b(y)$, and per-period payoffs (u, π). My aim is to prove that (IC_A), (IC_P), and (DEC) are necessary conditions to make the continuation contract $W(y)$ self-enforcing and the payoff pair (u, π) capable of sustaining PPE. To achieve this, I must prove that the parties' efforts satisfy conditions (i) - (v).

First, feasibility constraint (v) is met by the stage game setting. Second, each period the principal and the agent can choose any $e_P \in E_P$ and $e_A \in E_A$ respectively. Therefore, the (IC_A) and (IC_P) constraints serve as necessary conditions for self-enforcement, encouraging parties to choose efforts that facilitate the continuation of their relationship into the future. Thus, constraints (ii) and (iii) are met. Third, as either party can renege on discretionary payment and exit the relationship, (iv) must be satisfied:

$$(DEC_A) \qquad \delta(u-\overline{u}) \geq -(1-\delta)b(y), \qquad \text{and} \qquad (DEC_P) \qquad \delta(\pi-\overline{\pi}) \geq (1-\delta)b(y).$$

This is the constraint (iv) in stationary environment since $u(y) = u$ and $\pi(y) = \pi$. If (iv) is satisfied, then (i) is satisfied implicitly. Because the output varies, the bonus fluctuates accordingly. As a result, there exists a maximum value of the bonus that the principal is willing to pay ($\sup_{y\in Y} b(y)$), as well as a minimum value that the agent is willing to accept to refrain from quitting ($\inf_{y\in Y} b(y)$). Combining (DEC_P) and (DEC_A), in a stationary contract, I obtain (DEC):

$$(DEC) \qquad \frac{\delta}{1-\delta}(s-\bar{s}) \geq \sup_{y\in Y} b(y) - \inf_{y\in Y} b(y).$$

Without loss of generality, since $W(y) = w + b(y)$ I rewrite the constraint as:

$$\frac{\delta}{1-\delta}(s-\bar{s}) \geq \sup_{y\in Y} W(y) - \inf_{y\in Y} W(y).$$

Suppose there is a payment schedule $W(y)$ and efforts that satisfy (IC_A), (IC_P) and (DEC). Let define a fixed payment as $w \equiv \bar{u} - \mathbb{E}_y\left[b(y) - c(e_A)\right]$ and bonus payment as $b(y) \equiv W(y) - \inf_{y\in y} b(y)$. In a stationary contract with w and $b(y)$, and efforts e_P and e_A, and deviations punished with a reversion to the static equilibrium. This contract gives per-period payoffs $\bar{u}$ to the agent and $\pi \equiv s - \bar{u}$ to the principal. By (DEC), $s \geq \bar{s}$ and, therefore, $\pi \geq \bar{\pi}$, meaning both parties are willing to initiate the contract. Moreover (IC_P) and (IC_A) imply that the principal and the agent prefer e_P and e_A respectively to any other $e_P \in E_P$ and $e_A \in E_A$. Defining a fixed payment differently through the principal's reservation utility would provide the same results. $\square$

By the definition of stationary relational contract, in every period $e_{A_t} = e_A$, $e_{P_t} = e_P$, $b_t = b(y)$ and $w_t = w$ on the equilibrium path. That is, parties effort rules, the fixed wage and the output contingent bonus do not change over time. In other words, I fix a relational contract ($\sigma_A, \sigma_P, \zeta_A, \zeta_P$), letting all other variables be also fixed, i.e. w is the wage under this contract, $b(y)$ is the bonus under this contract, e_A, e_P - parties efforts, and u- agent's payoff π - principal's payoff under this contract and s - parties common joint value. Thus, following Proposition 1, the highest and the lowest contract or payment that would refrain parties can be expressed numerically, i.e. $\bar{W}$ and $\underline{W}$, namely,

$$(DEC) \qquad \frac{\delta}{1-\delta}(s-\bar{s}) \geq \bar{W} - \underline{W}.$$

Hence, given Proposition 1, under DMH I am able to rewrite the problem in a stationary environment. Particularly, a stationary optimal contract $\{e_P^*, e_A^*, W^*((y) = w + b(y)\}$ is the solution to the problem:

$$\max_{W(y)} s = \int_{\underline{y}}^{\overline{y}} \left[y - c(e_A^*) - \phi(e_P^*)\right] f(y|h)dy$$

$$\text{s.t.}$$

$$(IC_A) \qquad e_A^* \in \arg\max_{e_A} \int_{\underline{y}}^{\overline{y}} \left[W(y) - c(e_A)\right] f(y|h)dy,$$

$$(IC_P) \qquad e_P^* \in \arg\max_{e_P} \int_{\underline{y}}^{\overline{y}} \left[y - W(y) - \phi(e_P)\right] f(y|h)dy,$$

$$(DEC) \qquad \frac{\delta}{1-\delta}(s-\bar{s}) - \left(\bar{W} - \underline{W}\right) \geq 0,$$

Where (DEC) enforces the condition that the parties' mutual future gains outweigh their temptation to break the contract. Equation (DEC), along with (IC_P) and (IC_A), guarantees the existence of a perfect Bayesian equilibrium by safeguarding against reneging temptations.

Due to the FOA validity, the agent's incentive constraint (IC_A) and the principal's incentive constraint (IC_P) can be rewritten as follows:

$$e_A^* \in \arg\max_{e_A} \int_{\underline{y}}^{\overline{y}} \left[W(y) - c(e_A)\right] f(y|h)dy \iff$$

$$0 = \frac{d}{de_A} \int_{\underline{y}}^{\overline{y}} \left[W(y) - c(e_A)\right] f(y|h)dy \iff$$

$$c'(e_A^*) = \frac{d}{de_A} \int_{\underline{y}}^{\overline{y}} W(y) f(y|h)dy;$$

$$e_P^* \in \arg\max_{e_P} \int_{\underline{y}}^{\overline{y}} \left[y - W(y) - \phi(e_P)\right] f(y|h)dy \iff$$

$$0 = \frac{d}{de_P} \int_{\underline{y}}^{\overline{y}} \left[y - W(y) - \phi(e_P)\right] f(y|h)dy \iff$$

$$\phi'(e_P^*) = \frac{d}{de_P} \int_{\underline{y}}^{\overline{y}} [y - W(y)] \, f(y|h) dy.$$

Without loss of generality, since (IC_P) and (IC_A) bind, I combine them to obtain a common incentive constraint (CIC). (CIC) is satisfied, i.e. necessary and sufficient for the solution. Refer to Cong and Zhou (2021) for detailed explanations. This approach facilitates the solution for μ, the moral hazard rate of parties, and allows for the straightforward computation of parties' optimal efforts.[18] In contrast, having different incentive constraints for parties would result in their moral hazard levels depending on each other. I note that this issue does not arise when solving for the DMH in a static environment, as shown in Kim and Wang (1998). Hence, the combined (CIC) constraint can be expressed as follows:

$$(CIC) \qquad c'(e_A^*) + \phi'(e_P^*) = \frac{d}{de_A} \int_{\underline{y}}^{\overline{y}} W(y) f(y|h) dy + \frac{d}{de_P} \int_{\underline{y}}^{\overline{y}} [y - W(y)] \, f(y|h) dy.$$

Then, the optimal stationary contract $\{e_A^*, e_P^*, W^*(y) = w + b(y)\}$ solves the following problem:

$$\max_{W(\cdot)} s = \int_{\underline{y}}^{\overline{y}} [y - c(e_A^*) - \phi(e_P^*)] \, f(y|h) dy$$

s.t.

$$(CIC) \qquad c'(e_A^*) + \phi'(e_P^*) = \frac{d}{de_A} \int_{\underline{y}}^{\overline{y}} W(y) f(y|h) dy + \frac{d}{de_P} \int_{\underline{y}}^{\overline{y}} [y - W(y)] \, f(y|h) dy,$$

$$(DEC) \qquad \frac{\delta}{1-\delta} (s^* - \overline{s}) \geq \left(\bar{W} - \underline{W}\right).$$

Since $W(y)$ is a continuous function from $\mathbb{R}$ to $[\underline{W}, \bar{W}]$, this problem is linear in W. Therefore, I consider the Lagrangian function for the problem. Let L be the Lagrange function of the problem. Then,

$$L(W, e_A, e_P, \mu, \lambda) = \int_{\underline{y}}^{\overline{y}} [y - c(e_A^*) - \phi(e_P^*)] \, f\left(y|h\right) dy$$
$$+\mu[-c'(e_A^*) - \phi'(e_P^*) + \frac{d}{de_A} \int_{\underline{y}}^{\overline{y}} W(y) f(y|h) dy + \frac{d}{de_P} \int_{\underline{y}}^{\overline{y}} [y - W(y)] \, f(y|h) dy]$$
$$+ \lambda[\frac{\delta}{1-\delta}(s^* - \overline{s}) - \left(\overline{W} - \underline{W}\right)],$$

[18]Here, I refer to the solution of the Lagrangian introduced in Appendix B2.

where μ and λ are Kuhn-Tucker multipliers. μ represents the moral hazard degree of the parties. The pointwise differentiation yields necessary conditions for a solution to the problem:

$$\begin{aligned}\frac{\partial L}{\partial W} =& \mu h_{e_A} f_h(y|h) - \mu h_{e_P} f_h(y|h) \iff \\ =& \underbrace{\mu(h_{e_A} - h_{e_P}) f_h(y|h)}_{z(y,e_A,e_P)} = 0.\end{aligned} \tag{5}$$

Since we are in a scenario where both parties have equal moral hazard degrees, μ is always nonnegative. Given the assumption that the marginal rates are always greater than zero and never equal to each other ($h_{e_A} - h_{e_P} \neq 0$), the marginal effects of the agent's and principal's efforts on the output cannot be the same. Hence, there are four possible scenarios:

Scenario 1: when $h_{e_A} > h_{e_P}$ and $f_h(y|h)$ is increasing in y,

Scenario 2: when $h_{e_A} < h_{e_P}$ and $f_h(y|h)$ is increasing in y,

Scenario 3: when $h_{e_A} > h_{e_P}$ and $f_h(y|h)$ is decreasing in y,

Scenario 4: when $h_{e_A} < h_{e_P}$ and $f_h(y|h)$ is decreasing in y.

