# Efficient R&D Delegation *

Etienne Billette de Villemeur
Université de Toulouse (IDEI and GREMAQ)
and
Bruno Versaevel
Ecole de Management de Lyon and GATE

September 2006

**Abstract**

This paper constructs a model of R&D delegation, in which several firms may outsource new technology from the same for-profit laboratory. When R&D costs are uncertain at the contracting stage, necessary and sufficient conditions are provided that characterize Nash equilibria in the laboratory's outputs and the firms' non-negative transfer payments. A subset of these equilibria is identified in which firms' payments truthfully reflect their valuation of all possible alternatives vis-á-vis the expected net profits. These truthful Nash equilibria implement the first-best outcome when the support of the stochastic parameter is "not too large", in which case firms' net payoffs do not depend on the stochastic component and are undominated. When externalities impact the nature of competition among firms, both on the intermediate market for R&D services and the final market for goods, sufficient conditions are given for the laboratory to earn positive benefits or not, and for firms to choose either to cooperate horizontally or to acquire the laboratory. The latter two options are shown not to impact the R&D outputs supplied by the laboratory, nor the joint profits earnt by firms. A first policy implication is that no new regulatory tool is needed for aligning firms' interests with a social welfare objective. The results also support legal environments which do not prevent firms to include discriminatory clauses in their R&D contracts on the intermediate market for new knowledge, including in highly concentrated industries. This *laissez faire* message does not extend to final-market antitrust considerations.

*JEL classification:* C72; D72.
*Keywords:* R&D; common agency; adverse selection; limited liability; externalities; policy.

*Previous versions of the paper have circulated under the title "Conflict and Cooperation on R&D Markets". The authors wish to thank Joshua Graff Zivin, Michel Le Breton, Alessandro Pavan, Neeraj Sood and seminar participants at Northwestern U., Arizona State U., HEC Montreal, Marne la Vallée, Università di Padova, ESEM06, Université de Montpellier. None of them bears any responsibility for any error. We also like to thank the attendees of the "Pharmaceutical Economics and Policy Council 2006" (Pfizer) for introducing us to some specifics of the industry. The views presented here are nonetheless only those of the authors. Corresponding author: Etienne Billette de Villemeur (etienne.devillemeur@univ-tlse1.fr) – Tel.: +33+(0)5 61 12 85 68 – Fax: +33+(0)5 61 12 86 37.

Efficient R&D Delegation


**Abstract**

This paper constructs a model of R&D delegation, in which several firms may outsource new technology from the same for-profit laboratory. When R&D costs are uncertain at the contracting stage, necessary and sufficient conditions are provided that characterize Nash equilibria in the laboratory's outputs and the firms' non-negative transfer payments. A subset of these equilibria is identified in which firms' payments truthfully reflect their valuation of all possible alternatives *vis-à-vis* the expected net profits. These truthful Nash equilibria implement the first-best outcome when the support of the stochastic parameter is "not too large", in which case firms' net payoffs do not depend on the stochastic component and are undominated. When externalities impact the nature of competition among firms, both on the intermediate market for R&D services and the final market for goods, sufficient conditions are given for the laboratory to earn positive benefits or not, and for firms to choose either to cooperate horizontally or to acquire the laboratory. The latter two options are shown not to impact the R&D outputs supplied by the laboratory, nor the joint profits earnt by firms. A first policy implication is that no new regulatory tool is needed for aligning firms' interests with a social welfare objective. The results also support legal environments which do not prevent firms to include discriminatory clauses in their R&D contracts on the intermediate market for new knowledge, including in highly concentrated industries. This *laissez faire* message does not extend to final-market antitrust considerations.



JEL Classification: C72; D72.

Keywords: *Keywords:* R&D; common agency; adverse selection; limited liability; externalities; policy.

# 1 Introduction

Firms increasingly outsource new technology by delegating the production of R&D services to independent laboratories.[1] This phenomenon is documented by recent empirical data. The National Science Foundation (2006) indicates that "[t]he average annual growth rate of contracted-out R&D from 1993 to 2003 (9.4%) was about double the growth rate of in-house company-funded R&D (4.9%)".[2] The chemical industry leads this trend, as it reports the largest expenditures of all industries, with more than 15% of the total in all three years. In this sector, a fragmented set of small specialized for-profit laboratories – *i.e.*, new biotechnology firms (NBFs) – is serving large established firms that control distribution channels of pharmaceutical, agricultural, or other chemical products. In many instances, the same laboratory signs bilateral contractual agreements with a set of firms which compete on a vertically related final market. For example, Symyx (a U.S.-based private laboratory) uses a proprietary high-speed combinatorial technology – including instrumentation, software, and methods – to provide R&D services to Eli Lilly, Merck, and Pfizer, among others. Over the last two decades, business analyses have documented many changes that occurred in that sector on the intermediate market for R&D. While some firms substituted contracted out R&D for in-house projects, other players integrated backward by acquiring existing laboratories. More recently, horizontal mergers have increased the concentration of the downstream stage of the market.[3]

The objective of this paper is (i) to investigate the distribution of profits between a set of firms and an independent supplier of R&D services, (ii) to unveil firms' incentives either to coordinate their payment strategies on the intermediate market for R&D, or to integrate the R&D supplier, and (iii) to draw policy implications. To do that, we construct a model in which several firms may either rely on internal facilities or delegate the production of specified multi-dimensional R&D services to a profit-maximizing laboratory in exchange of positive payments.

Beyond particularities of all kinds, we focus on situations in which R&D operations are inherently uncertain, in the sense that difficulties may arise in a laboratory that were not anticipated in preparatory studies on the feasibility of an aimed outcome. The descriptive literature on the organization of R&D has frequently pointed to the uncertain character of R&D projects. An early example is Hamberg (1963), who emphasizes the gap between the theoretical design of a new product or process and successful experimentation in the laboratory, on the grounds that "unforeseen problems commonly arise that were not seen in preliminary investigations indicating the invention to be technically feasible" (p. 101). However, only a few sources evoke R&D contracts. An evergreen example is a text by Mowery and Rosenberg (1989), who claim that "[t]he effectiveness of contracts in the provision of research is undermined by the highly uncertain nature of the research enterprise", and more precisely by the "imperfect character of knowledge about a given project" (p. 85). Projects of unknown difficulty occur when R&D objectives – *i.e.*, the subject-matter of an R&D contract – impose the laboratory to experiment new research strategies, and not only to develop applications by following existing avenues. To realistically capture the fact that R&D projects are inherently uncertain for all parties – i.e. the contracting firms and the independent laboratory altogether – we assume that R&D contracts are written *ex ante*, that is before the laboratory tests technological options and thereby discovers the realization of a stochastic parameter. This is grounded on the observation that R&D contracts frequently refer to a preliminary stage of operations that consists in probing, screening, and experimenting possible protocols to assess the dificulty of a given project. Accordingly, the timing of events in our formal setup is borrowed from a well-known analysis of the production of innovations by Holmstrom (1989), we consider as a starting point. The latter paper is seminal in that it describes the production of innovations as a principal-agent problem for the first time. The specification that the agent in charge of a single innovation project with uncertain payoff "has no superior information about project returns *before* acting" is qualified as a "reasonable assumption if we are at the initial stages of a research undertaking" (p. 310, original emphasis). In what follows, the situation in which a single principal – *i.e.*, a firm – contracts with an agent

– we call a laboratory – is considered as a benchmark case. On this we build by turning to the case in which several firms are interested in contracting with the same laboratory in order to save on internal resources, transfer risks to an outside party, or benefit from economies of scale and/or scope.

The main features of our delegated R&D common agency game appear clearly when compared with related contributions to the literature on the organization of R&D. The assumption that firms and the laboratory contract for the production of specified R&D services, before the laboratory may observe the realization of its costs, contrasts with the specifications of two related contributions by Aghion and Tirole (1994) and Ambec and Poitevin (2001). Both papers focus on the impact of the non-deterministic character of R&D on the relative efficiency of a separated governance structure (in which a user buys an innovation from an independent unit) and an integrated structure (in which the user sources R&D internally). However, they fundamentally differ in their theoretical use of uncertainty. The first one emphasizes the non-contractibility of R&D outcome, which is assumed to be "ill defined *ex ante*" so that "parties cannot contract for delivery of a specific innovation" (p. 1186). The second paper stresses an informational problem as attached to the risky nature of R&D operations, which makes sense when the research unit has "more information about the cost of bringing the innovation to the market" (p. 2).

The present paper is complementary to the latter two contributions, in the sense that it captures the uncertain nature of R&D in a new way.

Firstly, we consider firm-specific R&D contracts which describe the nature of the activity to be conducted by the laboratory. The subject-matter of these contracts is the delivery of new but clearly identified tailor-made process or product technologies. This follows the general observation by innovation experts that, in the pharmaceutical industry, firms "contract out *specific* research tasks to an independent laboratory" (Tapon, 1989, p. 198, added emphasis). Indeed, examples of available contracts are many in which the client's objectives (say, the purchase of a new drug candidate with targeted biochemical parameters for the cure of a given pathology) are well specified.[4] This differs sharply from Aghion and Tirole (1994), in which the R&D contract is assumed to specify the allocation of property rights "on any forthcoming innovation" (p. 1186).

Moreover, we focus on situations in which all parties – including the agent in charge of R&D activities – are unaware of the true cost of an R&D program before it starts. This echoes quotations from researchers in the biotechnology sector, as reproduced by Tapon and Cadsby (1996, p. 389), according to whom "[y]ou're always going to have things that happen that *nobody* really foresaw" in the laboratory before getting hands dirty (added emphasis). In more technical terms, this is captured by the assumption that some unknown parameter is observed only in the course of R&D operations. This contrasts with Ambec and Poitevin (2001, p. 2), who assume that "the innovation quality is private information" when the research unit and the user reach the contracting stage toward a subsequent development phase.[5]

To summarize, the problem we consider is a purely organizational one, since the choice to outsource R&D services, and not to rely on proprietary resources, is not motivated by a relatively limited expertise of buyers *vis-à-vis* a technology provider. We seek a characterization of the distribution of profits between a set of firms and a common laboratory on a decentralized market for new knowledge, in which contracts are signed *ex ante*, and to compare them with the gains obtained in alternative market structures. This comes as a complement to the analysis of situations in which R&D activities are either non contractible or sourced from an *a priori* superiorly informed entity.

On the technical side, we construct a model in which several principals simultaneously address monetary transfers to an agent in order to influence its choice of an action, in the spirit of Bernheim and Whinston (1986a)'s framework of delegated common agency. A main result of this seminal contribution, in which information is complete and preferences are quasi-linear, is that a subset of Nash equilibria – referred to as "truthful" – induce an efficient choice, *i.e.* maximize the total payoffs of all principals and the agent. Our model, is an adverse selection variation of the same framework, with uncertainty at the contracting stage. More precisely, this is done by extending on several important fronts an existing setup by Laussel and Le Breton (1998). In their paper, principals offer transfer payments to influence the private provision of a public good by a common agent. The latter agent's production cost is impacted by a stochastic coefficient, which is unknown by all parties at the contracting stage. We extend the analysis on several important fronts for the analysis of a larger class of situations in which 1) the agent's output needs not be limited to a public good, 2) no specific form is introduced that limits the way the stochastic parameter relates to the agent's fixed or variable costs, and 3) externalities impact the nature of competition among principals for the use of the agent's resources.[6] In addition, we introduce a limited liability constraint on transfer payments. (This specification is motivated by the observation that independent specialized laboratories usually lack financial support, whereas firms that outsource R&D do not.)

The main theoretical results are as follows. We first unveil necessary and sufficient conditions that characterize Nash equilibria in the agent's choices and the principals' transfer payments. Then we characterize a subset of these equilibria in which principals' payments reflect their valuation of all possible alternatives *vis-à-vis* the *expected* net payoffs, that is an extension of the truthfulness refinement to our model with uncertainty. When the support of the stochastic parameter is "not too large", it is found that truthful Nash equilibria implement the first-best outcome. That is, efficiency is preserved *ex post*, in the sense that the agent's equilibrium choice and principals' non-negative transfers maximize the net total benefits of all parties, and for any realization of the random parameter. Moreover, for this subset of equilibria, principals' net equilibrium payoffs do not depend on the stochastic component and are undominated. This says that the agent only – not principals – bears the cost of uncertainty.

Eventually, we exploit the theoretical findings to investigate firms' incentives to coordinate their payment strategies (which can be achieved by merging horizontally), or to integrate the external laboratory. It is found that, when anti-complementarities dominate in the production of R&D services, delegating firms have strategic incentives to cooperate horizontally in their choices of payment schemes as addressed to the laboratory. Horizontal cooperation reduces competition on the market for R&D services, and thus drains up the laboratory's source of profits. On the other hand, by acquiring the laboratory a firm is entitled to ask outsiders a premium for the negative externalities it imposes by receiving R&D services. However, when complementarities dominate, delegating firms have no strategic incentive to shift to a more integrated structure, because they already appropriate all profits. This clear-cut opposition holds in the absence of efficiency gains or transaction costs, as specifically associated to particular governance structures.

In substance, incentives to merge vertically with the laboratory, or horizontally with a rival firm, are rooted in the simultaneous impact of two types of externalities on the nature of firm interactions. Externalities are of the *indirect* type when the laboratory's costs are not additively separable across users. If the cost of satisfying a firm's requirements depends on the level of efforts provided to meet another firm's needs, then firms interact through their respective specifications and associated payments. Externalities can be also of the *direct* type if each firm's gross profit function depends not only on the R&D services it receives, but also on the services received by another firm. This occurs, for example, when R&D results are not fully appropriable by users, and give rise to inter-firm spillovers.[7] Intuitively, negative externalities make competition tougher, whereas positive externalities make it softer, on the market for R&D services. Whether competition is relatively tough or soft is reflected by each firm's payment offers, and thus drives the distribution of innovation benefits between the laboratory and firms. In turn, the distribution of benefits impacts firms' incentives to merge horizontally, or to acquire the laboratory.

The remainder of the paper is organized as follows. In section 2, we construct a delegated R&D common agency game. In section 3, the uncertain character of R&D is introduced. In section 4, we offer a characterization of Nash equilibria. In section 5, examples illustrate the potential of the model. In section 6, the truthfulness refinement is introduced and truthful Nash equilibria are characterized. In section 7, sufficient conditions are given for the (non) appropriability of (some share of) total net equilibrium benefits by the laboratory. In section 8, situations are identified in which firms have an incentive to acquire the laboratory or to merge with another firm. Section 9 concludes the paper. Detailed proofs are in the Appendix.

# 2 The Model

In this section, we first present general specifications. Then evoking the taxation principle to justify the contractual form on which we focus. We also construct the contracting process as a delegated common agency situation. Eventually we examine the specific nature of uncertainty as introduced in this model.

## 2.1 General Specifications

Risk-neutral independent laboratories supply cost-reducing or/and demand-enhancing knowledge, *i.e.* R&D services, which are produced and delivered at some cost. These services may be used by risk-neutral profit-maximizing firms, we index by $i \in N$. These firms may choose either to operate in-house R&D operations, or to delegate the production of R&D services. In the latter case, each firm may tap R&D services from a dedicated laboratory, a situation we consider as a benchmark case. We focus on another situation in which firms may outsource R&D from the same independent laboratory – we identify by $\mathcal{L}$ – in order, say, to benefit from economies of scale or scope.

Uncertainty in the production of new knowledge implies that R&D costs cannot be known before some experimentation and tests are conducted. In slightly more general terms, the environment in which the laboratory operates can be impacted by stochastic events. We thus assume that the laboratory's technology is parametrized by a stochastic element $\theta$ in a (possibly multi-dimensional) set $\Theta$. As the firms may rely on proprietary state-of-the-art R&D resources, they are not *a priori* asymmetrically informed vis-à-vis the laboratory. Therefore the distribution of the stochastic parameter is described by a probability measure $\mu$ which is common knowledge to firms and $\mathcal{L}$. This means that, before the laboratory initiates R&D operations, all parties are equally aware of the difficulty of conducting R&D projects. The laboratory may learn more on the costs of R&D tasks, through the realization of the stochastic parameter, only by doing. The stochastic parameter is an argument of a function $r^S(\theta, \mathbf{x}) : 2^N \times \Theta \times X \rightarrow \Re_+$, which represents the cost in monetary units born by a laboratory when it accepts to contract with firms in $S$ in $2^N$ (which includes the empty set), in any state of nature, for the delivery of R&D services $\mathbf{x}$ in $X \subset R_+^{m \times n}$, where $m$ is the number of service dimensions, and $n$ is the number of firms.

R&D costs may include fixed and variable components. The level of fixed costs may depend on the number of firms with which the laboratory accepts to contract, and with their identity. We denote by $f_i^S(\theta, \mathbf{x})$ the incremental costs the laboratory must incur if it chooses to serve firm $i$ in addition to other firms in $S \backslash \{i\}$. Formally, let

$$f_i^S(\theta, \mathbf{x}) = r^S(\theta, \mathbf{x}) - r^{S \backslash \{i\}}(\theta, \mathbf{x}), \tag{1}$$

with $f_i^S(\theta, \mathbf{x}) \equiv 0$ if $i \notin S$, by convention. Remark that all fixed and variable cost components are possibly impacted by the stochastic parameter. (In the particular context of the pharmaceutical sector, the fixed costs associated to a given R&D project, *e.g.* for biochemical and cell-based assays, are likely to be very high and *a priori* uncertain.)

The specification of a multi-dimensional R&D output renders possible the contractual requirement by firms of specifically designed services. Each firm $i$ is interested in controlling the selection by the laboratory of an $x_i = \left(x_i^1, \ldots, x_i^m\right)$ in $\mathbf{x} = (x_1, \ldots, x_n)$ that fits its own needs. However, each firm's gross profit function depends not only on what it receives, but also on what the other firms receive. This is because contracting firms may be competitors on the same final product markets. We thus denote by $g_i : X \rightarrow \Re_+$ the gross monetary payoffs received by firm

$i$.

To capture the fact that a laboratory has less bargaining power than firms, we assume that the latter write contracts. This is a consequence of the structural conditions of the markets for R&D services we consider here. Veugelers (1997) remarks that when in-house facilities are available, as is typically the case in the pharmaceutical industry, the capacity to go for it alone increases a firm's bargaining power in negotiating with an external laboratory. Lerner and Merges (1998) evoke the financial constraints faced by specialized laboratories on the intermediate market for biotechnology, where R&D buyers are large pharmaceutical, agribusiness, or chemical firms. Argyres and Liebskind (2002), also in the context of the biotechnology sector, refer to a high rate of entry on the supply side, as opposed to a small set of established companies on the demand side.

We also assume that each firm $i$'s transfer be a function not only of $x_i$ but also of $x_j$, all $j \neq i$.[8] This is motivated by the observation that real-world contracts commonly feature complex clauses which elaborate a fine tuning between the received R&D services of firms in an industry and the exact payments of a given client. For example, in the late 1990s Tularik – a Californian independent laboratory that specializes in the research and development of therapeutic pharmaceutical products based on a proprietary technology – has signed a series of bilateral multi-annual contracts for the delivery of firm-specific and nevertheless technologically related R&D services to a set of American, Japanese, and European firms. The latter include Merck, Sumitomo Pharmaceuticals, and Roche Bioscience, which are potential rivals on final markets for pharmaceutical products. In publicly available contracts (see http://contracts.onecle.com), one reads clauses that explicitly acknowledge the existing contractual links which were signed in the past by the laboratory and third parties. Some other clauses also stipulate that, in the future, Tularik may not transfer any R&D output resulting from the contractual agreement without the prior written consent of the firm that originated it. In some cases, and in exchange of some specified payments, the firm has an option to purchase the rights to some particular

R&D output for a certain period of time from the date of the result. After expiration of this period, if the option is not exercised, the laboratory is under no further obligation to the firm with respect to the submitted R&D output.

## 2.2 The Delegation Principle

We also assume that, although transfers can be made contingent on the R&D outcome, they cannot depend on the state of nature $\theta \in \Theta$, which is unobservable by firms. Formally, a strategy for each firm $i$ is a function $t_i : X \rightarrow \Re_+$, which describes a single incentive contract. It takes the simple form of transfer payment offers that are contingent on the laboratory's deliveries. We denote the vector of strategies by $\mathbf{t} = (t_1, \ldots, t_n)$.

Unlike $\mathbf{x}$ and $\mathbf{t}$, the parameter $\theta$ which describes the state of nature is observable exclusively by the laboratory at the R&D operations stage. It cannot be observed by firms nor verified by a third party, and therefore cannot be contracted upon. This does not mean that firms will not organize for limiting the laboratory's ability to benefit from this informational asymmetry at their expenses. There is evidence that real-world firms do rely on a large variety of sophisticated communication mechanisms to keep control on the actions chosen after the contracting stage by an external laboratory. R&D contracts typically organize the formation of a research committee in which parties exchange information for an on-going monitoring of operations and a possible adjustment of payment schemes. These observed communication practices can be viewed through economic lenses as means to reveal the laboratory's ability to deliver a targeted outcome, *i.e.* its type $\theta$. The laboratory's type may relate not only to technological skills, but also to some exogenous state of nature, together with the set of contractual opportunities other firms may offer.

We see two possible ways of modelling the communication devices one observes on R&D markets. A first one would consist in defining the message sets available to each firm and the laboratory in the most general way, before specifying the mechanisms chosen by firms in their attempts to control the laboratory's decisions. Then it would be possible to capitalize on Epstein and Peters (1999). They establish the existence of a "universal types space" that renders possible the use the revelation principle in a multi-principal context, the same way as in a standard single-principal problem. Although conceptually attractive, this approach is weakened in practice by the difficult identification of the nature of the universal types space in a particular environment.[9]

Another more tractable approach, we follow, consists in imposing relevant restrictions on competition among firms. In a very general common agency setup, which encompasses our model, Peters (2001) demonstrates that all equilibrium allocations that can be supported with any negotiation mechanism available to the principles can also be supported by imposing principals to offer the agent a menu of contracts that associate actions (e.g., transfer payments) by the principal with observable choices by the single agent.[10] More specifically, the latter property still holds good in the present model in which each principal's strategy is restricted to the simplest possible menu, that is exactly one incentive scheme. Indeed, in our model the principals and the agent have symmetric information at the contracting stage, each principal's gains depend only on her own strategy and the agent's choices, and the agent's objective function is monotone in the payments he may receive from principals. In this case, we know from Peters (2003) that all equilibrium allocations that can be supported with a menu of contracts can also be supported as equilibria when each principal's strategy is restricted to a single transfer payment conditioned on the agent's choice.[11] In addition, we know again from Peters (2003) that the equilibria in single payment schemes we consider are robust to the possibility that principals might offer richer menus. This means that the equilibria we characterize in the following pages are also equilibria in a more realistic setup in which more sophisticated communication means are available to principals.

Consequently, in our delegated common agency setup, there is no loss of generality in imposing that principals compete only in single non-linear payment schemes. To anticipate, it will also be made apparent later that the contracts we consider are optimal, so that there is no restriction in focusing on them.

### 2.3 The Delegated R&D Common Agency Game

For each $\mathbf{x}$ it delivers to a subset $S$ of firms exclusively, the laboratory $\mathcal{L}$ obtains a net benefit given by the function

$$v_{\mathcal{L}}^{S}\left(\theta, \mathbf{t}, \mathbf{x}\right) = \sum_{i \in S} t_i(\mathbf{x}) - r^{S}(\theta, \mathbf{x}). \tag{2}$$

For each firm $i$, net profits are given by the function

$$v_i\left(t_i, \mathbf{x}\right) = g_i(\mathbf{x}) - t_i(\mathbf{x}). \tag{3}$$

In substance, we have a procurement market for R&D, in which the laboratory is an agent, and firms are principals. There is competition among firms, which are interrelated in their attempt to command the use of the laboratory's resources through transfer payments, that is incentive schemes. This constitutes an R&D common agency game. We specify the following timing.

- Stage 1: Proposed contracts.

  Firms consider all possible R&D outcomes to simultaneously and non-cooperatively propose contracts to the laboratory. As already mentioned, the state of nature $\theta$ is assumed to be unobservable by the firms. We assume that transfers depend only on $\mathcal{L}$'s delivered R&D outcome. Firm $i$'s contract is thus a commitment to transfer $t_i\left(\mathbf{x}\right)$, for any realization of $\theta$ (and for any possible choice of $\mathbf{x}$ by the laboratory), that is part of the set

$$T_i = \{t_i(\mathbf{x}) \geq 0 \mid \mathbf{x} \in X\}. \tag{4}$$

  This limited liability constraint on transfer payments reflects the observation that specialized laboratories typically lack financial support. The profile of a firm's payment function is left open, as it may for example include a fixed component, or be linear in the dimensions of R&D vectors.

- Stage 2: Accepted contracts.

Given transfer payment proposals $\mathbf{t}$ and the probability measure $\mu$, the laboratory considers its expected benefits to decide whether to accept or not each firm's proposed contract. This leads to define a subset, we denote by $A$, which exclusively includes the firms with which the laboratory accepts to contract. The laboratory signs at most $n$ R&D contracts. If the laboratory chooses not to sign with any firm, it does not receive any payment. Formally, a rationality condition writes

$$E_\theta \left[ \max_{\mathbf{x} \in X} \left( v_{\mathcal{L}}^{A}(\theta, \mathbf{t}, \mathbf{x}) \right) \right] \geq \sup \left\{ \underline{v}_{\mathcal{L}}, E_\theta \left[ \max_{\mathbf{x} \in X} \left( v_{\mathcal{L}}^{S}(\theta, \mathbf{t}, \mathbf{x}) \right) \right] \right\}, \tag{5}$$

for all $S \subseteq N$, where $\underline{v}_{\mathcal{L}}$ represents the value of an outside option to $\mathcal{L}$. Obviously, since transfer payments are supposed to be non-negative, there is no loss in generality in assuming that the set $A$ extends to the whole set $N$ at equilibrium, possibly with some "null contracts", *i.e.* firms offering transfers that are equal to zero for all equilibrium outcome. However, the very fact that the agent may refuse some contracts guarantees that firms cannot have a free lunch with R&D services. It is the manifestation of the laboratory bargaining power.

- Stage 3: R&D operations.

