View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Institutional Knowledge at Singapore Management University

Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection Lee Kong Chian School Of Business

Lee Kong Chian School of Business

2-2013

Step by Step: The Benefits of Stage-based R&D Licensing Contracts

Pascale Crama
Singapore Management University, pcrama@smu.edu.sg

Bert De Reyck
University College London

Zeger Degraeve
London Business School

Follow this and additional works at: https://ink.library.smu.edu.sg/lkcsb_research

Part of the Operations and Supply Chain Management Commons, and the Technology and Innovation Commons

Citation

Crama, Pascale; De Reyck, Bert; and Degraeve, Zeger. Step by Step: The Benefits of Stage-based R&D Licensing Contracts. (2013). *European Journal of Operational Research*. 224, (3), 572-582. Research Collection Lee Kong Chian School Of Business.
Available at: https://ink.library.smu.edu.sg/lkcsb_research/3228

This Journal Article is brought to you for free and open access by the Lee Kong Chian School of Business at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection Lee Kong Chian School Of Business by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email libIR@smu.edu.sg.

Step by step.

The benefits of stage-based R&D licensing contracts.

Pascale Crama †
Bert De Reyck ‡
Zeger Degraeve §¶

† Lee Kong Chian School of Business, Singapore Management University
50 Stamford Road, Singapore 178899
Corresponding Author. e-mail: pcrama@smu.edu.sg, Tel: +65 6828 0330
‡ Department of Management Science & Innovation, University College London
Gower Street, London WC1E 6BT, United Kingdom
§ Melbourne Business School
200 Leicester Street, Carlton Vic 3053, Melbourne, Australia
¶ Department of Management Science & Operations, London Business School
Regent's Park, London NW1 4SA, United Kingdom

September 7, 2012

European Journal of Operational Research, 2013 Feb, 224 (3), 572-582
Doi: 10.1016/j.ejor.2012.09.014
Available on ScienceDirect.com

Step by step.
The benefits of stage-based R&D licensing contracts.

Abstract

We examine how a licensor can optimally design licensing contracts for multi-phase R&D projects when he does not know the licensee's project valuation, leading to adverse selection, and cannot enforce the licensee's effort level, resulting in moral hazard. We focus on the effect of the phased nature typical of such projects, and compare single-phase and multi-phase contracts. We determine the optimal values for the upfront payment, milestone payments and royalties, and the optimal timing for outlicensing. Including multiple milestones and accompanying payments can be an effective way of discriminating between licensees holding different valuations, without having to manipulate the royalty rate, which induces licensees to invest less, resulting in lower project values and socially suboptimal solutions. Interestingly, we also find that multiple milestone payments are beneficial even when the licensor is risk-averse, contrary to standard contract theory results, which recommend that only an upfront payment should be used. In terms of licensing timing, we show that the optimal time depends on the licensor's risk aversion, the characteristics of the licensee and the project value.

Key words: Research & Development, Innovation; Contract Design, Asymmetric information; Industries, Pharmaceutical

1 Introduction

R&D projects typically consist of a series of phases, in the course of which technology and market risk is gradually resolved. A typical example can be found in the pharmaceutical industry, where regulatory requirements enforce a strict procedure of sequential phases with specific milestones at which the project is assessed. A similar pattern can be found in most new product development environments in high-technology industries. This phased nature of R&D projects is also mirrored in the contracts that govern licensing deals. In this paper, we examine how a licensor should design and structure such licensing contracts in the presence of information asymmetries to extract the highest possible value from the deal. We focus on a commonly used contract structure containing an upfront payment, lump sum payments at the successful completion of milestones, and royalties on the product's sales. This work extends the analysis presented in a paper by Crama et al. (2008) that focused exclusively on single-phase projects. In this paper, we demonstrate the importance of explicitly acknowledging the phased nature of R&D projects when designing licensing contracts, and present insights into how to structure such contracts. We examine at which stage it is optimal to license out a project, and whether the contract should contain an upfront payment, one or more milestone payments and/or royalties. We also derive insights into the impact of the licensor's risk aversion and the presence of moral hazard and adverse selection on these recommendations.

Our work with Phytopharm, a biotechnology company based in Cambridgeshire, England, introduced us the problems of R&D project valuation and licensing. Phytopharm had discovered appetite suppressant properties of a natural compound, and was looking for a partner to complete the development of the product and launch it in the meal replacement market. During the negotiations, various aspects of the project were scrutinized by their potential partners, including the structure of the project, the probability of technical success (PTS) of the project phases, development costs and sales forecasts. A major issue, however, was disagreement on the PTS of the project and its different phases, how the licensing contract should be structured, and which payments should be included at which phases in the contract. To support these negotiations, we developed a model to help value the project and different possible licensing deals (Crama et al., 2007). In the end, Phytopharm secured a deal with Unilever, which bought the exclusive rights to include the compound in its existing range of weight loss products, in exchange for an upfront payment, a series of milestone payments and royalties on the sales revenue of the product range.

Based on our experiences with Phytopharm, we investigate two sources of information asymmetries when designing licensing contracts, resulting in adverse selection and moral hazard. First, we examine the consequences of a licensor and licensee disagreeing on the PTS of a particular project and each of its phases, and thus also the likelihood that the product will reach the market. In fact, the licensor may not even know the licensee's PTS estimate for each of the project phases. For instance, when a biotech company is negotiating a licensing deal with a large pharmaceutical company, the latter can use its in-house experts to generate its own PTS estimates, adjusting them to the specificities of the project and its own expertise in the field. Macho-Stadler et al. (1996) mention that "the licensee is in some cases better acquainted with [...] the application of the innovation to his productive process". Conversely, a non-pharmaceutical licensee may have more limited knowledge about the project and the technology, and might thus make more conservative estimates than those presented by the biotechnology company. The PTS estimates directly affect the value of the project, which may expose the licensor to adverse selection: the licensee has an incentive to misrepresent her valuation and understate her estimates of the project's PTS to reduce the perceived project value (see also Du et al., 2006). Second, because the product sales and the licensor's royalty revenues depend on the licensee's development and marketing effort, a licensee might not invest enough in these activities to generate the best possible outcome for the licensor. We examine the impact of this type of moral hazard on the optimal contract design for the licensor.

We tackle these issues from the perspective of the licensor, who acts as the principal in the principal-agent models developed to optimize the design of the licensing contract. We have chosen the licensor as the principal because licensors in the pharmaceutical industry enjoy increasing bargaining power. Indeed, the biotech company offers a unique product in a market characterized by soaring demand for in-licensing from large pharmaceutical companies and a high level of maturity of the biotech industry. Biotech companies' rising market power is reflected in the increased value of recent deals (Financial Times, 12 January 2006), which has lead to the observation that the pharmaceutical licensing market has "become a sellers' market" in which the power balance is shifting to the benefit of the biotech companies. As the licensor suffers from incomplete information about the licensee and does not know the licensee's valuation of the project, determined by her PTS estimates, he formulates this as a screening problem.

This paper makes three contributions to the literature. First, we investigate whether the results and managerial implications of adverse selection and moral hazard for the licensor's optimal contract design presented by Crama et al. (2008) for a single-phase project can be extended to explicitly incorporate the phased nature of R&D projects. That phased nature creates multidimensional licensees types, i.e., the licensee is characterized by several different PTS estimates, one per phase, each of which affects her valuation of the project. To preserve analytical tractability, we restrict our analysis to two phases, as this is sufficient to capture the complexities and effects of the multi-stage nature of licensing contracts in practice, and the multidimensional nature of the problem. We examine whether the results and insights from the single-phase case extend to the multi-phase case, and find that this only happens under certain conditions, which define the additional complexity introduced by the multidimensional nature of phased projects. Second, we illustrate the impact on the contract structure when these conditions do not hold. We find that the optimal contract may include more than one milestone payment and distort the timing of those payments for licensees holding both high and low project valuations, contrary to the single-stage setting in which the contracts of high-value licensees are never distorted, i.e., their contracts are designed to maximize total value. Third, using a multi-phased setting allows us to examine in detail the role of royalties. We find that royalties should only be used for licensees having a low incentive to invest in the project, which also implies that royalties use the licensees' difference in their incentives to invest to discriminate. Finally, we analyze in which phase to out-license a project and we examine which elements have an effect on this, including the licensee's characteristics and the licensor's degree of risk aversion.

In the next section, we present a literature review on multidimensional adverse selection, one of the key issues when designing multi-phase contracts. Section 3 describes the licensor's problem in detail and introduces the relevant concepts. Section 4 presents optimal contract structures for licensing multi-phase R&D projects under different assumptions of information asymmetry. Section 5 examines the optimal timing of the licensor's licensing decision. Section 6 presents a number of managerial insights derived from our analysis, for both the licensor and the licensee. We conclude in Section 7 with some avenues for future research.

2 Literature Review

The literature in economics and management has long recognized the phased nature of R&D projects. Quinn and Mueller (1963) present a generic R&D project structure based on their observations of practice. They recommend that management prepare a decision plan determining the information needed to decide on the further life of the project at the end of each phase. Sturmey (1966) makes an argument for defining phases in an R&D project as points at which it may be "cheaper to stop than to go ahead". Kelm et al. (1995) discuss papers advocating the phased representation of projects. Huchzermeier and Loch (2001) show how the phased nature of R&D projects generates value by creating opportunities for flexibility in the project execution, which corresponds to a real option. We contribute to this stream of research by investigating the impact of phasing an R&D project on the licensor's revenue stream from licensing.

There is a large body of literature on contracting with information asymmetry of which Crama et al. (2008) provide an extensive review. Papers studying the contract structure for R&D or innovation licensing in particular typically include a combination of upfront payment, milestone payments and royalties, and study the impact of those contract elements on the value captured by the innovator (e.g. Decheneaux et al. 2009, Erat et al. 2007), the execution of the project by the licensee (e.g., Aghion and Tirole 1994, Decheneaux et al. 2009) or the innovator (Decheneaux et al. 2011), or the incentive to invest in research under competition (Kulatilaka and Lin 2006). Xiao and Xu (2009) look at a multi-stage R&D setting with informational asymmetry about the innovator's capability. They allow renegotiation after additional information about the product becomes available. The renegotiation centers around the level of royalties, and the conditions under which those royalties increase (decrease) are determined based on the relative technical and market uncertainty. Bhaskaran and Krishnan (2009) choose to go beyond contracts that share revenue to also allow the sharing of development cost. They model the shared decision-making problem and look at how uncertainty and differences in capabilities influence the optimal choice of cooperation.

All of these papers have at most one dimension of informational asymmetry. Only a few papers examine contracts in a setting with multidimensional types, i.e., where an agent is defined by its characteristics on several dimensions that affect its valuation for a (basket of) good(s). Salanié (1997) refers to several papers that have studied multidimensional types but warns that the effort becomes fairly involved. Rochet and Stole (2001) argue that multidimensionality is problematic because the incentive compatibility constraints, which ensure that each agent type chooses the contract that was designed for its type, may not only be locally binding for adjacent types. They present several restricted models for which analytical results can be obtained. Similarly, Armstrong and Rochet (1999) show that multidimensionality may require the introduction of upward incentive constraints, because low-type agents may otherwise have an incentive to pretend to be high-type agents. Sibley and Srinagesh (1997) contrast the impact of multidimensionality on the incentive design mechanism with the solution for the problem with a single dimension. Similarly, the complexity introduced in our problem setting by multidimensional types induces us to limit ourselves to two licensee types to obtain analytical results. We are thus able to identify the conditions required to be able to derive some important properties of the optimal licensing contract for multi-phase projects.

The following authors have solved contract design problems with multidimensionality using the Karush-Kuhn-Tucker (KKT) conditions to find incentive-compatible mechanisms for the problem of adverse selection with a finite number of multidimensional types. Spence (1980) looks at the profit maximization problem facing a monopolist offering multiple goods to consumers under adverse selection. He proposes a solution procedure based on iteratively changing the bundles of goods to improve the monopolist's profit. Dana (1993) discusses the regulator's welfare maximization problem in a situation with two goods, solving simultaneously for the optimal contract and industry organization under adverse selection over the cost function in both markets. He shows that the optimality of joint production depends on the correlation of the cost functions.

Matthews and Moore (1987) use a multidimensional contract to screen a one-dimensional type. Their results confirm that distortion is used to discriminate against the low types by changing some elements of the contract, such as quantity and/or quality, from their socially optimal level. This does not necessarily happen in a monotonic fashion but depends on the type distribution. McAfee and McMillan (1988) offer a solution method for a continuum of types in several dimensions, e.g., their preferences for functionality, quality, and design of the product. However, they assume that the dimensionality of the instruments in the contract, e.g., price and quality, is lower than the dimensionality of types, which does not hold in the problem we study in this paper. Armstrong (1996) solves a similar problem and concludes that it is possible to find an optimal solution analytically under specific restrictions, most notably on the consumer's utility function that is assumed to be homogeneous in the multiple characteristics governing her preferences for the different goods, which is not the case in our model. He also shows that it is always in the monopolist's interest to exclude some types from the market. Rochet and Choné (1998) similarly present a solution procedure for multidimensional adverse selection in which a multiproduct monopolist caters to a heterogeneous population of consumers. They prove existence of an optimal solution and indicate in several examples that bunching, i.e., offering the same contract to agents of different types, may be optimal under multidimensional adverse selection. Their analysis depends on the assumption that the consumer's utility functions are linear in the multiple dimensions. We also find evidence of bunching in the problem setting described in this paper.