The hint, MLRP assumption, provided by Innes (1990) in the problem of moral hazard with limited liability, similar to the one utilized by Levin (2003) in the moral hazard scenario within a dynamic environment decomposed into a stationary one would not work here. MLRP posits that the density function is always increasing in y for a fixed e. In my case, with the joint production function MLRP is: $\forall h \in \mathbb{R}$, $\frac{\partial}{\partial y}\left(\frac{f_h(y|h)}{f(y|h)}\right) > 0$, hence it does not guarantee that $f_h(y|h)$ is always increasing.

Under the same degree of moral hazard problem on the parties' sides, when FOA is valid, the optimal contract that generates a second-best optimal parties' efforts, e_P and e_A, depends on the trade-off between the marginal effect of parties' efforts on the output ($h_{e_A} \lesseqgtr h_{e_P}$) and the behavior of the density function $f_h(y|h) \lesseqgtr 0$.

Scenario 1 Under *Scenario 1*, equation (5) yields:

$$
\begin{aligned}
&z(y, e_A, e_P) = \mu(h_{e_A} - h_{e_P}) f_h(y|h) \\
&z(y, e_A, e_P) > 0 \rightarrow W(y) = \bar{W} = \underline{W} + \frac{\delta}{1-\delta}(s - \bar{s}) \\
&z(y, e_A, e_P) = 0 \rightarrow W(y) \in [\underline{W}, \bar{W}] \\
&z(y, e_A, e_P) < 0 \rightarrow W(y) = \underline{W}
\end{aligned}
\tag{6}
$$

In (6), one expects to have the proper value of surplus s. In this study, I prioritize the assumption $h_{e_A} > h_{e_P}$ due to its alignment with existing literature on moral hazard, where only the effort of the agent impacts the payoff distribution; its correspondence to the specific real-world scenario detailed in section 3; and its plausibility supported by the assumptions $F_h < 0$, indicating that greater joint effort h results in higher payoffs, and $f_h > 0$, suggesting a diminishing impact of joint effort h on higher payoffs. Nevertheless, I plan future extensions to encompass *Scenario 2* through *Scenario 4*.

Proposition 2. *When $h_{e_A} > h_{e_P}$ and $f_h(y|h(e_A, e_P)) > 0$, an optimal contract is "one-step", i.e. there is some $\hat{y}$ with $W(y) = \underline{W}$ if $y \leq \hat{y}$ and $W(y) = \bar{W}$ if $y \geq \hat{y}$.*

The proof of Proposition 2 is similar to Levin (2003)'s Theorem 5, where the optimal contract is a "one-step" contract. Follow the discussion below for the solution. A "one-step" contract allows to rewrite the problem with a step function as follows:[19]

$$
\max_{\underline{W}, \bar{W}, \hat{y}} s = \int_{\underline{y}}^{\bar{y}} \left[y - c(e_A^*) - \phi(e_P^*)\right] f(y|h) dy
$$

s.t.

$$
(CIC) \qquad c'\left(e_A^*\right) + \phi'\left(e_P^*\right) = h_{e_P} \int_{\underline{y}}^{\bar{y}} y f_h(y|h) dy + m(h_{e_P} - h_{e_A}) F_h\left(\hat{y}|h\right),
$$

$$
(DEC) \qquad \frac{\delta}{1-\delta}\left(s^* - \bar{s}\right) \geq m,
$$

where $m = \bar{W} - \underline{W}$.

[19] Refer to Appendix B1 for the detailed simplification of (CIC) constraint.

Let L_1 be the Lagrange function of the problem under *Scenario 1*:

$$L_1(\underline{W}, \bar{W}, \hat{y}, e_A^*, e_P^*, \mu, \lambda) = \int_{\underline{y}}^{\overline{y}} [y - c(e_A^*) - \phi(e_P^*)] f(y|h) \, dy$$
$$+\mu[-c'(e_A^*) - \phi'(e_P^*) + h_{e_P} \int_{\underline{y}}^{\bar{y}} y f_h(y|h) dy + m(h_{e_P} - h_{e_A}) F_h(\hat{y}|h)]$$
$$+ \lambda[\frac{\delta}{1-\delta}(s^* - \overline{s}) - m],$$

Refer to Appendix B2 for the solution of the Lagrangian. Therefore, the final program consists of four equations and four unknowns:

$$(I)\ f_h(\hat{y}|h) = 0,$$
$$(II)\ c'(e_A) + \phi'(e_P) = h_{e_P} \int_{\underline{y}}^{\overline{y}} y f_h(y|h) dy + m F_h(\hat{y}|h)(h_{e_P} - h_{e_A}),$$
$$(III)\ h_{e_A} \int_{\underline{y}}^{\overline{y}} y f_h(y|h) dy - c'(e_A) + \frac{1-\delta}{\delta(h_{e_P} - h_{e_A}) F_h(\hat{y}|h)}$$
$$\times \Bigg[- c''(e_A) + h_{e_P} h_{e_A} \int_{\underline{y}}^{\overline{y}} y f_{hh}(y|h) dy$$
$$+ m\left(F_{hh}(\hat{y}|h) h_{e_A}(h_{e_P} - h_{e_A}) - F_h(\hat{y}|h) h_{e_A e_A}\right) \Bigg] = 0,$$
$$(IV)\ h_{e_P} \int_{\underline{y}}^{\overline{y}} y f_h(y|h) \, dy - \phi'(e_P) + \frac{1-\delta}{\delta(h_{e_P} - h_{e_A}) F_h(\hat{y}|h)}$$
$$\times \Bigg[- \phi''(e_P) + h_{e_P e_P} \int_{\underline{y}}^{\overline{y}} y f_h(y|h) dy + h_{e_P} h_{e_P} \int_{\underline{y}}^{\overline{y}} y f_{hh}(y|h) dy$$
$$+ m\left(F_{hh}(\hat{y}|h) h_{e_P}(h_{e_P} - h_{e_A}) + F_h(\hat{y}|h) h_{e_P e_P}\right) \Bigg] = 0.$$

5 Comparative statistics

In this section, I provide a numerical example of the optimal contracts as outlined in Proposition 2, where $h_{e_A} > h_{e_P}$ and $f_h(y|h) > 0$. I utilize the distribution function from Spaeter (1998), which satisfies FOA.[20] I transform the distribution in two ways: multiplying it by a variable parameter $\eta > 0$ to understand the changes in the optimal contract with variations in the distribution sensitivity of the outcome to changes in parties' effort; using a

[20] Refer to the Appendix C1 for details on the FOA validity of the distribution and density function from Spaeter (1998).

joint production function instead of the effort level of a party.

Namely, I employ the distribution function, where $\underline{y} = 0$ and $\overline{y} = 1$,

$$F(y|h) = y\left(\frac{\eta(1-y)}{h+1} + 1\right),$$

and thus, the density function becomes

$$f(y|h) = \frac{d}{dy}F(y|h) = \frac{h+1+\eta(1-2y)}{h+1}.$$

With the aforementioned distribution, equation (I) provides the threshold level $\hat{y}$, where the density function changes its direction: the marginal benefit is negative for all $y < \hat{y}$ and positive for all $y > \hat{y}$. Thus,

$$(I)\ f_h(\hat{y}|h) = 0 \Leftrightarrow \frac{(2\hat{y}-1)\,\eta}{(h+1)^2} = 0 \Leftrightarrow \hat{y} = \frac{1}{2}.$$

Given $\hat{y} = \frac{1}{2}$ and using equations (II), (III), and (IV), I compute the optimal difference between the highest and lowest wages m^*, the agent's optimal effort e_A^* and the principal's optimal effort e_P^*:[21]

$$m^* = \frac{2(\eta h_{e_P} - 6(e_A^* + e_P^*)(h+1)^2)}{3\eta(h_{e_P} - h_{e_A})}, \tag{7}$$

$$e_A^* = \frac{\eta h_{e_A}}{6(h+1)^2} + \frac{1-\delta}{\delta}\times \left[\frac{12(h+1)^3 + 4\eta h_{e_P}h_{e_A} - 3m^*(h_{e_Ae_A}\eta(h+1) + (h_{e_P} - h_{e_A})2\eta h_{e_A})}{3(h+1)\eta(h_{e_P} - h_{e_A})}\right], \tag{8}$$

$$e_P^* = \frac{\eta h_{e_P}}{6(h+1)^2} + \frac{1-\delta}{\delta}\times \left[\frac{12(h+1)^3 - 2\eta h_{e_Pe_P} + 4\eta h_{e_P}h_{e_P} - 3m^*(h_{e_Pe_P}\eta(h+1) + (h_{e_P} - h_{e_A})2\eta h_{e_P})}{3(h+1)\eta(h_{e_P} - h_{e_A})}\right]. \tag{9}$$

Since m^* depends on δ through e_A^* or e_P^*, I discuss the sign of m^* for $\forall\delta$. By definition $\eta, e_A^*, e_P^*, h_{e_P}, h_{e_A} > 0$, and from Proposition 2 $h_{e_A} > h_{e_P}$, hence, from the equality (7), the sign of m^* depends on the trade-off between ηh_{e_P} and $6(e_A^* + e_P^*)(h+1)^2$. Thus, $m^* < 0$ if and only if $\eta^{-1}6(e_A^* + e_P^*)(h+1)^2 < h_{e_P} < h_{e_A}$. In contrast, $m^* > 0$ if and only if $h_{e_P} < h_{e_A}$

[21]Appendix C2 provides a detailed derivation of equations for the chosen distribution.

and $h_{e_P} < \eta^{-1} 6(e_A^* + e_P^*)(h+1)^2$.