Eventually, given transfer proposals $\mathbf{t}$ and the set of accepted contracts $A$, the laboratory learns the realization of the stochastic variable $\theta$. Then it produces the services $\mathbf{x}$ that maximize its net benefits, that is the difference between the transfer payments stipulated in the set of accepted contracts and the costs of R&D operations. The vectors $\mathbf{x}$ and $\mathbf{t}$, which constitute the subject-matter of the contracts, are both observable and verifiable by a Court. This means that a firm $i$ may not decide not to transfer the proposed payment

$t_i$ in response to the laboratory's chosen $\mathbf{x}$, and also that the laboratory cannot consider not to deliver the contractually agreed upon services in a state of nature that implies a negative benefit for any possible choice in $X$.[12]

*Remark 1: intrinsic vs. delegated common agency.* Bernheim and Whinston (1986b) distinguish two categories of common agency games. In an "intrinsic" common agency framework, the agent may either accept or refuse all contracts. In "delegated" common agency games, the agent may decide to accept or refuse any subset of the proposed contracts, as in the present case. We thus refer hereafter to a *delegated* R&D common agency game.

*Remark 2: contracting and decision stages.* In the first stage, each strategy $t_i$ is based on firm $i$'s expectation over $\mathcal{L}$'s possible R&D outcomes, which depend on the possible realizations of the stochastic parameter and the induced choices by the laboratory. By the same token, in the second stage, the laboratory's decision is based on its expectation over the possible realizations of the stochastic variable. For a clear understanding of the game, it is important to emphasize the distinction between these two former contracting stages, where decisions are based on expected payoffs, and the final R&D operations stage, where decisions are based on the specific realization of the stochastic variable.

*Remark 3: conditional transfer payments.* This specification contrasts with many contributions to the common agency literature which typically present theoretical results in the usual context of a market for consumer goods, in which each principal is described as a supplier of some (possibly differentiated) good and the common agent is a retailer or a consumer (e.g., Gal-Or (1991), Stole (1991), Martimort (1996), Bernheim and Whinston (1998), Calzolari and Scarpa (2004), Martimort and Stole (2004), *inter alia*). It is assumed in these papers that a given principal may not condition contracts on the agent's decision to accept another principal's contract, and *a fortiori* on the exact quantity supplied by another principal to the common agent. This assumption makes sense when contractual agreements between, say several suppliers and a given retailer, are concluded secretly, and thus not observable by third parties. The assumption is also well justified on institutional grounds, since the crux of contractual relationships that govern product market transactions is subject to no-discrimination rules. However, these considerations do not extend naturally to the context of the present paper. Indeed, independent laboratories usually advertise the contractual agreements they conclude with established firms in order to signal their expertise to potential clients. Moreover, antitrust rules toward R&D contractual agreements of all kinds clearly less stringent than the legal safeguards that prevail on product markets. In the US institutional context, Martin (2001, p. 464) evokes a "permissive attitude" in reference to the National Cooperative Research Act (NCRA) of 1984, which can be seen as a means to reduce legal disincentives to participate in contractual R&D agreements on all "properly defined, relevant research and development markets", including on the markets for R&D services we focus on in this paper.[13] Firms thus frequently exploit this favorable informational and institutional environment to tap R&D services from independent laboratories in exchange of payment schemes in connection to complex non-compete clauses or exclusivity conditions which explicitly refer to actual or potential rivals. Therefore our delegated R&D common agency game is of the *public* kind.[14]

### 2.4 Uncertainty in R&D

A interesting feature of the model is that uncertainty is introduced in a very general way. In particular, the state space $\Theta$ may be finite or infinite. Moreover, no monotonicity assumption is introduced that would restrict the nature of the impact of the stochastic parameter $\theta$ on the agent's costs. This parameter may appear in fixed and/or variable cost components.

To produce results, the only restriction we need introducing is that "radical uncertainty" is ruled out. This means that we focus on situations in which principals never face limited liability problems, in the sense that the liability constraint we (realistically) introduce may bind only with probability zero. Formally, this is done by introducing the set

$$I_\Gamma\left(\{\mathbf{x}(\theta)\}_{\theta\in\Theta}\right)=\{(v_1,\dots,v_n)\mid \forall i\in N,\ g_i(\mathbf{x}(\theta))\geq v_i,\text{ almost all }\theta\}. \tag{6}$$

If $\{\mathbf{x}(\theta)\}_{\theta\in\Theta}$ describes the agent's choices as a function of the stochastic parameter $\theta$, the assumption that the vector of principals' net payoff $\mathbf{v}=(v_1,\dots,v_N)$ is in $I_\Gamma\left(\{\mathbf{x}(\theta)\}_{\theta\in\Theta}\right)$ is equivalent to saying that each principal is almost always able to induce the agent to supply services that generate an amount of realized revenues bounded from below by $v_i$, all $i$. In other words, each principal's *gross* realized payoff must always be higher than the *net* expected payoff. Naturally, this does not say anything about the realized net payoff. However, it guaranties that a principal $i$ is always able to pay the agent up to $v_i$. Loosely speaking, the constraint in (6) is a "deep pocket" condition. We introduce it in order to reduce the occurrence of a limited liability problem on the principals' side. Conversely, the condition can be considered as describing an upper bound on each principal's expected net payoff $v_i$, given the set of agent choices $\{\mathbf{x}(\theta)\}_{\theta\in\Theta}$. This condition on $g_i(\mathbf{x}(\theta))$ obviously holds good when uncertainty is absent. If the agent's costs are relatively high (as compared with gross payoffs), which is a quite reasonable conjecture in an R&D context, the condition still holds true even with "quite a big amount" of uncertainty. We thus see it as a natural bound on the support of the stochastic parameter that leads us to concentrate on a delineated set of situations for any possible modelling of the problem under scrutiny.

The uncertain nature of R&D, in the present model, does not relate to situations in which the outcome of scientists inside the laboratory is unknown *a priori*. We focus on situations in which the R&D output is a clearly identified new process or product, that can be user-specific. What is unknown is the cost of obtaining that target. The model does not relate either to real-world instances asymetrically distributed knowledge at the contracting stage. We concentrate on situations in which firms and an external laboratory are endowed with the same (lack of) expertise *ex ante*.

## 3 Characterization of Nash Equilibria

In order to focus on the theoretical bones of the situation under scrutiny, in this section we describe the delegated R&D common agency game in the standard terminology of the theory of incentives. Then we define the Nash equilibrium solution concept of the game, before characterizing equilibria.

At the contracting stages, the parameter $\theta$ is unknown not only to the principals (firms) but also to the agent (laboratory). Principals thus face an adverse selection problem with *ex ante* uncertainty. An incentive compatible contract satisfies the following incentive and participation constraints.

**(IC)** For all $\theta$, and given principals' strategies $\mathbf{t}$, the agent's choices $\mathbf{x}^o(\theta)$ are such that:

$$v_{\mathcal{L}}^A(\theta, \mathbf{t}, \mathbf{x}^o(\theta)) \geq v_{\mathcal{L}}^A(\theta, \mathbf{t}, \mathbf{x}), \tag{7}$$

for all $\mathbf{x} \in X$. This says that for given principals' strategies $t_i$, and for all realizations of the stochastic parameter, the agent maximizes its individual benefits in equilibrium.

**(PC)** For all principals $i \in A$, given strategies $\mathbf{t}_{-i}^o$ as chosen by principals in $N \backslash \{i\}$, principal $i$'s strategy $t_i^o$ is such that

$$E_\theta\left[t_i^o(\mathbf{x}^o(\theta)) + f_i^A(\theta, \mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A\backslash\{i\}}(\theta, \mathbf{t}, \mathbf{x}^o(\theta))\right] \geq \sup\left\{0, E_\theta\left[\max_{\mathbf{x}} v_{\mathcal{L}}^S(\theta, \mathbf{t}, \mathbf{x})\right]\right\}, \tag{8}$$

for all $S \subseteq N \backslash \{i\}$. This says that the agent should earn at least as much expected benefits by serving principal $i$ (in addition to the others) as by serving any other set principals but principal $i$.

Principal $i$ takes into account the incentive constraint (7) together with its specifically related participation constraint in (8) to solve the maximization program

$$\max_{t_i(.)\in T_i} \{E_\theta [v_i(t_i, \mathbf{x}^o(\theta))]\} \equiv \max_{t_i(.)\in T_i} \{E_\theta [g_i(\mathbf{x}^o(\theta)) - t_i(\mathbf{x}^o(\theta))]\}. \tag{9}$$

Of course, the optimal contract $\{t_i^o(\theta), \mathbf{x}^o(\theta)\}_{\theta\in\Theta}$ offered by principal $i$ depends on the strategies (that is, transfer payments) chosen by all other principals in $N\backslash\{i\}$. This leads us to introduce the Nash equilibrium as a solution concept.

*Remark 4: agent participation constraints.* In contrast to the (unique) incentive constraint (7), in which all principals play the same role, there is a particular participation constraint (8) associated to each principal-agent relationship. The latter constraint is equivalent to saying that transfer payments $t_i^o$ are such that it is in the agent's interest to participate in the game with principal $i$ (that is to accept the contract offered by $i$). Observe however that both categories of constraints describe a strategic interaction among principals, since the choices made by the common agent are influenced by the transfers offered by all principals.

*Remark 5: principals' participation constraint.* Although not mentioned explicitly in the characterization of the Nash equilibria given above, the participation of a principal is also constrained by $v_i \geq \underline{v}_i$. In the present context, this minimum payoff $\underline{v}_i$ can be interpreted as a reservation payoff obtained either by sourcing substitutable services from another independent and exclusive agent, or by relying on proprietary assets.

**Definition 1 (Definition of the Nash equilibrium of a common agency game):** *The pair* $(\mathbf{t}^o, \mathbf{x}^o(\theta))$*, where* $\mathbf{t}^o$ *and* $\mathbf{x}^o$ *denote principals' strategies and the agent's choice, respectively, is a Nash equilibrium (NE) of the game if*

*i)* $\mathbf{x}^o(\theta) \in X(\theta, \mathbf{t}^o) \equiv \arg\max_{\mathbf{x}} v_{\mathcal{L}}(\theta, \mathbf{t}^o, \mathbf{x})$*;*

*ii)* $A \equiv \arg\max_S \left\{E_\theta\left[\max_{\mathbf{x}} v_{\mathcal{L}}^S(\theta, \mathbf{t}^o, \mathbf{x})\right]\right\}$*;*

*iii) there is no* $i \in N$*, no* $\widetilde{t}_i \in T$*, no* $\widetilde{A} \equiv \arg\max_S \left\{E_\theta\left[\max_{\mathbf{x}} v_{\mathcal{L}}^S\left(\theta, \widetilde{t}_i, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right]\right\}$ *and no* $\widetilde{\mathbf{x}}(\theta) \in X^o(\theta, \widetilde{t}_i, \mathbf{t}_{-i}^o) \equiv \arg\max_{\mathbf{x}} v_{\mathcal{L}}^{\widetilde{A}}\left(\theta, t_i, \mathbf{t}_{-i}^o, \mathbf{x}\right)$*, such that*

$$E_\theta\left[v_i\left(\widetilde{t}_i, \widetilde{\mathbf{x}}(\theta)\right)\right] > E_\theta\left[v_i\left(t_i^o, \mathbf{x}^o(\theta)\right)\right].$$

Assuming that the liability constraint $\mathbf{t} \geq \mathbf{0}$ is not binding, we are now able characterize the Nash equilibria of the game, as follows.

**Theorem 1 (Characterization of the Nash equilibria):** *A triplet* $(\mathbf{t}^o, A^o, \{\mathbf{x}^o(\theta)\})$ *is a Nash Equilibrium of the delegated common game* $\Gamma$ *if and only if:*

(1) *the action* $\mathbf{x}^o(\theta)$ *is in:*

$$X^o(\theta) = \arg\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}^o, \mathbf{x}) \quad \textit{for almost all } \theta$$

(2) *for all principals* $i$ *in* $A^o$*:*

$$E_\theta\left[v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}^o, \mathbf{x}^o(\theta))\right] = \sup\left\{0, \sup_{S\subseteq N\backslash\{i\}} E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{S}(\theta, \mathbf{t}_{-i}^o, \mathbf{x})\right]\right\};$$

(3) *for all principals* $i$ *in* $A^o$*:*

$$X^o(\theta) \subseteq \arg\max_{\mathbf{x}\in X}\left[g_i(\mathbf{x}) - f_i^{A^o}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{A^o\backslash\{i\}}(\theta, \mathbf{t}_{-i}^o, \mathbf{x})\right] \quad \textit{for almost all } \theta;$$

(with the convention $f_i^{A^o}(\theta, \mathbf{x}) \equiv 0$ if $i \notin A^o$)

$(4a)$ *for all principals* $i$ *in* $A^o$*:*

$$\begin{aligned} & E_\theta\left[\left(g_i(\mathbf{x}^o(\theta)) - f_i^{A^o}(\theta, \mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o\backslash\{i\}}(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta))\right)\right] \\ \geq \ & \sup\left\{0, E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x}) - f_i^{S\cup\{i\}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{S}(\theta, \mathbf{t}_{-i}^o, \mathbf{x})\right)\right]\right\} \quad \textit{all } S \subseteq N\backslash\{i\}; \end{aligned}$$

and

$$\begin{aligned} & E_\theta\left[\left(g_i(\mathbf{x}^o(\theta)) - f_i^{A^o}(\theta, \mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o\backslash\{i\}}(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta))\right)\right] \\ \geq \ & \sup\left\{0, E_\theta\left[g_i(\mathbf{x}_S^o(\theta)) + v_{\mathcal{L}}^{S}(\theta, \mathbf{t}_{-i}^o, \mathbf{x}_S^o(\theta))\right]\right\} \quad \textit{all } S \subseteq N\backslash\{i\}; \end{aligned}$$

with $\mathbf{x}_S^o(\theta) = \arg\max_{\mathbf{x}\in X} v_{\mathcal{L}}^S(\theta, \mathbf{t}_{-i}^o, \mathbf{x})$.

(4*b*) *for all principals $i$ in $N \backslash A^o$:*

$$\begin{aligned} & E_\theta \left[ \left( g_i \left( \mathbf{x}^o \left( \theta \right) \right) + v_{\mathcal{L}}^{A^o} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x}^o \left( \theta \right) \right) \right) \right] \\ \geq \quad & \sup \left\{ 0, E_\theta \left[ \max_{\mathbf{x} \in X} \left( g_i \left( \mathbf{x} \right) - f_i^{S \cup \{i\}} \left( \theta, \mathbf{x} \right) + v_{\mathcal{L}}^S \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right) \right] \right\} \quad \textit{all } S \subseteq N \backslash \{i\} . \end{aligned}$$

This proposition gives four conditions which, when all satisfied, characterize a Nash equilibrium in the principals' strategies and the agent's output. The first condition says that the agent maximizes its individual benefits for given non-negative transfers as proposed by the principals. The second condition says that all participation constraints are exactly binding. More precisely, either the agent's expected rent is zero, or it is strictly positive. However, in the latter case, the contractual relationship with principal $i$ does not increase the agent's expected rent, as obtained by contracting exclusively with principals in $N \backslash \{i\}$. The third condition states that, for almost all $\theta$, the equilibrium choices maximize the joint-payoff of the agent $\mathcal{L}$ and principal $i$, any $i$. In other words, the *bilateral* contractual relationships between the principals and the agent are efficient. The fourth condition states that, given the incremental costs $f_i^{A^o}(\theta, \mathbf{x})$, the contracting set $A^o$ is efficient, in that it maximizes the joint payoffs of the agent and all the contracting principals.

It is well-known that, in the standard (one) principal-agent setup, when the agent is risk-neutral, uncertainty at the contracting stage does not forbid the first-best to be achieved (See Harris and Raviv (1979)). In other words, when there is only one principal, optimal contracts are Pareto-efficient and maximize the joint payoff of the agent and the principal. This holds true also if one introduces risk aversion on the principal's side. This result however follows from the assumption that the agent may not breach the contract when the states of nature are such that it would have been in the agent's interest not to contract with the principal. In practice, there is often a limit on the maximum loss the agent can be forced to bear as a consequence of contracting with the principal. In this paper, for obvious scope of realism, we assume such limits to exist and introduce the limited liability constraint $\mathbf{t} \geq \mathbf{0}$. However, in order to focus on the strategic implications of the multi-principal setup (as opposed to the single principal case), all results are obtained by assuming that the latter constraint is not binding. We evidence afterward the conditions under which this is indeed the case.

*Remark 6: liability constraints.* A limited liability constraint may be of two types. It can either bear on the *ex post* level of the agent's utility (benefits), a case one would adopt to capture, say, bankruptcy laws. Alternatively, the limited liability may apply to the transfers from (to) the principal, a limit that reflects an upper bound on the rewards (fines) that may be imposed to the agent. We know from Sappington (1983) that, when a limited liability constraint is imposed on the agent's *ex post* utility level, and if this constraint is binding, then *ex post* Pareto efficiency no more holds. The optimal contract as offered by the principal thus does not require the agent to choose an efficient action. Similarly, a limited liability constraint that bears on transfers may forbid efficiency. Although the distortions introduced by both types of liability constraints highly differ (see Laffont and Martimort, chap. 3, pp. 118-121), there is an interesting link between them. Assuming $\mathbf{t} \geq \mathbf{0}$ implies indeed that $v_{\mathcal{L}}^{A}(\theta) \geq -\min_{\mathbf{x}\in X} r^{A}(\theta, \mathbf{x})$. The limited liability constraint imposed on transfers is thus equivalent to introducing a boundary condition on the agent's benefits, which is (possibly) state-dependent.

# 4 Examples

In this section, we add flesh to the analysis by constructing a pair of simple and contrasted examples. The objective is to illustrate the large possibilities of outcomes that can be obtained from the delegated R&D common agency model in equilibrium.

For simplicity, suppose that the stochastic parameter can take values in a discrete set $\{\underline{\theta}, \overline{\theta}\}$ with equiprobability, and that the possible R&D outcomes are in the finite set $X = \{0, y, z\}^2$. Denote by $Y = \{0, y\}^2$ (resp. by $Z = \{0, z\}^2$) the subset that does not include $z$ (resp. $y$). There are only two firms, that is $n = 2$. We now turn to two specifications for each firm's gross profit function and for the laboratory's costs.

■ **Example 1:** *On the laboratory's side, let*

$$r^S(\underline{\theta},\mathbf{x}) = \begin{cases} f & \text{if } \mathbf{x}=(0,0) \\ \tilde{r}+f_i^S(\underline{\theta}) & \text{if } \mathbf{x}\in X\cup Y\backslash(0,0) \\ 2\underline{r} & \text{otherwise;} \end{cases} \quad \textit{and } r^S(\overline{\theta},\mathbf{x}) = \begin{cases} f & \text{if } \mathbf{x}=(0,0) \\ \tilde{r}+f^S(\overline{\theta}) & \text{if } \mathbf{x}\in X\cup Y\backslash(0,0) \\ 2\overline{r} & \text{otherwise;} \end{cases}$$

*where* $0\leq f\leq\tilde{r}\leq\underline{r}\leq\overline{r}\leq\tilde{g}-(\underline{r}-\tilde{r})$ *and*

$$f_i^N(\underline{\theta})=\underline{r}-\tilde{r}; \qquad f_i^N(\overline{\theta})=\overline{r}-\tilde{r}; \qquad f_i^S(\theta)=0 \text{ otherwise.}$$

*R&D costs increase with the number of firms served by the laboratory. They also increase from a state of nature to the other. For simplicity, we assume however that providing services to only one firm induces a cost* $\tilde{r}$ *with no uncertainty. Incremental costs* $f_i^S(\theta)$ *make it clear that providing the same services to the two firms is more costly than serving a single firm. This may be the result, say, of a capacity investment. Supplying different services* $y$ *and* $z$ *simultaneously to the two firms costs more than* $2\tilde{r}$*, the cost of providing services to a single firm. This is a case of diseconomies of scope in R&D. On the firms' side, let*

$$g_1(\mathbf{x}) = \begin{cases} \tilde{g} & \text{if } x_1=y; \\ 0 & \text{otherwise;} \end{cases} \quad \text{and } g_2(\mathbf{x}) = \begin{cases} \tilde{g} & \text{if } x_2=z; \\ 0 & \text{otherwise;} \end{cases}$$

*where* $\tilde{g}>0$*. In words, firms are interested in receiving a specific R&D services, namely* $y$ *(resp.* $z$*) for firm* 1 *(resp.* 2*), otherwise they do not benefit from the laboratory's output. The 3-tuple* $(\mathbf{t}^o, A^o, \mathbf{x}^o)$ *as defined by*

$$t_1^o(y,\cdot)=t_2^o(\cdot,z)=\overline{r}+\underline{r}-\tilde{r} \quad \text{and} \quad t_i^o(\mathbf{x})=0 \text{ otherwise;}$$

$$A^o=\{1,2\};$$

$$\mathbf{x}^o(\theta)=(y,z), \quad \text{all } \theta;$$

*satisfies the four conditions displayed in Theorem 3. In this Nash equilibrium, each firm receives its most valuable option and earns a net profit equal to* $\tilde{g} - (\overline{r} + \underline{r} - \tilde{r})$. *By contrast, the laboratory's net equilibrium benefits depend on the state of nature. It earns* $2(\overline{r} - \tilde{r})$ *in the first state, and only* $2(\underline{r} - \tilde{r})$ *in the second state. Moreover, there is no other equilibrium that would guarantee to the principals as much as what they earn in this Nash equilibrium.* ■

■ **Example 2:** *Assume that the laboratory's cost function is*

$$r(\underline{\theta}, \mathbf{x}) = \begin{cases} \tilde{r} & \text{if } \mathbf{x} \in X; \\ +\infty & \text{otherwise;} \end{cases} \quad \textit{and } r(\overline{\theta}, \mathbf{x}) = \begin{cases} \tilde{r} & \text{if } \mathbf{x} \in Y; \\ +\infty & \text{otherwise;} \end{cases}$$

*where* $\tilde{r} \geq 0$. *That is, in the first (resp., second) state of nature, the cost of supplying* $y$ *(resp.,* $z$*) to any of the two firms or both is positive, whereas the cost of supplying* $z$ *(resp.* $y$*) is prohibitive. Firms' gross profit functions are*

$$g_1(\mathbf{x}) = \begin{cases} \overline{g} & \text{if } x_1 = y; \\ \underline{g} & \text{if } x_1 = z; \\ 0 & \text{otherwise;} \end{cases} \quad \textit{and } g_2(\mathbf{x}) = \begin{cases} \overline{g} & \text{if } x_2 = z; \\ \underline{g} & \text{if } x_2 = y; \\ 0 & \text{otherwise;} \end{cases}$$

*where* $\underline{g}$ *and* $\overline{g}$ *are positive parameters. Without loss of generality we assume that* $\underline{g} \leq \overline{g}$. *In addition, let*

$$\overline{g} + \underline{g} > \tilde{r} \quad \textit{and} \quad \overline{g} \leq \tilde{r},$$

*which means that the production of R&D services is profitable at the industry level, although no firm may afford it on its own. Moreover, we impose*

$$\overline{g} - \underline{g} < \tilde{r}/2,$$

*which can be viewed as an upper limit on the impact of uncertainty on wealth creation. Remark that both firms are interested in receiving either* $a$ *or* $b$*, but valuate the two services asymmetrically.*

*The 3-tuple* $(\mathbf{t}^o, A^o, \mathbf{x}^o)$ *as defined in the first state of nature (i.e.,* $\theta = \underline{\theta}$*) by*

$$\mathbf{t}^o\left(\mathbf{x}^o\left(\underline{\theta}\right)\right) = \left(\overline{g} - \alpha\left[\overline{g} + \underline{g} - \tilde{r}\right], \underline{g} - (1-\alpha)\left[\overline{g} + \underline{g} - \tilde{r}\right]\right),$$
$$A^o = \{1, 2\},$$
$$\mathbf{x}^o\left(\underline{\theta}\right) = (y, y),$$

*and in the second state of nature (i.e.,* $\theta = \overline{\theta}$*) by*

$$\mathbf{t}^o\left(\mathbf{x}^o\left(\overline{\theta}\right)\right) = \left(\underline{g} - \alpha\left[\overline{g} + \underline{g} - \tilde{r}\right], \overline{g} - (1-\alpha)\left[\overline{g} + \underline{g} - \tilde{r}\right]\right),$$
$$A^o = \{1, 2\},$$
$$\mathbf{x}^o\left(\overline{\theta}\right) = (z, z),$$

*with* $\alpha$ *in* $[0, 1]$*, satisfies the four conditions displayed in Theorem 3. In equilibrium the two firms receive the same R&D services, that is* $y$ *in the first state, and* $z$ *in the other state. Both firms earn a constant net profit in the two cases, namely* $\alpha(\overline{g} + \underline{g} - \tilde{r})$ *for firm* 1*, and* $(1-\alpha)(\overline{g} + \underline{g} - \tilde{r})$ *for firm* 2*. Here the number of equilibria is infinite, as it is driven by the value taken by the continuous parameter* $\alpha$*. In all equilibria the laboratory earns a net benefit equal to zero, for all states of nature.* ■

# 5 Truthful Equilibria

By choosing transfer payments, principals make strategic considerations in order to influence the agent's choice toward the production of an output that maximizes their own individual payoff. In this section, we compare the equilibrium strategies of this non-cooperative game with the strategies that would naturally emerge in a cooperative context, *i.e.* if the objective were to maximize the joint profits of the agent and all principals. To do that, we introduce the truthfulness concept, show that best responses in transfer payment offers can be truthful, and characterize equilibria in which all principals adopt truthful strategies.

### 5.1 Best Replies Do Not Exclude Truthfulness

**Definition 1** *A strategy $t_i(.)$ is truthful if there exists some $v_i^o \geq 0$ such that:*

$$t_i(\mathbf{x}) = \sup\{0, g_i(\mathbf{x}) - v_i^o\}. \tag{10}$$

Clearly, by transferring to the agent all gains in excess of a value $v_i^o$, each principal makes the agent a residual claimant of the individual benefits that accrue from received services. In a cooperative context, truthful strategies are natural instruments in that they convey the information the agent needs to behave as an efficiency maximizer. Of more interest is a non-cooperative context, in which the agent is not aware of the impact of his own decisions on each principal's gross payoff. This is the situation considered by Bernheim and Winston (1986). In a delegated common agency game where the agent "is poorly informed" (p. 2), they show that each principal's best-reply correspondence contains a strategy which reflects her marginal preferences, *i.e.* is truthful. However, this does not come that much as a surprise, because principals are assumed to have complete information, and thereby may share knowledge with the common agent without loosing control over his choices.

By contrast, when the knowledge of some parameter value is not shared by all players, as in the present model, it is well known that informed parties may find it profitable not to disclose private elements of information. More specifically, at first glance there is no reason to believe that principals will offer transfer payments which exactly reflect their marginal preferences to an agent who, beyond the contracting stage, is the only one to observe the realization of his type and has full control over the choice of $\mathbf{x}$.[15] Indeed, once principals have transferred incentive schemes, in the action stage they may only leave the agent freely make decisions, possibly at their expenses.

This makes salient the following result.