3 Model Description

Our model consists of a risk-averse or risk-neutral licensor (he) and a risk-neutral licensee (she). The licensor, often a small biotech company, may be risk-averse because of limited cash reserves and an erratic project pipeline. The licensee is typically a large pharmaceutical company with a large and well-diversified portfolio, which justifies our assumption of risk-neutrality. While our model is not limited to the pharmaceutical industry, it is a very suitable example as the project stages are well defined in drug development due to stringent regulatory requirements (e.g., see

Girotra et al. 2007). The licensor owns an R&D project, which he wants to license out. This R&D project requires further effort before being taken to the market. The project is divided into n consecutive stages, each of which can be a failure or a success. The product can only be launched if all stages in the project are successful. In each stage, the licensee has to invest a fixed amount in research activity. In the last stage, also called the development stage, the licensee can also invest in development activities, such as product positioning, manufacturing, and marketing, which increase the revenue of the product if it is successful. The licensor specifies the price for the project in terms of an upfront payment, milestone payments and a royalty rate. The upfront and milestone payments are fixed payments due at the beginning of the project and after successful completion of each stage, respectively. The royalty rate is a percentage of product sales.

This problem can be formulated within the principal-agent paradigm, with the licensor as the principal, i.e., offering the contract, and the licensee as the agent. This allows us to study different information regimes in which the licensee's type and/or action may be hidden to the licensor. We restrict ourselves to two discrete licensee types, which are labelled low (L) and high (H) depending on whether they have a low or high valuation of the project. Once we have defined the licensee's project valuation we will be able to formally define the licensee types. We also restrict our analysis to a project with two (remaining) stages, as this is sufficient to capture the multi-stage nature of the project and understand the difference between the research stages and the development stage. We summarize the notation below.

Parameters:

$\mathbf{p^o} = (p_1^o, p_2^o)$ – licensor's PTS estimate for each stage, $0 < p_i^o \leq 1$

$\mathbf{p^e} = (p_{1k}^e, p_{2k}^e)$ – PTS estimate of licensee type $k \in \{L, H\}$, $0 < p_{ik}^e \leq 1$

q_k – probability that licensee is of type $k \in \{L, H\}$, $q_L + q_H = 1$

$\mathbf{c} = (c_1, c_2)$ – research cost for each phase, $c_i \geq 0$

$s(x)$ – sales function, concave in x

Licensor's Decisions:

$\mathbf{m}_k = (m_{0k}, m_{1k}, m_{2k})$ – milestone payment for each licensee type $k = \{L, H\}$, $m_{ik} \geq 0$

r_k – royalty rate for each licensee type $k \in \{L, H\}$, $0 \leq r_k \leq 1$

Licensee's Decision:

x – development effort level, $x \geq 0$

All of the cash flows in our model are in present value terms. A sample timeline is shown in Figure 1. At time $t = 0$, the licensor offers a contract specifying the payment terms to the licensee. If the licensee accepts the contract, the upfront payment m_0 is made at $t = 1$ and the licensee commits the necessary research effort c_1 for the first R&D stage. At $t = 2$, the first stage of the project is completed and, if successful, the licensee makes another payment m_1. The licensee also commits to the second-stage research effort c_2 and chooses the development effort level x. Upon successful completion of the second stage, at $t = 3$, the licensee receives the sales income and makes the final payments to the licensor, a milestone payment m_2 and a royalty on sales r.

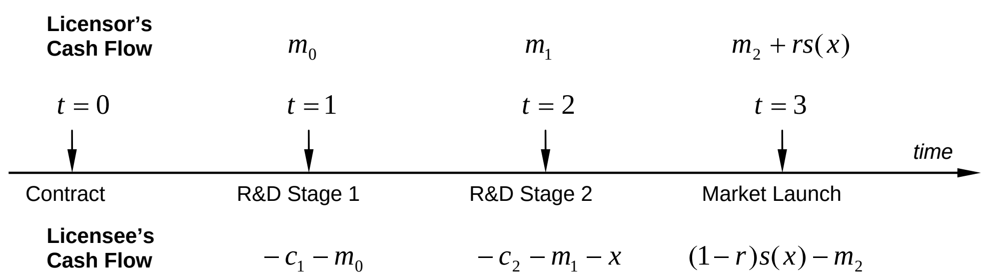


Figure 1: Timeline

Let us define the licensor's total utility:

$$V^o(\mathbf{m}, r, x) = (1 - p_1^o)u(m_0) + p_1^o(1 - p_2^o)u(m_0 + m_1) + p_1^o p_2^o u(m_0 + m_1 + m_2 + rs(x)),$$

where $u(z)$ is a concave Von Neumann-Morgenstern utility function. We will henceforth denote its first-order derivative in z by u_z.

The expected value for a licensee of type $k \in \{L, H\}$ who chooses a contract $l \in \{L, H\}$ and exerts development effort x is:

$$V_k^e(\mathbf{m}_l, r_l, x) = -c_1 - m_{0l} - p_{1k}^e(c_2 + m_{1l} + x) + p_{1k}^e p_{2k}^e((1 - r_l)s(x) - m_{2l}). \tag{1}$$

Note that we allow for the licensee of type k to choose any contract l, even if it does not correspond to the contract originally designed for her type k. Later we will show that it will be in the licensee's interest to choose the contract designed for her, so that $k = l$.

Eq. (1) shows that the licensee's valuation for the project, and hence her willingness to pay for the project, depends on her PTS estimates for the two remaining R&D stages. These PTS estimates are based on the licensee's analysis of the project and her ability to perform this type of research successfully. Thus, it is reasonable to assume that the licensee's PTS estimates are private information, unknown to the licensor.

Furthermore, the licensee needs to decide on the appropriate level of development effort. That level will be influenced by her valuation of the project and by the contract terms, particularly the royalty rate. Intuitively, the lower her valuation, or the higher the royalty rate, the less inclined the licensee will be to invest in the project. We assume that the development effort level is not directly observable to the licensor and that the licensor cannot enforce a specific effort level. Thus, the licensor has to take into account hidden action and information asymmetry when designing his optimal contract.

Let us briefly discuss the impact of the dimensionality of the licensee's private information. For a two-stage project, the licensee's private information, i.e., her type, is two-dimensional. If we assume that these two values are independent of each other and any pair p_{1k}^e and p_{2k}^e can occur, we can have a situation in which the licensee can be high on one dimension, e.g., the PTS estimate of the first stage, but low on the second dimension, e.g., the PTS estimate of the second stage. As the licensee's project valuation, i.e., the maximum price she is willing to pay, is determined jointly by both PTS estimates, we define the high type based on the project valuation that results from the PTS estimates instead of basing it on the PTS estimates themselves.

Definition 1 *The licensee with PTS estimates* (p_{1k}^e, p_{2k}^e) *is of high type iff* $V_k^e(0,0,x_k^*) > V_{k'}^e(0,0,x_{k'}^*)$, $k, k' \in \{L, H\}$ *and* $k \neq k'$.

4 The Optimal Contract

4.1 Complete Information

First, we examine the licensor's optimal contract choice under complete information, i.e., when he knows the licensee's type. We solve this rather straightforward case first to obtain a better understanding of the economic intuition of the optimal contract structure. The optimal solution to the complete information case yields an upper bound to the value under information asymmetry. We will examine the cases with and without hidden action.

Under complete information and without hidden action, the licensor optimizes his expected profit for each licensee type $k \in \{L, H\}$ by writing a contract $\{\mathbf{m}_k, r_k, x_k\}$ that solves the following problem:

$$\max_{\mathbf{m}_k \geq 0,\ r_k \geq 0,\ x_k \geq 0} V^o(\mathbf{m}_k, r_k, x_k) \tag{2}$$

$$\text{s.t.} \qquad V_k^e(\mathbf{m}_k, r_k, x_k) \geq 0 \tag{3}$$

The objective function, Eq. (2), maximizes the licensor's expected profit for a licensee of type k. The problem can be solved for each k separately, and the type subscript will henceforth be omitted for the complete information case. Eq. (3) is the licensee's participation constraint and guarantees that she receives her reservation profit, normalized to zero in our analysis. We can articulate the following lemma about the optimal contract under complete information and without moral hazard.

Lemma 1 *Under complete information and in the absence of moral hazard:*

(a) a milestone payment at market launch can be substituted for an equivalent royalty rate, and vice versa; and

(b) the optimal development effort level is set equal to the licensee's optimal development effort level given her type.

All proofs are in the Appendix. Lemma 1(a) enables us to reduce the dimensions of the licensor's problem and ignore the royalty rate, as a milestone payment at launch and a royalty rate are equivalent. Lemma 1(b) states that the optimal development effort level, regardless of the contract form, equals the licensee's desired effort level. Thus, the licensor does not try to impose his own preferred effort level, and does not distort the licensee's effort. The intuition behind this result is as follows. The price that the licensee is willing to pay, i.e., the total expected value of the milestone payments, is capped by her valuation of the project. For a given type, the licensor can thus set the highest price when the licensee's project valuation is largest. Thus, the licensor will choose the effort level that maximizes the licensee's value.

If the licensee's action cannot be verified, i.e., in the presence of moral hazard, the licensor's optimization problem changes slightly. First, the effort level x_{kl} becomes the licensee's decision variable rather than that of the licensor. The licensee's choice of effort level depends on her type and on the contract terms, and the licensee will choose her effort level to maximize her value: $x_{kl}^* = \arg\max_x \{V_k^e(\mathbf{m}_l, r_l, x)\}, k, l \in \{L, H\}$. This leads to the following lemma describing the impact of moral hazard under complete information.

Lemma 2 *In the presence of moral hazard under complete information, the optimal contract will not include a royalty rate.*

When the licensor can control the licensee's effort level, the royalty rate has no impact on the project value, as the licensor can always request the effort level to be set at its social optimum. In the presence of moral hazard, however, a royalty rate acts like a tax on the licensee's development effort, and reduces her incentive to invest in sales. This reduces the total value of the project to the licensee, which depresses the maximum price the licensor can ask for. Thus, it is better for the licensor to ask for milestone payments, which do not distort the licensee's incentive to invest. Under complete information, the licensor knows the licensee's project valuation and can set the milestone payment to the full project value, even in the presence of moral hazard. Thus, we observe that under complete information, moral hazard does not have an impact.

4.1.1 Risk-Neutral Licensor

A risk-neutral licensor maximizes the expected value of the milestone payments, and his optimal contract is characterized in the following theorem:

Theorem 1 *Under complete information, a risk-neutral licensor will include a single milestone payment* $m_i, i \in \{0, 1, 2\}$ *selected as follows:*

- m_0 *if* $p_1^e > p_1^o$ *and* $p_1^e p_2^e > p_1^o p_2^o$
- m_1 *if* $p_1^e < p_1^o$ *and* $p_2^e > p_2^o$
- m_2 *if* $p_1^e p_2^e < p_1^o p_2^o$ *and* $p_2^e < p_2^o$

There will always be such a $i \in \{0, 1, 2\}$*.*

The optimal contract singles out one milestone payment that is chosen to be as large as possible given the licensee's participation constraint. The licensor's choice of milestone payment is based on a comparison of the licensor's and the licensee's success probability for the phase associated with that payment. The different assessments of the PTS of each phase leads to different expected valuations of the same cash flow. The licensor exploits this difference to his advantage when choosing the milestone payment according to the rule laid down in Theorem 1. These conditions directly follow from the formulation of the dual to the licensor's problem (see proof in Appendix). Intuitively, the guiding principle for the licensor is to delay payment as long as he holds a higher combined probability of sucess for a (set of) stage(s), but no longer than until the licensee starts to hold a higher probability going forward. This rule can easily be generalized to any number of stages.

4.1.2 Risk-Averse Licensor

The optimal contract for a risk-averse licensor depends on his utility function and is determined by solving the KKT conditions. For a generic utility function, we can claim the following about the optimal contract.

Lemma 3 *Under complete information, a risk-averse licensor will choose his milestone payments as follows:*

- $m_1 = m_2 = 0$ *if* $p_1^e > p_1^o$ *and* $p_1^e p_2^e > p_1^o p_2^o$
- $m_2 = 0$ *if* $p_1^e < p_1^o$ *and* $p_2^e > p_2^o$

There will always be such a $i \in \{0, 1, 2\}$

Economic theory recommends that under complete information, if the principal is risk-averse and the agent is risk-neutral, the project should be sold for an upfront payment (Mas-Colell et al. 1995). Indeed, as the principal knows the agent's valuation, he can set the price accordingly, and this contract achieves optimal risk-sharing. Nonetheless, Lemma 3 shows that in our setting this does not necessarily hold. While the licensor's risk aversion suggests a contract with an upfront payment, the licensor's and licensee's respective PTS estimates may push toward a later milestone payment according to the conditions in Theorem 1. The optimal contract balances the risk of not having an upfront payment with the larger nominal value of the later milestone payment, and the resulting compromise may be a combination of several payments, up to the phase recommended by Theorem 1. Thus, the difference in valuation between licensor and licensee creates imperfect risk sharing, even under complete information.

4.1.3 Numerical Example

In Figure 2, we present an example of the composition of the optimal contract for a risk-neutral licensor under complete information. The graph shows the elements included in the optimal contract for a fixed licensee type (p_1^e, p_2^e) as a function of the licensor's PTS estimates (p_1^o, p_2^o). We observe the three different regions predicted by Theorem 1 in which each of the three different milestone payments is optimal. At the boundaries between regions, the licensor is indifferent between two or more milestone payments. The milestone payment is set such that the licensee's participation constraint is binding.