When the output's responsiveness to the parties' efforts is sufficiently high and the principal's contribution to the output is relatively small, it is advantageous to opt for a contract where the agent primarily receives the output and reimburses the principal with a designated bonus. Conversely, when the principal's contribution rate is comparatively low in relation to the sensitivity of the output, it is favorable for the principal to be the primary recipient of the output and share the bonus with the agent.

Further, I investigate the variation of (8) and (9) with respect to δ. Let consider two cases: A) theoretically, when the parties are very patient ($\delta \to 1$), and B) using simulations for all other scenarios due to the complexity of the equations.

Case A: $\delta \to 1$

As $\delta \to 1$, e_A^*, and e_P^* can be simplified as follows:

$$e_A^* = \frac{\eta h_{e_A}}{6\,(h+1)^2} \qquad \text{and} \qquad e_P^* = \frac{\eta h_{e_P}}{6\,(h+1)^2}. \tag{10}$$

I conclude that when parties are highly patient, their optimal effort levels, aimed at sustaining the relationship, vary in tandem with the sensitivity of the healthcare output to parties' efforts, i.e. as η enlarges, both e_A^* and e_P^* increase in the equilibrium. More precisely, with growth in output sensitivity, the optimum is reached at a higher level of parties' efforts. Nevertheless, I can not make any conclusion about the relation between the parties' optimal efforts and the marginal rate of their contribution. For example, if $h = ae_A^* + be_P^*$, where a and b are positive, that correspond to the complimentary of the parties efforts covered in this paper, an increase in either h_{e_A} or h_{e_P} leads to a decrease in the effort level of the respective party at the optimum. However, this is not always true, if the sign for a or b changes or a different function is introduced, the dependency can differ. Moreover, straightforward computations of optimal effort complicate the equations without producing reliable results.[22]

[22]Follow the Appendix C3 for a trial.

Case B: $\forall\delta$

For $\forall\delta$ and a linear joint production function $h = ae_A^* + be_P^*$, equations (III) and (IV) are as follows:

$$
\begin{aligned}
m^* &= \frac{2(\eta b - 6(e_A^* + e_P^*)(ae_A^* + be_P^* + 1)^2)}{3\eta(b-a)},\\
e_A^* &= \frac{a\eta}{6(ae_A^* + be_P^* + 1)^2} + \frac{1-\delta}{\delta}\left[\frac{12(ae_A^* + be_P^* + 1)^3 + 4\eta ab - 3m^*(b-a)2a\eta}{3(ae_A^* + be_P^* + 1)\eta(b-a)}\right],\\
e_P^* &= \frac{b\eta}{6(ae_A^* + be_P^* + 1)^2} + \frac{1-\delta}{\delta}\left[\frac{12(ae_A^* + be_P^* + 1)^3 + 4\eta b^2 - 3m^*(b-a)2b\eta}{3(ae_A^* + be_P^* + 1)\eta(b-a)}\right].
\end{aligned}
$$

Based on simulations conducted in Mathematica, I am able to provide the conclusions below. The figures in the Appendix (Fig. E1 and Fig. E2) correspond to the model discussed above. They reveal that the sign of m^*, which determines the party receiving the main output and deciding on bonus payment, depends not only on its marginal contribution (a classical interpretation in the trade where the principal's marginal contribution to output is higher than that of the agent results in $m^* > 0$, and vice versa for $m^* < 0$) but also on the inequality sign in the condition $\eta h_{e_P} \lesseqgtr 6(e_A^* + e_P^*)(h+1)^2$.

Both Fig. E1 and Fig. E2 in the Appendix yield $m^* > 0$. Therefore, I examine a scenario in which the principal determines the agent's bonus, where $h_{e_A} > h_{e_P}$ and $h_{e_P} < \eta^{-1}6(e_A^* + e_P^*)(h+1)^2$.

Let consider a numerical example from Fig. E1 in the Appendix, where the agent's contribution to the output is ten times higher than the principal's ($h = 50e_A^* + 5e_P^*$), parties are very patient ($\delta = 0.9$), and the sensitivity of the output to parties' efforts (η) varies. When the distribution sensitivity equals 10, the optimal relational contract establishes that the principal should exert 0.31 of her effort unit and pay to the agent either 13.56 wage units or 12.83 wage units, depending on whether the jointly generated output exceeds $1/2$. Increasing the sensitivity to 60 alters the requirements of the optimal relational contract. It now specifies that the principal should exert more effort in the optimum ($e_P^* = 0.57$) because, with increased sensitivity, the achieved output level is even more dependent on her efforts. The payment also changes: she would now pay 13.72 units if the output exceeds the threshold, and she should pay 12.98 units if it does not exceed. An interpretation is as follows, since the agent is the main contributor to the output, with the growth of output sensitivity to parties' efforts, the principal needs to motivate even more the agent, otherwise, his deviation will be very costly. It is important to note that the upper wage bound increases faster than the minimum wage bound, leading to a positive m^*. Since the agent is very valuable, the principal wants to pay a higher upper wage bound as the agent's contribution increases.

In other words, when an additional unit of external enforcement by the principal results in a lower healthcare output value than an additional unit of effort by the agent, in an environment where healthcare output does not react sensitively to the parties' efforts (such as facility management services supply), both external and internal enforcement should not be intense. In contrast, in more sensitive environments (such as surgery equipment supply), where outcomes directly impact people's lives and healthcare output, both types of enforcement should be increased.[23]

Nevertheless, the surplus achieved in equilibrium in a more sensitive environment is seven times lower than the surplus achieved in a less sensitive environment. This difference primarily relates to the fact that additional parties' efforts necessitate higher costs, thus decreasing the surplus.

Furthermore, Fig. E2 shows what happens as parties' patience changes when the agent's contribution to the output is ten times higher than the principal's ($h = 50e_A^* + 5e_P^*$), and parties' sensitivity to the output is fixed ($\eta = 50$). The result corresponds to the behavior of the dynamic enforcement constraint in Levin (2003). Specifically, the tightness of the restriction (m) depends on the discount factor (δ): as $\delta \to 1$, the range of payments (m) is unbounded, and as $\delta \to 0$, the m is limited but also cannot even be provided. This is what Fig. E2 produces. As δ increases, m^* enlarges, becoming indeed unbounded for δ approaching one. The relationship can only be sustained by very patient parties. The equilibrium does not exist for non-patient parties, i.e., Mathematica provides negative efforts. For example, for $\eta = 50$, $\delta = 0.95$ is the minimum patience level that sustains the relationship (Fig. E2 in the Appendix).

Fig. E2 also illustrates that as parties become more patient, both the principal's external enforcement level and internal enforcement increase, while the wage level paid to the agent, both over and below the threshold, decreases. Understanding that the agent values this relationship and is engaged in a long-term collaboration, the principal refrains from providing additional motivation, considering the agent's intrinsic motivation. Similarly, as parties exhibit more patience, they are rewarded more. Consequently, s^* increases with the parties' growing patience.

[23] Follow the Appendix D1 for the detailed real-world scenarios.

6 Conclusion

This study relates economic theory to real-world scenarios in the healthcare industry. In particular, it describes the trade-off between the principal's external enforcement effort level and internal enforcement in the environment, varying by healthcare output sensitivity to parties' efforts and parties' patience. Moreover, it computes the bonus rates paid above and below the agreed output threshold, which should be determined by the optimal relational contract in the equilibrium.

I establish that for highly patient parties, in an environment where parties' efforts are not very sensitive to the output, both external and internal enforcement can be less pressing. In contrast, in a very sensitive environment, the principal should intensify external and internal enforcement. I plan to expand the paper by adding three more scenarios, incorporating a decreasing density function in healthcare output and varying parties' contribution to the healthcare output. Additionally, I aim to compare the simulation results of DMH with those of moral hazard computed from Levin (2003)'s model.

The findings of this study can be useful for governments or healthcare providers when choosing contracts for their relationships with suppliers. It can also be an effective tool to correctly motivate the agent to maximize the joint surplus and reach equilibrium, establishing the agent's motivation depending on his contribution to output.