**Proposition 1** *For any $i$, and any vector of strategies $\mathbf{t}_{-i}$ chosen by other principals in $N\backslash\{i\}$, principal $i$'s best-reply correspondence contains a truthful strategy if and only if*

$$g_i\left(\mathbf{x}^o\left(\theta\right)\right) \geq v_i^o \quad \textit{for almost all } \theta, \tag{11}$$

*where $\mathbf{x}^o\left(\theta\right)$ is the induced agent's output, and*

$$v_i^o = E_\theta\left[g_i\left(\mathbf{x}^o\left(\theta\right)\right) + f_i^A\left(\theta, \mathbf{x}^o\left(\theta\right)\right) + v_{\mathcal{L}}^{A\backslash\{i\}}\left(\theta, \mathbf{t}, \mathbf{x}^o\left(\theta\right)\right)\right] - \sup\left\{0, \max_{S\subseteq N\backslash\{i\}} E_\theta\left[\max_{\mathbf{x}} v_{\mathcal{L}}^S(\theta, \mathbf{t}, \mathbf{x})\right]\right\}.$$

This result establishes two necessary and sufficient conditions for a principal's truthful strategy to be a best response to the given strategies of all other players. The first condition in (11) describes a floor case, as it imposes that the principal's lowest possible gross payoff resulting from the agent's choice $\mathbf{x}^o\left(\theta\right)$ is not lower than the net payoff $v_i^o$ she earns in equilibrium (the inequality may be violated only for subsets of $\Theta$ with zero measure). The second condition in (11) stipulates that the net equilibrium payoff $v_i^o$ is the highest value the principal may retain without violating the agent's participation constraint.[16] When the two conditions of proposition 5.1 hold, no limited liability constraint binds and no principal $i$ lets the agent extract any rent from the specific bilateral contract if offers on top of what is earned from all other principals in $N$. Consequently, if the agent obtains a positive rent from its contractual relationship with principals in $N\backslash\{i\}$, it is left with exactly the same expected rent $E_\theta\left[\max_{\mathbf{x}}\left(v_{\mathcal{L}}^{N\backslash\{i\}}(\theta, \mathbf{t}, \mathbf{x})\right)\right] > 0$ by contracting also with $i$. Otherwise the agent obtains no rent at all. Observe that situations in which the agent makes some positive rent by contracting with the other principals should not be excluded *a priori*. The delegated common agency setup is a non-cooperative game in which principals may possibly concede a positive rent to the agent by vying for the control of the latter's decision.

*Remark 7: implementation.* Condition (11) is an implementability condition on the agent's output that must be satisfied *in all states of nature*. This deserves some comments. Firstly, we conceive it as a natural condition. In the extreme situation in which uncertainty is absent, the condition boils down to specifying that a principal's gross payoff must be higher than the net payoff. This (usually implicit) assumption guarantees that the contractual relationship is beneficial to all parties by eliminating the trivial no-contract equilibrium. Secondly, the implementability condition makes more precise the role of uncertainty in the present model. Roughly, as long as the uncertain value taken by the stochastic parameter is limited to cases in which each principal $i$'s gross benefits are higher than equilibrium net payoffs, then all principals can rely on truthful strategies to maximize their gains. Intuitively, the higher the fixed costs incurred by the agent, the more the principals must pay to satisfy the agent's participation constraint, and the lower the value of each principal's net payoff. This means that, for given downstream conditions (as captured by principals' gross profit functions), the implementability condition is easily satisfied when the agent's technology is characterized by high fixed costs, which is very likely to be the case in most real-world R&D situations. Thirdly, the implementability condition amounts to imposing boundaries to the support of the stochastic parameter (without any limitation on its distribution). The size of this support increases with the magnitude of the fixed component of the costs the agent must incur to serve the principal(s) with which it has contracted. This is because these costs are sunk only *after* bilateral agreements have been made, that is when the agent has commited to produce contractually agreed upon actions in exchange of payments. By accepting contract(s), the agent anticipates it must deliver a sufficiently high level of output to cover fixed costs and satisfy the rationality constraint (5). While it makes sense to assume that these fixed costs can be (close to) zero in the usual context of a market for consumer products or standardized services, it is reasonable to suppose they are relatively high in the case of R&D services.

### 5.2 Characterization of Truthful Nash Equilibria

Proposition 5.1 offers (implicit) necessary and sufficient conditions under which there is no loss of generality in assuming that a principal chooses a truthful strategy, for any given strategies as chosen by all other principals. Although of interest for the understanding of the bilateral strategic interactions at play in the common agency game, in this claim the characterization of a given player's equilibrium action remains conditional on the choices made by other principals. A more complete characterization of truthful equilibria would refer to all strategies in equilibrium.

Toward this aim, we first introduce the notation

$$\Gamma^S \equiv \left[\Theta, X, F, r^S\left(.\right) - \sum_{i \in N \backslash S} t_i\left(.\right), \left\{g_i\left(.\right)\right\}_{i \in S}\right],$$

all $S \subseteq N$. Remark that our common agency game is fully specified by $\Gamma^N$. Then we define the joint payoff function

$$\Pi_S(\theta, \mathbf{x}) = \sum_{i \in S} g_i(\mathbf{x}) - r^S(\theta, \mathbf{x}), \tag{12}$$

for any subset $S$ of $N$ including the empty set, and any realization of the stochastic variable $\theta$. Given $S$ and $\theta$, the maximum of this function is denoted by $\overline{\Pi}_S(\theta) = \max_{\mathbf{x} \in X} \Pi_S(\theta, \mathbf{x})$, and a maximizer is denoted by $\mathbf{x}^*_S\left(\theta\right)$, an element of $X^*_S\left(\theta\right)$. If the agent does not contract with any principal, its expected benefits are $E_\theta\left[\overline{\Pi}_\emptyset(\theta)\right]$, we normalize to zero, whereas principals receive only a reservation payoff $\underline{v}_i$. If the agent contracts with at least one principal, we need identifying the possible distributions of net payoffs received by all principals in equilibrium.

Now, for any given action profile $\left\{\mathbf{x}\left(\theta\right)\right\}_{\theta \in \Theta}$, we define the following set

$$V_\Gamma\left(\left\{\mathbf{x}\left(\theta\right)\right\}_{\theta \in \Theta}\right) = \left\{(v_1, \ldots, v_n) \mid \forall S \subseteq N, \ \sum_{i \in S} v_i \leq E_\theta\left[\Pi_N\left(\theta, \mathbf{x}\left(\theta\right)\right) - \Pi_{N \backslash S}\left(\theta, \mathbf{x}\left(\theta\right)\right)\right]\right\},$$

which describes all payoff distributions $\mathbf{v} = (v_1, \ldots, v_n)$ for which no subset $S$ obtains more than its contribution to the expected joint payoffs of all principals in $N$. If a vector of equilibrium payoffs $\mathbf{v}^o$ is in the Pareto frontier of the distribution set $V_\Gamma\left(\left\{\mathbf{x}\left(\theta\right)\right\}_{\theta \in \Theta}\right)$, we denote by $\mathcal{V}_\Gamma$, then there exists no $\bar{\mathbf{v}}^o$ in $V_\Gamma\left(\left\{\mathbf{x}\left(\theta\right)\right\}_{\theta \in \Theta}\right)$ such that $\bar{\mathbf{v}}^o \geq \mathbf{v}^o$. In what follows the Pareto frontier of $V^*_\Gamma \equiv V_\Gamma(\left\{\mathbf{x}^*\left(\theta\right)\right\}_{\theta \in \Theta})$ is denoted by $\mathcal{V}^*_\Gamma$.

If $v^o_i \geq 0$ is a component of $\mathbf{v}^o$ in the implementation set $I_\Gamma\left(\left\{\mathbf{x}^o\left(\theta\right)\right\}_{\theta \in \Theta}\right)$, then recall from (6) that the limited liability constraint is not binding when principal $i$ adopts the related truthful strategy $t_i^o$ that induces the profile of choices $\{\mathbf{x}^o(\theta)\}_{\theta\in\Theta}$.[17] When the agent's output is in $X_N^*(\theta)$, we write $I_\Gamma^* \equiv I_\Gamma\left(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta}\right)$.

Assuming that all principals adopt truthful strategies, we are now able to characterize the corresponding (subset of) Nash equilibria:

**Theorem 2 (Characterization of Truthful Nash Equilibria):** *In all games $\Gamma^N$:*

1. *if $(\mathbf{t}^o, \mathbf{x}^o(\theta))$ is a truthful Nash equilibrium, then $\mathbf{x}^o(\theta)$ is in $X_N^*(\theta)$ for almost all $\theta$, and $(v_1^o, \ldots, v_n^o)$ is in $I_\Gamma^* \cap \mathcal{V}_\Gamma^*$;*

2. *if a payoff vector $(v_1, \ldots, v_n)$ is in $I_\Gamma^* \cap \mathcal{V}_\Gamma^*$, then it can be supported by a truthful Nash equilibrium.*

There are two claims in Theorem 2. In words, the first one says that truthful strategies are optimal (*i.e.* they lead the agent to produce an output that maximizes the joint payoffs of the agent and all principals), and also that truthful equilibria are Pareto optimal (*i.e.* there is no alternative equilibrium payoff distribution that can make all principals better-off). The second claim establishes that if a payoff distribution is both in the implementation set $I_\Gamma^*$ and the Pareto frontier $\mathcal{V}_\Gamma^*$ of the distribution set $V_\Gamma^*$, then it can be decentralized by the means of a truthful Nash equilibrium.

In addition to an extension of a result in the complete information case by Bernheim and Whinston (1986), Theorem 2 may be considered *a minima* as a generalization to the multi-principal case of a well-known result in the standard single principal-agent setup: efficiency is preserved when uncertainty is introduced at the contracting stage in the absence of limited liability concerns (see Laffont and Martimort, 2002, pp. 57-58). However, more is offered here since we establish that this result holds goods when principals' strategies are restricted to be truthful.

**Strategy of Proof:** *For the sake of clarity, the proof developed in the appendix is decomposed into four distinct results. We first establish in Proposition A2 that the output* $\mathbf{x}^o(\theta)$ *produced by the agent is such that no coalition of principals can obtain more at equilibrium than its marginal contribution, that is* $(v_1^o, .., v_n^o) \in V_\Gamma^* \equiv V_\Gamma(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta})$. *Then in Proposition A3, we claim that efficient outputs and equilibrium outputs coincide, that is* $\mathbf{x}^o(\theta) \in X^*(\theta)$ *and* $\mathbf{x}^*(\theta) \in X^o(\theta)$. *We demonstrate further that the principals' equilibrium payoffs are in the Pareto frontier of* $V_\Gamma^*$ *(Lemma A2 and Proposition A4) and, reciprocally, that any element of the Pareto frontier of* $V_\Gamma^*$ *such that the liability constraint is almost never binding (i.e.,* $(v_1, \ldots, v_n) \in I_\Gamma^* \equiv I_\Gamma\left(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta}\right)$*) can be supported by a truthful Nash equilibrium (Lemma A3 and Proposition A5).*

Although Proposition 5.1 makes it clear that, under well-defined conditions, a principal makes no strategic loss in adopting a truthful strategy, there is no reason to assume that it will always do so. A further argument in favour of the truthfulness refinement consists in demonstrating that truthful Nash equilibria are coalition-proof, *i.e.* stable to credible threats of deviations by subsets of principals. By "credible threat", it is meant that a deviating subset of principals is itself immune from the threat of deviations of a subset of players.[18] More formally:

**Definition 2: Coalition-proof Nash equilibrium**

1. *In all games* $\Gamma^N$ *with a single principal* $(n = 1)$, *the pair* $(\mathbf{t}^o, \mathbf{x}^o)$ *is a strictly coalition-proof Nash equilibrium if it is a Nash equilibrium.*

2. *In all games* $\Gamma^N$ *with several principals* $(n > 1)$, *the pair* $(\mathbf{t}^o, \mathbf{x}^o)$ *is a strictly coalition-proof Nash equilibrium if it is strictly self-enforcing and it is not Pareto dominated by another strictly self-enforcing pair* $(\bar{\mathbf{t}}^o, \bar{\mathbf{x}}^o)$.

3. *In all games* $\Gamma^N$ *with several principals* $(n > 1)$, *the pair* $(\mathbf{t}^o, \mathbf{x}^o)$ *is a strictly self-enforcing profile of strategies if it is a strictly coalition-proof Nash equilibrium of the restricted game*

*$\Gamma^S$, all $S \subset N$.*

Remark that coalition-proof Nash equilibria are defined recursively. A Nash equilibrium is coalition-proof if there is no subset $S$ of players that can deviate by choosing alternative strategies resulting in a higher payoff for each of them (for unchanged strategies as chosen by outsiders in $N \backslash S$), where these alternative strategies also constitute a coalition-proof equilibrium, *i.e.* are immune to deviations by a subset $T \subseteq S$, and so on.

**Theorem 3 (Equivalence between truthful Nash equilibria and coalition-proof Nash equilibria):** *Consider a common agency game $\Gamma$. Then all truthful Nash equilibria are (strictly) coalition-proof. All (strictly) coalition-proof Nash equilibria with payoffs in $I_\Gamma^*$ can be supported by a truthful Nash equilibrium.*

Interestingly, this says that, although the agent is more informed than principals when time comes to produce an output, in all circumstances in which the limited liability constraints are not binding the set of truthful Nash equilibria coincides with the set of coalition-proof Nash equilibria. This robustness result offers a justification for the use of truthful strategies in a delegated common agency set up in which the agent's type is uncertain at the contracting stage.[19]

**Strategy of Proof:** *The proof of Theorem 3 is adapted from Konishi et al. (1999). We first need results for any restricted common agency game $\Gamma^S$. That is, we consider a subset $S \subseteq N$ and take as given the strategies chosen by principals in $N \backslash S$. We demonstrate that a maximizer of the agent's benefit function is also a joint maximizing action for the agent and the coalition $S$ of principals (Proposition ???). Then we show that any truthful Nash equilibrium gives a vector of payoffs for principals in $S$ in the Pareto frontier $V_\Gamma^S(\theta, \mathbf{x}^*(\theta))$ of the set $V_\Gamma^S(\theta, \mathbf{x}^*(\theta))$ of all coalition-proof payoff distributions when the action is chosen by the agent. This implies that no subset of $S$ has any incentive to redistribute the joint-payoff differently*

*(Proposition 7). The proof of Theorem 3 then proceeds by induction on the number of principals $n$.*

## 6 Distribution of Equilibrium Payoffs

In what follows, we capitalize on the results displayed in the previous session to focus on truthful Nash equilibria (TNE). An important property of these equilibria is that equilibrium actions maximize the expected joint payoffs of the agent and all principals, that is $E_\theta \left[\Pi_N(\theta)\right]$. It follows that, as far as the aggregate payoff is concerned, the equilibrium outcome of this non-cooperative game is identical to the outcome of a cooperative game in which all individuals participate in the grand coalition in equilibrium. This motivates the use of cooperative TU game theory to characterize the distribution of payoffs. Toward this aim, we introduce a characteristic function $\overline{\Pi} : 2^N \to \Re_+$ which describes the expected maximum joint-payoff of the set of principals in $S$ together with the agent:

$$\overline{\Pi}(S) = E_\theta \left[ \max_{\mathbf{x} \in X} \left\{ \Pi_S(\theta, \mathbf{x}) \right\} \right] = E_\theta \left[ \max_{\mathbf{x} \in X} \left\{ \sum_{i \in S} g_i(\mathbf{x}) - r^S(\theta, \mathbf{x}) \right\} \right].$$

We denote by $v_{\mathcal{L}}^o$ the agent's expected net equilibrium profits. The next two propositions generalize two results by Laussel and Le Breton (2001) to the present setup, *i.e.* to situations in which the agent type is uncertain at the contracting stage. They demonstrate that the distribution of equilibrium payoffs is rooted in simple structural properties of the characteristic function.

**Proposition 3** *In all games $\Gamma^N$, if $\overline{\Pi}(.)$ is strongly subadditive, that is $\overline{\Pi}(N) - \overline{\Pi}(S) \leq \overline{\Pi}(T) - \overline{\Pi}(S \cap T)$ for all $S, T \in 2^N$ such that $S \cup T = N$, then there is a unique* TNE *in which the agent obtains a rent $v_{\mathcal{L}}^o > 0$, and principals' equilibrium payoffs are*

$$v_i^o = \overline{\Pi}(N) - \overline{\Pi}(N \backslash \{i\}), \text{ all } i \in N.$$

**Proposition 4** *In all games $\Gamma^N$, if $\overline{\Pi}(.)$ is convex, that is $\overline{\Pi}(T \cup \{i\}) - \overline{\Pi}(T) \geq \overline{\Pi}(S \cup \{i\}) - \overline{\Pi}(S)$ for all $S, T \in 2^N$, $i \in N$ with $S \subset T$, and $i \notin T$, then in all* TNE *the agent exactly breaks even,* i.e. $v^o_{\mathcal{L}} = 0$, *and all vectors of principals' equilibrium payoffs are such that*

$$\sum_{i \in N} v^o_i = \overline{\Pi}(N).$$

An intuitive interpretation of the two propositions consists in relating the structural properties of the characteristic function to the net impact of various kinds of externalities on the nature of competition between principals, for the control of the agent's output. These externalities can be of the *direct* type if each principal's gross payoff function depends not only on the output it receives, but also on the output received by another principal, a case of spill-overs. Externalities are of the *indirect* type when the agent's costs are not additively separable across principals. This means that the agent's cost of satisfying a principal's demand depends on the output supplied to meet another principal's needs.

To illustrate, let the set $X_i = \{0, 1\}$ represent the R&D outputs the laboratory may serve to each firm $i = 1, 2$. Consider the situation in which $\theta$ can take only values in $\{\underline{\theta}, \overline{\theta}\}$ with equiprobability. R&D costs $r(\theta, \mathbf{x})$ are

| | $x_2 = 0$ | $x_2 = 1$ |
|---|---|---|
| $x_1 = 0$ | $r_0$ | $r_1$ |
| $x_1 = 1$ | $r_1$ | $r_2$ |

with $0 \leq r_0 \leq r_1 \leq r_2$. Firms' gross profits $g_i(\mathbf{x})$ write also in matrix form as

| | $x_2 = 0$ | $x_2 = 1$ |
|---|---|---|
| $x_1 = 0$ | $(0, 0)$ | $(l, h)$ |
| $x_1 = 1$ | $(h, l)$ | $(b, b)$ |

where the first (second) payoff in each cell refers to firm 1 (firm 2). We assume that $b \geq 0$, and $l \leq h$. This says that 1) the two firms obtain more profits when they both receive services from the laboratory than when no R&D is produced, and 2) each firm is better off when it is the only beneficiary of R&D services. Remark that, in the latter matrix, the choice of $\mathbf{x}$ is made by the laboratory (which then decides to diffuse or not R&D services to firms 1 and/or 2), while the payoffs relate to firms.[20] Then consider the following two cases.

■ **Example 3:** *The R&D technology is characterized by a phenomenon of congestion, so that $E_\theta(r_2) > 2E_\theta(r_1)$. In addition, a firm's gross payoff is lowered when the rival benefits from R&D services, that is $l \leq 0$ and $b \leq h$. These inequalities are satisfied for the following scalar values, where the first two payoffs in each cell refer to the firms while the third value refers to the laboratory's gross payoff (i.e. $-r(\theta, \mathbf{x})$):*

| $\underline{\theta}$ | $x_2 = 0$ | $x_2 = 1$ | $\overline{\theta}$ | $x_2 = 0$ | $x_2 = 1$ |
|---|---|---|---|---|---|
| $x_1 = 0$ | $(0, 0, 0)$ | $(-1, 12, -1)$ | $x_1 = 0$ | $(0, 0, 0)$ | $(-1, 12, -3)$ |
| $x_1 = 1$ | $(12, -1, -1)$ | $(9, 9, -4)$ | $x_1 = 1$ | $(12, -1, -3)$ | $(9, 9, -8)$ |

*In the two states of nature, the joint profits of the laboratory and the two firms (the sum of gross payoffs) are maximized when the two firms receive R&D services, that is if $\mathbf{x} = (1, 1)$. In that case, and for the first state of nature, total profits are 14, as opposed to 10 if exactly one firm receives R&D, and 0 otherwise. In the second state of nature, the total profits are 10, 8, and 0, respectively. In both states, firm 1 will make no positive transfer on $(0, 1)$, nor firm 2 on $(1, 0)$. To avoid being refused access to the laboratory's expertise, a case of negative gross profits, firms compete frontally in transfer payments. To drive the direction of R&D services to their individual advantage, each firm must bid over the rival's proposal until both meet their budget constraint (i.e. the gross payoffs they receive), or until additional increments in payments only inflate the laboratory's revenues, without inducing any change in the direction of its R&D*

*efforts. Firms' transfer functions* $t_i^o(\mathbf{x})$ *are represented by*

| | $x_2 = 0$ | $x_2 = 1$ |
|---|---|---|
| $x_1 = 0$ | $(0, 0)$ | $(0, 10)$ |
| $x_1 = 1$ | $(10, 0)$ | $(7, 7)$ |

*where the first (second) value in each cell refers to firm 1 (firm 2)'s transfers. Together with* $A^o = \{1, 2\}$, $\mathbf{x}^o(\underline{\theta}) = \mathbf{x}^o(\overline{\theta}) = (1, 1)$, *transfers* $t_i^o$ *define a 3-tuple* $(\mathbf{t}^o, A^o, \mathbf{x}^o)$ *that satisfies the four conditions displayed in Theorem 3. In equilibrium both firms receive R&D services and earn a constant net profit in the two states of nature, namely* $v_1^o = v_2^o = 2$*. The laboratory receives a state-dependent net payoff. Its expected benefits are* $v_{\mathcal{L}}^o = 8$*. This strictly positive equilibrium gain results from the anti-complementarities that characterizes the functional form of this example. It is straightforward to show that, in this example, each firm has an incentive to cooperate with the other firm, or to acquire the laboratory. Indeed, an horizontal cooperation would turn the structure of the model into the simpler situation of a single entity (a joint-venture of the two firms) – which writes contracts – and a single laboratory. A transfer schedule* $t_{1+2}(\mathbf{x}) = 6$ *if* $\mathbf{x} = (1, 1)$ *and* $t_{1+2}(\mathbf{x}) = 0$ *otherwise would exactly satisfy the laboratory's participation constraint, and lead each party to the cooperative agreement to gain* $6$ *as a net payoff. Moreover, by acquiring the laboratory, a firm could benefit from exclusive R&D resources to guarantee itself an expected net payoff of* $10$*. In addition, by selling R&D outputs to the other firm, the integrated entity may improve on its benefits up to* $13$ *(the expected joint-profits minus the – negative – reservation profits of the non-integrated firm).* ■

■ **Example 4:** *There are economies of scope in the production of R&D services, so that* $E_\theta(r_2) < 2E_\theta(r_1)$*. Each firm's gross payoff increases if the other firm benefits from R&D services also, so that* $l \geq 0$ *and* $b - l \geq h$*. Again, these inequalities are satisfied for the following* *scalar values:*

| $\underline{\theta}$ | $x_2 = 0$ | $x_2 = 1$ | $\overline{\theta}$ | $x_2 = 0$ | $x_2 = 1$ |
|---|---|---|---|---|---|
| $x_1 = 0$ | $(0, 0, -1)$ | $(1, 9, -2)$ | $x_1 = 0$ | $(0, 0, -1)$ | $(1, 9, -4)$ |
| $x_1 = 1$ | $(9, 1, -2)$ | $(12, 12, -3)$ | $x_1 = 1$ | $(9, 1, -4)$ | $(12, 12, -5)$ |

*As in the previous example, in the two states of nature the unique efficient laboratory output is* $\mathbf{x}^* = (1, 1)$*, which yields an expected joint profit of* 20 *as opposed to* 8 *if exactly one firm receives R&D, and* 0 *otherwise (both states of nature are given the same probability). Here no firm is interested in being the sole buyer of the laboratory's output, since it benefits indirectly from economies of scope in the production of R&D, and also directly from the other firm's received services. This means that, in the two states of nature,* $t_1(\mathbf{x}) = t_2(\mathbf{x}) = 0$ *unless* $\mathbf{x} = (1, 1)$*. In the latter case only, a Nash equilibrium in transfer payments is obtained for any pair* $(t_1(1, 1), t_2(1, 1))$ *which satisfies the participation constraint, that is*

$$t_1^o(1, 1) + t_2^o(1, 1) = 4.$$

*There is a continuum of equilibria which are such that*

$$v_{\mathcal{L}}^o = 0 \text{ and } v_1^o + v_2^o = 20.$$

*In addition, recall that each firm i's expected equilibrium profits are bounded from below by* $\overline{\Pi}(\{i\})$*, so that* $v_i^o \geq g_i(1, 1) - E_\theta (r(\theta, 1, 1)) = 8$*. Even when* $\theta = \overline{\theta}$*, in which case the laboratory's realized profits are negative, the latter has no incentive not to serve the two firms. In this set-up, unless we introduce an informational asymmetry or some new distribution of bargaining power because of a change in the structure of the market, firms do not have any incentive to cooperate nor to acquire the laboratory.* ■

These two examples illustrate the impact of various specifications of indirect (i.e., through the agent's costs) and direct (i.e., through principals' gross payoffs) externalities on the distribution of benefits, and the related incentives to integrate.

Direct and indirect externalities can be negative or positive. Proposition 3 refers to cases in which negative externalities dominate. The strict subadditivity of $\overline{\overline{\Pi}}(.)$ implies decreasing returns in the size of the set of principals. Consequently, each principal is interested in limiting the number of other players. Competition for the control of the agent's output is tough, a situation from which the agent obtains benefits. Proposition 4 characterizes situations in which positive externalities prevail. The convexity of $\overline{\overline{\Pi}}(.)$ implies increasing returns in the number of principals. In that case, each principal is interested in seeing the other players contract also. Competition for the control of the agent's choice is soft, which is unfavorable to the agent.

In the latter case, truthful Nash equilibria are strong Nash equilibria, *i.e.* are stable to *any* deviation by a coalition of players.[21] More formally:

**Definition 3: Strong Nash equilibria** *The pair* $(\mathbf{t}^o, \mathbf{x}^o(\theta))$ *is a strong Nash equilibrium of the game if it is a Nash equilibrium and there exists no coalition* $S \subseteq N$*, strategy profile* $\bar{\mathbf{t}}^o_S \equiv \{\bar{t}^o_i\}_{i \in S}$*, output* $\bar{\mathbf{x}}^o(\theta)$ *in* $X^o\left(\theta, X(\theta, \bar{\mathbf{t}}^o_S, \mathbf{t}^o_{N\setminus S}),\right)$ *such that* $E_\theta\left[v_i(\bar{t}^o_i, \bar{\mathbf{x}}^o(\theta))\right] \geq E_\theta\left[v_i(t^o_i, \mathbf{x}^o(\theta))\right]$ *for all principals in* $S$ *with strict inequality for at least one of them, given* $\mathbf{t}^o_{N\setminus S} \equiv \{t^o_i\}_{i \in N\setminus S}$.