The optimal contract structure for a risk-averse licensor is illustrated in Figure 3. We use the constant absolute risk aversion utility function $u(x) = 1 - e^{-\rho x}$. Figure 3(a) has a

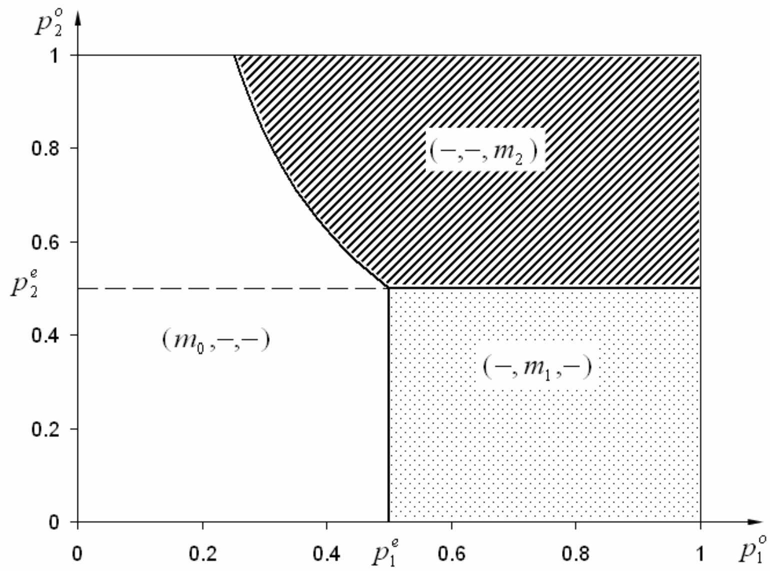
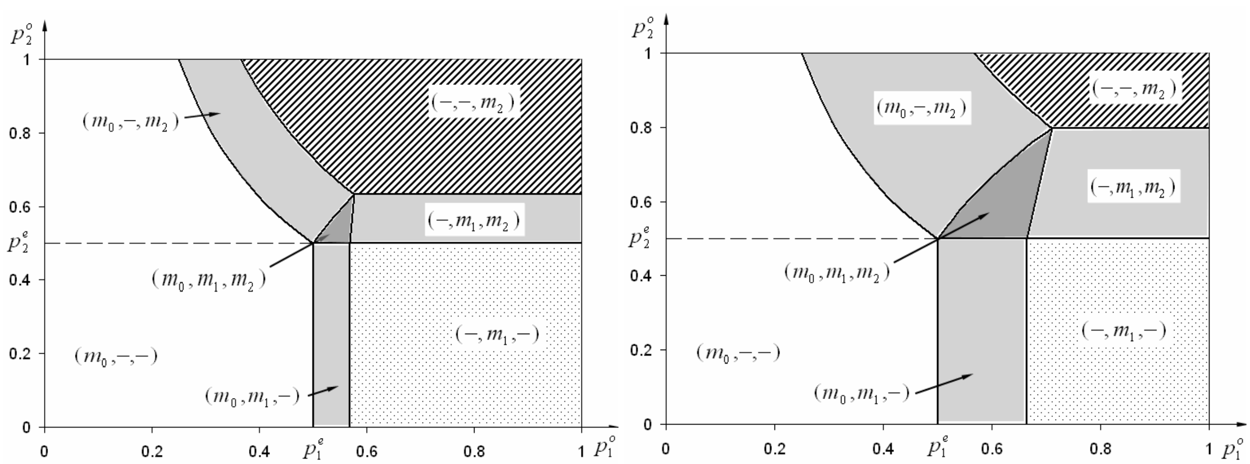


Figure 2: Optimal Contract Structure for Risk-Neutral Licensor Under Complete Information

lower risk aversion coefficient than Figure 3(b). There are three differences with the optimal contract structure for the risk-neutral licensor. First, as indicated by Lemma 3, we can now observe regions (in grey) where the licensor offers contracts that contain more than one milestone payment. Second, a risk-averse licensor prefers earlier payments than a risk-neutral licensor. Third, the more risk-averse the licensor, the more likely that a contract with multiple milestone payments will be chosen at the expense of lower expected milestone payments.



(a) Low Risk Aversion

(b) High Risk Aversion

Figure 3: Optimal Contract Structure for Risk-Averse Licensor Under Complete Information

4.2 Information Asymmetry

Under information asymmetry, the licensee's type, i.e., her PTS estimate for each stage, is her private information. In that case, it is easy to see that the licensee with the higher project valuation will pretend to have a low project valuation. Thus, the licensor cannot set a price that extracts all of the benefit from the licensee with the high valuation.

The licensor will design a contract menu $\{\mathbf{m}_k, r_k\}, k \in \{L, H\}$ that maximizes his expected revenue (utility) given his prior belief $\{q_L, q_H\}$. In general, the licensor can choose from a wide range of contract menu designs. However, the direct revelation principle (Fudenberg and Tirole 1991) restricts the category of contract menus by showing that there exists an optimal mechanism that induces truth telling. This allows formulating the licensor's contracting problem as follows:

$$\max_{\mathbf{m}_k \geq 0,\ r_k \geq 0} \sum_{k=L,H} q_k V^o\left(\mathbf{m}_k, r_k, x^*\left(r_k, \mathbf{p}_k^e\right)\right) \tag{4}$$

$$\text{s.t.} \qquad x^*\left(r_l, \mathbf{p}_k^e\right) = \arg\max_x \left\{V_k^e\left(\mathbf{m}_l, r_l, x\right)\right\} \qquad k, l \in \{L, H\} \tag{5}$$

$$V_k^e\left(\mathbf{m}_k, r_k, x^*\left(r_k, \mathbf{p}_k^e\right)\right) \geq 0 \qquad k \in \{L, H\} \tag{6}$$

$$V_k^e\left(\mathbf{m}_k, r_k, x^*\left(r_k, \mathbf{p}_k^e\right)\right) \geq V_k^e\left(\mathbf{m}_l, r_l, x^*\left(r_l, \mathbf{p}_k^e\right)\right) \qquad k, l \in \{L, H\} \tag{7}$$

The first constraint, Eq. (5), reflects the licensee's optimization problem in choosing the optimal effort level. The individual rationality constraint, Eq. (6), guarantees that the licensee, regardless of her type, receives her reservation utility. Finally, we observe a new set of constraints, the incentive compatibility constraints (Eq. 7), which are necessary to ensure that the contract menu will achieve truthful revelation of the licensee's type through her choice of contract.

4.2.1 Risk-Neutral Licensor

Before presenting the solution to the risk-neutral licensor's optimization problem, we repeat the definition of a high type licensee, i.e., the licensee with the highest valuation for the project in absence of any contractual payments. Because the licensee's value depends on the combination of two PTS estimates, the higher value is not necessarily linked to a higher PTS estimate for both stages. Of particular concern is the case in which the high type licensee has a lower valuation for the second stage than the low type licensee. That case is thus excluded in the theorem presenting the solution to a risk-neutral licensor's optimization problem, but will be touched upon later in the section.

Theorem 2 *If the high-type licensee holds a higher PTS estimate for the second stage, i.e., $p_{2H}^e > p_{2L}^e$, the contract for the high-type licensee in the optimal contract menu has the following characteristics:*

1. *the royalty rate is zero,* $r_H = 0$*;*

2. *a milestone payment as follows:*

   - m_{0H} *if* $p_1^e > p_1^o$ *and* $p_1^e p_2^e > p_1^o p_2^o$
   - m_{1H} *if* $p_1^e < p_1^o$ *and* $p_2^e > p_2^o$
   - m_{2H} *if* $p_1^e p_2^e < p_1^o p_2^o$ *and* $p_2^e < p_2^o$

Furthermore, the contract for the low-type licensee will include at most one milestone payment.

The main result of Theorem 2 confirms that the optimal contract for the high-type licensee is not distorted from its socially optimal structure: the licensee commits to the optimal development effort level as there is no royalty rate, and the milestone timing respects the licensor's optimality condition under complete information. The licensor will achieve discrimination between the two types by distorting the low-type licensee's effort level and/or milestone payment timing. This confirms the classical economic theory result that the contract for the high type should not be distorted: indeed, because the high-type licensee is more valuable to the licensor than the low-type licensee, it is better to distort the latter's effort level and contract structure to make it an unattractive contract for the former. Thus, the high-type licensee will be willing to pay extra to avoid the otherwise unfavorable contract terms. An interesting finding, however, is that discrimination between the two licensee types need not necessarily require the use of royalties; the choice of milestone payment amount and timing can be sufficient to distinguish between licensee types. This highlights one benefit of staged milestone payment contracts: they may obviate the use of a royalty rate as a tool for discrimination between licensing partners holding different valuations and thus avoid the inefficiencies caused by royalty payments.

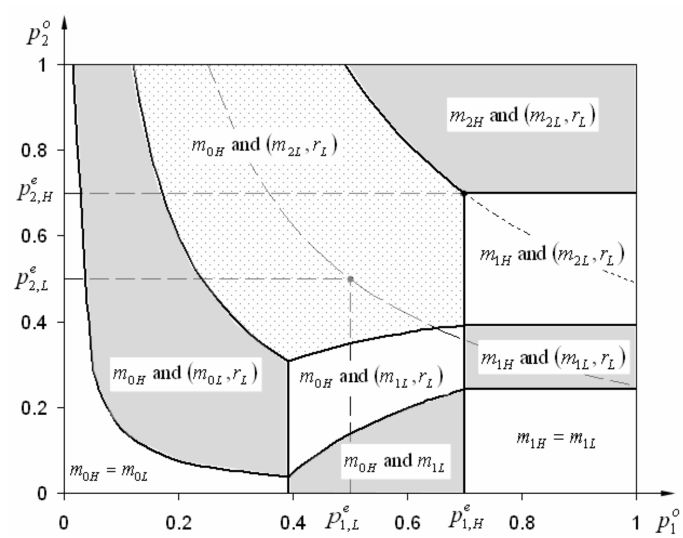
Figure 4 illustrates three possible pairs of licensee types $\mathbf{p}_H^e$ and $\mathbf{p}_L^e$ respecting the condition given in Theorem 2 and the resulting optimal contract structure for a varying $\mathbf{p}^o$. In Figure 4(a), the PTS estimates for both stages are higher for the high-type licensee than the low-type licensee. This is not the case in Figures 4(b) and 4(c). In Figure 4(b), however, the high-type licensee still has a higher total PTS estimate, whereas in Figure 4(c) the high-type licensee has a lower total PTS estimate (while still holding a higher valuation). The problem parameters were selected to produce a varied set of possible contract menus. The graphs confirm that the high-type licensee's effort is not distorted, because her contract never contains royalties, whereas the contract for the low-type licensee may contain royalties and distorted milestone payment timings. The direction of the timing distortion can go either way, i.e., the low-type licensee might be making earlier or later milestone payments than would be optimal under complete information. For example, whenever $p_1^o < p_{1L}^e$ and $p_1^o p_2^o < p_{1L}^e p_{2L}^e$, the socially optimal milestone payment for the low-type licensee should be an upfront payment; in Figures 4(a) and 4(b), however, we observe that the contract may contain a milestone payment at the first or second stage. Conversely, whenever $p_1^o > p_{1L}^e$ and $p_2^o < p_{2L}^e$, the socially optimal milestone payment for the low-type licensee should be a milestone payment at the first stage; however, Figures 4(b) and 4(c) show that the contract may sometimes contain an upfront payment.

Close observation of the solution structure described in Theorem 2 provides the intuition for the exclusion of the cases in which type and PTS for the second stage do not coincide. The condition listed in Theorem 2 guarantees that following the typical solution method to screening problems, i.e., relaxing the licensor's optimization problem by removing the individual rationality constraint for the high type and the incentive compatibility constraint for the low type (see Mas-Colell et al. (1995) for a discussion) will lead to the optimal solution to the original problem.

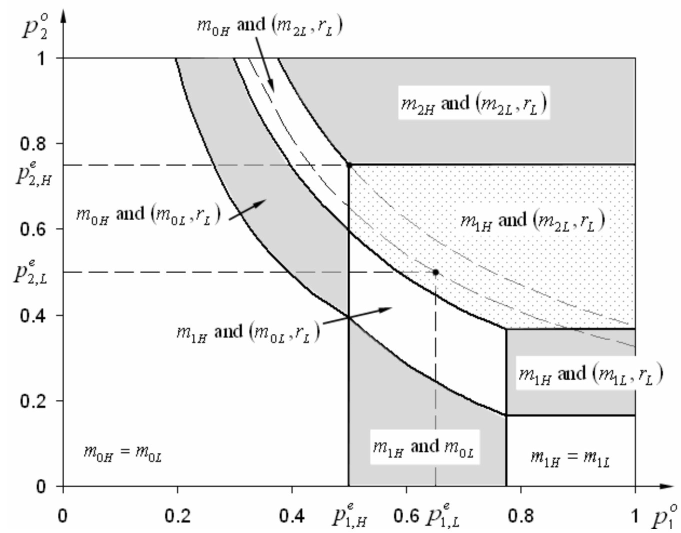
Intuitively, the licensee with the higher PTS estimate for the second stage will be more averse to paying royalties, because she would prefer to invest more in development effort. Therefore, the licensee with the higher PTS estimate for the second stage is willing to pay a higher lump-sum payment to avoid royalty payments. If that licensee also has a higher valuation for the project, she can afford to pay a higher milestone payment and still receive at least her reservation utility. The low-type licensee will not be able to afford the high-type contract, and her incentive constraint can be dropped. At the same time, the high-type licensee's constraint can be dropped because she always has a higher project valuation than the low-type licensee. Thus, the relaxation approach is appropriate under those conditions.