Nevertheless, the model has two strong assumptions, the levelling of which could impact the results: first, an identical moral hazard degree on both parties' sides, and second, the Grimm-trigger strategy. Relaxation of the first assumption, for example, could accelerate the necessity of external enforcement efforts from the principal with the growth of the agent's hidden actions, leveraging it with a downgrade on internal enforcement. Relaxation of the second assumption would allow to examine what happens when, with the deviation of one party, the relationship does not end, but continues with possible contract renegotiation.

Furthermore, improving the model involves solving it in a dynamic environment without simplifying it to stationarity. The solution of the model with simplification to a stationary environment, as examined in this study, fixes a relational contract within a specific time period and allows all other variables to remain unchanged over time. Finally, another point to note is that the principal's effort complements the agent's effort rather than impacting it. Considering the opposite could lead to a more realistic discussion but would require a completely different theoretical model to develop.

7 Appendix

Appendix A

A1 Validity of the first order approach

The MLRP and the CDFC are known as the Mirrlees-Rogerson conditions. These conditions should be satisfied to make the FOA valid.[24] Levin (2003) uses this approach to simplify the solution of the program, allowing for the agent's incentive constraint to bind, i.e. equalise it to zero.[25]

MLRP property: given any two effort levels, $e, e' \in E$ with $e > e'$, the ratio $\frac{f(y|e)}{f(y|e')}$ is increasing in y, i.e. higher output is more indicative of the higher effort (Milgrom, 1981). Given continuous output y, continues effort e and twice continuously differentiable distribution $F(\cdot)$, which satisfies MLRP in e if

$$\frac{\partial}{\partial y}\frac{f_e(y|e)}{f(y|e)} = \frac{\partial}{\partial e}\ln f(y|e) \geq 0$$

for all $e > 0$ and $y \geq 0$, where $E\{y|e = 0\}$.[26] The condition says that for any outcomes $y' > y$, enlarges effort increases the log density at y' more than at y. Roughly, more effort makes low outcomes more likely. MLRP states that the likelihood ratio $f_e(y|e)/f(y|e)$ must be non-decreasing in the output y: it is more likely to observe large revenues for a high level of effort.

CDFC property: A distribution $F(y|e)$ satisfies CDFC condition if

$$F_{ee}(y|e) \geq 0$$

for all (y, e). CDFC requires the distribution function to be convex in effort.

[24]Follow Rogerson (1985) for the proofs on FOA validity.

[25]Jewitt (1988) suggests an alternative set of conditions that should be satisfied to make FOA valid. Compared to Levin (2003), these conditions do not require strong output distributional assumption while leaving the agent's utility unconstrained. Instead, they allow for milder restrictions on output distribution and the utility function.

[26]MLRP implies First Order Stochastic Dominance (FOSD) property, $F_e(y|e) \leq 0$, that tells that higher effort makes the higher output more likely, and it guarantees that there is always a benefit of higher effort levels, gross of effort costs.

Appendix B: Proposition 2

B1 Detailed simplification of (CIC) constraint

$$(CIC) \qquad c'(e_A^*) + \phi'(e_P^*) = \frac{d}{de_A}\int_{\underline{y}}^{\overline{y}} W(y)f(y|h)dy + \frac{d}{de_P}\int_{\underline{y}}^{\overline{y}} \left[y - W(y)\right] f(y|h)dy \iff$$

$$c'(e_A^*) + \phi'(e_P^*) = \frac{d}{de_A}\int_{\underline{y}}^{\hat{y}} \underline{W} f dy + \frac{d}{de_A}\int_{\hat{y}}^{\overline{y}} \bar{W} f dy + \frac{d}{de_P}\int_{\underline{y}}^{\hat{y}} (y - \underline{W}) f dy + \frac{d}{de_P}\int_{\hat{y}}^{\overline{y}} (y - \bar{W}) f dy.$$

Detailed simplification of the two parts of (CIC) constraint:

$\frac{d}{de_A}\int_{\underline{y}}^{\hat{y}} \underline{W} f(y|h)dy + \frac{d}{de_A}\int_{\hat{y}}^{\overline{y}} \overline{W} f(y|h)dy = \frac{d}{de_A}[\underline{W}F(\hat{y}|h)] + \frac{d}{de_A}[\overline{W}(1 - F(\hat{y}|h))] = \underline{W}h_{e_A}F_h(\hat{y}|h) - \overline{W}h_{e_A}F_h(\hat{y}|h) = [\underline{W} - \overline{W}]h_{e_A}F_h(\hat{y}|h)$. Given that $m = \overline{W} - \underline{W},$ then $[\underline{W} - \overline{W}]h_{e_A}F_h(\hat{y}|h) = -mh_{e_A}F_h(\hat{y}|h)$;

$\frac{d}{de_P}\int_{\underline{y}}^{\hat{y}} [y - \underline{W}] f(y|h)dy + \frac{d}{de_P}\int_{\hat{y}}^{\overline{y}} [y - \overline{W}] f(y|h)dy = \frac{d}{de_P}\int_{\underline{y}}^{\hat{y}} y f(y|h)dy + \frac{d}{de_P}\int_{\hat{y}}^{\overline{y}} y f(y|h)dy - \left[\frac{d}{de_P}[\underline{W}F(\hat{y}|h)] + \frac{d}{de_P}[\overline{W}(1 - F(\hat{y}|h))]\right] = h_{e_P}\int_{\underline{y}}^{\overline{y}} y f_h(y|h)dy - \left[\underline{W}h_{e_P}F_h(\hat{y}|h) - \overline{W}h_{e_P}F_h(\hat{y}|h)\right] = h_{e_P}\int_{\underline{y}}^{\overline{y}} y f_h(y|h)dy - [\underline{W} - \overline{W}]h_{e_P}F_h(\hat{y}|h)$. Given that $m = \overline{W} - \underline{W},$ then $h_{e_P}\int_{\underline{y}}^{\overline{y}} y f_h(y|h)dy + mh_{e_P}F_h(\hat{y}|h)$.

Thus, (CIC) constraint can be rewritten as:

$$c'\left(e_A^*\right) + \phi'\left(e_P^*\right) = h_{e_P}\int_{\underline{y}}^{\overline{y}} y f_h(y|h)dy + m(h_{e_P} - h_{e_A})F_h\left(\hat{y}|h\right).$$

B2 Solution of the Lagrangian

$$
\begin{aligned}
L_1(\underline{W}, \bar{W}, \hat{y}, e_A^*, e_P^*, \mu, \lambda) = & \int_{\underline{y}}^{\overline{y}} [y - c(e_A^*) - \phi(e_P^*)]\, f(y|h)\, dy \\
& + \mu[-c'(e_A^*) - \phi'(e_P^*) + h_{e_P} \int_{\underline{y}}^{\bar{y}} y f_h(y|h) dy + m(h_{e_P} - h_{e_A}) F_h(\hat{y}|h)] \\
& + \lambda[\frac{\delta}{1-\delta}(s^* - \overline{s}) - m],
\end{aligned}
$$

From the Envelop theorem on s^*,

$$
\frac{\partial L}{\partial s^*} = \lambda \frac{\delta}{1-\delta} = 1. \tag{11}
$$

Pointwise differentiation provides the following equations:

$$
\frac{\partial L}{\partial m} = \mu F_h(\hat{y}|h)(h_{e_P} - h_{e_A}) - \lambda = 0, \tag{12}
$$

$$
\frac{\partial L}{\partial \hat{y}} = \mu m f_h(\hat{y}|h)(h_{e_P} - h_{e_A}) = 0. \tag{13}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial e_A} = & h_{e_A} \int_{\underline{y}}^{\overline{y}} y f_h(y|h) dy - c'(e_A) - \mu c''(e_A) \\
& + \mu(h_{e_P} \frac{\partial \int_{\underline{y}}^{\overline{y}} y f_h(y|h) dy}{\partial e_A} \\
& + m(\frac{\partial F_h(\hat{y}|h)}{\partial e_A}(h_{e_P} - h_{e_A}) + F_h(\hat{y}|h)\frac{\partial(-h_{e_A})}{\partial e_A})) = 0,
\end{aligned} \tag{14}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial e_P} = & h_{e_P} \int_{\underline{y}}^{\overline{y}} y f_h(y|h)\, dy - \phi'(e_P) - \mu \phi''(e_P) \\
& + \mu(h_{e_P e_P} \int_{\underline{y}}^{\overline{y}} y f_h(y|h) dy + h_{e_P} \frac{\partial \int_{\underline{y}}^{\overline{y}} y f_h(y|h) dy}{\partial e_P} \\
& + m(\frac{\partial F_h(\hat{y}|h)}{\partial e_P}(h_{e_P} - h_{e_A}) + F_h(\hat{y}|h)\frac{\partial h_{e_P}}{\partial e_P})) = 0,
\end{aligned} \tag{15}
$$