In words, a Nash equilibrium is strong if there is no coalition of principals who can deviate by choosing alternative strategies that result in a higher payoff for each of them, for unchanged strategies as chosen by other principals.

**Theorem 4 (Equivalence between TNE and strong NE)** *In all games* $\Gamma$ *where* $\overline{\overline{\Pi}}(.)$ *is convex, all truthful Nash equilibria with payoffs are strong Nash equilibria.*

This claim can be seen as a further argument in favour of the truthfulness refinement, on top of Theorem 3, since the strong Nash equilibrium concept describes a superior degree of stability. Indeed, strong Nash equilibria – whenever they exist – form a subset of coalition-proof Nash equilibria. However, this stability property applies only to a subclass of characteristic functions that capture situations of positive externalities.

# 7 Sufficient Conditions

We now provide a simple generalization of two existing results in common agency theory (Villemeur and Versaevel, 2003). (The extension to the *ex ante* contracting setting proceeds by considering the expected version of expressions in proofs). They offer conditions on $r^S(\theta, \mathbf{x})$ and $g_i(\mathbf{x})$ which are sufficient for $\overline{\Pi}(.)$ to be strictly subadditive, or convex.

**Proposition** 5 *In all games $\Gamma^N$, if $g_i(\mathbf{x}_i, \mathbf{x}'_j) \leq g_i(\mathbf{x}_i, \mathbf{x}_j)$, all $\mathbf{x}'_j \geq \mathbf{x}_j$, and $r^S(\theta, \mathbf{x})$ is strictly supermodular in $\mathbf{x}$, then $\overline{\Pi}(.)$ is strictly subadditive.*

In this proposition, the condition on $g_i$ describes negative *direct* externalities, and the condition on $r^S(\theta, \mathbf{x})$ describes negative *indirect* externalities. The supermodularity condition formalizes a case of decreasing returns in the dimensions of $\mathbf{x}$ on the laboratory's side. Supplying more R&D services to one firm makes it more costly to serve the other firms. This can be interpreted as a phenomenon of congestion. A variation on Proposition 7 is obtained by observing that when the monotonicity property of $g_i$ in $\mathbf{x}_j$ holds true, and $r^S(\theta, \mathbf{x})$ is additive separable across firms (so that indirect externalities are nil), then it is sufficient to verify that $g_i$ is strictly submodular in $\mathbf{x}$, all $i$, to conclude again that $\overline{\Pi}(.)$ is strictly subadditive.if $\overline{\Pi}(.)$.[22] The latter property, combined with Theorem 3.1 in Laussel and Le Breton (2003, p. 102), leads to the conclusion that the laboratory obtains positive benefits in all equilibria. When strict subadditivity and strong subadditivity coincide, as in the case $\#N = 2$, then it remains to evoke Proposition 6 to compute the equilibrium payoffs of all game participants.

■ **Example 5:** *$N = \{1, 2\}$, and $\mathbf{x} = (x_1, x_2)$ describes the diffusion possibilities of a patent attached to some process R&D output, with a "winner-take-all" feature, with $X_i = \{0, 1\}$. The stochastic parameter $\theta$ can be either $\underline{\theta}$ or $\overline{\theta}$ with probability $1/2$. R&D cost are $r^S(\theta, \mathbf{x}) = 0$ if*

*$x_1 = x_2 = 0$, $r^S(\theta, \mathbf{x}) = \theta$ if $x_1 + x_2 = 1$, and $r^S(\theta, \mathbf{x}) = +\infty$ otherwise, a cost specification borrowed from Laussel and Le Breton* (2001)*. The function $r^S$ is strictly supermodular in $\mathbf{x}$. Each firm i's unit production cost is a positive constant $c_i(x_i)$, with $c_i(0) = c_H$ and $c_i(1) = c_L < c_H$. The two firms produce a homogeneous good, compete in prices, and total demand is $q = 1 - p$. We obtain $g_i(\mathbf{x}) = (c_H - c_L)(1 - c_H) > 0$ if $x_i = 1$ and $x_j = 0$, and $g_i(\mathbf{x}) = 0$ otherwise. Clearly, $g_i(x_i, x') \leq g_i(x_i, x_j)$, all $x'_j \geq x_j$. In equilibrium, the laboratory maximizes benefits by contracting with one firm only (say, firm 1, indifferently), so that $A^o = \{1\}$. Firm* 1 *then receives $x_1^o(\theta) = 1$ in exchange of $t_1^o((1,0)) = g_1((1,0)) > 0$, and $v_1^o = 0$. Firm* 2 *receives $x_2^o(\theta) = 0$ in exchange of $t_2^o((1,0)) = g_2((1,0)) = 0$, and $v_1^o = 0$ also. The laboratory earns $(c_H - c_L)(1 - c_H) - \left(\underline{\theta} + \overline{\theta}\right)/2 > 0$.*■

**Proposition** 6 *In all games $\Gamma^N$, if $g_i(\mathbf{x}_i, \mathbf{x}'_j) \geq g_i(\mathbf{x}_i, \mathbf{x}_j)$, all $\mathbf{x}'_j \geq \mathbf{x}_j$, $r^S(\theta, \mathbf{x})$ is submodular in* $\mathbf{x}$, *and $g_i(\mathbf{x})$ is supermodular in $\mathbf{x}$, all $i$, then $\overline{\Pi}(.)$ is convex.*

Here the monotonicity condition on $g_i$ captures a case of positive *direct* externalities. This condition is not sufficient to obtain a characterization of $\overline{\Pi}(.)$. We need complementarities on the firms' side, as described by the supermodularity condition on $g_i$, and also on the laboratory's side, as captured by the submodularity of $r^S(\theta, \mathbf{x})$ in $\mathbf{x}$. The latter property is a case positive *indirect* externalities. This means that delivering more to a given firm makes it less costly to satisfy the other firms. When specific algebraic expressions satisfy the conditions displayed in Proposition 7, it remains to refer to Proposition 6 to conclude that $v_{\mathcal{L}}^o = 0$ and $\sum_{i \in N} v_i^o = \overline{\Pi}(N)$.

■ **Example 6:** *$N = \{1, 2\}$, and the laboratory's costs are $r^S(\theta, \mathbf{x}) = x_1 + x_2 - \theta$, with $x_i \in X_i = \{0, 1\}$, and $\theta$ can be either $\underline{\theta}$ or $\overline{\theta}$ with probability* 1/2*. Each firm i's variable costs are normalized to zero, and demand is $q_i(\mathbf{p}, \mathbf{x}) = 1/(1+\gamma) - 1/(1-\gamma^2)p_i + \gamma/(1-\gamma^2)p_j$, where the degree of substitutability $\gamma$ is a function of $\mathbf{x}$, as in Lambertini and Rossini* (1998)*, that is $\gamma \equiv \gamma(\mathbf{x}) \in [0, 1]$. R&D aims at enhancing symmetric horizontal differentiation that occurs* *only if the two firms buy services from the laboratory. Formally, let* $\gamma(\mathbf{x}) = 0$ *if* $x_1 x_2 > 0$*, and* $\gamma(\mathbf{x}) \rightarrow 1$ *otherwise. Non-cooperative profit maximization in prices in the market stage yields* $g_i(x_i, x_j) = (1-\gamma) / \left((1+\gamma)(2-\gamma)^2\right)$*, and one finds* $g_i(x_i, x_j') \geq g_i(x_i, x_j)$*, all* $x_j' \geq x_j$*, and* $g_i(\mathbf{x})$ *is supermodular in* $\mathbf{x}$*. In equilibrium, the laboratory maximizes benefits by contracting with the two firms, so that* $A^o = \{1, 2\}$*. Each firm receives* $x_i^o(\theta) = 1$ *in exchange of* $t_i^o((1,1)) = g_i((1,1)) > 0$*, and* $v_1^o + v_2^o = \theta - 3/2$*. The laboratory's expected benefits are zero.* ■

# 8 Incentives to Cooperate and Integrate on the Market for Technology

In this section we exploit the properties of the common agency game to investigate the incentives firms face to coordinate their decisions to purchase R&D services, or to integrate vertically by acquiring the external laboratory. Toward this aim, we assume that no efficiency gain obtains, in the sense that firms do not upgrade their expertise by cooperating in R&D choices with a rival, nor by acquiring the external laboratory.

Consider first the extreme situation in which firms and the laboratory *all* participate in some form of coordination mechanism. This occurs if all $n$ firms share the ownership of the laboratory and control it as a joint venture. There is no gain in joint profits to be obtained in this scenario. To see that, recall that equilibrium R&D outcomes $\mathbf{x}^*$ belong to the set of efforts which yield the highest joint profits for all parties (the joint profits maximization property). The net residual share of joint profits accruing to each buyer of another firm's equity would thus not improve on the amount of net profits received at equilibrium in the delegated R&D common agency game. Indeed, complete forward integration (*i.e.*, all firms become subsidiaries) would imply the payment of $v_i^o$ by the laboratory to the owners of downstream assets. By the same token, backward integration (*i.e.*, the laboratory becomes a joint venture) would require the total payment of $v_{\mathcal{L}}^o$ by firms for the ownership of R&D assets. The equality $v_{\mathcal{L}}^o + \sum_N v_i^o = \overline{\Pi}(N)$ holds in all cases, unless further assumptions are introduced (*e.g.*, costs or gross benefits depend on the governance structure).

If, in an alternative scenario, all firms and the laboratory merge (*i.e.*, firms merge horizontally also), then gross profits increase because firms become divisions of a $n$-product monopolist and thereby internalize downstream competition. This effect is orthogonal to the concern of this paper, since it is a consequence of a change in the final market structure only, not in the upstream conditions of delivery of R&D services. Henceforth we assume that firms may contractually commit on the transfer payments they offer to the laboratory only, or choose to integrate the external laboratory, but may not coordiante their decisions on the final market for goods.[23]. This allows us to focus on the impact on the distribution of R&D benefits of the participation of any subset of firms (including $N$) in such a horizontal coordination mechanism, or of the shift to a more vertically integrated structure that leads any subset of firms (excluding $N$) to control the laboratory.

We demonstrate that the existence of strategic incentives for more horizontal of vertical coordination depends essentially on the nature of competition on the market for R&D services, *i.e.* on the existence of anti-complementarities *vs.* complementarities, in the sense of Propositions 3 and 4.

### 8.1 Cases of anti-complementarities

We focus in this section on all cases characterized by anti-complementarities in the production or use of R&D services, as described in Proposition 3.

**Horizontal cooperation** Consider first the impact of a horizontal agreement on the distribution of R&D benefits. For all $n \geq 2$, suppose that two firms cooperate in their choices of transfer payment offers to the independent laboratory. This leaves the structure of the delegated R&D common agency game unchanged, although the distribution of profits does change to the advantage of the cooperating parties.

To see that, recall that precise results on the distribution of benefits depend on the structural properties of $\overline{\Pi}$. If anti-complementarities dominate, so that the characteristic function is strongly subadditive, we know that there is a unique TNE, and that firm $i$'s (expected) profits are $v_i^o = \overline{\Pi}(N) - \overline{\Pi}(N \backslash \{i\})$, all $i$. Without any loss of generality, assume that firms 1 and 2 coordinate their choices, as if they were a unified entity, we denote by a subscript $u$. It follows that the joint equilibrium profits are $v_u^o = \overline{\Pi}(N) - \overline{\Pi}(N \backslash \{1, 2\})$. The gains from coordination are thus given by the difference

$$v_u^o - (v_1^o + v_2^o) = \overline{\Pi}(N \backslash \{1\}) + \overline{\Pi}(N \backslash \{2\}) - \overline{\Pi}(N \backslash \{1, 2\}) - \overline{\Pi}(N),$$

which is positive as a consequence of the strong subadditivity of $\overline{\Pi}$. Therefore firms have strategic incentives to cooperate horizontally on the intermediate market for technology in the case of anti-complementarities in the production and use of R&D services. Now recall that firm $i$'s equilibrium profits $v_i^o$ remain equal, all $i \geq 3$, in the absence of post-coordination efficiency gains. As a result, the laboratory contemplates a reduction in its individual equilibrium profits. A limit case is $n = 2$ (a situation illustrated by Figure 1).[24] If the two firms coordinate their strategies, everything happens as if there were only one player left on the intermediate market for technology. As a principal, this player writes contracts in order to make the agent exactly break even, so that $v_{\mathcal{L}}^o = 0$.

[**insert Figure 1 here**]

**Vertical integration** Consider now the impact of (partial) vertical integration, where the laboratory merges with only one out of $n$ users, say firm 1 (without loss of generality). The assumption that firms can choose to source R&D services internally, by relying on proprietary resources, leads to discuss two possible cases.[25] Assume first that firm 1 is endowed with superior capabilities, that is $\underline{v}_1 > \overline{\Pi}(\{1\})$. Then the participation constraint on the firm's side, which imposes $v_1^o \geq \underline{v}_1$, leads to $v_1^o = \overline{\Pi}(N) - \overline{\Pi}(N \backslash \{1\}) \geq \overline{\Pi}(\{1\})$, which contradicts the strict subadditivity property. We can thus assume that the alternative case holds, in which the laboratory is endowed with superior capabilities, that is $\underline{v}_1 \leq \overline{\Pi}(\{1\})$. In other words, when anti-complementarities dominate, a firm delegates the production of R&D services to a specialized laboratory only if the latter can compensate for diseconomies in the number of contracting users by offering a sufficiently high level of expertise.

In this context, consider the situation in which the merged entity does not contract with any other potential user of R&D services we label $j$, a case of foreclosure (see Figure 2 in the case $n = 2$). This choice does not make parties to the merger appropriate a higher level of joint profits than earned without integrating. To see that, note that the merged entity which does not sell R&D services to other firms earns $\max_{\mathbf{x}} \left(g_1(\mathbf{x}) - r^{\{1\}}(\mathbf{x})\right) \equiv \overline{\Pi}(\{1\})$. Then recall from Theorem 2 that $\sum_{j \in N \backslash \{1\}} v_j^o \leq \overline{\Pi}(N) - \overline{\Pi}(\{1\})$ (principals' payoff vector is in $V_\Gamma^*$) and also that $v_{\mathcal{L}}^o + v_1^o = \overline{\Pi}(N) - \sum_{j \in N \backslash \{1\}} v_j^o$ (Pareto efficiency). It follows that $\overline{\Pi}(\{1\}) \leq v_{\mathcal{L}}^o + v_1^o$. It is therefore not possible for the laboratory and a given firm to receive more than the reservation values $v_{\mathcal{L}}^o$ and $v_1^o$, respectively, by simply merging vertically and not serving outsiders with R&D services.

**[insert Figure 2 here]**

Assume now that, although anti-complementarities dominate, the merged entity does sell R&D services to other firms. In this case the structure of the game changes, since non-merging parties may not write contracts and address them on a take-it or leave-it basis to the merged entity. This is because the latter combines the R&D capabilities of the laboratory with the firm's bargaining position. We see this as a situation of bilateral bargaining for the choice of $\mathbf{x}$ between the integrated party $\mathcal{L} + 1$ on the one side, and each outsider $j$ in $N \backslash \{1\}$ on the other side. The payoff function of the merged unit is given by $g_1(\mathbf{x}) - r^{\{1\}}(\mathbf{x}) + t_j(\mathbf{x})$, while the payoff function of a separate firm $j$ remains unchanged, that is $g_j(\mathbf{x}) - t_j(\mathbf{x})$. Given the lump-sum transferability of payments, Pareto optimality of the bargaining process induces efficiency of the bargaining equilibria. Formally,

$$\tilde{v}_{\mathcal{L}+1} + \sum_{j \in N \backslash \{1\}} \tilde{v}_j = \max_{\mathbf{x} \in X} \left( \sum_{j \in N} g_j(\mathbf{x}) - r^N(\mathbf{x}) \right) \equiv \overline{\Pi}(N),$$

where $\tilde{v}_{\mathcal{L}+1}$ and $\tilde{v}_j$ denote the profits of the merged pair and each outsider, respectively, in *all* bargaining equilibria. Moreover, the disagreement point is $(\underline{v}_{\mathcal{L}+1}, \underline{\mathbf{v}}_j)$, where $\underline{v}_{\mathcal{L}+1}$ is the minimum level of expected profits the merged entity may obtain by choosing to behave as an agent (and thereby not to use firm 1's bargaining power), and $\underline{\mathbf{v}}_j$ is the vector of outsiders' reservation profits. By definition, $\underline{v}_{\mathcal{L}+1}$ is equal to the joint expected profits minus the sum of expected profits earned by all firms $j$ in $N \backslash \{1\}$ in the unique TNE of the truncated delegated R&D common agency game $\Gamma^{N \backslash \{1\}}$. Since $\overline{\Pi}(N)$ and $v_j^o = \overline{\Pi}(N) - \overline{\Pi}(N \backslash \{j\})$ remain the same in the latter game as in the pre-merger setup $\Gamma^N$, we have

$$\underline{v}_{\mathcal{L}+1} = v_{\mathcal{L}}^o + v_1^o.$$

This implies in turn that, in all equilibria $\tilde{\mathbf{x}}$ of the bargaining game, related profits $\tilde{v}_{\mathcal{L}+1}$ and $\tilde{v}_j$ are such that

$$\tilde{v}_{\mathcal{L}+1} = g_1(\tilde{\mathbf{x}}) - r^{\{1\}}(\tilde{\mathbf{x}}) + t_j(\tilde{\mathbf{x}}) \geq v_{\mathcal{L}}^o + v_1^o \text{ and } \tilde{v}_j \geq \underline{v}_j,$$

all $j$, $j \neq 1$.

As a result, the vertically integrated entity cannot be worse-off post integration, including in the very extreme situation in which it has no bargaining power with respect to other firms. Note that the latter claim implies that foreclosure is always a dominated choice for the integrated structure.

The following proposition summarizes the discussion.

**Proposition 7** *In all delegated R&D common agency games $\Gamma^N$, with $n \geq 2$, if $\overline{\Pi}(.)$ is strongly subadditive then in all* TNE *firms have strategic incentives either to coordinate their transfer payment strategies or to acquire the laboratory.*

## 8.2 Cases of complementarities

We focus hereafter on all cases in which complementarities dominate. In that case, recall from Proposition 4 that, although potential users of R&D services compete through monetary offers, which truly reflect their needs, the laboratory exactly breaks even.

**Horizontal cooperation** Suppose now that firms 1 and 2 (without loss of generality) may coordinate their transfer payment strategies, in which case they behave as a unified entity on the market for R&D services.

In the absence of cooperation, we know from Proposition 4 that the sum of the two firms' expected profits is the difference $\overline{\Pi}(N) - \sum_{j \in N \backslash \{1,2\}} v_j^o$, and from Theorem 2 that no firm can get more than its marginal contribution to the coalition payoff ($\mathbf{v}^o$ is in $V_\Gamma^*$), implying that $-\sum_{j \in N \backslash \{1,2\}} v_j^o \geq \overline{\Pi}(\{1,2\}) - \overline{\Pi}(N)$. In equilibrium, the two firms they may thus not earn less than the joint level of expected profits $\hat{v}_{1,2} = \overline{\Pi}(\{1,2\})$. Recalling that $\underline{v}_1$ and $\underline{v}_2$ denote reservation profits, it follows that the *total* expected equilibrium profits of the two firms are bounded from *below* by

$$\underline{v}_{1,2} = \sup \{\underline{v}_1 + \underline{v}_2, \hat{v}_{1,2}\}.$$

Remark now that, in the absence of efficiency gain, or in the nature of competition on the final market for goods, the cooperation of firms 1 and 2 has no impact on $\overline{\Pi}(N)$ nor on any outsider's marginal contribution $\overline{\Pi}(N) - \overline{\Pi}(N \backslash \{j\})$, for all $j$ in $N \backslash \{1,2\}$. It follows that the lowest value of the set of the cooperating parties' *joint* expected equilibrium profits $\underline{v}_u$ remains the same as in the absence of cooperation, or more formally that

$$\underline{v}_u = \underline{v}_{1,2}.$$

We may also identify an upper bound for the interval of expected profits as earned by the two firms if they do not cooperate. To see that, recall from Theorem 2 again that the *sum* of the two firms' expected equilibrium profits are bounded from *above* by the sum of their marginal contribution, that is

$$\overline{v}_{1,2} = 2\overline{\Pi}\left(N\right) - \overline{\Pi}\left(N\backslash\left\{1\right\}\right) - \overline{\Pi}\left(N\backslash\left\{2\right\}\right).$$

Next, if firms 1 and 2 cooperate, and thus behave as a unified entity, the latter's expected equilibrium profits may not exceed the marginal contribution

$$\overline{v}_{u} = \overline{\Pi}\left(N\right) - \overline{\Pi}\left(N\backslash\left\{1,2\right\}\right),$$

all other firms' individual expected profits remaining the same. A simple comparison of the latter two displayed expressions leads to

$$\overline{v}_{u} - \overline{v}_{1,2} = \left[\overline{\Pi}\left(N\backslash\left\{2\right\}\right) - \overline{\Pi}\left(N\backslash\left\{1,2\right\}\right)\right] - \left[\overline{\Pi}\left(N\right) - \overline{\Pi}\left(N\backslash\left\{1\right\}\right)\right] < 0,$$

where the negative sign results from the convexity of $\overline{\Pi}$. To conclude, horizontal coordination leads the ordered set of joint equilibrium profits, as available to cooperating firms, to be truncated from above. Firms have thus no strategic incentives to cooperate horizontally on the intermediate market for R&D in the case of complementarities.

**Vertical integration** We know from Proposition 4 that the convexity of $\overline{\Pi}$ leads to $v_i^o = \overline{\Pi}(N) - \sum_{j\in N\backslash\{i\}} v_j^o$, and from Theorem 2 that no firm can get more than its marginal contribution to the coalition payoff ($\mathbf{v}^o$ is in $V_\Gamma^*$), implying that $-\sum_{j\in N\backslash\{i\}} v_j^o \geq \overline{\Pi}(\{i\}) - \overline{\Pi}(N)$. It follows that $v_i^o \geq \overline{\Pi}(\{i\})$, which implies that firm $i$'s expected equilibrium profits are bounded from *below* by

$$\sup\{\underline{v}_i, \overline{\Pi}(\{i\})\}.$$

Toward the identification of an upper bound to the set of expected equilibrium profits, recall that the participation constraints of individual firms in $N\backslash\{i\}$ imposes that $v_j^o \geq \underline{v}_j$, and from Theorem 2 again that in all equilibria of the delegated R&D common agency game $v_i^o = \overline{\Pi}(N) -$

$\sum_{j \in N \setminus \{i\}} v_j^o$ (efficiency property). It follows that $v_i^o$ is bounded from *above* by

$$\overline{\Pi}(N) - \sum_{j \in N \setminus \{i\}} \underline{v}_j.$$

Suppose now that the laboratory is acquired by, say, firm 1, so that the unified entity may benefit from the R&D capabilities of the laboratory together with the firm's bargaining power.[26] Since outsiders may not address take-it or leave-it contracts to the integrated party, again we turn to a situation of bilateral bargaining for the choice of $\mathbf{x}$ between the unified entity which receives $g_1(\mathbf{x}) - r^{\{1\}}(\mathbf{x}) + t_j(\mathbf{x})$, and each outsider $j$ which obtains $g_j(\mathbf{x}) - t_j(\mathbf{x})$.

Consider first the situation of foreclosure in which the vertically integrated entity does not transact with any outsider. The integrated entity may either choose to use the laboratory's facilities, in which case it earns an expected profit equal to $v_{\mathcal{L}+1}^o = \overline{\Pi}(\{1\})$, or decide to close down the laboratory and obtain $v_{\mathcal{L}+1}^o = \underline{v}_1$. This says that the integrated entity may guarantee for itself the disagreement value $\sup\{\underline{v}_1, \overline{\Pi}(\{1\})\}$.

Now consider the situation in which the vertically integrated structure does sell R&D services to any separate firm $j$. In all equilibria of the bargaining game, we know that efficiency holds (in the sense of joint profits maximization), so that $v_{\mathcal{L}+1}^o + \sum_{j \in N \setminus \{1\}} v_j^o = \overline{\Pi}(N)$. Since each outsider's disagreement profit is $\underline{v}_j$, the merged entity's expected profits may not exceed $\overline{\Pi}(N) - \sum_{j \in N \setminus \{1\}} \underline{v}_j$.

Eventually, both in the bargaining and agency games, we have

$$\sup\{\underline{v}_1, \overline{\Pi}(\{1\})\} \leq v_1^o \leq \overline{\Pi}(N) - \sum_{j \in N \setminus \{1\}} \underline{v}_j.$$

The set of equilibria in the common agency game is thus identical to the set of equilibria in the bargaining game. This leads to the conclusion that, unless some additional refinement is introduced that justifies the selection of distinct equilibria in the two games, firms have no incentive to integrate vertically.

The next proposition concludes the discussion.

**Proposition 8** *In all delegated R&D common agency games* $\Gamma^N$, *with* $n \geq 2$, *if* $\overline{\overline{\Pi}}(.)$ *is convex then in all* TNE *firms have no strategic incentive either to cooperate or to acquire the laboratory.*

# 9 Policy implications

The discussion in the previous paragraphs, as summarized in Propositions 7 and 8, establishes a relationship between the distribution of total profits – as earned by all firms and the laboratory – and alternative structures or collaborative patterns on the intermediate market for R&D services. Depending on structural properties of the characteristic function $\overline{\overline{\Pi}}$, which reflect a combination of factors – including the laboratory's technology, and direct externalities that impact firms' gross payoffs – firms may find in individually profitable or not to coordinate their choices of payment schemes, or to acquire the laboratory, all other things remaining equal. We find that neither the coordination of transfer payments nor the acquisition of the laboratory impact the equilibrium levels of R&D outputs received by firms. Indeed, we know from Theorem 2 that firms' truthful strategies lead the laboratory to produce a multidimensional output that maximizes total profits, and also that there is no alternative distribution of equilibrium profits that can make all firms more profitable. Summarizing:

**Proposition 9** *In all delegated R&D common agency games* $\Gamma^N$, *with* $n \geq 2$, *R&D outputs and joint profits do not depend on firms' decision either to cooperate horizontally or to acquire the laboratory.*

We now whish to examine policy implications. A first interesting point is that the latter result is clearly supportive of legal environments that do not inhibit firms either to conclude R&D cooperative agreements with rivals, or to integrate a specialized laboratory, including in highly concentrated industries. Recall also that, in our model, each firm may relate its payments to the laboratory to the exact amount of R&D received by other firms. This comes in favour of a legal environment that allows parties to include discriminatory clauses in the R&D contracts that govern their outsourcing practices.