When the low-type licensee has the higher PTS estimate for the second stage the relaxation approach fails. The relaxed version of the licensor's optimization problem would propose to impose a royalty rate on the low-type licensee, which would reduce the effort level for the licensee who is willing to invest the most. Thus, the usual solution procedure does not apply and it is unclear ex ante which constraint will be binding. We have performed a numerical analysis to study the contract structure when the conditions of Theorem 2 do not hold. The contract terms and constraint structure are shown in Figures 5(a) and 5(b) respectively.

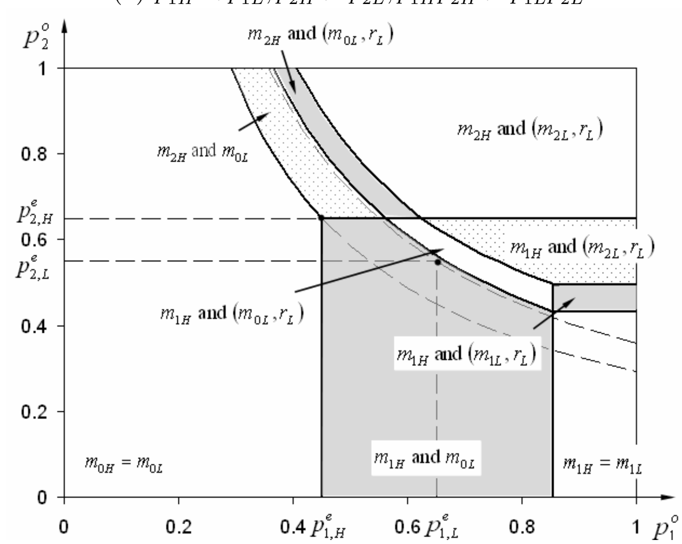
Figure 5(a) confirms our intuition that the royalty rate should always be set to zero for the licensee with the higher PTS estimate for the second stage. A striking difference with



(a) $p^e_{1H} > p^e_{1L}, p^e_{2H} > p^e_{2L}$

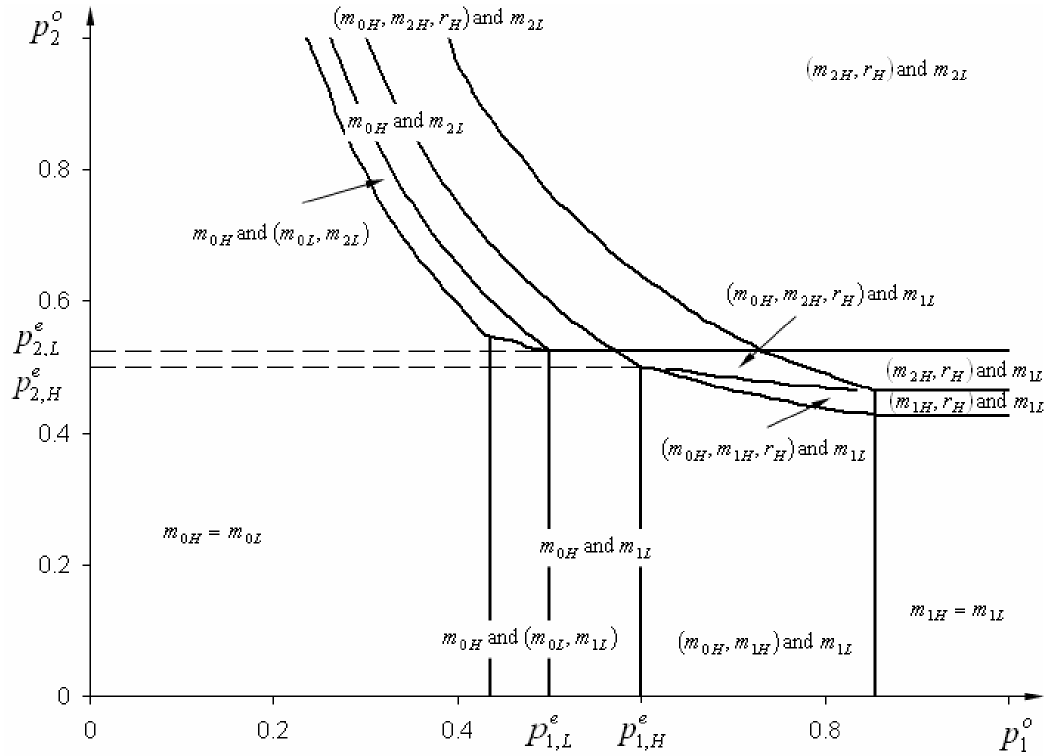


(b) $p^e_{1H} < p^e_{1L}, p^e_{2H} > p^e_{2L}, p^e_{1H}p^e_{2H} > p^e_{1L}p^e_{2L}$

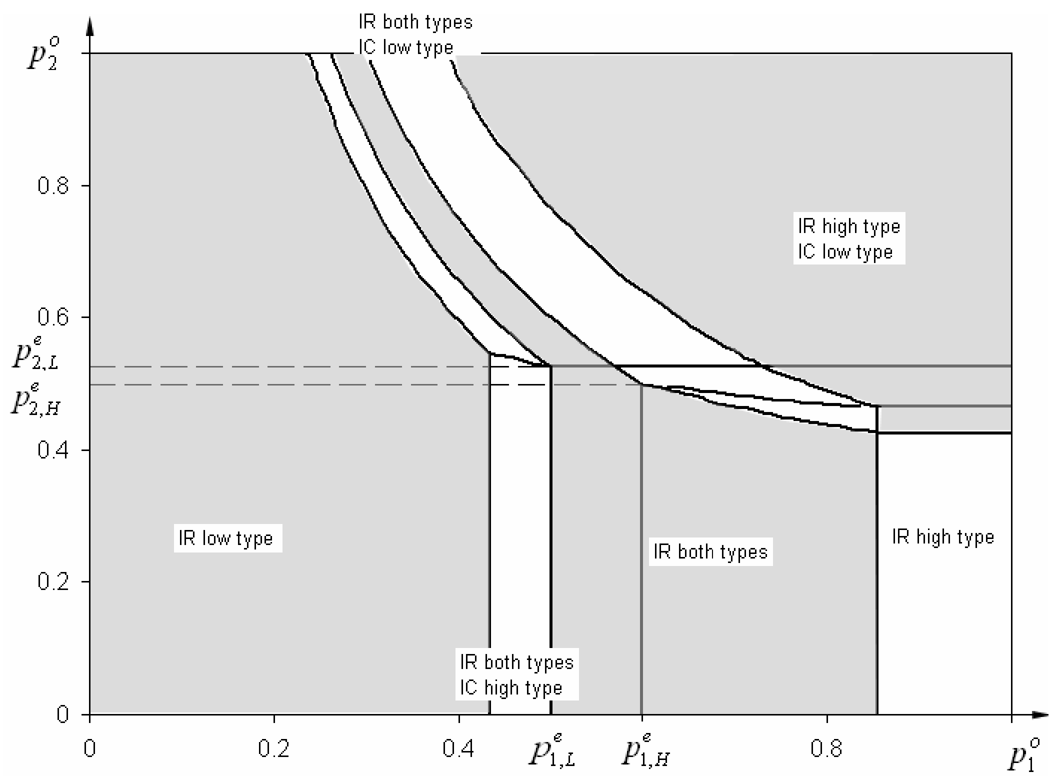


(c) $p^e_{1H} < p^e_{1L}, p^e_{2H} > p^e_{2L}, p^e_{1H}p^e_{2H} < p^e_{1L}p^e_{2L}$

Figure 4: Optimal Contract Structure for a Risk-Neutral Licensor Under Information Asymmetry



(a) Optimal contracts



(b) Binding constraints

Figure 5: Optimal Contract Structure for a Risk-Neutral Licensor Under Information Asymmetry with Varying Binding Constraints

the results under Theorem 2 is that the contract for either licensee type may combine two milestone payments. This immediately implies that the timing of the milestone payments can be distorted for either licensee type. Figure 5(b) shows that it is impossible to predict ex ante which constraints will be binding. At most one of the incentive compatibility constraints is binding, but it can be either the low- or high-type licensee's constraint. (Note that when a common pooling contract is offered to both licensee types, the incentive compatibility constraints are redundant.) It is interesting to note, however, that incentive compatible contracts can be achieved such that neither incentive compatibility constraint is binding but both individual rationality constraints are.

Thus, we find that although under some circumstances the results from the single-stage case are maintained, i.e., the licensor includes at most one milestone payment in the optimal contract and the royalty rate for the high-type licensee is zero, this does not always hold for multi-stage R&D projects. In a multi-dimensional setting, the distinction between a high- and low-type licensee is not obvious because the higher PTS estimate for the last stage does not necessarily correspond with the higher project valuation. If the high-type licensee holds a higher PTS estimate for the second stage than the low-type licensee, the insights from the single-stage licensing problem apply. If the low-type licensee holds a higher PTS estimate for the second stage than the high-type licensee, licensing a project with multiple stages can lead to more complex contract structures that could not be predicted based on single-stage models. In particular, we find that the direction of the incentive compatibility constraint can change and that the optimal contract may include several milestone payments.

4.2.2 Risk-Averse Licensor

Intuitively, one would expect a risk-averse licensor to prefer earlier payments than a risk-neutral licensor. We confirm this result.

Lemma 4 *If the high-type licensee holds the higher PTS estimate for the second stage, i.e., $p^e_{2H} > p^e_{2L}$, the contract for the high-type licensee in the optimal contract menu has the following characteristics:*

1. *the royalty rate is zero, $r_H = 0$; and*

2. *the milestone payments are determined such that:*

   - $m_{1H} = m_{2H} = 0$ *if* $p^e_1 > p^o_1$ *and* $p^e_1 p^e_2 > p^o_1 p^o_2$; *and*

- $m_{2H} = 0$ *if* $p_1^e < p_1^o$ *and* $p_2^e > p_2^o$

Lemma 4 is a logical extension of the results of Theorem 2 to the risk-averse licensor case. The royalty rate for the high-type licensee is zero to avoid hurting the licensee who is most likely to want to invest in the product. Furthermore, a risk-averse licensor prefers not to ask for royalties as they are delayed and risky payments. In relation to milestone payments, it is clear that risk-aversion will drive the licensor to prefer earlier rather than later payments. Thus, if we know the timing of the milestone payment the risk-neutral licensor prefers, the risk-averse licensor will choose the same and/or an earlier timing. This also holds for the low-type licensee and it is possible that the optimal contract for the low-type licensee will contain several milestone payments.

5 Timing of the Licensing Contract

In addition to designing an optimal licensing contract, another crucial issue is to determine *when* to offer the licensing opportunity, i.e., in which project stage. We again distinguish between different cases, with and without adverse selection, for a risk-neutral and a risk-averse licensor.

5.1 Complete Information

5.1.1 Risk-Neutral Licensor

The risk-neutral licensor's optimal licensing timing is described in Theorem 3.

Theorem 3 *Under complete information, when a risk-neutral licensor can choose the licensing timing, the optimal contract will not include a deferred payment, unless licensing occurs in the last stage.*

The result of Theorem 3 stems from the conditions in Theorem 1, which state that a milestone payment is optimal if the licensor has a higher PTS estimate than the licensee for all stages before that milestone payment, but a lower PTS estimate than the licensee for all stages after that milestone payment. In other words, the licensee overestimates the risk of the early stages, which reduces the licensee's valuation of the project and also the price she is willing to pay. For all stages after the milestone payment, however, the licensee holds a higher estimate of the probability or reaching the market and thus has a higher project valuation than the licensor. Hence, it is better for the licensor to bear the cost of development up to the stage of the optimal milestone payment determined by Theorem 1, as he would receive less than the project is worth to him by selling early. Once the licensee values the project more than the licensor, it is optimal to sell the project at the maximum price the licensee is willing to pay. This means that the joint decision of timing and contract terms ideally leads to a contract with an upfront payment only. The sole exception to this result occurs if the licensor would prefer delaying licensing based on the comparison of his PTS estimates with those of the licensee, but cannot do so because the licensee's contribution, i.e., her development effort, is needed. Thus, we find that introducing the freedom to choose when to outlicense brings us closer to the theoretical claim that the market is the most efficient way to organize economic transactions in the absence of market failures (Mas-Colell et al. 1995), with the exception of the case in which the need for the licensee's development effort forces the licensor to license earlier than he would otherwise choose.

Using Theorem 3 to determine the optimal licensing timing, we can restrict ourselves to comparing contracts with upfront payments only, along with a contract for licensing in the last stage that includes a payment at market launch. For a project with n stages, this requires solving at most n optimization problems.

5.1.2 Risk-Averse Licensor

We can make the following statement about the risk-averse licensor's optimal timing decision.