From (11) $\lambda = \frac{1-\delta}{\delta}$. Therefore,

$$(12) \Rightarrow \mu = \frac{1-\delta}{\delta(h_{e_P} - h_{e_A})F_h(\hat{y}|h)} \tag{16}$$

$$\begin{aligned}
&(14) \text{ and } (16) \Rightarrow \\
&h_{e_A}\int_{\underline{y}}^{\overline{y}} yf_h(y|h)dy - c'(e_A) + \frac{1-\delta}{\delta(h_{e_P} - h_{e_A})F_h(\hat{y}|h)} \\
&\times[-c''(e_A) + h_{e_P}h_{e_A}\int_{\underline{y}}^{\overline{y}} yf_{hh}(y|h)dy \\
&+ m\left(F_{hh}(\hat{y}|h)h_{e_A}(h_{e_P} - h_{e_A}) - F_h(\hat{y}|h)h_{e_Ae_A}\right)] = 0
\end{aligned} \tag{17}$$

$$\begin{aligned}
&(15) \text{ and } (16) \Rightarrow \\
&h_{e_P}\int_{\underline{y}}^{\overline{y}} yf_h\left(y|h\right)dy - \phi'(e_P) + \frac{1-\delta}{\delta(h_{e_P} - h_{e_A})F_h(\hat{y}|h)} \\
&\times[-\phi''(e_P) + h_{e_Pe_P}\int_{\underline{y}}^{\overline{y}} yf_h(y|h)dy + h_{e_P}h_{e_P}\int_{\underline{y}}^{\overline{y}} yf_{hh}(y|h)dy \\
&+ m\left(F_{hh}(\hat{y}|h)h_{e_P}(h_{e_P} - h_{e_A}) + F_h(\hat{y}|h)h_{e_Pe_P}\right)] = 0
\end{aligned} \tag{18}$$

$$\text{From } (13) \Rightarrow \text{ either } \mu = 0, \text{ or } m = 0,\ h_{e_P} - h_{e_A} = 0, \text{ or } f_h(\hat{y}|h) = 0. \tag{19}$$

Therefore, I have the final program with four equations and four unknowns:

$$\begin{aligned}
&(I)\ f_h(\hat{y}|h) = 0, \\
&(II)\ c'(e_A) + \phi'(e_P) = h_{e_P}\int_{\underline{y}}^{\overline{y}} yf_h(y|h)dy + mF_h(\hat{y}|h)(h_{e_P} - h_{e_A}), \\
&(III)\ h_{e_A}\int_{\underline{y}}^{\overline{y}} yf_h(y|h)dy - c'(e_A) + \frac{1-\delta}{\delta(h_{e_P} - h_{e_A})F_h(\hat{y}|h)} \\
&\times[-c''(e_A) + h_{e_P}h_{e_A}\int_{\underline{y}}^{\overline{y}} yf_{hh}(y|h)dy \\
&+ m\left(F_{hh}(\hat{y}|h)h_{e_A}(h_{e_P} - h_{e_A}) - F_h(\hat{y}|h)h_{e_Ae_A}\right)] = 0, \\
&(IV)\ h_{e_P}\int_{\underline{y}}^{\overline{y}} yf_h\left(y|h\right)dy - \phi'(e_P) + \frac{1-\delta}{\delta(h_{e_P} - h_{e_A})F_h(\hat{y}|h)} \\
&\times[-\phi''(e_P) + h_{e_Pe_P}\int_{\underline{y}}^{\overline{y}} yf_h(y|h)dy + h_{e_P}h_{e_P}\int_{\underline{y}}^{\overline{y}} yf_{hh}(y|h)dy \\
&+ m\left(F_{hh}(\hat{y}|h)h_{e_P}(h_{e_P} - h_{e_A}) + F_h(\hat{y}|h)h_{e_Pe_P}\right)] = 0.
\end{aligned}$$

Appendix C: Comparative statistics

C1 FOA validity with a modified Spaeter (1998)'s distribution

The FOA is valid if the CDFC and MLRP conditions are satisfied (Rogerson, 1985).

1) CDFC:

$$F_h(y|h) = \frac{d}{dh}\left(y\left(\frac{\eta(1-y)}{(h+1)}+1\right)\right) = \frac{\eta y\,(y-1)}{(h+1)^2},$$
$$F_{hh}(y|h) = \frac{d}{dh}\left(\frac{\eta y\,(y-1)}{(h+1)^2}\right) = \frac{2\eta y\,(1-y)}{(h+1)^3},$$

where $\eta > 0$, $y \in [0,1]$, and $h \in \mathbb{R}$. However, $h = ae_A + be_P$, where a and b are positive constants ($a, b > 0$), leading to $h > 0$. Therefore, $F_{hh}(y|h) \geq 0$, and CDFC is satisfied.

2) MLRP:

$$f_h(y|h) = \frac{d}{dh}\left(\frac{h+1+\eta(1-2y)}{h+1}\right) = \frac{\eta\,(2y-1)}{(h+1)^2},$$
$$\frac{f_h(y|h)}{f(y|h)} = \left(\frac{\eta(2y-1)}{(h+1)^2}\right) \times \left(\frac{h+1}{h+1+\eta(1-2y)}\right) = \frac{\eta(2y-1)}{(h+1)(1+h+\eta-2\eta y)},$$
$$\frac{d}{dy}\left(\frac{f_h(y|h)}{f(y|h)}\right) = \frac{d}{dy}\left(\frac{\eta(2y-1)}{(h+1)(1+h+\eta-2\eta y)}\right) = \frac{2\eta}{(1+h+\eta-2\eta y)^2},$$

where $\eta > 0$, $y \in [0,1]$ and $h > 0$. Hence, $\frac{d}{dy}\left(\frac{f_h(y|h)}{f(y|h)}\right) > 0$, and MLRP is satisfied.

C2 Derivation of Equations (II), (III), and (IV) with a modified Spaeter (1998)'s distribution

Given the distribution function $F(y|h) = y\left(\frac{\eta(1-y)}{h+1}+1\right)$ and the density function $f(y|h) = \frac{h+1+\eta(1-2y)}{h+1}$, I compute:

$$
\begin{aligned}
F_h(y|h) &= \frac{\eta y(y-1)}{(h+1)^2} \\
F_h(\hat{y}|h) &= -\frac{\eta}{4(h+1)^2} \\
F_{hh}(y|h) &= \frac{\eta 2y(1-y)}{(h+1)^3} \\
F_{hh}(\hat{y}|h) &= \frac{\eta}{2(h+1)^3} \\
f_h(y|h) &= \frac{\eta(2y-1)}{(h+1)^2} \\
f_{hh}(y|h) &= -\frac{2\eta(2y-1)}{(h+1)^3} \\
\int_{\underline{y}}^{\overline{y}} y f_h(y|h)dy &= \int_0^1 y\frac{\eta(2y-1)}{(h+1)^2}dy = \frac{\eta}{6(h+1)^2} \\
\int_{\underline{y}}^{\overline{y}} y f_{hh}(y|h)dy &= \int_0^1 -\frac{y2\eta(2y-1)}{(h+1)^3}dy = -\frac{\eta}{3(h+1)^3} \\
\int_{\underline{y}}^{\overline{y}} y f(y|h)dy &= \int_0^1 y\frac{h+1+\eta(1-2y)}{h+1}dy = \frac{3h-\eta+3}{6(h+1)}
\end{aligned}
$$

Therefore, Equations (II), (III), and (IV) transform as follows:

$$
\begin{aligned}
&(II)\ c'(e_A)+\phi'(e_P) = h_{e_P}\int_{\underline{y}}^{\overline{y}} y f_h(y|h)dy + mF_h(\hat{y}|h)(h_{e_P}-h_{e_A}) \Leftrightarrow \\
&e_A+e_P = h_{e_P}\frac{\eta}{6\left(h+1\right)^2} - m\frac{\eta}{4\left(h+1\right)^2}(h_{e_P}-h_{e_A}) \Leftrightarrow \\
&m\frac{\eta}{4\left(h+1\right)^2}(h_{e_P}-h_{e_A}) = -e_A-e_P+h_{e_P}\frac{\eta}{6\left(h+1\right)^2} \Leftrightarrow \\
&m\frac{\eta\left(h_{e_P}-h_{e_A}\right)}{4(h+1)^2} = \frac{-(e_A+e_P)6(h+1)^2+\eta h_{e_P}}{6(h+1)^2} \Leftrightarrow \\
&m = \frac{2(\eta h_{e_P}-6(e_A+e_P)(h+1)^2)}{3\eta\left(h_{e_P}-h_{e_A}\right)}.
\end{aligned}
$$

$$(III)\ h_{e_A}\int_{\underline{y}}^{\overline{y}} y f_h(y|h)dy - c'(e_A) + \frac{1-\delta}{\delta(h_{e_P}-h_{e_A})F_h(\hat{y}|h)}\times$$

$$\left[-c''(e_A) + h_{e_P}h_{e_A}\int_{\underline{y}}^{\overline{y}} y f_{hh}(y|h)dy + m\left(F_{hh}(\hat{y}|h)h_{e_A}(h_{e_P}-h_{e_A}) - F_h(\hat{y}|h)h_{e_Ae_A}\right)\right] = 0 \Leftrightarrow$$

$$0 = h_{e_A}\frac{\eta}{6\left(h+1\right)^2} - e_A - \frac{1-\delta}{\delta\left(h_{e_P}-h_{e_A}\right)}\frac{4(h+1)^2}{\eta}\times$$

$$\left[-1 - \frac{\eta h_{e_P}h_{e_A}}{3\left(h+1\right)^3} + m\left(\frac{\left(h_{e_P}-h_{e_A}\right)\eta h_{e_A}}{2\left(h+1\right)^3} + \frac{h_{e_Ae_A}\eta}{4\left(h+1\right)^2}\right)\right] \Leftrightarrow$$