Second, the preservation of joint-profit maximization property gives no theoretical reason to forge new regulatory tools that are specific to the delegation of R&D by independent firms to a for-profit independent laboratory. This is because, although R&D costs are uncertain to all parties at the contracting stage, and firms design payment schemes non-cooperatively, the equilibrium R&D outputs we obtain coincide with the ones an integrated entity – a joint venture – would choose. Note however that R&D outputs, as driven by private incentives, may fall short of their social value. Therefore existing considerations on the relevant incentives for a monopolist to invest in R&D, as first investigated by Arrow (1962), are still valid when applied to all firms considered as a whole.

The problem faced by a regulator, or a competition authority, remains thus to align the incentives to invest in R&D – as faced here by all firms and the laboratory – with a social welfare objective. This policy issue is also the object of a series of theoretical studies from which we may draw interesting lessons, as they compare easily with the present analysis. The most commented contributions to this literature (e.g., d'Aspremont and Jacquemin (1988), Kamien, Muller, and Zang (1992), Suzumura (1992)) offer models which have in common the assumption that information is complete, and firms first engage in R&D by relying on proprietary resources (i.e., there is no R&D delegation), before competing on a product market. The most relevant benchmark is Leahy and Neary (1997), in which the model specifications encompass most of the previous settings. It leads to a ranking of welfare levels, as obtained in several scenarios, including some government intervention in the form of a per unit subsidy to R&D. Some conditions are unveiled which, when satisfied, lead to a higher welfare level when firms are assumed to cooperate in R&D than when they conduct R&D non-cooperatively. This clearly applies to a particular version of our delegated R&D common agency game, as obtained by specifying that R&D costs, together with gross profit expressions, adopt the relevant algebraic forms. In that case, our optimality result implies that welfare outcomes coincide with the ones that can be obtained in a counterpart model *à la* Leahy and Neary where firms tap new technology in-house and cooperate horizontally in R&D choices to maximize joint profits. More can be said by

It follows that in the present model, as in all models of R&D cooperation mentioned above, R&D levels depend on the nature of competition on the final market. Although equilibrium R&D outputs on the intermediate market for new technology maximize total profits, a change in final market conditions will result in a change in R&D levels. For example, an increase in the concentration of the industry, or the entry of new competitors, will motivate firms to adapt their payment strategies, and thus lead the laboratory to deliver adjusted outputs. This in turn will impact firms' profits, the consumer surplus, and welfare. In other words, the final market structure conditions firms' ability to drive upstream R&D operations by the means of transfer payments through the laboratory. The technology of the latter, and resulting R&D outputs, in turn impact the benefits firms earn by selling goods to consumers. This implies that policies of all kinds, including subsidies and antitrust intervention, when they focus on product market conditions only, leaving aside the intermediate market for R&D, may result in harmful welfare consequences.

To illustrate the latter point, we construct and compare two specific examples which exploit the fact that our model specifications may represent a large spectrum of intermediate and final market conditions. In each of them, the set $X_i = [0, c)$, where $c < 1$, represents the range of cost-reducing R&D outputs the laboratory may deliver at some cost $r^S(\theta, \mathbf{x})$ to a subset $S$ of identical firms, with $S \subseteq N$. The stochastic parameter $\theta$ takes values in $[\underline{\theta}, \overline{\theta}]$. Firms compete over two periods. In the first one, they compete in transfer payment offers on the intermediate market for R&D, by participating in the delegated R&D common agency game $\Gamma^N$. In the second period, they compete in quantities to sell a homogeneous product on a final market, *à la* Cournot. Each firm $i$'s marginal production cost $c - x_i$ is decreasing in the R&D service it receives from the laboratory, and there is a fixed cost component $f$. The inverse final market demand is $P = 1 - \sum_{i \in N} q_i$. Toward subgame-perfection, we consider the two periods backward. By computing the Cournot-Nash equilibrium of the second period, one obtains each firm's gross profit function of $\mathbf{x}$, that is

$$g_i(\mathbf{x}) = \left(1 - c + nx_i - \sum_{j \neq i} x_j\right)^2 / (1+n)^2 - f.$$

Given the latter expression, firms' equilibrium payoffs depend on the R&D outputs supplied in the first period. We consider now two variations, which relate to alternative cost conditions inside the laboratory.

■ **Example 7:**

*The laboratory's costs are*

$$r^S(\theta, \mathbf{x}) = (1-\sigma)\frac{\theta}{2}\sum_{i \in S} x_i^2 + F,$$

*where* $\sigma \in [0,1]$ *represents a government per-unit subsidy, and* $F$ *is a fixed component. We focus on symmetric interior solutions to the joint-profit maximization problem* $\max_{\mathbf{x} \in X} \Pi_N(\theta, \mathbf{x})$, *implying that the laboratory contracts with all firms, so that* $A^o = N$.[27] *Remark that R&D costs are additive separable across firms, and that* $g_i(\mathbf{x})$ *is strictly submodular in* $\mathbf{x}$, *so that* $\overline{\Pi}(.)$ *is strictly subadditive (see comments below Proposition* 7*). It follows that, in equilibrium, each firm* $i$ *receives* $x_i^o(\theta) = 2(1-c)/\left(\theta(1-\sigma)(1+n)^2 - 2\right)$ *in exchange of* $t_i^o(\mathbf{x}^o(\theta)) = g_i(\mathbf{x}^o(\theta)) - v_i^o$, *where* $v_i^o = \overline{\Pi}(N) - \overline{\Pi}(N \backslash \{i\})$, *from Proposition* 6*. Note that the R&D outputs are monotone increasing in* $\sigma$, *and decreasing in* $n$. *It is easy to check that both industry profits* $\overline{\Pi}$ *and the consumer surplus* $CS$ *are increasing with* $\sigma$, *and that total welfare, net of subsidies, reaches a maximum at* $\sigma = n/(n+2)$. *On the other hand, the impact on welfare of a change in the number of firms is ambiguous, since*

$$\mathrm{d}\overline{\Pi}(N)/\mathrm{d}n < 0,$$

*whereas the change in consumer surplus depends on the value of parameters, as*

$$\mathrm{d}CS/\mathrm{d}n > 0 \text{ iff } \theta\,(1-\sigma) < 2(2n+1)/\,(1+n)^2\,. \tag{13}$$

*This means that entry is beneficial to consumers only if* $\sigma$ *is sufficiently* high.

Remark that the specifications of Example 7 are a particular case of the linear set-up described by Leahy and Neary (1997). More precisely, the additive separability of costs implies that the symmetric R&D outcomes, total profits, and consumer surplus we compute would also be obtained if firms were assumed to cooperate in R&D instead of delegating R&D non-cooperatively to a common independent laboratory. We may thus appropriate the policy implications developed in the latter paper.

In the following example, we introduce anti-complementarities in the laboratory's R&D production process, so that serving one firm implies an increase in the costs of serving other firms.

■ **Example 8:**

*The laboratory's costs are*

$$r^S\,(\theta, \mathbf{x}) = (1-\sigma)\,\frac{\theta}{2}\left(\sum_i x_i\right)^2 + F,$$

*with the same notation as in the previous example. The only difference is that R&D outputs* $x_i$ *are now substitutable in the expression of* $r^S$*. Again, we focus on symmetric interior solutions to the joint-profit maximization problem* $\max_{\mathbf{x}\in X} \Pi_N(\theta, x)$*, implying that* $A^o = N$*.*[28] *Each firm i's received equilibrium R&D service is* $x_i^o\,(\theta) = 2\,(1-c)\,/\,\left(n\theta\,(1-\sigma)\,(n+1)^2 - 2\right)$*, in exchange of payment* $t_i^o\,(\mathbf{x}^o\,(\theta)) = g_i(\mathbf{x}^o\,(\theta)) - v_i^o$*, where* $v_i^o = \overline{\Pi}(N) - \overline{\Pi}(N\backslash\{i\})$ *from Proposition* 6*. As in the previous example, both industry profits and the consumer surplus are increasing in* $\sigma$*, and total welfare reaches a maximum at* $\sigma = n/\,(n+2)$*. Now, although one still finds*

$$\mathrm{d}\overline{\Pi}(N)/\mathrm{d}n < 0,$$

*a change in the number of firms will not always impact welfare in the same direction as in the previous example. This is because the change in consumer surplus that results from a rise in $n$ is now positive only if $\sigma$ sufficiently* low, *as*

$$\mathrm{d}CS/\mathrm{d}n > 0 \text{ if and only if } \theta\left(1-\sigma\right) > 2\left(3n+2\right)/n\left(1+n\right)^{2}. \tag{14}$$

The comparison of the two examples leads to the conclusion that policy measures should not be designed by considering product market conditions in isolation, that is by excluding the intermediate market for R&D from the analysis. To see that, suppose first that R&D is not subsidized (i.e., $\sigma = 0$). Then examine the impact of greater rivalry – as captured by the number of firms – on welfare. In both examples, when the stochastic parameter takes values in a range of finite dimension, one finds that policy intervention in favour of *less* concentration impacts negatively industry profits and positively the consumer surplus.[29] However, if $\theta$ is relatively small (in the sense of (13)), then a rising number of firms results in more consumer surplus in Example 7 only. On the other hand, if $\theta$ is relatively high (as in (14)), entry results in more consumer surplus in Example 8 only.

These comments clearly echoe a recent analysis by Katz and Shelanski (2004) of the potentially conflicting economic effects an increase in industry concentration might have on research and development. They describe "the impact of innovation on economic welfare and the impact of market structure on innovation" (p. 67), and encourage antitrust agencies to take into account the industrial organization of R&D of merging parties in factual inquiries specific to a given case. The comparison of Examples 7 and 8, in which product market conditions are identical, points to the same direction.[30] In the two cases, a reduction in the number of firms results in higher R&D levels and more joint profits. However, depending on R&D cost conditions inside the external laboratory, the impact on consumer surplus can be negative in the first example (when $\theta$ is relatively small), and positive in the second one. A merger policy that retains focus on product market concerns exclusively may therefore reach opposite welfare consequences when enforced in real world contexts represented by one or the other setting.

In addition, by introducing R&D subsidies a regulator may impact the relationship between market concentration and the welfare components. The outcome will differ across examples, although final market conditions are symmetric. Indeed, suppose that parameter values are such that entry is beneficial to consumers in the two examples. This occurs if $\theta$ is intermediate, in the sense that it satisfies conditions (13) and (14). Then $\mathrm{d}CS/\mathrm{d}n$ remains positive in Example 7 if $\sigma$ departs from 0 and rises, while its sign changes in Example 8 if $\sigma$ reaches a sufficiently high value. For a specific example, let $\theta = 1$ and $n = 2$, and suppose that the R&D subsidies parameter rises. Then in Example 7, in which R&D costs are additively separable across firms, the consumer surplus increases monotonically when $n$ rises for all levels of $\sigma$. In Example 8, the consumer surplus increases as a result of entry only if $\sigma$ is less than 1/9 (which is lower than the welfare maximizing level), and decreases otherwise. The same regulatory intervention results in different welfare consequences in the latter example only because of the existence of anti-complementarities inside the laboratory.

## 10 References


1. Aghion, P. and Tirole, J. 1994. On the management of innovation. Quarterly Journal of Economics 109, 1185-1209.

2. Ambec, S., and Poitevin, M. 2001. Organizational design of R&D activities. Working Paper 38, CIRANO.

3. Argyres, N. S. and Liebeskind, J. P. 2002. Governance inseparability and the evolution of US biotechnology industry. Journal of Economic Behavior and Organization 47, 197-219.

4. Armour, H. 0. and Teece, D. J. 1978. Vertical integration and technological innovation. Review of Economics and Statistics 60, 470-474.

5. Arrow, K. 1962. Economic welfare and the allocation of resources for invention, in Nelson, R. (ed.), The Rate and Direction of Incentive Activity: Economic and Social Factors, Princeton University Press.

6. Bernheim, B. D. and Whinston, M. D. 1986a. Menu auctions, resource allocation, and economic influence. Quarterly Journal of Economics 101, 1-31.

7. Bernheim, B. D. and Whinston, M. D. 1986b. Common agency. Econometrica 54, 923-942.

8. Bernheim, B. D. and Whinston, M. D. 1998. Exclusive dealing. Journal of Political Economy 106, 64-103.

9. Bernheim, B.D., Peleg, B. and Whinston, M. D. 1987. Coalition-proof Nash equilibria: I concepts. Journal of Economic Theory 42(1), pp. 1-12.

10. Billette de Villemeur, E. and Versaevel, B. 2003. From private to public common agency. Journal of Economic Theory 111, 305-309.

11. Bousquet, A., H. Crémer, M. Ivaldi, and M. Wolkowicz, M. 1998. Risk sharing in licensing. International Journal of Industrial Organization 16, 535-554.

12. Calzolari, G. and Scarpa, C. 2004. Non-intrinsic common agency. mimeo.

13. Cassiman, B. and Veugelers, R. 2000. External technology sources: embodied or disembodied technology acquisition. DTEW, KULeuven, mimeo.

14. Crémer, J. and Riordan, M. H. 1987. On governing multilateral transactions with bilateral contracts. RAND Journal of Economics 18, 436-451.

15. d'Aspremont, C. and Jacquemin, A. 1988. Cooperative and noncooperative R&D in duopoly with spillovers. American Economic Review 78, 1133-1137.

16. Dobler, D. W. and Burt, D. N. 1996. Purchasing and Supply Management: Text and Cases. 6th ed. McGraw-Hill, New-York.

17. Epstein, L. and Peters, M. 1999. A revelation principle for competing mechanisms. Journal of Economic Theory 88, 119-160.

18. Gal-Or, E. 1991. A common agency with incomplete information. RAND Journal of Economics 22, 274-286.

19. Hamberg, D. 1963. Invention in the industrial research laboratory. Journal of Political Economy 61, 95-115.

20. Harris, M. and Raviv, A. 1979. Optimal incentive contracts with imperfect information. Journal of Economic Theory 20, 231-259.

21. Helfat, C. E. 1994. Firm-specificity in corporate applied R&D. Organization Science 5, 173-184.

22. Howells, J. 1999. Research and technology outsourcing. Technology Analysis and Strategic Management 11, 17-29.

23. Jensen, R. 1992. Reputational spillovers, innovation, licensing, and entry. International Journal of Industrial Organization 10, 193-212.

24. Jorde, T. and Teece, D. 1990. Innovation and cooperation: implications for competition and antitrust. Journal of Economic Perspectives 4, 75-96.

25. Katz, M. L. and Shapiro, C. 1986. How to license intangible property. Quarterly Journal of Economics 16, 504-520.

26. Katz, M. L. and Ordover, J. A. 1990. R&D cooperation and competition. Brookings Papers: Microeconomics, 137-203.

27. Katz, M. L. and Shelanski, H. A. 2004. Merger policy and innovation: must enforcement change to account for technological change? Working Paper 10710, NBER.

28. Laffont, J.-J. and Martimort, D. 1997. The firm as a multicontract organization. Journal of Economics and Management Strategy 6, 201-234.

29. Lambertini, L. and Rossini, G. 1998. Product homogeneity as a prisoner's dilemma in duopoly with R&D. Economics Letters 58. 297-301.

30. Laussel, D. and Le Breton, M. 1998. Efficient private production of public goods under common agency. Games and Economic Behavior 25, 194-218.

31. Laussel, D. and Le Breton, M. 2001. Conflict and cooperation: the structure of equilibrium payoffs in common agency. Journal of Economic Theory 100, 93-128.

32. Leahy, D. and Neary, J. P. 1997. Public policy towards R&D in oligopolistic industries. American Economic Review 87, 642-662.

33. Leahy, D. and Neary, J. P. 2005. Symmetric research joint ventures: cooperative substitutes and complements. International Journal of Industrial Organization 23, 381-397.

34. Majewski, S. 2004. How do consortia organize collaborative R&D?: evidence from the national cooperative research act. Discussion Paper 483, Harvard Law School.

35. Martimort, D. 1992. Multi-principaux avec anti-sélection. Annales d'Economie et de Statistiques 28, 1-38.

36. Martimort, D. 1996. Exclusive dealing, common agency, and multiprincipals incentive theory. RAND Journal of Economics 27, 1-31.

37. Martimort, D. 2005. Multi-contracting mechanism design. mimeo.

38. Martimort, D. and Stole. L. 2001. Contractual externalities and common agency equilibria. Advances in Theoretical Economics 3, Issue 1, Article 4. http://www.bepress.com/bejte.

39. Martimort, D. and Stole, L. 2002. The revelation and delegation principles in common agency games. Econometrica 70, 1659-1673.

40. Martimort, D. and Stole, L. 2004. Market participation under delegated and intrinsic common agency games. mimeo.

41. Martin, S. 2001. Advanced Industrial Economics. Blackwell.

42. Mowery, D. C. 1990. The development of industrial research in U.S. manufacturing. American Economic Review 80, 345-349.

43. Mowery, D. C. and Rosenberg, N. 1989. Technology and the Pursuit of Economic Growth. Cambridge University Press, Cambridge.

44. Narula, R. 2001. Choosing between internal and non-internal R&D activities: some technological and economic factors. Technology Analysis & Strategic Management 13, 365-387.

45. National Science Foundation. 2006. Research and development funds and technology linkage. Science and Engineering Indicators Arlington, VA, Chapter 4.

46. Oxley, J. 1997. Appropriability hazards and governance in strategic alliances: a transaction cost approach. Journal of Law, Economics, and Organization 13, 387-411.

47. Peters, M. 2001. Common agency and the revelation principal. Econometrica 69(5), 1349-1372.

48. Peters, M. 2003. Negotiation and take it or leave it in common agency. Journal of Economic Theory 111, 88-109.

49. Pisano, G. 1989. Using equity to support exchange: evidence from the biotechnology industry. Journal of Law, Economics, and Organization 5, 109-126.

50. Pisano, G. 1991. The governance of innovation: vertical integration and collaborative arrangements in the biotechnology industry. Research Policy 20, 237-249.

51. Poyago-Theotoky, J. 1997. Research joint ventures and product innovation: some welfare aspects. Economics of Innovation and New Technology 5, 51-73.

52. Poyago-Theotoky, J. 1999. A note on endogenous spillovers in a non-tournament R&D duopoly. Review of Industrial Organization 13, 249-276.

53. Sappington, D. 1983. Limited Liability Contracts between Principal and Agent. Journal of Economic Theory 29, 1-21.

54. Schmitz, P. W. 2001. On monopolistic licensing strategies under asymmetric information. Journal of Economic Theory 106, 177-189.

55. Sinclair-Desgagné. B. 2001. Incentives in common agency. Working Paper 2001s-66, CIRANO.

56. Stole, L. 1991. Mechanism design under common agency. mimeo.

57. Suzumura, K. 1992. Cooperative and noncooperative R&D in an oligopoly with spillovers. American Economic Review 82, 1307-1320.

58. Tapon, F. 1989 A transaction costs analysis of innovations in the organization of pharamaceutical R&D. Journal of Economic Behavior and Organization 12, 197-213.

59. Teece, D. J. 1986. Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. Research Policy 15, 285-305.

60. Teece, D. J. 1989. Inter-organizational requirements of the innovation process. Managerial and Decision Economics, Special Issue, 35-45.

61. Veugelers, R. 1997. Internal R&D expenditures and external technology sourcing, Research Policy 26, 303-316.

62. Vonortas, N. S. 1994. Inter-firm cooperation with imperfectly appropriable research. International Journal of Industrial Organization 12, 413-435.

63. Bernheim, B. D. and Whinston, M. D. 1986a. Menu auctions, resource allocation, and economic influence. Quarterly Journal of Economics 101, 1-31.

## Notes

[1]In this paper, we interchangeably use the terms "outsourcing", "contracting out", and "delegation", in order to refer to some R&D tasks undertaken by a laboratory on behalf of one or several firms under conditions laid out in a contract agreed formally beforehand.

[2]The National Science Foundation (2006) uses the term "contract R&D" to denote a transaction with external parties involving R&D payments or income, regardless of the actual legal form of the transaction. The quoted figures do not include contract R&D expenses by U.S. companies that do not perform internal R&D, or that contract out R&D to companies located overseas.

[3]For analyses of changes in the industrial organization of R&D in the chemical industry over the last two decades, see Rosemberg (1990), Pisano (1991), Lerner and Merges (1998), Martin (2001), Argyres and Liebskind (2002), Katz and Shelanski (2004), among others.

[4]For example, in a contract signed by Tularik (a US laboratory) and Japan Tobacco, the objective of the R&D program is "to agonize or antagonize Orphan Nuclear Receptors for the treatment of disease in humans ("Field"). Subject to the fourth sentence of this paragraph, 'Orphan Nuclear Receptors' shall mean: (i) [*], (ii) any protein containing a [*] domain of [*] amino acid residues [*] of which are [*] that is further characterized by [*] of the [*] type and to the [*] of which [*] domain is a region with [*] to the [*] domain of any of the [*] set forth in (i) above, for which a [*] shall not have been identified, but excluding [*] and the [*] (as defined in [*]; and (iii) other [*] added to the research component of the Program (...)" Although detailed technological specifications that appear in the original version of the contract are private information (and thus marked by brackets [*] in the quotation), the publicly available clauses support our specification that outputs of outsourced R&D projects are targeted in very precise directions.

[5]In the context of the pharmaceutical industry, and before focusing on situations in which the laboratory is better informed than the user, Ambec and Poitevin (2001) emphasize that the properties of a new drug "are never known *ex ante*". They also recognize that the value of an innovation depends on the efficiency of the new process or the quality of the new product, an information "difficult to obtain *before* the innovation is developed" (p. 2, added emphasis). This is the reason why the contract they consider is signed *after* the research phase, but before development operations. By contrast, we look at situations in which the laboratory may contract with one or several firms before research and development operations are engaged.

[6]Several other contributions to the theoretical economic literature introduce changes in the specifications of Bernheim and Whinston (1986a) in order to explain a large variety phenomena. Among others, Mitra (1999) adds a coalition formation stage, in which individual principals may decide to bear the cost of forming a group or not, with an application to trade policy. Another

example is a paper by Dixit, Grossman and Helpman (1999), which relaxes the assumption that preferences are quasi-linear, in a model of tax policy. More recently, Martimort (2004) modifies the basic framework by introducing moral hazard (i.e., the agent 's action cannot be verifiable), in a political economy environment.

[7]In the terminology of Laffont and Martimort (1997), indirect externalities are of "type 1", and direct externalities of "type 2".

[8]In an example, Sinclair-Desgagné (2001) also refers to the analysis of R&D contracts as a possible application of a common agency setup in which principals condition their payments to the agent on the quantities received by other principals.

[9]To elaborate, the set of types defined in Epstein and Peters (1999) includes the mechanisms used by the other principals, and also informs whether the latter mechanisms depend on other principals' mechanisms, and so on.

[10]Peters (2001) also shows that any equilibrium allocation that can be obtained with this set of menus is "weakly robust", in the sense that the allocation persists as an equilibrium allocation – among possible others – when the set of mechanisms available to the principals is enlarged. In a related contribution, Martimort and Stole (2002) also investigate the possibility of substituting menus of contracts for more complex communication mechanisms. However, the common agency model they construct specifies that a given principal may not contract over the set of allocations controlled by another principal, whereas we are interested in situations in which each principal may condition payment transfers to the agent on the output received by rival principals.

[11]In Peters (2003)'s terminology, the strategies $t_i$ we consider in the present model are "take-it or leave-it offers", that is "degenerate menus consisting of a single incentive contract" (p. 89).

[12]This does not exclude the no R&D output case from the set of possible outcomes. When it finds it profitable to do so, the laboratory may choose $\mathbf{x} = 0$ in stage 3.

[13]This act limits antitrust recoveries against registered agreements to reduced damages if the terms of the submitted agreement are found to violate the law. It also gives firms the possibility of limiting antitrust analysis to the rule of reason in lieu of the *per se* illegality rule. The most frequently mentioned cooperative agreements which registered under the NCRA come in support to existing models of R&D cooperation, in which rival firms collaborate *horizontally* at the pre-competitive stage by coordinating the use of proprietary R&D facilities. However, agreements of a very different kind apply. For example, Katz *et al.* (1990, pp. 186-190) describe the case of MCC (Microelectronics and Computer Technology Corporation), which became "an open-market supplier of contract R&D" to supply application-oriented "deliverables" to large firms (*e.g.*, Hewlett-Packard, Motorola, Honeywell, Westinghouse Electric). In a more systematic way, Majewski (2004), analyzes detailed contract-level data to present stylized facts on all R&D

projects that filed for protection under the NCRA. A large subset of these projects refer to competitors in product markets that choose to outsource the production of R&D services to a third-party contractor. These agreements are better matched by the specifications of the present delegated R&D setup than by models of horizontal R&D cooperation.

[14]This qualification appears in Martimort (2005). This says that each principal may observe and verify the output level delivered to the other principal, and thus condition its offered payments on it (by contrast, in a *private* common agency set-up, a principal may contract only on the output it specifically receives from the agent).

[15]Remark that an agent who received truthful payment offers in the contracting stage has strictly *more* information than principals after the realization of the stochastic parameter in the action stage.

[16]To see that, observe that $g_i(\mathbf{x}^o(\theta)) \geq v_i^o$ implies $t_i(\mathbf{x}^o(\theta)) = g_i(\mathbf{x}^o(\theta)) - v_i^o$ from (10), then substitute $g_i(\mathbf{x}^o(\theta)) - v_i^o$ for $t_i^o(\mathbf{x}^o(\theta))$ in (8).

[17]More precisely, the limited liability constraints can be binding, but only over subsets of $\Theta$ with zero measure.

[18]The notion of coalition-proof Nash equilibrium is introduced by Bernheim *et al.* (1987).

[19]Remark that Theorem 3 does not preclude the existence of coalition-proof Nash equilibria outside $I_\Gamma^*$.

[20]This unusual payoff matrix presentation is borrowed from Prat and Rustichini (2003). It illustrates the fact that the two firms, as principals, play a game through the laboratory, an agent.

[21]This concept was introduced by Aumann (1959). In contrast to coalition-proof Nash equilibria, the definition of strong Nash equilibria does not introduce any restriction on the set of possible deviations.

[22]This can be proved by rewritting the agent's costs as $\tilde{r}^S(\theta, \mathbf{x}) = r^S(\theta, \mathbf{x}) - \sum_{i \in S} g_i(\mathbf{x})$. Then Proposition 7 applies.