Lemma 5 *Under complete information, a risk-averse licensor will never license a project at a stage later than optimal for an otherwise equivalent risk-neutral licensor.*

The result of Lemma 5 is intuitive: a risk-averse licensor will prefer not to finance the research cost of a project with uncertain future revenue any longer than necessary. As risk-neutral licensors ask for upfront payments – with the possible exception of licensing delayed until the last stage – a risk-averse licensor will also ask for an upfront payment. Depending on his level of risk aversion, however, a risk-averse licensor may also opt to license earlier than a risk-neutral licensor would choose, in which case the contract can either consist of an upfront payment only, or a combination of upfront and milestone payments. Advancing licensing and requesting an upfront payment reduces the licensor's risk exposure, whereas the inclusion of milestone payments increases the total value of the payments. Note that risk-aversion reduces the probability that the project will be sold for an upfront payment only. Indeed, if it is optimal for the risk-neutral licensor to wait until the end of the first stage and sell for an upfront payment, risk-aversion may lead a licensor to sell the project immediately for a mix of upfront and milestone payments. Thus, in the case of multi-stage R&D projects, and even with the inclusion of choice in licensing timing, the potential for valuation differences leads us to qualify the expectation based on economic theory (Mas-Colell et al. 1995) that a risk-averse licensor should sell to a risk-neutral licensee for an upfront payment.

According to Lemma 5, the optimal licensing timing for a risk-averse licensor with a project that has n stages involves solving at most n optimization problems. The licensor then selects the licensing timing and contract with the highest utility.

5.2 Information Asymmetry

For a risk-neutral licensor and a two-stage R&D project, we have the following result.

Lemma 6 *Under information asymmetry and for a risk-neutral licensor:*

(a) if the optimal contract menu for immediate licensing includes a milestone payment for both licensee types, it is optimal to delay licensing; and

(b) if the optimal contract menu for immediate licensing includes an upfront payment for both licensee types, it is optimal to license immediately.

Lemma 6 offers sufficient conditions that create two distinct sets of licensor PTS values (p_1^o, p_2^o) for which the optimal licensing timing decision is known. These two sets are mutually exclusive but not jointly exhaustive and they are separated by a region in which lie PTS licensor values that do not fall under the sufficient conditions listed above. In that region, the optimal timing is determined by verifying the licensor's profit with and without delaying. This allows us to establish the licensor's indifference curve, i.e., the combination of PTS values (p_1^o, p_2^o) such that licensor is indifferent between delaying or not. This indifference curve is the frontier between the region in which the licensor prefers immediate licensing and the region in which he prefers to delay licensing.

We have shown in Theorem 3 that under complete information the licensor's optimal timing depends on the comparison between the licensor's and the licensee's PTS estimates. Lemma 6 implies that under information asymmetry the same principle holds and a risk-neutral licensor will be more likely to delay licensing if his PTS estimates (p_1^o, p_2^o) are high relative to the two licensee types, and less likely to delay licensing if his PTS estimates are low relative to the two licensee types. It follows that the indifference curve depends on the PTS estimates of the high- and low-type licensees.

The indifference curve will also be shaped by other problem parameters such as the prior probability for the high and low type and the licensor's risk-aversion. First, as the probability of a high (low) type, q_H (q_L), increases, the region for which delaying becomes optimal will gradually move to coincide with the region for which delaying for the high (low) type is optimal. Second, it is intuitive to expect that risk aversion will encourage the licensor to license early rather than delay licensing in order to reduce his risk exposure. Our numerical results confirm that as the licensor's risk aversion increases, the region of licensor PTS estimates for which delaying is optimal gradually shrinks.

Finally, numerical results have also shown the importance of the net present value of the project: the lower the value of the project, the more likely that the licensor will prefer to delay licensing. Indeed, the licensor's contract menu discriminates between licensee types by distorting, i.e. reducing, the project value for the low-type licensee, which encourages the high-type licensee to pay more in milestone payments to avoid costly distortion. If the value for the low-type licensee is small, however, the licensor has very little freedom to create the distortion necessary to increase his revenue from the high-type licensee. In that case, waiting for the successful completion of the first stage increases the value of the project to the licensee and thus the licensor's ability to capture value from the high-type licensee through distortion of the low-type licensee's value.

The final decision to delay licensing or sell the project immediately will obviously depend on the combination of all the factors identified above. For example, if the project has a low net present value and the licensor is risk-averse, it may be within his interests to delay licensing the project. Indeed, after successful completion of the first research stage, the licensee's project value will increase, allowing the licensor more leeway in designing the contract, enabling him to better discriminate to obtain a higher utility that compensates for the additional research cost and risk. We summarize how each factor influences the licensor's decision in Figure 6.

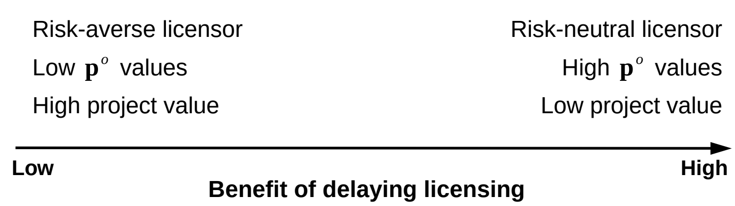


Figure 6: Impact of Problem Parameters on the Value of Delaying

6 Managerial Implications

The licensor's optimal contract depends on his risk attitude, the information regime and the licensee types he is facing. Under complete information, the results from the single-stage R&D licensing model are preserved. In particular, we find that the optimal contract for a risk-neutral licensor contains a single upfront or milestone payment, the choice of which is based on the relative valuation of the licensor and the licensee for that particular payment. The optimal contract also avoids using a royalty rate to achieve optimal project execution, which maximizes the value that the licensor can extract from the licensee. Risk-aversion induces the licensor to prefer earlier payment, and may lead him to combine upfront and milestone payments in the optimal contract in an effort to balance risk exposure and nominal value of the payments.

Under information asymmetry, the insights of the single-stage case hold if the project's valuation, i.e., the licensee type, and the PTS estimate for the second stage are aligned. In particular, we can show that it is not optimal to combine upfront and milestone payments in the optimal contract for a risk-neutral licensor, and that the optimal contract menu will introduce distortions for the low-type licensee's contract in the timing of the milestone payments or the use of a royalty rate. These distortions reduce the value that the licensor obtains from the low-type licensee to enable him to capture more value from the high-type licensee. The contract for the high-type licensee, however, should never be distorted. In case the low- and high-type licensee project valuation and PTS estimate for the second stage are not aligned, we find a very different pattern of optimal contracts, with multiple milestone payments for either licensee type. This also means that there can be distortion in the timing of the milestone payments for both the low- and the high-type licensee.

Third, the use of the royalty rate is determined not by the licensee's type or project valuation, but rather by her PTS estimate for the second stage. The licensee with the higher PTS estimate for the second stage will never be required to pay royalties: it is always the licensee who will be least willing to invest – regardless of her total project valuation – whose effort should be distorted in order to create an incentive for the licensee who wishes to invest to pay more to avoid royalties. Note that this also means that the royalty rate can only discriminate between licensee types who have a different PTS estimate for the second, or development, stage. If the PTS estimate for the development stage is common knowledge in the industry, the royalty rate cannot achieve discrimination, and a limited degree of discrimination can only be achieved through the appropriate timing of the milestone payments. In the pharmaceutical industry, for instance, this could happen if the licensee were to wait until the FDA submission to invest in the product, because there exists reliable historical information on the outcome of the FDA review process.

Finally, a multi-stage R&D project offers the licensor the possibility to choose when to outlicense his project. Our analysis shows that the right choice of timing and optimal contract can increase the licensor's value. Under complete information, the optimal licensing point is determined by a comparison between the licensor's and the licensee's valuation, as implied by their respective PTS estimates. It is interesting to note that a risk-neutral licensor should ideally wait until it is optimal to sell the project for an upfront payment, unless the licensee's participation in the project is required before that point, in which case the licensor may license out for a milestone payment. A risk-averse licensor will never license later than a risk-neutral licensor. If the risk-averse licensor chooses to license earlier, his optimal contract will probably include a combination of upfront and milestone payments.

Asymmetric information complicates the situation, and the licensor's optimal licensing timing depends on a combination of factors. The licensor is more likely to benefit from delaying licensing if the licensee's PTS estimates are relatively low compared to the licensor's and if the value of the project is low. Increasing risk aversion obviously pushes the licensor toward earlier licensing. Thus, the licensor's decision when to license the project should be considered carefully, taking all of these factors simultaneously into account.

7 Future Research

Out-licensing R&D projects is a complex proposition. Our research points out that to make the most of the licensing opportunity, the licensor's work must start before the licensing negotiations. Indeed, the licensor should first carefully plan the project to identify all of its phases, and then decide on the optimal timing to start with the licensing negotiations. Once the licensing deal is under negotiation, designing the optimal contract remains difficult because of the licensee's incentive to understate her valuation of the project. In that case, the licensor can use a varying royalty rate to induce the licensee to reveal her type.

Many issues remain to be studied in this context. We can identify two avenues of future research. First, it would be interesting to incorporate the possibility of the licensor's continuing involvement in the project, with an impact on product quality, and examine how that changes the issue of incentives. Second, we could investigate what happens when at the end of each phase new information on the project's market potential is revealed. This may change the execution of the project, which may be abandoned not because of technical failure but because the market forecast shows inadequate returns.

Acknowledgments

We thank participants in seminars at Singapore Management University, the University of California San Diego, the University of North Carolina Chapel Hill, the National University of Singapore, Erasmus University Rotterdam, and IESE for their feedback.

References

Aghion, P., J. Tirole. 1994. The Management of Innovation. *The Quarterly Journal of Economics* **109** 1185-1209.

Armstrong, M. 1996. Multiproduct Nonlinear Pricing. *Econometrica* **64** 51-75.

Armstrong, M., J-C. Rochet. 1999. Multi-dimensional screening: A User's Guide. *European Economic Review* **43** 959-979.

Bhaskaran, S.R., V. Krishnan. 2009. Effort, Revenue, and Cost Sharing Mechanisms for Collaborative New Product Development. *Management Science* **55** 1152-1169.

Boer, F.P. 2002. Financial Management of R&D Projects. *Research Technology Management* **45** 23-36.

Crama, P., B. De Reyck, Z. Degraeve, W. Chong. 2007. R&D Project Valuation and Licensing Negotiations at Phytopharm plc. *Interfaces* **37** 472-487.

Crama, P., B. De Reyck, Z. Degraeve. 2008. Milestone Payments or Royalties? Contract Design for R&D Licensing. *Operations Research* **56** 1539-1552.

Dana, J.D. 1993. The Organization and Scope of Agents: Regulating Multiproduct Industries. *Journal of Economic Theory* **59** 288-310.

Decheneaux, E., M. Thursby, J. Thursby. 2009. Shirking, Sharing Risk and Shelving: The Role of University Licensing. *International Journal of Industrial Organization* **27** 80-91.

Decheneaux, E., J. Thursby, M. Thursby. 2011. Inventor moral hazard in university licensing: The role of contracts. *Research Policy* **40** 94-104.

Du, L., Q. Hu, L. Liu. 2006. A profit sharing scheme for a two-firm joint venture. *European*

Journal of Operational Research **170** 277-229.

Erat, S., S. Kavadias, C. Gaimon. 2007. Licensing of Component Technologies. Georgia Institute of Technology Working Paper.

Financial Times. 2006. Change of Culture: How Big Pharma is Picking the Best of Biotech as a Sector Starts to Mature. (12 January.)

Girotra, K., C. Terwiesch, K. Ulrich. 2007. Valuing R&D Projects in a Portfolio: Evidence from the Pharmaceutical Industry. *Management Science* **53** 1452-1466.

Huchzermeier, A., C.H. Loch. 2001. Project Management Under Risk: Using the Real Options Approach to Evaluate Flexibility in R&D. *Management Science* **47** 85-101.

Kelm, K.M., V.K. Narayanan, G.E. Pinches. 1995. Shareholder Value Creations during R&D Innovation and Commercialization Stages. *The Academy of Management Journal* **38** 770-786.

Kulatilaka, N., L. Lin. 2006. Impact of Licensing on Investment and Financing of Technology Development. *Management Science* **52** 1824-1837.

Macho-Stadler, I., X. Martinez-Giralt, J. Pérez-Castrillo. 1996. The Role of Information in Licensing Contract Design. *Research Policy* **25** 43-57.

Mas-Colell, A., M.D. Whinston, J.R. Green. 1995. *Micro-Economic Theory.* Oxford University Press.

Matthews, S., J. Moore. 1987. Monopoly Provision of Quality and Warranties: An Exploration in the Theory of Multidimensional Screening. *Econometrica* **55** 552-467.

McAfee, R.P., J. McMillan. 1988. Multidimensional Incentive Compatibility and Mechanism Design. *Journal of Economic Theory* **46** 334-354.

Pennings, E., O. Lint. 1997. The Option Value of Advanced R&D. *European Journal of Operational Research* **103** 83-94.

Quinn, J.B., J.A. Mueller. 1963. Transferring Research Results to Operations. *Harvard Business Review* **41** 49-66.

Roberts, K., M.L. Weitzman. 1981. Funding Criteria for Research, Development, and Exploration Projects. *Econometrica* **49** 1261-1288.

Rochet, J-C., P. Choné. 1998. Ironing, Sweeping, and Multidimensional Screening. *Econometrica* **66** 783-826.

Rochet, J-C., L.A. Stole. 2001. The Economics of Multidimensional Screening. Working Paper - Chapter in Advances in Economics and Econometrics, Theory and Applications 8^{th} World Congress.

Salanié, B. 1997. *The Economics of Contracts: A Primer* The MIT Press, Cambridge, Massachusetts.

Sibley, D.S., P. Sriganesh. 1997. Multiproduct Nonlinear Pricing with Multiple Taste Characteristics. *The RAND Journal of Economics* **28** 684-707.