$$0 = h_{e_A}\frac{\eta}{6\left(h+1\right)^2} - e_A + \frac{1-\delta}{\delta\left(h_{e_P}-h_{e_A}\right)}\frac{4(h+1)^2}{\eta}\times$$

$$\left[1 + \frac{\eta h_{e_P}h_{e_A}}{3\left(h+1\right)^3} - m\left(\frac{\left(h_{e_P}-h_{e_A}\right)\eta h_{e_A}}{2\left(h+1\right)^3} + \frac{h_{e_Ae_A}\eta}{4\left(h+1\right)^2}\right)\right] \Leftrightarrow$$

$$0 = h_{e_A}\frac{\eta}{6\left(h+1\right)^2} - e_A + \frac{1-\delta}{\delta\left(h_{e_P}-h_{e_A}\right)}\frac{4(h+1)^2}{\eta}\times$$

$$\left[1 + \frac{\eta h_{e_P}h_{e_A}}{3\left(h+1\right)^3} - m\left(\frac{h_{e_Ae_A}\eta\left(h+1\right) + \left(h_{e_P}-h_{e_A}\right)2\eta h_{e_A}}{4\left(h+1\right)^3}\right)\right] \Leftrightarrow$$

$$0 = h_{e_A}\frac{\eta}{6\left(h+1\right)^2} - e_A + \frac{1-\delta}{\delta\left(h_{e_P}-h_{e_A}\right)}\frac{4(h+1)^2}{\eta}\times$$

$$\left[1 + \frac{4\eta h_{e_P}h_{e_A}}{12\left(h+1\right)^3} - \frac{3m(h_{e_Ae_A}\eta\left(h+1\right) + \left(h_{e_P}-h_{e_A}\right)2\eta h_{e_A})}{12\left(h+1\right)^3}\right] \Leftrightarrow$$

$$0 = h_{e_A}\frac{\eta}{6\left(h+1\right)^2} - e_A + \frac{1-\delta}{\delta\left(h_{e_P}-h_{e_A}\right)}\frac{4(h+1)^2}{\eta}\times$$

$$\left[1 + \frac{4\eta h_{e_P}h_{e_A} - 3m(h_{e_Ae_A}\eta\left(h+1\right) + \left(h_{e_P}-h_{e_A}\right)2\eta h_{e_A})}{12\left(h+1\right)^3}\right] \Leftrightarrow$$

$$0 = h_{e_A}\frac{\eta}{6\left(h+1\right)^2} - e_A + \frac{1-\delta}{\delta\left(h_{e_P}-h_{e_A}\right)}\frac{4(h+1)^2}{\eta}\times$$

$$\left[\frac{12(h+1)^3 + 4\eta h_{e_P}h_{e_A} - 3m(h_{e_Ae_A}\eta\left(h+1\right) + \left(h_{e_P}-h_{e_A}\right)2\eta h_{e_A})}{12\left(h+1\right)^3}\right] \Leftrightarrow$$

$$0 = h_{e_A}\frac{\eta}{6\left(h+1\right)^2} - e_A + \frac{1-\delta}{\delta\eta\left(h_{e_P}-h_{e_A}\right)}\times$$

$$\left[\frac{12(h+1)^3 + 4\eta h_{e_P}h_{e_A} - 3m(h_{e_Ae_A}\eta\left(h+1\right) + \left(h_{e_P}-h_{e_A}\right)2\eta h_{e_A})}{3\left(h+1\right)}\right] \Leftrightarrow$$

$$0 = \frac{\eta h_{e_A}}{6\left(h+1\right)^2} - e_A + \frac{1-\delta}{\delta}\left[\frac{12(h+1)^3 + 4\eta h_{e_P}h_{e_A} - 3m(h_{e_Ae_A}\eta\left(h+1\right) + \left(h_{e_P}-h_{e_A}\right)2\eta h_{e_A})}{3\left(h+1\right)\eta\left(h_{e_P}-h_{e_A}\right)}\right].$$

$$(IV)\ h_{e_P}\int_{\underline{y}}^{\overline{y}} y f_h\left(y|h\right)dy-\phi'(e_P)+\frac{1-\delta}{\delta(h_{e_P}-h_{e_A})F_h(\hat{y}|h)}$$

$$\times\left[-\phi''(e_P)+h_{e_Pe_P}\int_{\underline{y}}^{\overline{y}} y f_h(y|h)dy+h_{e_P}h_{e_P}\int_{\underline{y}}^{\overline{y}} y f_{hh}(y|h)dy\right.$$

$$\left.+\,m\left(F_{hh}(\hat{y}|h)h_{e_P}(h_{e_P}-h_{e_A})+F_h(\hat{y}|h)h_{e_Pe_P}\right)\right]=0\Leftrightarrow$$

$$0=h_{e_P}\frac{\eta}{6(h+1)^2}-e_P+\frac{1-\delta}{\delta}\frac{1}{(h_{e_P}-h_{e_A})F_h(\hat{y},h)}\times$$

$$\left[-1+\frac{\eta h_{e_Pe_P}}{6(h+1)^2}-\frac{\eta h_{e_P}h_{e_P}}{3(h+1)^3}+m\left(\frac{\eta h_{e_P}}{2(h+1)^3}(h_{e_P}-h_{e_A})-\frac{\eta h_{e_Pe_P}}{4(h+1)^2}\right)\right]\Leftrightarrow$$

$$0=h_{e_P}\frac{\eta}{6(h+1)^2}-e_P-\frac{1-\delta}{\delta(h_{e_P}-h_{e_A})}\frac{4(h+1)^2}{\eta}\times$$

$$\left[-1+\frac{\eta h_{e_Pe_P}}{6(h+1)^3}-\frac{\eta h_{e_P}h_{e_P}}{3(h+1)^3}+m\left(\frac{(h_{e_P}-h_{e_A})\eta h_{e_P}}{2(h+1)^3}+\frac{h_{e_Pe_P}\eta}{4(h+1)^2}\right)\right]\Leftrightarrow$$

$$0=h_{e_P}\frac{\eta}{6(h+1)^2}-e_P+\frac{1-\delta}{\delta(h_{e_P}-h_{e_A})}\frac{4(h+1)^2}{\eta}\times$$

$$\left[1-\frac{\eta h_{e_Pe_P}}{6(h+1)^3}+\frac{\eta h_{e_P}h_{e_P}}{3(h+1)^3}-m\left(\frac{h_{e_Pe_P}\eta(h+1)+(h_{e_P}-h_{e_A})2\eta h_{e_P}}{4(h+1)^3}\right)\right]\Leftrightarrow$$

$$0=h_{e_P}\frac{\eta}{6(h+1)^2}-e_P+\frac{1-\delta}{\delta(h_{e_P}-h_{e_A})}\frac{4(h+1)^2}{\eta}\times$$

$$\left[1-\frac{2\eta h_{e_Pe_P}}{12(h+1)^3}+\frac{4\eta h_{e_P}h_{e_P}}{12(h+1)^3}-\frac{3m(h_{e_Pe_P}\eta(h+1)+(h_{e_P}-h_{e_A})2\eta h_{e_P})}{12(h+1)^3}\right]\Leftrightarrow$$

$$0=h_{e_P}\frac{\eta}{6(h+1)^2}-e_P+\frac{1-\delta}{\delta(h_{e_P}-h_{e_A})}\frac{4(h+1)^2}{\eta}\times$$

$$\left[\frac{12(h+1)^3-2\eta h_{e_Pe_P}+4\eta h_{e_P}h_{e_P}-3m(h_{e_Pe_P}\eta(h+1)+(h_{e_P}-h_{e_A})2\eta h_{e_P})}{12(h+1)^3}\right]\Leftrightarrow$$

$$0=h_{e_P}\frac{\eta}{6(h+1)^2}-e_P+\frac{1-\delta}{\delta\eta(h_{e_P}-h_{e_A})}\times$$

$$\left[\frac{12(h+1)^3-2\eta h_{e_Pe_P}+4\eta h_{e_P}h_{e_P}-3m(h_{e_Pe_P}\eta(h+1)+(h_{e_P}-h_{e_A})2\eta h_{e_P})}{3(h+1)}\right]\Leftrightarrow$$

$$0=\frac{\eta h_{e_P}}{6(h+1)^2}-e_P+\frac{1-\delta}{\delta}\times$$

$$\left[\frac{12(h+1)^3-2\eta h_{e_Pe_P}+4\eta h_{e_P}h_{e_P}-3m(h_{e_Pe_P}\eta(h+1)+(h_{e_P}-h_{e_A})2\eta h_{e_P})}{3(h+1)\eta(h_{e_P}-h_{e_A})}\right].$$