[23]We may equivalently consider the horizontal merger of firms if we maintain the assumption that divisions of a merged entity non-cooperatively maximize their individual profits in their own control variables on the final market (price, quantity, quality, advertising, and so on).

[24]The strong subadditivity property is closely related to the strict subadditivity property. The characteristic function $\overline{\Pi}(.)$ is subadditive if $\overline{\Pi}(N) \leq \overline{\Pi}(S) + \overline{\Pi}(T)$ for all $S, T \in 2^N$ such that $S \cup T = N$ and $S \cap T = \emptyset$. There is strict subbaditivity if at least one of the inequalities above is strict. For $n = 2$, strong and strict subadditivity coincide. However, the strong subadditivity property is more demanding than strict subadditivity for all $n > 2$.

[25]Recall that $\underline{v}_i$ can be interpreted as reservation profits obtained by sourcing R&D services from a dedicated laboratory or by relying on in-house capabilities. The latter interpretation is favoured in the following, for simplicity.

[26]We exclude here all incentives to merge that may stem from the acquisition of some exogenously assumed superior ability to supply R&D outputs. For well-known empirically grounded accounts on efficiency gains as a motive for vertical integration between R&D and production stages, see Armour and Teece (1978) and Pisano (1989, 1991).

[27]For simplicity, we treat firms equally *ex ante*, in the words of Leahy and Neary (2005). In that case, a sufficient second-order condition for a symmetric optimum is $\partial^2\Pi_S(\theta,\mathbf{x})/\partial x_i^2 + (n-1)\partial^2\Pi_S(\theta,\mathbf{x})/\partial x_i \partial x_j < 0$, all $i, j$, $i \neq j$. In this example, this is equivalent to $\theta(1-\sigma) > 2/(1+n)^2 - 1$, which holds true.

[28]As in the previous example, we treat firms equally *ex ante*. From Leahy and Leary (2005), a sufficient second-order condition for a symmetric optimum is $\theta(1-\sigma) > 2/\left(n(1+n)^2\right)$.

[29]This holds true when $\theta$ is in $\left(2(3n+2)/n(1+n)^2, 2(2n+1)/(1+n)^2\right)$, for all $n \geq 2$. Note that the latter interval is always defined.

[30]Katz and Shelanski (2004) do not evoke R&D contracts. In short analyzes of recent merger cases, with a special emphasis on the pharmaceutical industry, they report the existence of R&D agreements among competitors.

# 11 Appendix

## 11.1 Notation and Definitions

Recall that, for any given distribution of $\theta$, principal $i$'s expected net payoff is $E_\theta\left[g_i(\mathbf{x}(\theta)) - t_i(\mathbf{x}(\theta))\right]$, where $t_i(\mathbf{x})$ is principal $i$'s contingent transfer payment, and $g_i(\mathbf{x})$ is principal $i$'s gross payoff.

**Principals' total transfers and total gross payoffs** For any realization of the stochastic variable $\theta$, define:

$$\mathbf{t}(\mathbf{x}(\theta)) = \sum_{i \in N} t_i(\mathbf{x}(\theta)), \tag{15}$$

and

$$\mathbf{g}(\mathbf{x}(\theta)) = \sum_{i \in N} g_i(\mathbf{x}(\theta)). \tag{16}$$

**Agent's total cost and incremental costs** The (research) costs *a priori* depends on the exact set of principals that actually contract with the agents *i.e.* the set of accepted contracts $A$.

For any set of principal $S \subseteq N$, define

$$r(S, \theta, \mathbf{x}) \tag{17}$$

the costs associated to the production of $\mathbf{x} \in X$ in the state of nature $\theta$ when the agent accept to contract with $S$ principals. For any $i \in S \subseteq N$, we shall denote $f_i^S(\theta, \mathbf{x})$ the incremental cost defined by:

$$f_i^S(\theta, \mathbf{x}) = r(S, \theta, \mathbf{x}) - r(S \backslash \{i\}, \theta, \mathbf{x}). \tag{18}$$

By convention, if $i \notin S$, then $f_i^S(\theta, \mathbf{x}) \equiv 0$.

More generally, for any set $T \subseteq S$,we shall denote $f_T^S$ the incremental cost defined by

$$f_T^S(\theta, \mathbf{x}) = r(S, \theta, \mathbf{x}) - r(S \backslash T, \theta, \mathbf{x}). \tag{19}$$

**Limited liability** We shall assume that principal transfers are non-negative and bounded from below. More precisely, we will focus on the cases where, at equilibrium, for all $S \subseteq A^o$

$$\sum_{j\in S} t_j\left(\mathbf{x}\right) \geq f_S^{A^o}(\theta, \mathbf{x}),$$

almost all $\theta$. If research costs are $r(S, \theta, \mathbf{x})$

**Agent's chosen actions set and efficient actions set** For any realization of the stochastic variable $\theta$, define:

$$X^o\left(\theta\right) = \arg\max_{\mathbf{x}\in X} \left\{v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}, \mathbf{x}) = \mathbf{t}(\mathbf{x}) - r(A^o, \theta, \mathbf{x})\right\}, \tag{20}$$

where $A^o$ is the equilibrium set of accepted contracts. $X^o\left(\theta\right)$ is the agent's profit-maximizing set of actions associated to state of nature $\theta$. Let

$$X^*\left(\theta\right) = \arg\max_{\mathbf{x}\in X} \left(W\left(\theta, \mathbf{x}\right) = \mathbf{g}(\mathbf{x}) - r(A^*, \theta, \mathbf{x})\right), \tag{21}$$

where $A^*$ is the set of accepted contracts.that allow to maximize the expected joint-profits. $X^*\left(\theta\right)$ is the joint-profit maximizing set of actions associated to state of nature $\theta$, also referred to as the set efficient actions.

**Joint profits of the agents and a subset of principals** Let $2^N$ represent the set of subsets of $N$ including the empty set. For any set $S \in 2^N$, and any realization $\theta$ of the stochastic variable, define:

$$W(S, \theta) = \max_{\mathbf{x}\in X} \left(W(S, \theta, \mathbf{x}) = \sum_{i\in S} g_i(\mathbf{x}\left(\theta\right)) - r(S, \theta, \mathbf{x})\right), \tag{22}$$

that is the joint profit of the agent and principals in $S$. A maximizer of $W(S, \theta, \mathbf{x})$ is denoted $\mathbf{x}_S^*\left(\theta\right)$, an element of:

$$X_S^*\left(\theta\right) = \arg\max_{\mathbf{x}\in X} W\left(S, \theta, \mathbf{x}\right). \tag{23}$$

We assume that the expected profits of the agent in isolation is zero: $E_\theta\left[W(\emptyset, \theta)\right] = 0$.

**Payoff distributions sets** For any realization of the stochastic variable $\theta$, define:

$$V_\Gamma(\theta, \mathbf{x}) = \left\{ \mathbf{v} = (v_1, v_2, .., v_n) : \text{for all } S \subseteq N, \ \sum_{i \in S} v_i \leq \left[ \sum_{i \in N} g_i(\mathbf{x}) - r^N(\theta, \mathbf{x}) - W(\theta, N \backslash S) \right] \right\}, \tag{24}$$

that is the set of all possible payoff distributions when the action $\mathbf{x}$ is chosen by the agent for the realisation $\theta$ of the stochastic variable.

Of interest is also the set of the possible payoff distributions at the *ex-ante* stage, that is when the realisation of the stochastic variable is not yet known. Let $\mathbf{x}(\theta)$ be the action chosen by the agent for the realisation $\theta$ of the stochastic variable, and define:

$$V_\Gamma(\{\mathbf{x}(\theta)\}_{\theta \in \Theta}) = \left\{ \begin{array}{c} \mathbf{v} = (v_1, \ldots, v_n) : \text{for all } S \subseteq N, \\ \sum_{i \in S} v_i \leq E_\theta \left[ \sum_{i \in N} g_i(\mathbf{x}(\theta)) - r^N(\theta, \mathbf{x}(\theta)) - W(\theta, N \backslash S) \right] \end{array} \right\}. \tag{25}$$

We denote $V_\Gamma^*$ the set of the possible payoff distribution when the action $\mathbf{x}(\theta)$ chosen by the agent is in $X^*(\theta)$:

$$V_\Gamma^* \equiv V_\Gamma(\{\mathbf{x}^*(\theta)\}_{\theta \in \Theta}) = \left\{ \mathbf{v} = (v_1, \ldots, v_n) : \text{for all } S \subseteq N, \ \sum_{i \in S} v_i \leq E_\theta \left[ W(\theta, N) - W(\theta, N \backslash S) \right] \right\}.$$

Remark indeed that, by definition of $X^*(\theta)$,

$$W(\theta, N) = \sum_{i \in N} g_i(\mathbf{x}^*(\theta)) - r^N(\theta, \mathbf{x}^*(\theta))$$

for all $\mathbf{x}^*(\theta)$ in $X^*(\theta)$ so that $V_\Gamma^* \equiv V_\Gamma(\{\mathbf{x}^*(\theta)\}_{\theta \in \Theta})$ does not depend on the specific set $\{\mathbf{x}^*(\theta)\}_{\theta \in \Theta}$ considered.

*Observe implementation of the payoff distribution is robust to uncertainty if*

$$\mathbf{v} \in \cap_{\theta \in \Theta} V_\Gamma(\theta, \mathbf{x}^*(\theta))$$

*It is ex-post implementable is* $\mathbf{v} \in V_\Gamma(\theta, \mathbf{x}^*(\theta))$, *while it is ex-ante implementable (i.e. implementable in expected terms if* $\mathbf{v} \in V_\Gamma(\{\mathbf{x}^*(\theta)\}_{\theta \in \Theta})$

**Implementable payoff set**

For any $\mathbf{x} \in \mathbf{X}$, define:

$$I_\Gamma(\mathbf{x}) = \{\mathbf{v} = (v_1, \ldots, v_n) : \text{for all } i \in N,\, v_i \leq g_i(\mathbf{x})\},$$

that is the set of all payoff distributions that is implementable *by the means of non-negative transfers* when the action $\mathbf{x}$ is chosen by the agent.

Of interest is also the set of payoff distributions that is implementable *by the means of non-negative transfers* at the *ex-ante* stage, that is when the realisation of the stochastic variable, hence the action chosen by the agent is not yet known. Let $\mathbf{x}(\theta)$ be the action chosen by the agent for the realisation $\theta$ of the stochastic variable, and define:

$$I_\Gamma\left(\{\mathbf{x}(\theta)\}_{\theta\in\Theta}\right) = \{\mathbf{v} = (v_1, \ldots, v_n) : \text{for all } i \in N,\, v_i \leq g_i(\mathbf{x}(\theta))\}, \tag{26}$$

almost all $\theta$.

We denote $I_\Gamma^*$ the set of the implementable payoff distribution when there exists a set of action $\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta}$ in $\{X^*(\theta)\}_{\theta\in\Theta}$ such that $\mathbf{v}$ is in $I_\Gamma\left(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta}\right)$. Formally:

$$I_\Gamma^* = \left\{\mathbf{v} = (v_1, \ldots, v_n) : \mathbf{v} \in I_\Gamma\left(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta}\right) \text{ for some } \{\mathbf{x}^*(\theta)\}_{\theta\in\Theta}\right\}.$$

Remark indeed that, in contrast to $V_\Gamma(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta})$, the set $I_\Gamma\left(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta}\right)$ depends on the specific set $\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta}$ considered.

## 11.2 Proof of Theorem 3

### 11.2.1 Theorem 11.2.1 (Characterization of the Nash equilibria):

*A triplet* $(\mathbf{t}^o, A^o, \{\mathbf{x}^o(\theta)\})$ *is a Nash Equilibrium of the delegated common game* $\Gamma$ *if and only if:*

(1) *the action* $\mathbf{x}^o(\theta)$ *is in:*

$$X^o(\theta) = \arg\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}^o, \mathbf{x}) \quad \textit{for almost all } \theta$$

(2) *for all principals* $i$ *in* $A^o$*:*

$$E_\theta\left[v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}^o, \mathbf{x}^o(\theta))\right] = \sup\left\{0, \sup_{S\subseteq N\setminus\{i\}} E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{S}(\theta, \mathbf{t}_{-i}^o, \mathbf{x})\right]\right\};$$

(3) *for all principals* $i$ *in* $A^o$*:*

$$X^o(\theta) \subseteq \arg\max_{\mathbf{x}\in X}\left[g_i(\mathbf{x}) - f_i^{A^o}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{A^o\setminus\{i\}}(\theta, \mathbf{t}_{-i}^o, \mathbf{x})\right] \quad \textit{for almost all } \theta;$$

(with the convention $f_i^{A^o}(\theta, \mathbf{x}) \equiv 0$ if $i \notin A^o$)

(4$a$) *for all principals* $i$ *in* $A^o$*:*

$$\begin{aligned} & E_\theta\left[\left(g_i(\mathbf{x}^o(\theta)) - f_i^{A^o}(\theta, \mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o\setminus\{i\}}(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta))\right)\right] \\ \geq\ & \sup\left\{0, E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x}) - f_i^{S\cup\{i\}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{S}(\theta, \mathbf{t}_{-i}^o, \mathbf{x})\right)\right]\right\} \quad \textit{all } S \subseteq N\setminus\{i\}; \end{aligned}$$

and

$$\begin{aligned} & E_\theta\left[\left(g_i(\mathbf{x}^o(\theta)) - f_i^{A^o}(\theta, \mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o\setminus\{i\}}(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta))\right)\right] \\ \geq\ & \sup\left\{0, E_\theta\left[g_i(\mathbf{x}_S^o(\theta)) + v_{\mathcal{L}}^{S}(\theta, \mathbf{t}_{-i}^o, \mathbf{x}_S^o(\theta))\right]\right\} \quad \textit{all } S \subseteq N\setminus\{i\}; \end{aligned}$$

with $\mathbf{x}_S^o(\theta) = \arg\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{S}(\theta, \mathbf{t}_{-i}^o, \mathbf{x})$.

(4b) *for all principals $i$ in $N \backslash A^o$:*

$$E_\theta \left[ \left( g_i \left( \mathbf{x}^o \left( \theta \right) \right) + v_{\mathcal{L}}^{A^o} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x}^o \left( \theta \right) \right) \right) \right]$$
$$\geq \sup \left\{ 0, E_\theta \left[ \max_{\mathbf{x} \in X} \left( g_i \left( \mathbf{x} \right) - f_i^{S \cup \{i\}} \left( \theta, \mathbf{x} \right) + v_{\mathcal{L}}^S \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right) \right] \right\} \quad \text{all } S \subseteq N \backslash \{i\} .$$

### 11.2.2 Proof of Theorem 11.2.1:

$\Leftarrow$ *Sufficiency:*

Suppose that the triplet $(\mathbf{t}^o, A^o, \{\mathbf{x}^o(\theta)\})$ is *not* a Nash equilibrium of the game. There exists a principal $i \in N$, a strategy $\hat{t}_i(\mathbf{x})$, a set $\widehat{A}$ of accepted contracts and a set of actions $\{\hat{\mathbf{x}}(\theta)\}_{\theta \in \Theta}$ chosen by the agent, that make the principal $i$ strictly better off. There are four possible cases:

**a)** Principal $i \in A^o \cap \widehat{A}$ :

By definition,

$$v_{\mathcal{L}}^{\widehat{A}}(\theta, \mathbf{t}, \mathbf{x}) \equiv t_i(\mathbf{x}) - f_i^{\widehat{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A} \backslash \{i\}}(\theta, \mathbf{t}_{-i}, \mathbf{x}) . \tag{27}$$

If the agent contracts with principal $i \in \widehat{A}$ when it is offered $\mathbf{t} = \left( \hat{t}_i, \mathbf{t}_{-i}^o \right)$, we know that

$$E_\theta \left[ \hat{t}_i \left( \hat{\mathbf{x}} \left( \theta \right) \right) - f_i^{\widehat{A}} \left( \theta, \hat{\mathbf{x}} \left( \theta \right) \right) + v_{\mathcal{L}}^{\widehat{A} \backslash \{i\}} \left( \theta, \mathbf{t}_{-i}^o, \hat{\mathbf{x}} \left( \theta \right) \right) \right] \geq E_\theta \left[ \max_{\mathbf{x} \in X} v_{\mathcal{L}}^S \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right], \tag{28}$$

all $S \subseteq N \backslash \{i\}$.

If principal $i$ is strictly better off by adopting the strategy $\hat{t}_i(\mathbf{x})$ it does also mean that:

$$E_\theta \left[ g_i \left( \hat{\mathbf{x}} \left( \theta \right) \right) - \hat{t}_i \left( \hat{\mathbf{x}} \left( \theta \right) \right) \right] > E_\theta \left[ g_i \left( \mathbf{x}^o \left( \theta \right) \right) - t_i^o \left( \mathbf{x}^o \left( \theta \right) \right) \right] . \tag{29}$$

Now condition $(4a)$ implies that

$$E_\theta \left[ g_i \left( \mathbf{x}^o \left( \theta \right) \right) - f_i^{A^o} \left( \theta, \mathbf{x}^o \left( \theta \right) \right) + v_{\mathcal{L}}^{A^o \backslash \{i\}} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x}^o \left( \theta \right) \right) \right]$$
$$\geq E_\theta \left[ g_i \left( \mathbf{x} \left( \theta \right) \right) - f_i^{S \cup \{i\}} \left( \theta, \mathbf{x} \left( \theta \right) \right) + v_{\mathcal{L}}^S \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \left( \theta \right) \right) \right] \tag{30}$$

all $\{\mathbf{x}(\theta)\}_{\theta\in\Theta} \in X^{\Theta}$, all $S \subseteq N\backslash\{i\}$. In particular, by setting $\mathbf{x}(\theta) \equiv \hat{\mathbf{x}}(\theta)$ and $S = \widehat{A}\backslash\{i\}$ in (30) we get that:

$$\begin{aligned} & E_\theta\left[g_i\left(\mathbf{x}^o(\theta)\right) - f_i^{A^o}\left(\theta, \mathbf{x}^o(\theta)\right) + v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right)\right] \\ \geq \;& E_\theta\left[g_i\left(\hat{\mathbf{x}}(\theta)\right) - f_i^{\widehat{A}}\left(\theta, \hat{\mathbf{x}}(\theta)\right) + v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \hat{\mathbf{x}}(\theta)\right)\right]. \end{aligned} \tag{31}$$

Adding (29) and (31), we find

$$\begin{aligned} & E_\theta\left[-f_i^{A^o}\left(\theta, \mathbf{x}^o(\theta)\right) + v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right) - \hat{t}_i\left(\hat{\mathbf{x}}(\theta)\right)\right] \\ > \;& E_\theta\left[-f_i^{\widehat{A}}\left(\theta, \hat{\mathbf{x}}(\theta)\right) + v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \hat{\mathbf{x}}(\theta)\right) - t_i^o\left(\mathbf{x}^o(\theta)\right)\right]. \end{aligned} \tag{32}$$

Collecting terms:

$$E_\theta\left[\hat{t}_i\left(\hat{\mathbf{x}}(\theta)\right) - f_i^{\widehat{A}}\left(\theta, \hat{\mathbf{x}}(\theta)\right)\right] < E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}^o, \mathbf{x}^o(\theta)\right)\right] - E_\theta\left[v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \hat{\mathbf{x}}(\theta)\right)\right]. \tag{33}$$

Now, we know from agent's $\mathcal{L}$ participation constraint that $E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}^o, \mathbf{x}^o(\theta)\right)\right]$ must be non-negative. More precisely, since $i \in A^o$, we know from condition (2) that there are two cases:

- Either $E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}^o, \mathbf{x}^o(\theta)\right)\right] = 0$, in which case equation (33) writes directly

$$E_\theta\left[v_{\mathcal{L}}^{\widehat{A}}\left(\theta, \hat{t}_i, \mathbf{t}_{-i}^o, \hat{\mathbf{x}}(\theta)\right)\right] = E_\theta\left[\hat{t}_i\left(\hat{\mathbf{x}}(\theta)\right) - f_i^{\widehat{A}}\left(\theta, \hat{\mathbf{x}}(\theta)\right) + v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \hat{\mathbf{x}}(\theta)\right)\right] < 0. \tag{34}$$

  It cannot be the case that the agent accepts to contract with principals in $\widehat{A}$ and chooses the set of actions $\{\hat{\mathbf{x}}(\theta)\}_{\theta\in\Theta}$ since it yields negative profits.

- Or $E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}^o, \mathbf{x}^o(\theta)\right)\right] > 0$, in which case we know from condition (2) that there exists $S \subseteq N\backslash\{i\}$ and $\mathbf{x}_{-i}^o(\theta) \in \arg\max_{\mathbf{x}\in X} v_{\mathcal{L}}^S\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)$ such that

$$E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}^o, \mathbf{x}^o(\theta)\right)\right] = E_\theta\left[v_{\mathcal{L}}^S\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}_{-i}^o(\theta)\right)\right]. \tag{35}$$

Hence inequality (33) rewrites:

$$E_\theta\left[\hat{t}_i\left(\hat{\mathbf{x}}\left(\theta\right)\right)-f_i^{\widehat{A}}\left(\theta,\hat{\mathbf{x}}\left(\theta\right)\right)\right]<E_\theta\left[v_{\mathcal{L}}^{S}\left(\theta,\mathbf{t}_{-i}^{o},\mathbf{x}_{-i}^{o}\left(\theta\right)\right)\right]-E_\theta\left[v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^{o},\hat{\mathbf{x}}\left(\theta\right)\right)\right]. \tag{36}$$

that is:

$$E_\theta\left[\hat{t}_i\left(\hat{\mathbf{x}}\left(\theta\right)\right)-f_i^{\widehat{A}}\left(\theta,\hat{\mathbf{x}}\left(\theta\right)\right)+v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^{o},\hat{\mathbf{x}}\left(\theta\right)\right)\right]<E_\theta\left[v_{\mathcal{L}}^{S}\left(\theta,\mathbf{t}_{-i}^{o},\mathbf{x}_{-i}^{o}\left(\theta\right)\right)\right], \tag{37}$$

which contradicts (28). In words, the agent would not accept to contract with principal $i\in\widehat{A}$ if it could make a strictly better profit by contracting with the set $S\subset N\backslash\{i\}$ of principals. □

**b)** Principal $i$ in $A^o$ and *not* in $\widehat{A}$ :

If principal $i$ is strictly better off by offering the strategy $\hat{t}_i(\mathbf{x})$ it does mean that:

$$E_\theta\left[g_i\left(\hat{\mathbf{x}}\left(\theta\right)\right)\right]>E_\theta\left[g_i\left(\mathbf{x}^o\left(\theta\right)\right)-t_i^o\left(\mathbf{x}^o\left(\theta\right)\right)\right], \tag{38}$$

From condition $(4a)$ however we know that

$$E_\theta\left[g_i\left(\mathbf{x}^o\left(\theta\right)\right)-f_i^{A^o}\left(\theta,\mathbf{x}^o\left(\theta\right)\right)+v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^{o},\mathbf{x}^o\left(\theta\right)\right)\right]\geq E_\theta\left[g_i\left(\hat{\mathbf{x}}\left(\theta\right)\right)+v_{\mathcal{L}}^{\widehat{A}}\left(\theta,\mathbf{t}_{-i}^{o},\hat{\mathbf{x}}\left(\theta\right)\right)\right]. \tag{39}$$

Summing both inequalities gives

$$E_\theta\left[t_i^o\left(\mathbf{x}^o\left(\theta\right)\right)-f_i^{A^o}\left(\theta,\mathbf{x}^o\left(\theta\right)\right)+v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^{o},\mathbf{x}^o\left(\theta\right)\right)\right]>E_\theta\left[v_{\mathcal{L}}^{\widehat{A}}\left(\theta,\mathbf{t}_{-i}^{o},\hat{\mathbf{x}}\left(\theta\right)\right)\right], \tag{40}$$

that is

$$E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}^o,\mathbf{x}^o\left(\theta\right)\right)\right]>E_\theta\left[v_{\mathcal{L}}^{\widehat{A}}\left(\theta,\mathbf{t}_{-i}^{o},\hat{\mathbf{x}}\left(\theta\right)\right)\right]. \tag{41}$$

From agent's $\mathcal{L}$ participation constraint, we know that $E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}_{-i}^{o},\mathbf{x}^o\left(\theta\right)\right)\right]$ is non negative. More precisely, since $i\in A^o$, we know from condition (2) that there are two cases:

- Either $E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}^o,\mathbf{x}^o\left(\theta\right)\right)\right]=0$, in which case equation (41) writes directly

$$E_\theta\left[v_{\mathcal{L}}^{\widehat{A}}\left(\theta,\mathbf{t}_{-i}^o,\hat{\mathbf{x}}\left(\theta\right)\right)\right]<0. \tag{42}$$

  It cannot be the case that the agent accepts to contract with principals in $\widehat{A}$ and chooses the set of actions $\{\hat{\mathbf{x}}\left(\theta\right)\}_{\theta\in\Theta}$ since it yields negative profits.