Spence, A.M. 1980. Multi-Product Quantity-Dependent Prices and Profitability Constraints. *The Review of Economic Studies* **47** 821-841.

Sturmey, S.G. 1966. The Management of Research and Development. *The Journal of Industrial Economics* **14** 91-100.

Xiao, W., Y. Xu. 2009. The Impact of Royalty Contract Revision in a Multi-stage Strategic R&D Alliance. Working Paper.

Technical Appendix to "Step by step. The benefits of stage-based R&D licensing contracts."

Proof of Lemma 1.

(a) In the absence of adverse selection and moral hazard, consider a feasible contract $(\mathbf{m}, r, x)$ that respects the licensee's IR constraint and the non-negativity constraints. Since the contract sets both the royalty rate and the effort level, the licensor's royalties revenue conditional on market launch is $rS(x)$. This contract is then equivalent to the following contract: $(\overline{\mathbf{m}}, 0, x)$, with all milestone payments identical except the last milestone payment $\overline{m}_2 = m_2 + rS(x)$. It is immediately possible to confirm that this new contract is similarly feasible and yields the same value to the licensor. Thus the licensor is indifferent to the use of royalties.

(b) Under (a), we have shown that the optimal contract can always be formulated as a contract with milestone payments only. Thus the individual rationality constraint defines the maximum amount that the licensor can extract from the licensee, for any combination of upfront and milestone payments:

$$m_0 + \sum_{i=1}^{n} \prod_{j=1}^{i} p_j^e m_i = -c_1 - \sum_{i=2}^{n} \prod_{j=1}^{i-1} p_j^e c_i - \prod_{j=1}^{n-1} p_j^e x + \prod_{j=1}^{n} p_j^e s(x)$$

The right hand side of this equation is the licensee's expected value from the project and is maximized when the effort level x is set to the licensee's optimal development effort level, i.e., x^* such that $p_n^e ds(x)/dx = 1$. This maximizes the expected value of the milestone payments, i.e., the left hand side of the equation, which the licensor can claim in payment from the licensee.

Proof of Lemma 2

In the presence of moral hazard, a non-zero royalty rate reduces the licensee's incentive to invest in the project, thus reducing the project value. Given Lemma 1, we know that the licensor maximizes his value when the licensee's development effort is optimal given her type (Lemma 1 (b)) and that a royalty contract can be transformed into an equivalent contract with milestone payments only (Lemma 1 (a)). Therefore, the optimal contract and its value to the licensor without adverse selection and moral hazard can be achieved if a contract with milestone payments only is offered in the presence of moral hazard.

Proof of Theorem 1

First, we invoke Lemma 2 to restrict our search to contracts with milestone payments only. This transforms the optimization problem into a linear problem. The dual of the licensor's problem is:

$$\min_{y} V_k^e(0, 0, x_k)y$$

$$\begin{aligned} \text{subject to } y &\geq 1 \\ p_1^o y &\geq p_1^e \\ p_1^o p_2^o y &\geq p_1^e p_2^e \end{aligned}$$

The optimal solution puts y to its lower bound jointly allowed by the three constraints of the dual. Note that each constraint corresponds to a different timing of the milestone payment. This directly yields the conditions listed in Theorem 1. The shadow price of the binding constraint is the optimal milestone payment for the licensor's original optimization problem.

Furthermore, for any licensee estimates vector $\mathbf{p}^e$, there is always a i such that the above equations hold. To prove this, we start by assuming a vector $\mathbf{p}^e$ such that there is a i satisfying the conditions above. We show that by changing the probability estimates one stage at a time, there will always be a new κ satisfying the set of conditions. From our starting point, we generate all possible licensee estimates, with at least one stage satisfying the conditions above.

First, assume that the probability estimate of a stage $l > i$ is changed. An increase in p_l^e will never invalidate the inequalities holding for i. Thus, let us decrease p_l^e until a first condition is violated, i.e. for one $\kappa \leq l$ we have $\prod_{j=1}^{\kappa} p_{i+j}^e \leq \prod_{j=1}^{\kappa} p_{i-j}^o$. If $\kappa < n$, we know by definition of i that $\prod_{j=1}^{s-i} p_{i+j}^e \geq \prod_{j=1}^{s-i} p_{i+j}^o$, $\forall s \in \{\kappa+1, ..., n\}$, which can only hold if $\prod_{j=1}^{s} p_{\kappa+j}^e \geq \prod_{j=1}^{s} p_{\kappa+j}^o$, $\forall j \in \{0, ..., n-\kappa\}$. Furthermore, by definition of i we also know that $\prod_{j=1}^{s} p_{i+j}^e \geq \prod_{j=1}^{s} p_{i+j}^o$, $\forall j \in \{0, ..., n-\kappa\}$, which can only hold if $\prod_{j=0}^{s} p_{\kappa-j}^e \leq \prod_{j=0}^{s} p_{\kappa-j}^o$. Together with $\prod_{j=0}^{s} p_{i-j}^e \leq \prod_{j=0}^{s} p_{i-j}^o$, $\forall j \in \{0, ..., i-1\}$, κ now satisfies the necessary set of conditions.

A similar reasoning can be applied for a change in the estimate of stage $l \leq k$.

Proof of Lemma 3

Suppose that the risk-neutral licensor's optimal milestone payment is m_i^*, scheduled at the end of research stage i. Set any milestone payment $\overline{m}_{i+k} > 0$, $k > 0$ and correspondingly set $\overline{m}_i = m_i^* - \prod_{j=1}^{k} p_{j+i}^e \overline{m}_{i+k}$ such that the contract remains feasible.

We need to show that this new contract $\overline{m}_i, \overline{m}_{i+k}$ will never outperform m_i^* for a risk-averse licensor, or:

$$\left(\prod_{j=1}^{i} p_j^o - \prod_{j=1}^{i+k} p_j^o\right) u(\overline{m}_i) + \prod_{j=1}^{i+k} p_j^o u(\overline{m}_i + \overline{m}_{i+k}) \leq \prod_{j=1}^{i} p_j^o u(m_i^*)$$

This follows from the concavity of the objective function:

$$\begin{aligned}
\left(1 - \prod_{j=1}^{k} p_{i+j}^o\right) u(\overline{m}_i) + \prod_{j=1}^{k} p_{i+j}^o u(\overline{m}_i + \overline{m}_{i+k}) &\leq u\left(\left(1 - \prod_{j=1}^{k} p_{i+j}^o\right) \overline{m}_i + \prod_{j=1}^{k} p_{i+j}^o (\overline{m}_i + \overline{m}_{i+k})\right) \\
\iff \left(1 - \prod_{j=1}^{k} p_{i+j}^o\right) u(\overline{m}_i) + \prod_{j=1}^{k} p_{i+j}^o u(\overline{m}_i + \overline{m}_{i+k}) &\leq u\left(\overline{m}_i + \prod_{j=1}^{k} p_{i+j}^o \overline{m}_{i+k}\right) \\
\iff \left(1 - \prod_{j=1}^{k} p_{i+j}^o\right) u(\overline{m}_i) + \prod_{j=1}^{k} p_{i+j}^o u(\overline{m}_i + \overline{m}_{i+k}) &\leq u\left(m_i^* + \left(\prod_{j=1}^{k} p_{i+j}^o - \prod_{j=1}^{k} p_{i+j}^e\right) \overline{m}_{i+k}\right) \leq u(m_i^*)
\end{aligned}$$

The last inequality follows from the optimality of m_i^* for the risk-neutral licensor, i.e., $\prod_{j=1}^{k} p_{i+j}^o \leq \prod_{j=1}^{k} p_{i+j}^e$.

Conversely, set any milestone payment $\overline{m}_{i-k} > 0$, $0 < k \leq i$ and correspondingly set $\overline{m}_i = m_i^* - \overline{m}_{i-k} / \prod_{j=1}^{k} p_{i-k+j}^e$ such that the contract remains feasible.

We need to show that there exists conditions under which this new contract $\overline{m}_{i-k}, \overline{m}_i$ outperforms m_i^* for a risk-averse licensor:

$$\begin{aligned}
\left(\prod_{j=1}^{i-k} p_j^o - \prod_{j=1}^{i} p_j^o\right) u(\overline{m}_{i-k}) + \prod_{j=1}^{i} p_j^o u(\overline{m}_{i-k} + \overline{m}_i) &\geq \prod_{j=1}^{i} p_j^o u(m_i^*) \\
\iff \left(1 - \prod_{j=1}^{k} p_{i-k+j}^o\right) u(\overline{m}_{i-k}) + \prod_{j=1}^{k} p_{i-k+j}^o u(\overline{m}_{i-k} + \overline{m}_i) &\geq \prod_{j=1}^{k} p_{i-k+j}^o u(m_i^*) \\
\iff \left(1 - \prod_{j=1}^{k} p_{i-k+j}^o\right) u(\overline{m}_{i-k}) + \prod_{j=1}^{k} p_{i-k+j}^o u(\overline{m}_{i-k} + \overline{m}_i) \geq u(\overline{m}_{i-k}) &\geq \prod_{j=1}^{k} p_{i-k+j}^o u(m_i^*)
\end{aligned}$$

This will hold for all $\prod_{j=1}^{k} p_{i-k+j}^o \leq u(\overline{m}_{i-k})/u(m_i^*)$, i.e., it depends on the concavity of the utility function.

Proof of Theorem 2

Using standard economic theory approach, we relax the licensor's problem by dropping the individual rationality constraint for the high-type licensee and the incentive compatibility constraint for the low-type licensee. This yields the licensor's relaxed optimization problem:

$$\max_{\mathbf{m}_k \geq 0,\ r_k \geq 0,} q_L(m_{0L} + p_1^o m_{1L} + p_1^o p_2^o (m_{2L} + r_L s(x_{LL}^*)) + q_H(m_{0H} + p_1^o m_{1H} + p_1^o p_2^o (m_{2H} + r_H s(x_{HH}^*))$$

$$\text{s.t.} \quad -c_1 - m_{0L} - p^e_{1L}(c_2 + m_{1L} + x^*_{LL}) + p^e_{1L}p^e_{2L}((1-r_L)s(x^*_{LL}) - m_{2L}) \geq 0$$

$$-m_{0H} - p^e_{1H}(m_{1H} + x^*_{HH}) + p^e_{1H}p^e_{2H}((1-r_H)s(x^*_{HH}) - m_{2H}) \geq$$

$$-m_{0L} - p^e_{1H}(m_{1L} + x^*_{HL}) + p^e_{1H}p^e_{2H}((1-r_L)s(x^*_{HL}) - m_{2L})$$

$$m^j_i \geq 0, r^j \geq 0$$

Using the low-type licensee's individual rationality constraint and the high-type licensee's incentive compatibility constraint, we substitute for m_{0L} and m_{0H} in the objective function, and then write the KKT constraints of the Lagrangian W:

$$\partial W/\partial m_{1L}: \quad q_L(p^o_1 - p^e_{1H}) + (p^e_{1H} - p^e_{1L}) - \lambda_1 p^e_{1L} + \lambda_2(p^e_{1H} - p^e_{1L}) + \lambda_3 = 0$$

$$\partial W/\partial m_{1H}: \quad (1-q_L)(p^o_1 - p^e_{1H}) - \lambda_2 p^e_{1H} + \lambda_4 = 0$$

$$\partial W/\partial m_{2L}: \quad q_L(p^o_1 p^o_2 - p^e_{1H}p^e_{2H}) + (p^e_{1H}p^e_{2H} - p^e_{1L}p^e_{2L}) - \lambda_1 p^e_{1L}p^e_{2L} + \lambda_2(p^e_{1H}p^e_{2H} - p^e_{1L}p^e_{2L}) + \lambda_5 = 0$$

$$\partial W/\partial m_{2H}: \quad (1-q_L)(p^o_1 p^o_2 - p^e_{1H}p^e_{2H}) - \lambda_2 p^e_{1H}p^e_{2H} + \lambda_6 = 0$$

$$\partial W/\partial r_L: \quad p^e_{1L}p^e_{2L}s(x^*_{LL}) + (1-q_L)p^e_{1H}p^e_{2H}s(x^*_{HL}) + q_L p^o_1 p^o_2(s(x^*_{LL}) + r_L s'_{r_L})$$
$$- \lambda_1 p^e_{1L}p^e_{2L}s(x^*_{LL}) + \lambda_2(p^e_{1H}p^e_{2H}s(x^*_{HL}) - p^e_{1L}p^e_{2L}s(x^*_{LL})) + \lambda_7 = 0$$

$$\partial W/\partial r_H: \quad (1-q_L)p^o_1 p^o_2(s(x^*_{HH}) + r_H s'_{r_H}) - (1-q_L)p^e_{1H}p^e_{2H}s(x^*_{HH}) - \lambda_2 p^e_{1H}p^e_{2H}s(x^*_{HH}) + \lambda_8 = 0$$

The characteristics of the optimal contract menu for the high-type licensee immediately follow from the solution to the second, fourth and last KKT condition of this relaxed problem. This creates three different scenarios for the high-type licensee, i.e., the optimal contract contains either an upfront payment, a milestone payment after the first phase or a milestone payment at product launch. Solving the first, third and fifth KKT condition under each of these three different scenarios separately, we obtain boundaries for p^o_1 and p^o_2 which determine areas within which different contracts for the low-type licensee will be written.