C3 Derivation of e_A^* and e_P^*

For instance, let consider a linear joint production function $h = ae_A^* + be_P^*$ that satisfies the conditions in the setup ($h_e > 0, h_{ee} \leq 0$). For $e_A^*, e_P^*, h > 0$, from (8), the patient principal's optimal effort level is $e_P^* = \frac{\sqrt{\eta a} - (2 + ae_A^*)\sqrt{6e_A^*}}{\sqrt{6e_A^*}b}$. Hence, inserting this equation into (9) provides

$$
\begin{aligned}
e_A^* &= \frac{\sqrt{\eta b} - (1 + be_P^*)\sqrt{6e_P^*}}{a\sqrt{6e_P^*}} = \frac{\sqrt{\eta b} - (1 + b\frac{\sqrt{\eta a} - (2+ae_A^*)\sqrt{6e_A^*}}{\sqrt{6e_A^*}b})\sqrt{6\frac{\sqrt{\eta a} - (2+ae_A^*)\sqrt{6e_A^*}}{\sqrt{6e_A^*}b}}}{a\sqrt{6\frac{\sqrt{\eta a} - (2+ae_A^*)\sqrt{6e_A^*}}{\sqrt{6e_A^*}b}}} \\
&= e_A^* + \frac{1 - \frac{\sqrt{a\eta}}{\sqrt{6}\sqrt{e_A^*}} + \frac{b\sqrt{e_A^*}\sqrt{b\eta}\sqrt{\frac{-2\sqrt{6} - \sqrt{6}ae_A^* + \frac{\sqrt{a\eta}}{\sqrt{e_A^*}}}{b}}}{6^{1/4}\left(-2\sqrt{6}\sqrt{e_A^*} - \sqrt{6}ae_A^{3/2} + \sqrt{a\eta}\right)}}{a} \iff \\
0 &= \frac{1 - \frac{\sqrt{a\eta}}{\sqrt{6}\sqrt{e_A^*}} + \frac{b\sqrt{e_A^*}\sqrt{b\eta}\sqrt{\frac{-2\sqrt{6} - \sqrt{6}ae_A^* + \frac{\sqrt{a\eta}}{\sqrt{e_A^*}}}{b}}}{6^{1/4}\left(-2\sqrt{6}\sqrt{e_A^*} - \sqrt{6}ae_A^{3/2} + \sqrt{a\eta}\right)}}{a}.
\end{aligned}
$$

Appendix D

D1 Applying conclusions to reality

In this Appendix, I relate the findings to real-world scenarios. Let $(\eta_1, \eta_2) \in \eta$, where $\eta_1 < \eta_2$.

$h_{e_A} > h_{e_P}$, $h_{e_P} < \eta^{-1}6(e_A^* + e_P^*)(h+1)^2$ **and** η_2

In this example, the sensitivity of the distribution is higher, indicating that the parties' deviations from average output have a greater impact. Moreover, the pharmaceutical company's (agent's) effort contributes more to the output than the hospital's (principal's) effort.

Real-world example: A pharmaceutical company, operating without monopolistic control in the market, provides drugs to hospitals. The company's endeavors to maintain quality, ensure availability, and facilitate timely delivery have a substantial influence on healthcare output. Although government regulations and external enforcement efforts also contribute, the direct impact of the pharmaceutical company's efforts is more pronounced, owing to the intricacies and particularities of drug production and delivery.

$h_{e_A} > h_{e_P}$**,** $h_{e_P} < \eta^{-1}6(e^*_A + e^*_P)(h+1)^2$ **and** η_1

In this example, the sensitivity of the distribution is lower than that of the previous scenario. However, the facility management company (agent) contributes more to the output than the facility administration (principal).

Real-world example: Healthcare facility administration frequently outsources facility management tasks, including cleaning, maintenance, and security, to specialized companies. While the efforts of the facility administration to oversee these tasks are crucial, the cleanliness of the healthcare facility primarily hinges on the quality and frequency of the efforts made by the facility management company.

References

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica*, *56*(2), 383–396.

Abreu, D., Pearce, D., & Stacchetti, E. (1986). Optimal cartel monitoring with imperfect information. *Journal of Economic Theory*, *39*(1), 251–269.

Abreu, D., Pearce, D., & Stacchetti, E. (1990). Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, *58*(5), 1041–1063.

Al-Najjar, N. I., Anderlini, L., & Felli, L. (2002). *Unforeseen contingencies* (STICERD Discussion Paper TE/02/431). London School of Economics.

Baker, G., Gibbons, R., & Murphy, K. J. (1994). Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics*, *109*(4), 1125–1156.

Baldi, S., Bottasso, A., Conti, M., & Piccardo, C. (2016). To bid or not to bid: That is the question: Public procurement, project complexity and corruption. *European Journal of Political Economy*, *43*, 89–106.

Bandiera, O., Barankay, I., & Rasul, I. (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica*, *77*(4), 1047–1094.

Bhattacharya, S., Gaba, V., & Hasija, S. (2015). A comparison of milestone-based and buyout options contracts for coordinating R&D partnerships. *Management Science*, *61*(5), 963–978.

Bhattacharya, S., Gupta, A., & Hasija, S. (2014). Joint product improvement by client and customer support center: The role of gain-share contracts in coordination. *Information Systems Research*, *25*(1), 137–151.

Bhattacharyya, S., & Lafontaine, F. (1995). Double-sided moral hazard and the nature of share contracts. *The RAND Journal of Economics*, *26*(4), 761–781.

Bull, C. (1987). The existence of self-enforcing implicit contracts. *The Quarterly Journal of Economics*, *102*(1), 147–159.

Cavalieri, M., Guccio, C., Lisi, D., & Rizzo, I. (2020). Does institutional quality matter for infrastructure provision? A non-parametric analysis for italian municipalities. *Italian Economic Journal*, *6*, 521–562.

Chong, E., Huet, F., Saussier, S., & Steiner, F. (2006). Public-private partnerships and prices: Evidence from water distribution in France. *Review of Industrial Organization*, *29*, 149–169.

Cong, J., & Zhou, W. (2021). Optimal contract under double moral hazard and limited liability. *Journal of Economics*, *134*, 49–71.

Cooper, R., & Ross, T. W. (1985). Product warranties and double moral hazard. *The RAND Journal of Economics*, 103–113.

Corbett, C. J., & DeCroix, G. A. (2001). Shared-savings contracts for indirect materials in supply chains: Channel profits and environmental impacts. *Management science*, *47*(7), 881–893.

Corbett, C. J., DeCroix, G. A., & Ha, A. Y. (2005). Optimal shared-savings contracts in supply chains: Linear contracts and double moral hazard. *European journal of operational research*, *163*(3), 653–667.

Coviello, D., & Gagliarducci, S. (2017). Tenure in office and public procurement. *American Economic Journal: Economic Policy*, *9*(3), 59–105.

Coviello, D., Moretti, L., Spagnolo, G., & Valbonesi, P. (2018). Court efficiency and procurement performance. *The Scandinavian Journal of Economics*, *120*(3), 826–858.

Culyer, A. J., & Newhouse, J. P. (2000). *Handbook of health economics*. Elsevier.

De Fraja, G. (2000). Contracts for health care and asymmetric information. *Journal of Health Economics*, *19*(5), 663–677.

Appendix E: Graphs

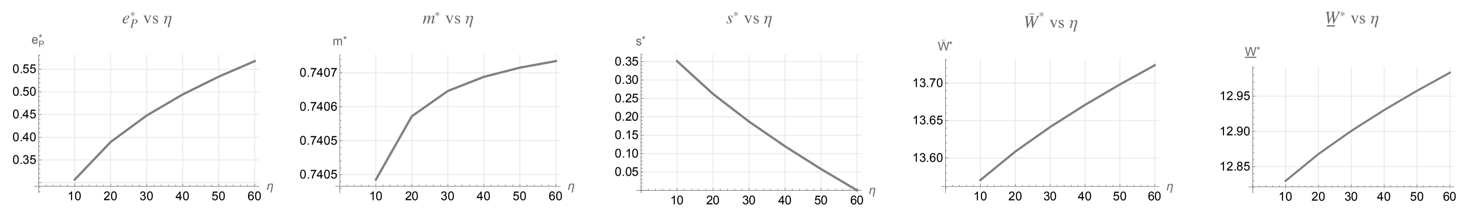

Figure E1: Changes of the distribution sensitivity η for very patient parties, where $\delta = 0.9$ and $h = 50e_A^* + 5e_P^*$

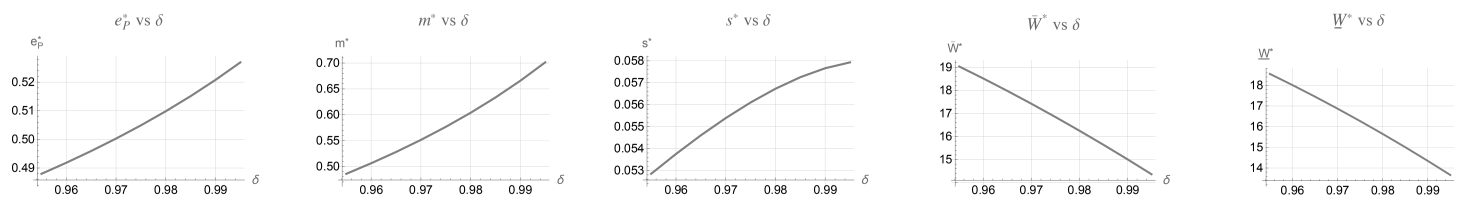

Figure E2: Changes of parties' patience δ, where $\eta = 50$ and $h = 50e_A^* + 5e_P^*$

Domingues, S., & Sarmento, J. M. (2016). Critical renegotiation triggers of european transport concessions. *Transport Policy*, *48*, 82–91.

Doornik, K. (2010). Incentive contracts with enforcement costs. *The Journal of Law, Economics & Organization*, *26*(1), 115–143.