- Or $E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}^o,\mathbf{x}^o\left(\theta\right)\right)\right]>0$, in which case we know from condition (2) that there exists $S\subseteq N\backslash\{i\}$ and $\mathbf{x}_{-i}^o\left(\theta\right)\in\arg\max_{\mathbf{x}\in X} v_{\mathcal{L}}^S\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)$ such that

$$E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}^o,\mathbf{x}^o\left(\theta\right)\right)\right]=E_\theta\left[v_{\mathcal{L}}^S\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}_{-i}^o\left(\theta\right)\right)\right]. \tag{43}$$

  Hence inequality (41) rewrites:

$$E_\theta\left[v_{\mathcal{L}}^{\widehat{A}}\left(\theta,\mathbf{t}_{-i}^o,\hat{\mathbf{x}}\left(\theta\right)\right)\right]<E_\theta\left[v_{\mathcal{L}}^S\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}_{-i}^o\left(\theta\right)\right)\right], \tag{44}$$

  which contradicts the assumption that agent $\mathcal{L}$ choices $\left(\widehat{A},\{\hat{\mathbf{x}}\left(\theta\right)\}_{\theta\in\Theta}\right)$ when it is offered $\mathbf{t}=\left(\hat{t}_i,\mathbf{t}_{-i}^o\right)$ maximize its profits. $\square$

**c)** Principal $i$ is *not* in $A^o$ but in $\widehat{A}$ :

If principal $i$ is strictly better off by offering the strategy $\hat{t}_i\left(\mathbf{x}\right)$ it does mean that:

$$E_\theta\left[g_i\left(\hat{\mathbf{x}}\left(\theta\right)\right)-\hat{t}_i\left(\hat{\mathbf{x}}\left(\theta\right)\right)\right]>E_\theta\left[g_i\left(\mathbf{x}^o\left(\theta\right)\right)\right], \tag{45}$$

From condition ($4b$) however we know that

$$E_\theta\left[g_i\left(\mathbf{x}^o\left(\theta\right)\right)+v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o\left(\theta\right)\right)\right]\geq E_\theta\left[g_i\left(\mathbf{x}\left(\theta\right)\right)-f_i^{S\cup\{i\}}\left(\theta,\mathbf{x}\left(\theta\right)\right)+v_{\mathcal{L}}^S\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\left(\theta\right)\right)\right], \tag{46}$$

all $\{\mathbf{x}\left(\theta\right)\}_{\theta\in\Theta}\in X^\Theta$, all $S\subseteq N\backslash\{i\}$. In particular, by setting $\mathbf{x}\left(\theta\right)\equiv\hat{\mathbf{x}}\left(\theta\right)$ and $S=\widehat{A}\backslash\{i\}$ in (46) we get that:

$$E_\theta\left[g_i\left(\mathbf{x}^o\left(\theta\right)\right)+v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o\left(\theta\right)\right)\right]\geq E_\theta\left[g_i\left(\hat{\mathbf{x}}\left(\theta\right)\right)-f_i^{\widehat{A}}\left(\theta,\hat{\mathbf{x}}\left(\theta\right)\right)+v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\hat{\mathbf{x}}\left(\theta\right)\right)\right]. \tag{47}$$

Summing both inequalities gives:

$$E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o(\theta)\right)\right] > E_\theta\left[\hat{t}_i\left(\hat{\mathbf{x}}(\theta)\right) - f_i^{\widehat{A}}\left(\theta,\hat{\mathbf{x}}(\theta)\right) + v_{\mathcal{L}}^{\widehat{A}\setminus\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\hat{\mathbf{x}}(\theta)\right)\right], \tag{48}$$

that is

$$E_\theta\left[v_{\mathcal{L}}^{\widehat{A}}\left(\theta,\hat{t}_i,\mathbf{t}_{-i}^o,\hat{\mathbf{x}}(\theta)\right)\right] < E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o(\theta)\right)\right]$$

which contradicts the assumption that agent $\mathcal{L}$ choices $\left(\widehat{A},\{\hat{\mathbf{x}}(\theta)\}_{\theta\in\Theta}\right)$ when it is offered $\mathbf{t}=\left(\hat{t}_i,\mathbf{t}_{-i}^o\right)$ maximize its profit. □

**d)** Principal $i\notin A^o\cup\widehat{A}$ :

If the agent does not find profitable to contract with principal $i$ in either cases (when the later offers $t_i^o(\mathbf{x})$ or $\hat{t}_i(\mathbf{x})$), it does mean that the change of strategy has no impact on agent $\mathcal{L}$ choices $\left(\widehat{A},\{\hat{\mathbf{x}}(\theta)\}_{\theta\in\Theta}\right)$. This contradicts the assumption that offering $\hat{t}_i(\mathbf{x})$ result in a strict improvement of principal $i$'s payoff. □

(Remind that $i\notin A\cup\widehat{A}$ implies that $f_i^A(\theta,\mathbf{x})\equiv f_i^{\widehat{A}}(\theta,\mathbf{x})\equiv 0$).

$\Rightarrow$ *Necessity:*

Condition (1) follows from the assumption, given the contracting set $A^o$, the agent chooses for all $\theta\in\Theta$ the action $\mathbf{x}^o(\theta)$ that maximizes its own profits;

Condition (2) is Lemma 11.2.3 below and its implication displayed as Corollary **??**. It has a twofold meaning. First, it states that the set of accepted contracts $A^o$ maximizes the expected profits of the agent. Second, it evidences that the participation constraint of the agent to the contractual relationship with any principal $i\in A^o$ is binding;

Condition (3) is Lemma 11.2.5 below that evidences the fact that, given the contracting set $A^o$, the strategy $t_i^o(\mathbf{x})$ of any principal $i$ in the contracting set is such that, for almost all $\theta\in\Theta$ , the agent chooses the action $\mathbf{x}^*(\theta)$ that maximizes her joint-profits with this principal;

Conditions $(4a)$ and $(4b)$ are Lemma 11.2.4 below. The later evidences the fact that, given the cost functions $f_i^S(\theta, \mathbf{x})$, the set of accepted contracts $A^o$ maximizes the joint-profits of the agent and any principal $i$ in the contracting set $A^o$. Moreover, for any principal $i \in N$, contracting takes place if and only if, it improves on the sum of his expected equilibrium payoff and the agent's expected equilibrium payoff.

### 11.2.3 Lemma 11.2.3 (Expected marginal contribution of a single principal to the equilibrium profits of the agent):

*In any Nash equilibrium, each contracting principal $i$ in $A^o$ extract all the expected rent of the agent given the other principals equilibrium strategies $\mathbf{t}_{-i}^o$. In other words, the strategy $t_i^o$ is such that the agent's expected equilibrium profits remain the same if $i$ does not contract. Formally,*

$$E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}^o, \mathbf{x})\right] = \sup\left\{0, \sup_{S\subseteq N\backslash\{i\}} E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^S\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right]\right\}; \tag{49}$$

*for all principals $i$ in $A^o$.*

**Proof:** By assumption, all contractual relationships are beneficial to the agent. Formally, for any $i \in A^o$,

$$E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}^o, \mathbf{x})\right] \geq \sup\left\{0, \sup_{S\subseteq N\backslash\{i\}} E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^S\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right]\right\}. \tag{50}$$

Assume that there is some $i \in A^o$ for which there is a strict inequality and look for a contradiction. Denote $\widehat{\varepsilon}_i$ the difference between the agent's expected net gains and what she can obtains if she does not contract with principal $i$ :

$$\widehat{\varepsilon}_i = E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}^o, \mathbf{x})\right] - \sup\left\{0, \sup_{S\subseteq N\backslash\{i\}} E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^S\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right]\right\} > 0. \tag{51}$$

Consider the strategy

$$\widehat{t}_i(\mathbf{x}) = t_i^o(\mathbf{x}) - \widehat{\varepsilon}_i/2. \tag{52}$$

If such a strategy is offered by principal $i$, the agent still find it profitable to contract with him. Indeed, by keeping unchanged the contracting set $A^o$ and the equilibrium actions $\mathbf{x}^o(\theta)$, her profits are already strictly higher than the profits she would get by contracting with any set $S \subseteq N\backslash\{i\}$. To see that, observe that by definition

$$E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta, \widehat{t}_i, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right)\right] = E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}^o, \mathbf{x})\right] - \widehat{\varepsilon}_i/2 \tag{53}$$

so that from (51)

$$E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta, \widehat{t}_i, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right)\right] = \sup\left\{0, \sup_{S\subseteq N\backslash\{i\}} E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{S}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right]\right\} + \widehat{\varepsilon}_i - \widehat{\varepsilon}_i/2. \tag{54}$$

Thus, given the equilibrium strategies of the other principals $\mathbf{t}_{-i}^o$, the agent is strictly better off by contracting with principal $i$ even when the later offers the modified strategy $\widehat{t}_i(\mathbf{x})$ :

$$E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o}\left(\theta, \widehat{t}_i, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right] > \sup\left\{0, \sup_{S\subseteq N\backslash\{i\}} E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{S}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right]\right\}. \tag{55}$$

It is easy to show that the agent cannot improve upon her expected payoff by contracting with a set $A \subset (\{i\} \cup N\backslash\{i\})$ different from $A^o$. It is also obvious that, for any realization of the parameter $\theta$, the agent cannot improve upon her realized payoff by choosing an action $\mathbf{x} \in X$ which is not in $X^o(\theta)$. Indeed, for both decision sets $A^o$ and $\{\mathbf{x}^o(\theta)\}$, a possible improvement would mean that the triplet $(\mathbf{t}^o, A^o, \{\mathbf{x}^o(\theta)\})$ is not a Nash equilibrium (The agent would not be profit-maximising when offered $\mathbf{t}^o$). It follows that, by offering $\widehat{t}_i(\mathbf{x})$ rather than $t_i^o(\mathbf{x})$, principal $i$ does not modify the decisions of the agent.

Observe that, if the agent does not modify its decisions, principal $i$ is realising a strictly higher payoff with strategy $\widehat{t}_i(\mathbf{x})$ than with $t_i^o(\mathbf{x})$, for any realization of the parameter $\theta$:

$$\begin{aligned} v_i\left(\theta, \widehat{t}_i, \mathbf{t}_{-i}^o\right) &= g_i(\mathbf{x}^o(\theta)) - \widehat{t}_i(\mathbf{x}^o(\theta)) \\ &= g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta)) + \widehat{\varepsilon}_i/2 \\ &> v_i\left(\theta, t_i^o, \mathbf{t}_{-i}^o\right). \end{aligned} \tag{56}$$

As a result, principal $i$ expected payoff is strictly higher when he offers $\widehat{t}_i(\mathbf{x})$ rather than $t_i^o(\mathbf{x})$. This contradicts the fact that the triplet $(\mathbf{t}^o, A^o, \{\mathbf{x}^o(\theta)\})$ is a Nash equilibrium of the game.□

### 11.2.4 Lemma 11.2.4 (The contracting set maximizes the sum of the expected profits of the agent and any principal $i$):

*In any Nash equilibrium $(\mathbf{t}^o, A^o, \{\mathbf{x}^o(\theta)\})$, the equilibrium contracting set $A^o$ maximizes the expected joint-profits of the agent and any principal $i$ in $A^o$. Formally, for all principals $i$ in $A^o$:*

$$\begin{aligned} & E_\theta\left[g_i(\mathbf{x}^o(\theta)) - f_i^{A^o}(\theta, \mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right)\right] \\ \geq\ & \sup\left\{0, E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x}) - f_i^{S\cup\{i\}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^S\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right)\right]\right\} \quad \textit{all } S \subseteq N\setminus\{i\}; \end{aligned}$$

*Furthermore, contracting occurs if and only if, given the incremental costs $f_i^S(\theta, \mathbf{x})$, it improves over the sum of their equilibrium payoffs. Formally, for all principals $i$ in $A^o$:*

$$\begin{aligned} & E_\theta\left[\left(g_i(\mathbf{x}^o(\theta)) - f_i^{A^o}(\theta, \mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right)\right)\right] \\ \geq\ & \sup\left\{0, E_\theta\left[g_i(\mathbf{x}_S^o(\theta)) + v_{\mathcal{L}}^S\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}_S^o(\theta)\right)\right]\right\} \quad \textit{all } S \subseteq N\setminus\{i\}; \end{aligned}$$

with $\mathbf{x}_S^o(\theta) = \arg\max_{\mathbf{x}\in X} v_{\mathcal{L}}^S(\theta, \mathbf{t}_{-i}^o, \mathbf{x})$. *For all principals $i$ in $N\setminus A^o$:*

$$\begin{aligned} & E_\theta\left[\left(g_i(\mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right)\right)\right] \\ \geq\ & \sup\left\{0, E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x}) - f_i^{S\cup\{i\}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^S\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right)\right]\right\} \quad \textit{all } S \subseteq N\setminus\{i\}. \end{aligned}$$

**Proof:** Assume this is not true, *i.e.* there exists either (Case A) a principal $i \in A^o$ and a contracting set $\widehat{A} \ni i$ such that

$$\begin{aligned} & E_\theta\left[\left(g_i(\mathbf{x}^o(\theta)) - f_i^{A^o}(\theta, \mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right)\right)\right] \\ <\ & \sup\left\{0, E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x}) - f_i^{\widehat{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A}\setminus\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right)\right]\right\}, \end{aligned} \tag{57}$$

or (Case B) a principal $i \in A^o$ and a contracting set $\widehat{A} \subseteq N \backslash \{i\}$ such that

$$\begin{aligned} & E_\theta \left[ \left( g_i \left( \mathbf{x}^o (\theta) \right) - f_i^{A^o} \left( \theta, \mathbf{x}^o (\theta) \right) + v_{\mathcal{L}}^{A^o \backslash \{i\}} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x}^o (\theta) \right) \right) \right] \\ < \ & \sup \left\{ 0, E_\theta \left[ g_i \left( \widehat{\mathbf{x}} (\theta) \right) + v_{\mathcal{L}}^{\widehat{A}} \left( \theta, \mathbf{t}_{-i}^o, \widehat{\mathbf{x}} (\theta) \right) \right] \right\}, \end{aligned} \tag{58}$$

with $\widehat{\mathbf{x}}(\theta) = \arg\max_{\mathbf{x} \in X} v_{\mathcal{L}}^{\widehat{A}}(\theta, \mathbf{t}_{-i}^o, \mathbf{x})$ or (Case C) a principal $i \notin A^o$ and a contracting set $\widehat{A} \ni i$ such that

$$\begin{aligned} & E_\theta \left[ \left( g_i \left( \mathbf{x}^o (\theta) \right) + v_{\mathcal{L}}^{A^o} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x}^o (\theta) \right) \right) \right] \\ < \ & \sup \left\{ 0, E_\theta \left[ \max_{\mathbf{x} \in X} \left( g_i (\mathbf{x}) - f_i^{\widehat{A}} (\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A} \backslash \{i\}} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right) \right] \right\}. \end{aligned} \tag{59}$$

We examine all three cases and look for a contradiction.

Without any loss of generality, we assume in the remaining of the demonstration that, whenever $i \in \widehat{A}$ (*i.e.* for Cases A and C) the set $\widehat{A}$ is the set $S \subseteq N$ for which the expected joint-profits

$$E_\theta \left[ \max_{\mathbf{x} \in X} \left( g_i (\mathbf{x}) - f_i^S (\theta, \mathbf{x}) + v_{\mathcal{L}}^{S \backslash \{i\}} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right) \right]$$

takes the highest value; while, whenever $i \in \widehat{A}$ (*i.e.* for Case B), the set $\widehat{A}$ is the set $S \subseteq N \backslash \{i\}$ for which the agents profits

$$E_\theta \left[ \max_{\mathbf{x} \in X} \left( v_{\mathcal{L}}^S \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right) \right]$$

takes the highest value.

**Case A** Principal $i \in A^o \cap \widehat{A}$ :

If $i \in A^o$, it cannot be the case that

$$E_\theta \left[ g_i \left( \mathbf{x}^o (\theta) \right) - f_i^{A^o} \left( \theta, \mathbf{x}^o (\theta) \right) + v_{\mathcal{L}}^{A^o \backslash \{i\}} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x}^o (\theta) \right) \right] < 0. \tag{60}$$

Indeed, for contracting to be profitable for principal $i$, it must be the case that

$$v_i = E_\theta \left[ g_i \left( \mathbf{x}^o (\theta) \right) - t_i^o \left( \mathbf{x}^o (\theta) \right) \right] \geq 0. \tag{61}$$

Substracting the later equation to equation (60) gives

$$\begin{aligned}&E_\theta\left[t_i^o\left(\mathbf{x}^o\left(\theta\right)\right)-f_i^{A^o}\left(\theta,\mathbf{x}^o\left(\theta\right)\right)+v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o\left(\theta\right)\right)\right]\\ \equiv\;&E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}^o,\mathbf{x}^o\left(\theta\right)\right)\right]<0,\end{aligned}\tag{62}$$

which means that agent would make negative profits at equilibrium. We therefore know from equation (57) that

$$E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i\left(\mathbf{x}\right)-f_i^{\widehat{A}}\left(\theta,\mathbf{x}\right)+v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right)\right]>0\tag{63}$$

so that the equation (57) itself rewrites

$$\begin{aligned}&E_\theta\left[\left(g_i\left(\mathbf{x}^o\left(\theta\right)\right)-f_i^{A^o}\left(\theta,\mathbf{x}^o\left(\theta\right)\right)+v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o\left(\theta\right)\right)\right)\right]\\ <\;&E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i\left(\mathbf{x}\right)-f_i^{\widehat{A}}\left(\theta,\mathbf{x}\right)+v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right)\right].\end{aligned}\tag{64}$$

Let $\widehat{\varepsilon_i}>0$ be the expected difference between the pair of joint-profits defined as follows:

$$\begin{aligned}\widehat{\varepsilon_i}\;=\;&E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i\left(\mathbf{x}\right)-f_i^{\widehat{A}}\left(\theta,\mathbf{x}\right)+v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right)\right]\\ &-E_\theta\left[g_i\left(\mathbf{x}^o\left(\theta\right)\right)-f_i^{A^o}\left(\theta,\mathbf{x}^o\left(\theta\right)\right)+v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o\left(\theta\right)\right)\right].\end{aligned}\tag{65}$$

Let $k_i$ be the expected difference between maximum joint-profit of the agent and principal $i$ and the equilibrium profits of the agent:

$$\begin{aligned}k_i\;=\;&E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i\left(\mathbf{x}\right)-f_i^{\widehat{A}}\left(\theta,\mathbf{x}\right)+v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right)\right]\\ &-E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}^o,\mathbf{x}^o\left(\theta\right)\right)\right].\end{aligned}\tag{66}$$

Consider the strategy $\widehat{t_i}\left(\mathbf{x}\right)$ defined as follows:

$$\widehat{t_i}\left(\mathbf{x}\right)\equiv g_i\left(\mathbf{x}\right)-k_i+\widehat{\varepsilon_i}/2.\tag{67}$$

If the agent contracts with the set $\widehat{A}$ when offered $\left(\widehat{t_i},\mathbf{t}_{-i}^o\right)$, its expected payoff writes

$$\begin{aligned}&E_\theta\left[\max_{\mathbf{x}\in X}v_{\mathcal{L}}^{\widehat{A}}\left(\theta,\widehat{t_i},\mathbf{t}_{-i}^o,\mathbf{x}\right)\right]\\ \equiv\;&E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i\left(\mathbf{x}\right)-f_i^{\widehat{A}}\left(\theta,\mathbf{x}\right)+v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right)\right]-k_i+\widehat{\varepsilon_i}/2\\ =\;&E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}^o,\mathbf{x}^o\left(\theta\right)\right)\right]+\widehat{\varepsilon_i}/2.\end{aligned}\tag{68}$$

We thus know that, if the agent participation constraint is satisfied when the agent is offered $\mathbf{t}^o$, it will still be the case when she is offered $\left(\widehat{t}_i, \mathbf{t}^o_{-i}\right)$. However, there is *a priori* no reason for the agent to contract with the set $\widehat{A}$ when offered $\left(\widehat{t}_i, \mathbf{t}^o_{-i}\right)$. It may well exist a set that would yield a higher profit. when the agent is offered $\left(\widehat{t}_i, \mathbf{t}^o_{-i}\right)$. We now show it cannot be the case.

Assume $\widehat{A}$ is not a contracting set when the agent is offered $\left(\widehat{t}_i, \mathbf{t}^o_{-i}\right)$ and look for a contradiction. Let $\widetilde{A}$ be the contracting set when the agent is offered $\left(\widehat{t}_i, \mathbf{t}^o_{-i}\right)$. The very fact that $\widehat{A}$ is not a contracting set means that

$$E_\theta \left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widehat{A}} \left(\theta, \widehat{t}_i, \mathbf{t}^o_{-i}, \mathbf{x}\right)\right] < E_\theta \left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widetilde{A}} \left(\theta, \widehat{t}_i, \mathbf{t}^o_{-i}, \mathbf{x}\right)\right]. \tag{69}$$

If $i \notin \widetilde{A}$, the expected profits can be rewritten as

$$E_\theta \left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widetilde{A}} \left(\theta, \widehat{t}_i, \mathbf{t}^o_{-i}, \mathbf{x}\right)\right] \equiv E_\theta \left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widetilde{A}} \left(\theta, \mathbf{t}^o_{-i}, \mathbf{x}\right)\right]. \tag{70}$$

From Lemma 11.2.3, we know that the agent's expected equilibrium profits when offered $\mathbf{t}^o$ are at least as high as the expected profits when the contracting set is $\widetilde{A} \subseteq N\backslash\{i\}$. Formally,

$$E_\theta \left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widetilde{A}} \left(\theta, \mathbf{t}^o_{-i}, \mathbf{x}\right)\right] \leq E_\theta \left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o} \left(\theta, \mathbf{t}^o, \mathbf{x}\right)\right] \tag{71}$$

Combining inequalities (69) and (71) with the equality (70) leads to

$$E_\theta \left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widehat{A}} \left(\theta, \widehat{t}_i, \mathbf{t}^o_{-i}, \mathbf{x}\right)\right] < E_\theta \left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o} \left(\theta, \mathbf{t}^o, \mathbf{x}\right)\right], \tag{72}$$

which clearly contradicts (68) that states that, when the agent is offered $\left(\widehat{t}_i, \mathbf{t}^o_{-i}\right)$ she can guarantee herself a strictly higher payoff than $E_\theta \left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o} \left(\theta, \mathbf{t}^o, \mathbf{x}\right)\right]$. It follows that, if $\widehat{A} \ni i$ is not a contracting set when the agent is offered $\left(\widehat{t}_i, \mathbf{t}^o_{-i}\right)$, the later must also contain principal $i$.

If $i \in \widetilde{A}$ then, by the very definition (67) of the strategy $\widehat{t}_i(\mathbf{x})$,

$$\begin{aligned}
& E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widetilde{A}}\left(\theta, \widehat{t}_i, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right] \\
= \;& E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x}) - f_i^{\widetilde{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widetilde{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right)\right] - k_i + \widehat{\varepsilon}_i/2. \qquad (73)
\end{aligned}$$

From the definition of $\widehat{A}$, we also know that

$$\begin{aligned}
& E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x}) - f_i^{\widetilde{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widetilde{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right)\right] \\
\leq \;& E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x}) - f_i^{\widehat{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right)\right]. \qquad (74)
\end{aligned}$$

Combining the later inequality with the two expected payoffs defined in (68) and (73), one gets

$$E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widetilde{A}}\left(\theta, \widehat{t}_i, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right] \leq E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widehat{A}}\left(\theta, \widehat{t}_i, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right]$$

which contradicts the fact that $\widehat{A}$ is not a contracting set when the agent is offered $\left(\widehat{t}_i, \mathbf{t}_{-i}^o\right)$. Thus we know that, $\widehat{A}$ is a contracting set when the agent is offered $\left(\widehat{t}_i, \mathbf{t}_{-i}^o\right)$. From the definition (67) of the strategy $\widehat{t}_i(\mathbf{x})$, we know that principal $i$ payoff writes

$$\begin{aligned}
v_i \;=\;& k_i - \widehat{\varepsilon}_i/2 \\
=\;& E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x}) - f_i^{\widehat{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right)\right] - E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}^o, \mathbf{x}^o(\theta)\right)\right] - \widehat{\varepsilon}_i/2 \\
=\;& E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x}) - f_i^{\widehat{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right)\right] + E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right] \\
& - E_\theta\left[g_i(\mathbf{x}^o(\theta)) - f_i^{A^o}(\theta, \mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right)\right] - \widehat{\varepsilon}_i/2 \\
=\;& E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right] + \widehat{\varepsilon}_i/2. \qquad (75)
\end{aligned}$$

It follows that principal $i$ would get a strictly higher payoff by offering $\widehat{t}_i$ rather $t_i^o$, a clear contradiction with assumption that $(\mathbf{t}^o, A^o, \{\mathbf{x}^o(\theta)\})$ is a Nash equilibrium.

To sum up, Case A is excluded: if principal $i \in A^o \cap \widehat{A}$, equation (57) cannot hold true.

**Case B** Principal $i$ in $A^o$ but not in $\widehat{A}$:

As already shown in case A, since $i \in A^o$, we know that

$$E_\theta\left[g_i(\mathbf{x}^o(\theta)) - f_i^{A^o}(\theta, \mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right)\right] \geq 0.$$

Thus equation (58) rewrites

$$\begin{aligned} & E_\theta\left[g_i\left(\mathbf{x}^o(\theta)\right)-f_i^{A^o}\left(\theta,\mathbf{x}^o(\theta)\right)+v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o(\theta)\right)\right] \\ < \quad & E_\theta\left[g_i\left(\widehat{\mathbf{x}}(\theta)\right)+v_{\mathcal{L}}^{\widehat{A}}\left(\theta,\mathbf{t}_{-i}^o,\widehat{\mathbf{x}}(\theta)\right)\right]. \end{aligned} \tag{76}$$

where $\widehat{\mathbf{x}}(\theta)\in\arg\max_{\mathbf{x}}\left[v_{\mathcal{L}}^{\widehat{A}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right]$.

From the very definition of a Nash-Equilibrium, principal $i$ strategy $t_i^o(\mathbf{x})$ is a best-response to other principals' strategies. In particular the net equilibrium payoff $v_i^o = E_\theta\left[g_i\left(\mathbf{x}^o(\theta)\right)-t_i^o\left(\mathbf{x}^o(\theta)\right)\right]$ is greater than the one that would be obtained by offering the null strategy $\widehat{t_i}(\mathbf{x})\equiv 0$, that would induce the agent to contract with $\widehat{A}$ and adopt the actions $\widehat{\mathbf{x}}(\theta)$. Formally

$$E_\theta\left[g_i\left(\widehat{\mathbf{x}}(\theta)\right)\right]\leq E_\theta\left[g_i\left(\mathbf{x}^o(\theta)\right)-t_i^o\left(\mathbf{x}^o(\theta)\right)\right]. \tag{77}$$

Adding the later inequality to (76) gives

$$\begin{aligned} & E_\theta\left[t_i^o\left(\mathbf{x}^o(\theta)\right)-f_i^{A^o}\left(\theta,\mathbf{x}^o(\theta)\right)+v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o(\theta)\right)\right] \\ < \quad & E_\theta\left[v_{\mathcal{L}}^{\widehat{A}}\left(\theta,\mathbf{t}_{-i}^o,\widehat{\mathbf{x}}(\theta)\right)\right] \end{aligned}$$

which contradicts the fact that contracting set $A^o$ is the equilibrium set.

To sum up, Case B is excluded: if principal $i$ is in $A^o$ but not in $\widehat{A}$, equation (58) cannot hold true.