1. $p^o_1 \leq \min\{p^e_{1H}, p^e_{1H} - \frac{p^e_{1H} - p^e_{1L}}{q_L}\}$ and $p^o_1 p^o_2 \leq \max\{0, \frac{p^e_{1L}(p^e_{2L})^2 - (1-q_L)p^e_{1H}(p^e_{2H})^2}{q_L p^e_{2L}}\}$

   Pooling contract with $m_{0L} = m_{0H}$; m_{0L} as per the low-type licensee's project value.

2. $p^o_1 \leq \min\{p^e_{1H}, p^e_{1H} - \frac{p^e_{1H} - p^e_{1L}}{q_L}\}$ and $\frac{p^e_{1L}(p^e_{2L})^2 - (1-q_L)p^e_{1H}(p^e_{2H})^2}{q_L p^e_{2L}} \leq p^o_1 p^o_2 \leq \min\{p^e_{1H}p^e_{2H}, \frac{p^e_{1L}p^e_{2L} - (1-q_L)p^e_{1H}p^e_{2H}}{q_L}\}$

   Low-type licensee contract (m_{0L}, r_L) and high-type licensee contract m_{0H}.

3. $p^e_{1H} - \frac{p^e_{1H} - p^e_{1L}}{q_L} \leq p^o_1 \leq p^e_{1H}$ and $0 \leq p^o_2 \leq \max\{0, p^e_{2L} - \frac{(1-q_L)p^e_{1H}((p^e_{2H})^2 - (p^e_{2L})^2)}{q_L p^o_1 p^e_{2L}}\}$

   Low-type licensee contract m_{1L} and high-type licensee contract m_{0H}.

4. $p^e_{1H} - \frac{p^e_{1H} - p^e_{1L}}{q_L} \leq p^o_1 \leq p^e_{1H}$ and $p^e_{2L} - \frac{(1-q_L)p^e_{1H}((p^e_{2H})^2 - (p^e_{2L})^2)}{q_L p^o_1 p^e_{2L}} \leq p^o_2 \leq \max\{0, p^e_{2L} - \frac{(1-q_L)p^e_{1H}(p^e_{2H} - p^e_{2L})}{q_L p^o_1}\}$

Low-type licensee contract (m_{1L}, r_L) and high-type licensee contract m_{0H}.

5. $p^e_{1H} \leq p^0_1 \leq \frac{q_L p^e_{1L} p^e_{1H}}{p^e_{1H} - (1-q_L)p^e_{1L}}$ and $0 \leq p^o_2 \leq \max\{0, \min\{p^e_{2H}, \frac{q_L p^e_{1L}(p^e_{2L})^2 p^e_{1H} - (1-q_L)p^o_1(p^e_{1H}(p^e_{2H})^2 - p^e_{1L}(p^e_{2L})^2)}{q_L p^o_1 p^e_{1H} p^e_{2L}}\}\}$

   Low-type licensee contract m_{0L} and high-type licensee contract m_{1H}.

6. $p^e_{1H} \leq p^0_1 \leq \frac{q_L p^e_{1L} p^e_{1H}}{p^e_{1H} - (1-q_L)p^e_{1L}}$ and $\min\{p^e_{2H}, \frac{q_L p^e_{1L}(p^e_{2L})^2 p^e_{1H} - (1-q_L)p^o_1(p^e_{1H}(p^e_{2H})^2 - p^e_{1L}(p^e_{2L})^2)}{q_L p^o_1 p^e_{1H} p^e_{2L}}\} \leq p^o_2 \leq \max\{0, \min\{p^e_{2H}, \frac{q_L p^e_{1L} p^e_{2L} p^e_{1H} - (1-q_L)p^o_1(p^e_{1H} p^e_{2H} - p^e_{1L} p^e_{2L})}{q_L p^o_1 p^e_{1H}}\}\}$

   Low-type licensee contract (m_{0L}, r_L) and high-type licensee contract m_{1H}.

7. $p^0_1 \geq \max\{p^e_{1H}, \frac{q_L p^e_{1L} p^e_{1H}}{p^e_{1H} - (1-q_L)p^e_{1L}}\}$ and $p^o_2 \leq \max\{0, \frac{(p^e_{2L})^2 - (1-q_L)(p^e_{2H})^2}{q_L p^e_{2L}}\}\}$

   Pooling contract $m_{1L} = m_{1H}$; m_{1L} as per the low-type licensee's project value.

8. $p^0_1 \geq \max\{p^e_{1H}, \frac{q_L p^e_{1L} p^e_{1H}}{p^e_{1H} - (1-q_L)p^e_{1L}}\}$ and $\frac{(p^e_{2L})^2 - (1-q_L)(p^e_{2H})^2}{q_L p^e_{2L}} \leq p^o_2 \leq \max\{0, \frac{p^e_{2L} - (1-q_L)p^e_{2H}}{q_L}\}\}$

   Low-type licensee contract (m_{1L}, r_L) and high-type licensee contract m_{1H}.

9. $p^0_1 \leq p^e_{1H}$ and $p^o_2 \geq \max\{0, \frac{p^e_{1L} p^e_{2L} - (1-q_L)p^e_{1H} p^e_{2H}}{q_L p^o_1}, p^e_{2L} - \frac{(1-q_L)p^e_{1H}(p^e_{2H} - p^e_{2L})}{q_L p^o_1}\}$

   Low-type licensee contract (m_{2L}, r_L) and high-type licensee m_{0H}.

10. $p^0_1 \geq p^e_{1H}$ and $\max\{\frac{q_L p^e_{1L} p^e_{2L} p^e_{1H} - (1-q_L)p^o_1(p^e_{1H} p^e_{2H} - p^e_{1L} p^e_{2L})}{q_L p^o_1 p^e_{1H}}, \frac{p^e_{2L} - (1-q_L)p^e_{2H}}{q_L}\} \leq p^o_2 \leq p^e_{2H}$

    Low-type licensee contract (m_{2L}, r_L) and high-type licensee contract m_{1H}.

11. $p^o_2 \geq p^e_{2H}$ and $p^e_{1H} p^e_{2H} \leq p^o_1 p^o_2 \leq \max\{p^e_{1H} p^e_{2H}, \frac{q_L p^e_{1L} p^e_{2L} p^e_{1H} p^e_{2H}}{q_L p^e_{1H} p^e_{2H} - (1-q_L)(p^e_{1L} p^e_{2L} - p^e_{1H} p^e_{2H})}\}$

    Low-type licensee contract (m_{0L}, r_L) and high-type licensee contract m_{2H}.

12. $p^o_2 \geq p^e_{2H}$ and $p^o_1 p^o_2 \geq \max\{p^e_{1H} p^e_{2H}, \frac{q_L p^e_{1L} p^e_{2L} p^e_{1H} p^e_{2H}}{q_L p^e_{1H} p^e_{2H} - (1-q_L)(p^e_{1L} p^e_{2L} - p^e_{1H} p^e_{2H})}\}$

    Low-type licensee contract (m_{2L}, r_L) and high-type licensee contract m_{2H}.

Finally, using the structure of the optimal contract menu, it is easy to verify that the optimal contracts satisfy the constraints that were previously dropped.

Proof of Lemma 4

Using the relaxation approach outlined in Theorem 2, we drop the high-type licensee's individual rationality constraint and the low-type licensee's incentive compatibility constraint. We can observe from the resulting KKT conditions that the contract for the high-type licensee is not distorted, i.e., follows the optimal structure defined under Lemma 3.

Proof of Theorem 3

In the absence of adverse selection, suppose it is optimal for the risk-neutral licensor to delay licensing for $k \leq n-1$ phases and offer a contract with a milestone payment $m_l, l \in \{1, ..., n-k\}$. To determine the unique optimal contract, we use Theorem 1.

For notational convenience, let us introduce $W^e(k)$, the value of the project to the licensee at the start of phase $k+1$. Then the milestone payment will be $m_l = \frac{W^e(k)}{\prod_{i=k+1}^{k+l} p_i^e}$.

The licensor's value from the contract at the time of licensing is: $\prod_{i=k+1}^{k+l} p_i^o m_l$.

Should the licensor delay the contract one more phase, to $k+1$, Theorem 1 specifies that the milestone payment would still be located at the end of the same phase, and would thus be $\overline{m}_0 = W^e(k+1)$ if $l = 1$, and $\overline{m}_{l-1} = \frac{W^e(k+1)}{\prod_{i=k+2}^{k+l} p_i^e}$ otherwise.

The licensor's new value from the contract if $l > 1$ is: $-c_k + \prod_{i=k+1}^{k+l} p_i^o \overline{m}_l$, which we can prove is larger than $\prod_{i=k+1}^{k+l} p_i^o m_l$:

$$-c_k + \prod_{i=k+1}^{k+l} p_i^o \frac{W^e(k+1)}{\prod_{i=k+2}^{k+l} p_i^e} \geq \prod_{i=k+1}^{k+l} p_i^o \frac{W^e(k)}{\prod_{i=k+1}^{k+l} p_i^e}$$

$$-\prod_{i=k+1}^{k+l} p_i^e c_k + p_{k+1}^e \prod_{i=k+1}^{k+l} p_i^o W^e(k+1) \geq \prod_{i=k+1}^{k+l} p_i^o [-c_k + p_{k+1}^e W^e(k+1)]$$

$$c_k \left[\prod_{i=k+1}^{k+l} p_i^o - \prod_{i=k+1}^{k+l} p_i^e \right] \geq 0$$

since $\prod_{i=k+1}^{k+l} p_i^o \geq \prod_{i=k+1}^{k+l} p_i^e$ by Theorem 1.

Prove similarly for $l = 1$.

Thus it is not optimal to license the project at a phase such that the optimal contract would contain a milestone payment. Rather, it can be successively further delayed until the optimal contract contains an upfront payment, thus increasing its value, or until no further delay is possible because the licensee's development effort is needed.

Proof of Lemma 5

Suppose that for a risk-neutral licensor it is optimal to delay the licensing of the project until phase $k < n-1$, at which phase he receives an upfront payment m_0^*. This generates a series of cash flows for the licensor $(-c_1, ..., -c_{k-1}, m_0^*)$ with their associated probabilities, the product of his PTS estimates.

Should a risk-averse licensor delay the project beyond phase k, he will face the same cash outflows up until phase k. After that, depending on the delay chosen and the corresponding optimal contract terms, the cash flows will change to $(z_1, ..., z_{n-k})$. This generates the following series of cash flows: $(-c_1, ..., -c_{k-1}, -c_k, z_1, ..., z_{n-k})$ with their associated probabilities, the product of his PTS estimates.

Since the first part is identical, it can be ignored. We know from the fact that delaying the licensing of the project until phase $k < n-1$ for an upfront payment m_0^* is optimal for the risk-neutral licensor that by Theorem 1:

$$\begin{aligned}
m_0^* &\geq -c_k + \sum_{i=1}^{n-k} \prod_{j=k+1}^{k+i} p_j^o z_i \\
&\geq -(1-p_{k+1}^o)c_k + \sum_{i=1}^{n-k-1} (1-p_{k+i+1}^o) \prod_{j=k+1}^{k+i} p_j^o \left(-c_k + \sum_{j=1}^{i} z_i\right) + \prod_{j=k+1}^{n} p_j^o \left(-c_k + \sum_{j=1}^{n} z_i\right)
\end{aligned}$$

Since $u^o(.)$ is a monotonically increasing concave function:

$$\begin{aligned}
u^o(m_0^*) &\geq u^o\left(-(1-p_{k+1}^o)c_k + \sum_{i=1}^{n-k-1} (1-p_{k+i+1}^o) \prod_{j=k+1}^{k+i} p_j^o \left(-c_k + \sum_{j=1}^{i} z_i\right)\right. \\
&\quad \left. + \prod_{j=k+1}^{n} p_j^o \left(-c_k + \sum_{j=1}^{n} z_i\right)\right) \\
\Longleftrightarrow u^o(m_0^*) &\geq (1-p_{k+1}^o)u^o(-c_k) + p_{k+1}^o u^o\left((1-p_{k+2}^o)\left(-c_k + \sum_{j=1}^{i} z_i\right)\right. \\
&\quad \left. + \sum_{i=2}^{n-k-1} (1-p_{k+i+1}^o) \prod_{j=k+2}^{k+i} p_j^o \left(-c_k + \sum_{j=1}^{i} z_i\right) + \prod_{j=k+2}^{n} p_j^o \left(-c_k + \sum_{j=1}^{n} z_i\right)\right) \\
\Longleftrightarrow u^o(m_0^*) &\geq (1-p_{k+1}^o)u^o(-c_k) + \sum_{i=1}^{n-k-1} (1-p_{k+i+1}^o) \prod_{j=k+1}^{k+i} p_j^o u^o\left(-c_k + \sum_{j=1}^{i} z_i\right) \\
&\quad + \prod_{j=k+1}^{n} p_j^o u^o\left(-c_k + \sum_{j=1}^{n} z_i\right)
\end{aligned}$$

The concavity is applied in successive steps until the coefficients are fully extracted from the utility function to obtain the last inequality.