Dumav, M., Fuchs, W., & Lee, J. (2022). Self-enforcing contracts with persistence. *Journal of Monetary Economics*, *128*, 72–87.

Estache, A. (2006). PPI partnerships vs PPI divorces in LDCs. *Review of Industrial Organization*, *29*(1/2), 3–26.

Estache, A., Serebrisky, T., & Wren-Lewis, L. (2015). Financing infrastructure in developing countries. *Oxford Review of Economic Policy*, *31*(3/4), 279–304.

Finocchiaro Castro, M., Guccio, C., Pignataro, G., & Rizzo, I. (2018). Is competition able to counteract the inefficiency of corruption? The case of Italian public works. *Economia e Politica Industriale*, *45*(1), 55–84.

Finocchiaro Castro, M., Guccio, C., & Rizzo, I. (2014). An assessment of the waste effects of corruption on infrastructure provision. *International Tax and Public Finance*, *21*(4), 813–843.

Frank, R. G., Glazer, J., & McGuire, T. G. (2000). Measuring adverse selection in managed health care. *Journal of Health Economics*, *19*(6), 829–854.

Fuchs, W. (2007). Contracting with repeated moral hazard and private evaluations. *American Economic Review*, *97*(4), 1432–1448.

Fudenberg, D., & Levine, D. K. (1994). Efficiency and observability with long-run and short-run players. *Journal of Economic Theory*, *62*(1), 103–135.

Garcia, A., Reitzes, J. D., & Benavides, J. (2005). Incentive contracts for infrastructure, litigation and weak institutions. *Journal of Regulatory Economics*, *27*, 5–24.

Giacomelli, S., & Menon, C. (2017). Does weak contract enforcement affect firm size? Evidence from the neighbour's court. *Journal of Economic Geography*, *17*(6), 1251–1282.

Gibbons, R., & Roberts, J. (2013). *The handbook of organizational economics*. Princeton University Press.

Gil, R., & Zanarone, G. (2014). Testing for the interaction of formal and informal contracts [Accessed June 3, 2022. Retrieved from https://extranet.sioe.org/uploads/isnie2014/gil_zanarone.pdf]. Mimeo.

Gil, R., & Marion, J. (2013). Self-enforcing agreements and relational contracting: Evidence from California highway procurement. *The Journal of Law, Economics & Organization*, *29*(2), 239–277.

Girth, A. M. (2014). A closer look at contract accountability: Exploring the determinants of sanctions for unsatisfactory contract performance. *Journal of Public Administration Research and Theory*, *24*(2), 317–348.

Guasch, J.-L. (2004). *Granting and renegotiating infrastructure concessions: Doing it right*. World Bank Publications.

Guasch, J.-L., Laffont, J.-J., & Straub, S. (2008). Renegotiation of concession contracts in Latin America: Evidence from the water and transport sectors. *International Journal of Industrial Organization*, *26*(2), 421–442.

Guccio, C., Lisi, D., & Rizzo, I. (2017). Institutional and social quality of local environment and efficiency in public works execution. *Global public procurement theories and practices*, 199–212.

Halac, M. (2012). Relational contracts and the value of relationships. *American Economic Review*, *102*(2), 750–779.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics*, *13*(2), 324–340.

Innes, R. D. (1990). Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory*, *52*(1), 45–67.

Jain, N., Hasija, S., & Popescu, D. G. (2013). Optimal contracts for outsourcing of repair and restoration services. *Operations Research*, *61*(6), 1295–1311.

Jewitt, I. (1988). Justifying the first-order approach to principal-agent problems. *Econometrica*, *56*(5), 1177–1190.

Kim, S. K., & Wang, S. (1998). Linear contracts and the double moral-hazard. *Journal of Economic Theory*, *82*(2), 342–378.

Klein, B., & Leffler, K. B. (1981). The role of market forces in assuring contractual performance. *Journal of Political Economy*, *89*(4), 615–641.

Kvaløy, O., & Olsen, T. E. (2009). Endogenous verifiability and relational contracting. *American Economic Review*, *99*(5), 2193–2208.

Kvaløy, O., & Olsen, T. E. (2010). *Incentive provision when contracting is costly* (Discussion Paper No. 2010/10). NHH Dept. of Finance & Management Science.

Laffont, J.-J. (2003). Enforcement, regulation and development. *Journal of African Economies*, *12*(suppl_2), ii193–ii211.

Laffont, J.-J., & Meleu, M. (2000). *Enforcement of contracts with adverse selection in LDCs*. Institut d'Économie Industrielle.

Leonard, K. L., & Zivin, J. G. (2005). Outcome versus service based payments in health care: Lessons from african traditional healers. *Health Economics*, *14*(6), 575–593.

Levin, J. (2003). Relational incentive contracts. *American Economic Review*, *93*(3), 835–857.

Li, J., & Matouschek, N. (2013). Managing conflicts in relational contracts. *American Economic Review*, *103*(6), 2328–51.

Macaulay, S. (2018). Non-contractual relations in business: A preliminary study. In *The sociology of economic life* (pp. 198–212). Routledge.

Macchiavello, R. (2022). Relational contracts and development. *Annual Review of Economics*, *14*, 337–362.

MacLeod, W. B. (2003). Optimal contracting with subjective evaluation. *American Economic Review*, *93*(1), 216–240.

MacLeod, W. B. (2007). Reputations, relationships, and contract enforcement. *Journal of Economic Literature*, *45*(3), 595–628.

MacLeod, W. B., & Malcomson, J. M. (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica*, *57*(2), 447–480.

MacLeod, W. B., & Malcomson, J. M. (2023). *Implicit contracts, incentive compatibility, and involuntary unemployment: Thirty years on* (IZA Discussion Paper No. No. 15881). IZA Institute of Labour Economics.

Major, I. (2019). Two-sided information asymmetry in the healthcare industry. *International Advances in Economic Research*, *25*(2), 177–193.

Makris, M., & Siciliani, L. (2013). Optimal incentive schemes for altruistic providers. *Journal of Public Economic Theory*, *15*(5), 675–699.

Malcomson, J. M. (2016). Relational incentive contracts with persistent private information. *Econometrica*, *84*(1), 317–346.

McCullough, J. S., & Snir, E. M. (2010). Monitoring technology and firm boundaries: Physician–hospital integration and technology utilization. *Journal of Health Economics*, *29*(3), 457–467.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, *12*(2), 380–391.

Noorderhaven, N. G. (1992). The problem of contract enforcement in economic organization theory. *Organization Studies*, *13*(2), 229–243.

Pearce, D. G., & Stacchetti, E. (1998). The interaction of implicit and explicit contracts in repeated agency. *Games and Economic Behavior*, *23*(1), 75–96.

Radner, R. (1985). Repeated principal-agent games with discounting. *Econometrica*, *53*(5), 1173–1198.

Roels, G., Karmarkar, U. S., & Carr, S. (2010). Contracting for collaborative services. *Management Science*, *56*(5), 849–863.

Rogerson, W. P. (1985). The first-order approach to principal-agent problems. *Econometrica*, *53*(6), 1357–1367.

Roussey, L., & Soubeyran, R. (2018). Overburdened judges. *International Review of Law and Economics*, *55*, 21–32.

Rubinstein, A. (1979). An optimal conviction policy for offenses that may have been committed by accident. In *Applied game theory* (pp. 406–413). Physica Heidelberg.

Saussier, S., Staropoli, C., & Yvrande-Billon, A. (2009). Public–private agreements, institutions, and competition: When economic theory meets facts. *Review of Industrial Organization*, *35*(1/2), 1–18.

Schneider, U. (2004). *Asymmetric information and the demand for health care - the case of double moral hazard* (Economics Discussion Paper No. No. 02-03). University of Bayreuth.

Siciliani, L. (2006). Selection of treatment under prospective payment systems in the hospital sector. *Journal of Health Economics*, *25*(3), 479–499.

Spaeter, S. (1998). *The principal-agent relationship: Two distributions satisfying MLRP and CDFC* (Working Paper 98-11). Centre for Research on Transportation.

Spear, S. E., & Srivastava, S. (1987). On repeated moral hazard with discounting. *The Review of Economic Studies*, *54*(4), 599–617.

Telser, L. G. (1980). A theory of self-enforcing agreements. *Journal of Business*, *53*(1), 27–44.

Thomas, J., & Worrall, T. (1988). Self-enforcing wage contracts. *The Review of Economic Studies*, *55*(4), 541–554.

Wang, K. F. (2015). *Moral hazard and hospital physician integration* [Doctoral dissertation, University of Washington].

Watson, J. (2021). Theoretical foundations of relational incentive contracts. *Annual Review of Economics*, *13*(1), 631–659.

Watson, J., Miller, D. A., & Olsen, T. E. (2020). Relational contracting, negotiation, and external enforcement. *American Economic Review*, *110*(7), 2153–2197.

Zanarone, G. (2013). Contract adaptation under legal constraints. *The Journal of Law, Economics & Organization*, *29*(4), 799–834.

Zhao, R. R. (2007). Dynamic risk-sharing with two-sided moral hazard. *Journal of Economic Theory*, *136*(1), 601–640.

Zhu, T., & Wang, S. (2005). Control allocation, revenue sharing, and joint ownership. *International Economic Review*, *46*(3), 895–915.