Proof of Lemma 6

(a) Given Theorem 2, the optimal two-stage contract will not include an upfront payment for either licensee type in the following four cases, for which we need to show that delaying is optimal:

1. $p_1^0 \geq \max\{p_{1H}^e, \frac{q_L p_{1L}^e p_{1H}^e}{p_{1H}^e - (1-q_L)p_{1L}^e}\}$ and $p_2^o \leq \max\{0, \frac{(p_{2L}^e)^2 - (1-q_L)(p_{2H}^e)^2}{q_L p_{2L}^e}\}\}$

   Under immediate licensing, the contract consists of $m_{1L} = m_{1H} = [-c_1 - p_{1L}^e(c_2 + x_{LL}^*) + p_{1L}^e p_{2L}^e s(x_{LL}^*)]/p_{1L}^e$. Assume that we delay licensing and ask for a feasible contract with upfront payment only, i.e., $\overline{m}_{1L} = \overline{m}_{1H} = -(c_2 + x^*_{LL}) + p^e_{2L}s(x^*_{LL})$. Then,

$$\frac{p^o_1}{p^e_{1L}}[-c_1 - p^e_{1L}(c_2 + x^*_{LL}) + p^e_{1L}p^e_{2L}s(x^*_{LL})] \leq -c_1 + p^o_1(-c_2 - x^*_{LL} + p^e_{2L}s(x^*_{LL}))$$
$$\iff p^e_{1L} \leq p^o_1,$$

which is true by assumption.

2. $p^0_1 \geq \max\{p^e_{1H}, \frac{q_L p^e_{1L}p^e_{1H}}{p^e_{1H}-(1-q_L)p^e_{1L}}\}$ and $\frac{(p^e_{2L})^2-(1-q_L)(p^e_{2H})^2}{q_L p^e_{2L}} \leq p^o_2 \leq \max\{0, \frac{p^e_{2L}-(1-q_L)p^e_{2H}}{q_L}\}\}$

   Under immediate licensing, the low-type licensee contract is (m_{1L}, r_L) and the high-type licensee contract m_{1H}, with $m_{1L} = [-c_1 - p^e_{1L}(c_2 + x^*_{LL}) + p^e_{1L}p^e_{2L}(1 - r_L)s(x^*_{LL})]/p^e_{1L}$ and $m_{1H} = m_{1L} - (x^*_{HH} - x^*_{HL}) + p^e_{2H}(s(x^*_{HH}) - (1 - r_L)s(x^*_{HL})) = m_{1L} + \Delta$. Assume that we delay licensing and ask for a feasible contract menu $(\overline{m}_{0L}, r_L)$ and $\overline{m}_{0H}$, with $\overline{m}_{0L} = -c_2 - x^*_{LL} + p^e_{2L}(1 - r_L)s(x^*_{LL})$ and $\overline{m}_{0H} = \overline{m}_{0L} + \Delta$. Then,

$$\frac{p^o_1}{p^e_{1L}}[-c_1 - p^e_{1L}(c_2 + x^*_{LL}) + p^e_{1L}p^e_{2L}(1 - r_L)s(x^*_{LL})] + q_L p^o_1 p^o_2 r_L s(x^*_{LL}) + (1 - q_L)p^o_1\Delta$$
$$\leq -c_1 + p^o_1(-c_2 - x^*_{LL} + p^e_{2L}(1 - r_L)s(x^*_{LL})) + q_L p^o_1 p^o_2 r_L s(x^*_{LL}) + (1 - q_L)p^o_1\Delta$$
$$\iff p^e_{1L} \leq p^o_1,$$

   which is true by assumption.

3. $p^0_1 \geq p^e_{1H}$ and $\max\{\frac{q_L p^e_{1L}p^e_{2L}p^e_{1H}-(1-q_L)p^o_1(p^e_{1H}p^e_{2H}-p^e_{1L}p^e_{2L})}{q_L p^o_1 p^e_{1H}}, \frac{p^e_{2L}-(1-q_L)p^e_{2H}}{q_L}\} \leq p^o_2 \leq p^e_{2H}$

   Under immediate licensing, the low-type licensee contract is (m_{2L}, r_L) and the high-type licensee contract is m_{1H}, with $m_{2L} = [-c_1 - p^e_{1L}(c_2 + x^*_{LL}) + p^e_{1L}p^e_{2L}(1 - r_L)s(x^*_{LL})]/(p^e_{1L}p^e_{2L})$ and $m_{1H} = p^e_{2H}m_{2L} - (x^*_{HH} - x^*_{HL}) + p^e_{2H}(s(x^*_{HH}) - (1 - r_L)s(x^*_{HL})) = p^e_{2H}m_{2L} + \Delta$. Assume that we delay licensing and ask for a feasible contract menu $(\overline{m}_{1L}, r_L)$ and $\overline{m}_{0H}$, with $\overline{m}_{1L} = [-c_2 - x^*_{LL} + p^e_{2L}(1 - r_L)s(x^*_{LL})]/p^e_{2L}$ and $\overline{m}_{0H} = p^e_{2H}\overline{m}_{1L} + \Delta$. Then,

$$\frac{p^o_1}{p^e_{1L}p^e_{2L}}(q_L p^o_2 + (1 - q_L)p^e_{2H})[-c_1 - p^e_{1L}(c_2 + x^*_{LL}) + p^e_{1L}p^e_{2L}(1 - r_L)s(x^*_{LL})]$$
$$+ q_L p^o_1 p^o_2 r_L s(x^*_{LL}) + (1 - q_L)p^o_1\Delta$$
$$\leq -c_1 + \frac{p^o_1}{p^e_{2L}}(q_L p^o_2 + (1 - q_L)p^e_{2H})(-c_2 - x^*_{LL} + p^e_{2L}(1 - r_L)s(x^*_{LL}))$$
$$+ q_L p^o_1 p^o_2 r_L s(x^*_{LL}) + (1 - q_L)p^o_1\Delta$$
$$\iff p^e_{1L}p^e_{2L} \leq p^o_1(q_L p^o_2 + (1 - q_L)p^e_{2H}),$$

   which is true by assumption.

4. $p^o_2 \geq p^e_{2H}$ and $p^o_1 p^o_2 \geq \max\{p^e_{1H}p^e_{2H}, \frac{q_L p^e_{1L}p^e_{2L}p^e_{1H}p^e_{2H}}{q_L p^e_{1H}p^e_{2H}-(1-q_L)(p^e_{1L}p^e_{2L}-p^e_{1H}p^e_{2H})}\}$

Under immediate licensing, the low-type licensee contract is (m_{2L}, r_L) and the high-type licensee contract is m_{2H}, with $m_{2L} = [-c_1 - p^e_{1L}(c_2 + x^*_{LL}) + p^e_{1L}p^e_{2L}(1-r_L)s(x^*_{LL})]/(p^e_{1L}p^e_{2L})$ and $m_{2H} = m_{2L} + [-(x^*_{HH} - x^*_{HL}) + p^e_{2H}(s(x^*_{HH}) - (1-r_L)s(x^*_{HL}))]/p^e_{2H} = m_{2L} + \Delta/p^e_{2H}$. Assume that we delay licensing and ask for a feasible contract menu $(\overline{m}_{1L}, r_L)$ and $\overline{m}_{1H}$, with $\overline{m}_{1L} = [-c_2 - x^*_{LL} + p^e_{2L}(1-r_L)s(x^*_{LL})]/p^e_{2L}$ and $\overline{m}_{1H} = \overline{m}_{1L} + \Delta/p^e_{2H}$. Then,

$$\begin{aligned}
&\frac{p^o_1 p^o_2}{p^e_{1L}p^e_{2L}}[-c_1 - p^e_{1L}(c_2 + x^*_{LL}) + p^e_{1L}p^e_{2L}(1-r_L)s(x^*_{LL})] + q_L p^o_1 p^o_2 r_L s(x^*_{LL}) + (1-q_L)p^o_1\Delta \\
&\qquad \le -c_1 + \frac{p^o_1 p^o_2}{p^e_{2L}}(-c_2 - x^*_{LL} + p^e_{2L}(1-r_L)s(x^*_{LL})) + q_L p^o_1 p^o_2 r_L s(x^*_{LL}) + (1-q_L)p^o_1\Delta \\
&\iff p^e_{1L}p^e_{2L} \le p^o_1 p^o_2,
\end{aligned}$$

which is true by assumption.

(b) Given Theorem 2, the optimal two-stage contract will only include an upfront payment for both licensee types in the following two cases, for which we need to show that delaying will never be optimal:

1. $p^o_1 \le \min\{p^e_{1H}, p^e_{1H} - \frac{p^e_{1H} - p^e_{1L}}{q_L}\}$ and $p^o_1 p^o_2 \le \max\{0, \frac{p^e_{1L}(p^e_{2L})^2 - (1-q_L)p^e_{1H}(p^e_{2H})^2}{q_L p^e_{2L}}\}$

   Under immediate licensing, the licensor obtains $m_{0L} = m_{0H} = -c_1 - p^e_{1L}(c_2 + x^*_{LL}) + p^e_{1L}p^e_{2L}s(x^*_{LL})$.

   After delaying, the optimal contract could be either of the following menus:

   (a) $\overline{m}_{0L} = \overline{m}_{0H} = -c_2 - x^*_{LL} + p^e_{2L}s(x^*_{LL})$. It is immediately obvious that this yields a lower value for the licensor than immediate licensing.

   (b) $(\overline{m}_{0L}, r_L)$ and $\overline{m}_{0H}$, with $\overline{m}_{0L} = -c_2 - x^*_{LL} + p^e_{2L}(1-r_L)s(x^*_{LL})$ and $\overline{m}_{0H} = \overline{m}_{0L} - (x^*_{HH} - x^*_{HL}) + p^e_{2H}(s(x^*_{HH}) - (1-r_L)s(x^*_{HL})) = \overline{m}_{0L} + \Delta$. Assuming we license immediately with the same (suboptimal) royalty rate r_L and the corresponding feasible milestone payments m_{0L} and m_{0H}, we get:

   $$\begin{aligned}
   &p^e_{1L}(-c_2 - x^*_{LL} + p^e_{2L}(1-r_L)s(x^*_{LL})) + p^e_{1H}(1-q_L)\Delta \\
   &\qquad \ge p^o_1(-c_2 - x^*_{LL} + p^e_{2L}(1-r_L)s(x^*_{LL}) + (1-q_L)\Delta)
   \end{aligned}$$

   This always holds since $p^o_1 \le \min\{p^e_{1H}, p^e_{1H} - \frac{p^e_{1H} - p^e_{1L}}{q_L}\}$.

   (c) $(\overline{m}_{1L}, r_L)$ and $\overline{m}_{0H}$, with $\overline{m}_{1L} = [-c_2 - x^*_{LL} + p^e_{2L}(1-r_L)s(x^*_{LL})]/p^e_{2L}$ and $\overline{m}_{0H} = p^e_{2H}\overline{m}_{1L} - (x^*_{HH} - x^*_{HL}) + p^e_{2H}(s(x^*_{HH}) - (1-r_L)s(x^*_{HL}))$. Assuming we license immediately with the same (suboptimal) royalty rate r_L and the corresponding feasible milestone payments m_{0L} and m_{0H}, we get:

$$\begin{aligned}&p_{1L}^e(-c_2 - x_{LL}^* + p_{2L}^e(1-r_L)s(x_{LL}^*)) + p_{1H}^e(1-q_L)\Delta \\ &\quad \geq (p_{1L}^e - C)(-c_2 - x_{LL}^* + p_{2L}^e(1-r_L)s(x_{LL}^*)) + p_1^o(1-q_L)\Delta\end{aligned}$$

where $C \leq (1-q_L)p_{1H}^e(p_{1H}^e p_{2H}^e - p_{1L}^e p_{2L}^e)/p_1^o p^{eL_2}$. This always holds since $p_1^o \leq \min\{p_{1H}^e, p_{1H}^e - \frac{p_{1H}^e - p_{1L}^e}{q_L}\}$.

(d) $(\overline{m}_{1L}, r_L)$ and $\overline{m}_{1H}$, with $\overline{m}_{1L} = [-c_2 - x_{LL}^* + p_{2L}^e(1-r_L)s(x_{LL}^*)]/p_{2L}^e$ and $\overline{m}_{1H} = \overline{m}_{1L} + [-(x_{HH}^* - x_{HL}^*) + p_{2H}^e(s(x_{HH}^*) - (1-r_L)s(x_{HL}^*))]/p_{2H}^e$. Assuming we license immediately with the same (suboptimal) royalty rate r_L and the corresponding feasible milestone payments m_{0L} and m_{0H}, we get:

$$\begin{aligned}&p_{1L}^e(-c_2 - x_{LL}^* + p_{2L}^e(1-r_L)s(x_{LL}^*)) + p_{1H}^e(1-q_L)\Delta \\ &\quad \geq \frac{p_1^o p_2^o}{p_{2L}^e}(-c_2 - x_{LL}^* + p_{2L}^e(1-r_L)s(x_{LL}^*)) + \frac{p_1^o p_2^o}{p_{2H}^e}(1-q_L)\Delta\end{aligned}$$

where $\frac{p_1^o p_2^o}{p_{2L}^e} \leq p_{1L}^e$ and $\frac{p_1^o p_2^o}{p_{2H}^e} \leq p_{1H}^e$.

2. $p_1^o \leq \min\{p_{1H}^e, p_{1H}^e - \frac{p_{1H}^e - p_{1L}^e}{q_L}\}$ and $\frac{p_{1L}^e(p_{2L}^e)^2 - (1-q_L)p_{1H}^e(p_{2H}^e)^2}{q_L p_{2L}^e} \leq p_1^o p_2^o \leq \min\{p_{1H}^e p_{2H}^e, \frac{p_{1L}^e p_{2L}^e - (1-q_L)p_{1H}^e p_{2H}^e}{q_L}\}$

   Low-type licensee contract (m_{0L}, r_L) and high-type licensee contract m_{0H}.

   The analysis is similar to the 4 cases elaborated above for the pooling contract at immediate licensing.