**Case C** Principal $i$ is not $A^o$ but is in $\widehat{A}$ :

We assume that there is no principal $i\in N$ for which participation to the R&D market is *a priori* excluded. Thus, there exists at least one set $S\ni i$ such that:

$$E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x})-f_i^S(\theta,\mathbf{x})+v_{\mathcal{L}}^{S\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right)\right]\geq 0.$$

It follows that (59) may be rewritten as

$$\begin{aligned} & E_\theta\left[\left(g_i\left(\mathbf{x}^o(\theta)\right)+v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o(\theta)\right)\right)\right] \\ < \quad & E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x})-f_i^{\widehat{A}}(\theta,\mathbf{x})+v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right)\right]. \end{aligned} \tag{78}$$

Define again $\widehat{\varepsilon_i} > 0$ as the expected difference

$$\begin{aligned} \widehat{\varepsilon_i} &= E_\theta \left[ \max_{\mathbf{x}\in X} \left( g_i(\mathbf{x}) - f_i^{\widehat{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right)\right] \\ &\quad - E_\theta \left[ g_i(\mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right)\right], \end{aligned} \tag{79}$$

and consider the strategy $\widehat{t_i}(\mathbf{x})$ defined as follows:

$$\widehat{t_i}(\mathbf{x}) \equiv g_i(\mathbf{x}) - k_i + \widehat{\varepsilon_i}/2, \tag{80}$$

where $k_i$ is the difference between maximum expected joint-profit of principal $i$ and the agent and the expected equilibrium profits of the agent:

$$\begin{aligned} k_i &= E_\theta \left[ \max_{\mathbf{x}\in X} \left( g_i(\mathbf{x}) - f_i^{\widehat{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right)\right] \\ &\quad - E_\theta \left[ v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}^o, \mathbf{x}^o(\theta)\right)\right]. \end{aligned} \tag{81}$$

If the agent contracts with the set $\widehat{A}$ when offered $\left(\widehat{t_i}, \mathbf{t}_{-i}^o\right)$, its expected payoff writes

$$\begin{aligned} & E_\theta \left[ \max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widehat{A}}\left(\theta, \widehat{t_i}, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right] \\ \equiv\ & E_\theta \left[ \max_{\mathbf{x}\in X} \left( g_i(\mathbf{x}) - f_i^{\widehat{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right)\right] - k_i + \widehat{\varepsilon_i}/2 \\ =\ & E_\theta \left[ v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}^o, \mathbf{x}^o(\theta)\right)\right] + \widehat{\varepsilon_i}/2. \end{aligned} \tag{82}$$

We thus know that, if the agent find profitable to contract with $A^o$ when she is offered $\mathbf{t}^o$, her participation constraint will still hold true when she is offered $\left(\widehat{t_i}, \mathbf{t}_{-i}^o\right)$. However, there is *a priori* no reason for the agent to contract with the set $\widehat{A}$ when offered $\left(\widehat{t_i}, \mathbf{t}_{-i}^o\right)$. It may well exist a set that would yield a higher profit. when the agent is offered $\left(\widehat{t_i}, \mathbf{t}_{-i}^o\right)$. We now show it cannot be the case.

Assume $\widehat{A}$ is not a contracting set when the agent is offered $\left(\widehat{t_i}, \mathbf{t}_{-i}^o\right)$ and look for a contradiction. Let $\widetilde{A}$ be the contracting set when the agent is offered $\left(\widehat{t_i}, \mathbf{t}_{-i}^o\right)$. The very fact that $\widehat{A}$ is not a contracting set means that

$$E_\theta \left[ \max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widehat{A}}\left(\theta, \widehat{t_i}, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right] < E_\theta \left[ \max_{\mathbf{x}\in X} v_{\mathcal{L}}^{\widetilde{A}}\left(\theta, \widehat{t_i}, \mathbf{t}_{-i}^o, \mathbf{x}\right)\right]. \tag{83}$$

If $i \notin \widetilde{A}$, the expected profits can be rewritten as

$$E_\theta \left[ \max_{\mathbf{x} \in X} v_{\mathcal{L}}^{\widetilde{A}} \left( \theta, \widehat{t}_i, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right] \equiv E_\theta \left[ \max_{\mathbf{x} \in X} v_{\mathcal{L}}^{\widetilde{A}} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right]. \tag{84}$$

And by definition of the equilibrium contracting set $A^o$, we know (this is actually Lemma 11.2.3) that

$$E_\theta \left[ \max_{\mathbf{x} \in X} v_{\mathcal{L}}^{\widetilde{A}} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right] \leq E_\theta \left[ \max_{\mathbf{x} \in X} v_{\mathcal{L}}^{A^o} \left( \theta, \mathbf{t}^o, \mathbf{x} \right) \right], \tag{85}$$

since contracting with $\widetilde{A}$ was already possible when the agent was offered $\mathbf{t}^o$. Combining inequalities (83) and (85) with the equality (84) leads to

$$E_\theta \left[ \max_{\mathbf{x} \in X} v_{\mathcal{L}}^{\widehat{A}} \left( \theta, \widehat{t}_i, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right] < E_\theta \left[ \max_{\mathbf{x} \in X} v_{\mathcal{L}}^{A^o} \left( \theta, \mathbf{t}^o, \mathbf{x} \right) \right], \tag{86}$$

which clearly contradicts (82) that states that, when the agent is offered $\left( \widehat{t}_i, \mathbf{t}_{-i}^o \right)$ she can guarantee herself a strictly higher payoff than $E_\theta \left[ \max_{\mathbf{x} \in X} v_{\mathcal{L}}^{A^o} \left( \theta, \mathbf{t}^o, \mathbf{x} \right) \right]$. It follows, if $\widehat{A} \ni i$ is not the contracting set, $\widetilde{A}$ must also contain principal $i$.

If $i \in \widetilde{A}$ then, by the very definition (80) of the strategy $\widehat{t}_i(\mathbf{x})$,

$$\begin{aligned} & E_\theta \left[ \max_{\mathbf{x} \in X} v_{\mathcal{L}}^{\widetilde{A}} \left( \theta, \widehat{t}_i, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right] \\ = \; & E_\theta \left[ \max_{\mathbf{x} \in X} \left( g_i(\mathbf{x}) - f_i^{\widetilde{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widetilde{A} \setminus \{i\}} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right) \right] - k_i + \widehat{\varepsilon}_i / 2. \end{aligned} \tag{87}$$

From the definition of $\widehat{A}$, we also know that

$$\begin{aligned} & E_\theta \left[ \max_{\mathbf{x} \in X} \left( g_i(\mathbf{x}) - f_i^{\widetilde{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widetilde{A} \setminus \{i\}} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right) \right] \\ \leq \; & E_\theta \left[ \max_{\mathbf{x} \in X} \left( g_i(\mathbf{x}) - f_i^{\widehat{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A} \setminus \{i\}} \left( \theta, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right) \right]. \end{aligned} \tag{88}$$

Combining the later inequality with the two expected payoffs defined in (82) and (87), one gets

$$E_\theta \left[ \max_{\mathbf{x} \in X} v_{\mathcal{L}}^{\widetilde{A}} \left( \theta, \widehat{t}_i, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right] \leq E_\theta \left[ \max_{\mathbf{x} \in X} v_{\mathcal{L}}^{\widehat{A}} \left( \theta, \widehat{t}_i, \mathbf{t}_{-i}^o, \mathbf{x} \right) \right]$$

which contradicts the fact that $\widehat{A}$ is not a contracting set when the agent is offered $\left( \widehat{t}_i, \mathbf{t}_{-i}^o \right)$. Thus we know that, $\widehat{A}$ is a contracting set when the agent is offered $\left( \widehat{t}_i, \mathbf{t}_{-i}^o \right)$. From the definition (80) of the strategy $\widehat{t_i}(\mathbf{x})$, we know that, if the agent contracts with $\widehat{A}$, principal $i$ payoff writes

$$\begin{aligned}
v_i &= k_i - \widehat{\varepsilon}_i/2 \\
&= E_\theta \left[ \max_{\mathbf{x}\in X} \left( g_i(\mathbf{x}) - f_i^{\widehat{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right) \right) \right] - E_\theta\left[ v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}^o, \mathbf{x}^o(\theta)) \right] - \widehat{\varepsilon}_i/2 \\
&= E_\theta \left[ \max_{\mathbf{x}\in X} \left( g_i(\mathbf{x}) - f_i^{\widehat{A}}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{\widehat{A}\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right) \right) \right] + E_\theta\left[ g_i(\mathbf{x}^o(\theta)) \right] \\
&\quad - E_\theta\left[ g_i(\mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right) \right] - \widehat{\varepsilon}_i/2 \\
&= E_\theta\left[ g_i(\mathbf{x}^o(\theta)) \right] + \widehat{\varepsilon}_i/2. \qquad (89)
\end{aligned}$$

It follows that principal $i$ would get a strictly higher payoff by offering $\widehat{t_i}$ rather $t_i^o$, a clear contradiction with assumption that $(\mathbf{t}^o, A^o, \{\mathbf{x}^o(\theta)\})$ is a Nash equilibrium.

To sum up, Case C is excluded: if principal $i$ is not $A^o$ but is in $\widehat{A}$, equation (59).

All three cases A, B and C are excluded. □

### 11.2.5 Lemma 11.2.5 (Maximization of the joint profits of the agent and principal $i$ at equilibrium):

*In any Nash equilibrium $(\mathbf{t}^o, A^o, \{\mathbf{x}^o(\theta)\})$, the action equilibrium $\mathbf{x}^o(\theta)$ chosen by the agent is such that the joint profit of the agent $\mathcal{L}$ and any principal i in the contracting set $A^o$ is maximized for almost all $\theta$. Formally,*

$$g_i(\mathbf{x}^o(\theta)) - f_i^{A^o}(\theta, \mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}^o(\theta)\right) \geq g_i(\mathbf{x}) - f_i^{A^o}(\theta, \mathbf{x}) + v_{\mathcal{L}}^{A^o\backslash\{i\}}\left(\theta, \mathbf{t}_{-i}^o, \mathbf{x}\right) \qquad (90)$$

*for almost all $\theta$, for all $i \in N$, and for all feasible action $\mathbf{x} \in X$.*

**Proof:** This follows almost straightforwardly from Lemma 11.2.4. We know indeed that, for all $i \in A^o$

$$\begin{aligned} & E_\theta\left[g_i\left(\mathbf{x}^o(\theta)\right)-f_i^{A^o}\left(\theta,\mathbf{x}^o(\theta)\right)+v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o(\theta)\right)\right] \\ \geq\ & E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x})-f_i^{S\cup\{i\}}(\theta,\mathbf{x})+v_{\mathcal{L}}^S\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right)\right] \quad \textit{all } S\subseteq N\setminus\{i\}\,; \end{aligned} \tag{91}$$

This holds true in particular for $S \equiv A^o \setminus \{i\}$, so that

$$\begin{aligned} & E_\theta\left[g_i\left(\mathbf{x}^o(\theta)\right)-f_i^{A^o}\left(\theta,\mathbf{x}^o(\theta)\right)+v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o(\theta)\right)\right] \\ \geq\ & E_\theta\left[\max_{\mathbf{x}\in X}\left(g_i(\mathbf{x})-f_i^{A^o}(\theta,\mathbf{x})+v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right)\right]. \end{aligned} \tag{92}$$

However, by definition of the maximum,

$$\begin{aligned} & g_i\left(\mathbf{x}^o(\theta)\right)-f_i^{A^o}\left(\theta,\mathbf{x}^o(\theta)\right)+v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o(\theta)\right) \\ \leq\ & \max_{\mathbf{x}\in X}\left(g_i(\mathbf{x})-f_i^{A^o}(\theta,\mathbf{x})+v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right) \end{aligned} \tag{93}$$

so that for (92) to hold true, it must be the case that

$$\begin{aligned} & g_i\left(\mathbf{x}^o(\theta)\right)-f_i^{A^o}\left(\theta,\mathbf{x}^o(\theta)\right)+v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o(\theta)\right) \\ =\ & \max_{\mathbf{x}\in X}\left(g_i(\mathbf{x})-f_i^{A^o}(\theta,\mathbf{x})+v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right), \end{aligned}$$

almost all $\theta$. □

## 11.3 Truthfull Nash Equilibria Characterisation3

### 11.3.1 Remark 11.3.1 (Constant payoffs for the contracting principals at equilibrium):

*In all TNE, for all principal $i$ in the contracting set $A^o$*

$$g_i(\mathbf{x}^o(\theta)) - E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right] \geq 0, \tag{94}$$

*for almost all $\theta$. As a result, contracting principals' payoffs do not depend on $\theta$, that is:*

$$v_i^o = g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta)) = E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right], \tag{95}$$

*for all $i \in A^o$, almost all $\theta$.*

**Proof:** Assume equation (94) does not hold, *i.e.* that there exists a principal $i \in A^o$ and a set $\underline{\Theta} \subset \Theta$ of strictly positive measure such that $g_i(\mathbf{x}^o(\theta)) < E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right]$. Clearly the expected transfer can be decomposed over $\underline{\Theta}$ and its complementary set $\Theta \backslash \underline{\Theta}$.

From the definition of truthfulness, for any $\theta \in \underline{\Theta}$ principal $i$'s equilibrium transfer is zero, that is:

$$t_i^o((\mathbf{x}^o(\theta)) = 0 > g_i(\mathbf{x}^o(\theta)) - E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right].$$

Hence, the net payoff of the principal is strictly smaller than the expected payoff over the set $\underline{\Theta}$.

From the definition of truthfulness, for any $\theta \in \Theta \backslash \underline{\Theta}$ principal $i$'s equilibrium transfer is:

$$t_i^o((\mathbf{x}^o(\theta)) = g_i(\mathbf{x}^o(\theta)) - E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right].$$

Hence, the net payoff of the principal is exactly equal to the expected payoff over the set $\Theta \backslash \underline{\Theta}$.

It follows that:

$$\begin{aligned}
E_\theta\left[v_i(\mathbf{x}^o(\theta))\right] &= E_{\theta \in \underline{\Theta}}\left[g_i(\mathbf{x}^o(\theta)) - 0\right] + E_{\theta \in \Theta \backslash \underline{\Theta}}\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right] \\
&< E_{\theta \in \underline{\Theta}}\left[g_i(\mathbf{x}^o(\theta)) - g_i(\mathbf{x}^o(\theta)) + E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right]\right] \\
&\quad + E_{\theta \in \Theta \backslash \underline{\Theta}}\left[g_i(\mathbf{x}^o(\theta)) - g_i(\mathbf{x}^o(\theta)) + E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right]\right] \\
&< E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right],
\end{aligned}$$

a contradiction.

Now, equation (95) is a simple restatement of (94) obtained by applying the definition of truthfulness, that is:

$$t_i^o(\mathbf{x}^o(\theta)) = \sup\{0, g_i(\mathbf{x}^o(\theta)) - E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right]\} = g_i(\mathbf{x}^o(\theta)) - E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right],$$

for almost all $\theta$. Hence:

$$v_i(\mathbf{x}^o(\theta)) = g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta)) = E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right],$$

for *all* $i \in A^o$, almost all $\theta$. $\square$

### 11.3.2 Proposition 11.3.2 (Upper-bound on equilibrium payoffs, or $\mathbf{v}^o \in V_\Gamma^* \equiv V_\Gamma(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta})$):

*The equilibrium action $\mathbf{x}^o(\theta) \in X^o(\theta)$ of the common agency game gives rise to equilibrium payoffs $\mathbf{v}^o = (v_1^o, ..., v_n^o)$ in $V_\Gamma^* \equiv V_\Gamma(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta})$, that is:*

$$\sum_{k\in S} v_k^o \leq E_\theta\left[W(\theta, A^o) - W(\theta, A^o \backslash S)\right], \text{ all } S \subseteq A^o. \tag{96}$$

**Proof:** Recall that $\mathbf{x}^o(\theta) \in X^o(\theta)$ denotes the equilibrium action chosen by the agent, when the realization of the stochastic variable is $\theta$. Then for all $S \subseteq A^o$, define:

$$X_S^*(\theta) = \arg\max_{\mathbf{x}\in X}\left(\sum_{i\in S} g_i(\mathbf{x}) - r(S, \theta, \mathbf{x})\right). \tag{97}$$

Let $\mathbf{x}_{A^o\backslash S}^*(\theta)$ an element of $X_{A^o\backslash S}^*(\theta)$. By definition of the agent's equilibrium choice $\mathbf{x}^o(\theta)$, we know that:

$$\begin{aligned} E_\theta\left[\sum_{i\in A^o\backslash S} t_i^o(\mathbf{x}_{A^o\backslash S}^*(\theta)) - r(A^o\backslash S, \theta, \mathbf{x}_{A^o\backslash S}^*(\theta))\right] &\leq E_\theta\left[\max_{\mathbf{x}\in\mathbf{X}}\left(\sum_{i\in A^o\backslash S} t_i^o(\mathbf{x}) - r(A^o\backslash S, \theta, \mathbf{x})\right)\right] \\ &\leq E_\theta\left[\sum_{i\in A^o} t_i^o(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^o(\theta))\right]. \end{aligned} \tag{98}$$

Then $S \cup A^o \backslash S = A^o$ and $A^o \cap A^o \backslash S = \emptyset$ imply that:

$$E_\theta\left[\sum_{i \in A^o} t_i^o(\mathbf{x}^o(\theta))\right] = E_\theta\left[\sum_{k \in S} t_k^o(\mathbf{x}^o(\theta))\right] + E_\theta\left[\sum_{l \in A^o \backslash S} t_l^o(\mathbf{x}^o(\theta))\right], \tag{99}$$

which yields:

$$\begin{aligned} E_\theta\left[\sum_{l \in A^o N \backslash S} t_l^o(\mathbf{x}^*_{A^o \backslash S}(\theta)) - r(A^o \backslash S, \theta, \mathbf{x}^*_{A^o \backslash S}(\theta))\right] \quad &\leq \quad E_\theta\left[\sum_{k \in S} t_k^o(\mathbf{x}^o(\theta))\right] - E_\theta\left[r\left(A^o, \theta, \mathbf{x}^o(\theta)\right)\right] \\ &\quad + E_\theta\left[\sum_{l \in A^o \backslash S} t_l^o(\mathbf{x}^o(\theta))\right]. \end{aligned} \tag{100}$$

In order to transform the left-hand side of the latter expression, recall from the definition of truthful strategies that, for all $l \in A^o \backslash S$, we have:

$$g_l(\mathbf{x}^*_{N \backslash S}(\theta)) - E_\theta\left[g_l(\mathbf{x}^o(\theta)) - t_l^o(\mathbf{x}^o(\theta))\right] \leq t_l^o(\mathbf{x}^*_{N \backslash S}(\theta)). \tag{101}$$

Substitute the latter inequality in the previous one to obtain:

$$\begin{aligned} &E_\theta\left[\sum_{l \in A^o \backslash S}\left[g_l(\mathbf{x}^*_{A^o \backslash S}(\theta)) - g_l(\mathbf{x}^o(\theta)) + t_l^o(\mathbf{x}^o(\theta))\right] - r\left(A^o \backslash S, \theta, \mathbf{x}^*_{A^o \backslash S}(\theta)\right)\right] \\ &\qquad \leq E_\theta\left[\sum_{k \in S} t_k^o(\mathbf{x}^o(\theta))\right] - E_\theta\left[r\left(A^o, \theta, \mathbf{x}^o(\theta)\right)\right] + E_\theta\left[\sum_{l \in A^o \backslash S} t_l^o(\mathbf{x}^o(\theta))\right]. \end{aligned}$$

The term $\sum_{l \in A^o \backslash S} t_l^o(\mathbf{x}^o(\theta))$ cancels out on each side of the inequality sign, and a simple reorganization gives:

$$\begin{aligned} -E_\theta\left(\sum_{k \in S} t_k^o(\mathbf{x}^o(\theta))\right) \quad &\leq \quad -E_\theta\left[r\left(A^o, \theta, \mathbf{x}^o(\theta)\right)\right] + E_\theta\left(\sum_{l \in A^o \backslash S} g_l\left(\mathbf{x}^o(\theta)\right)\right) \\ &\quad - E_\theta\left(\sum_{l \in A^o \backslash S} g_l(\mathbf{x}^*_{A^o \backslash S}(\theta)) - r(A^o \backslash S, \theta, \mathbf{x}^*_{A^o \backslash S}(\theta))\right). \end{aligned} \tag{103}$$

By adding $\sum_{k \in S} g_k(\mathbf{x}^o(\theta))$ on each side, one finds that:

$$\begin{aligned} &E_\theta\left(\sum_{k \in S}\left[g_k(\mathbf{x}^o(\theta)) - t_k^o(\mathbf{x}^o(\theta))\right]\right) \\ &\leq E_\theta\left(\sum_{i \in A^o} g_i(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^o(\theta))\right) - E_\theta\left(\sum_{l \in A^o \backslash S} g_l(\mathbf{x}^*_{A^o \backslash S}(\theta)) - r(A^o \backslash S, \theta, \mathbf{x}^*_{A^o \backslash S}(\theta))\right). \end{aligned}$$

Let $v_k^o = E_\theta\left[g_k(\mathbf{x}^o(\theta)) - t_k^o(\mathbf{x}^o(\theta))\right]$ denotes principal $k$'s net equilibrium expected profits. By using $\sum_{i \in A^o} g_i(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^o(\theta)) \leq \sum_{i \in A^o} g_i(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta))$, all $\theta$, we finally obtain:

$$\sum_{k \in S} v_k^o \leq E_\theta\left[W(\theta, A^o) - W(\theta, A^o \backslash S)\right], \tag{105}$$

for all $S \subseteq A^o$, as required. $\square$

### 11.3.3 Proposition 11.3.3 (Efficient actions and equilibrium actions almost always coincide):

*Almost all elements $\mathbf{x}^o(\theta)$ of the agent's profit-maximizing set $X^o(\theta)$ are elements of the efficient set $X^*(\theta)$, and almost all elements $\mathbf{x}^*(\theta)$ of the efficient set $X^*(\theta)$ are elements of the agent's profit-maximizing set $X^o(\theta)$.*

**Proof:**

$(\mathbf{x}^o(\theta) \in X^*(\theta))$ : Assume that there exists $\widehat{\Theta} \subset \Theta$ of strictly positive measure such that there exists $\mathbf{x}^o(\theta)$ in $X^o(\theta)$ and *not* in $X^*(\theta)$ for all $\theta$ in $\widehat{\Theta}$, and look for a contradiction. By definition the truthfulness of the strategy $t_i^o$ implies that:

$$g_i(\mathbf{x}) - E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right] \leq t_i^o(\mathbf{x}).$$

We know from Remark **11.3.1** that, in equilibrium, the strategy $t_i^o$ of any principal in the contracting set $A^o$ is truthful relative to $\mathbf{x}^o(\theta)$ for almost all realizations of $\theta$, that is:

$$g_i(\mathbf{x}) - g_i(\mathbf{x}^o(\theta)) + t_i^o(\mathbf{x}^o(\theta)) \leq t_i^o(\mathbf{x}), \tag{106}$$

for all $i \in A^o$, all $\mathbf{x}$, and for almost all $\theta$. Take in particular $\mathbf{x} = \mathbf{x}^*(\theta) \in X^*(\theta)$. We obtain:

$$g_i(\mathbf{x}^*(\theta)) - g_i(\mathbf{x}^o(\theta)) + t_i^o(\mathbf{x}^o(\theta)) \leq t_i^o(\mathbf{x}^*(\theta)), \tag{107}$$

for almost all $\theta$. Summing through $A^o$, and substracting $r(A^o, \theta, \mathbf{x}^*(\theta))$ on each side of the inequality sign, yields:

$$\sum_{i\in A^o} g_i(\mathbf{x}^*(\theta)) - \sum_{i\in A^o} g_i(\mathbf{x}^o(\theta)) + \sum_{i\in A^o} t_i^o(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta)) \leq \sum_{i\in A^o} t_i^o(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta)), \tag{108}$$

for almost all $\theta$. Then introduce $r(A^o, \theta, \mathbf{x}^o(\theta))$ on the left-hand side, and reorganize terms, to obtain:

$$\begin{aligned}&\left[\sum_{i\in A^o} g_i(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta))\right] - \left[\sum_{i\in A^o} g_i(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^o(\theta))\right]\\ &\quad + \sum_{i\in A^o} t_i^o(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^o(\theta)) \leq \sum_{i\in A^o} t_i^o(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta)),\end{aligned} \tag{109}$$

for almost all $\theta$.

Observe that $\mathbf{x}^*(\theta) \in X^*(\theta)$ and $\mathbf{x}^o(\theta) \notin X^*(\theta)$ imply that $\sum_{i\in A^o} g_i(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta)) > \sum_{i\in A^o} g_i(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^o(\theta))$, which in turn implies that the difference $\sum_{i\in N} t_i^o(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^o(\theta))$ is strictly smaller than the whole algebraic expression on the left-hand side of the inequality sign. Therefore:

$$\sum_{i\in A^o} t_i^o(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^o(\theta)) < \sum_{i\in A^o} t_i^o(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta)), \tag{110}$$

almost all $\theta$. Since $\widehat{\Theta}$ is a set of strictly positive measure, there exists a non-empty subset of $\widehat{\Theta}$ for which we have both $\mathbf{x}^o(\theta)$ in $X^o(\theta)$ and the inequality (110) verified, a contradiction.

$(\mathbf{x}^*(\theta) \in X^o(\theta))$ : Assume that there exists $\widehat{\Theta} \subset \Theta$ of strictly positive measure such that there exists $\mathbf{x}^*(\theta)$ in $X^*(\theta)$ and *not* in $X^o(\theta)$ for all $\theta$ in $\widehat{\Theta}$, and look for a contradiction.

By definition

$$\sum_{i\in A^o} t_i^o(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^o(\theta)) \geq \sum_{i\in A^o} t_i^o(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta)), \tag{111}$$

all $\theta$. The inequality is strict if $\mathbf{x}^*(\theta) \notin X^o(\theta)$. For all $i \in A^o$ and all $\mathbf{x} \in X$, we know from the truthfulness of strategies $t_i^o$ that:

$$t_i^o(\mathbf{x}) \geq g_i(\mathbf{x}) - E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right]. \tag{112}$$

In the latter expression, substitute $\mathbf{x}^*(\theta)$ for $\mathbf{x}$ on both sides, and $v_i^o$ for $E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i^o(\mathbf{x}^o(\theta))\right]$ on the right-hand side. Then sum through $A^o$, and substract $r(A^o, \theta, \mathbf{x}^*(\theta))$ on each side of the inequality sign, to obtain:

$$\sum_{i\in A^o} t_i^o\left(\mathbf{x}^*(\theta)\right) - r\left(A^o, \theta, \mathbf{x}^*(\theta)\right) \geq \sum_{i\in A^o}\left[g_i\left(\mathbf{x}^*(\theta)\right) - v_i^o\right] - r\left(A^o, \theta, \mathbf{x}^*(\theta)\right), \tag{113}$$

all $\theta$. By transitivity, taking (111) and (113) together leads to:

$$\sum_{i\in A^o} t_i^o(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^o(\theta)) \geq \sum_{i\in A^o}\left[g_i(\mathbf{x}^*(\theta)) - v_i^o\right] - r(A^o, \theta, \mathbf{x}^*(\theta)).$$

The inequality is strict if $\mathbf{x}^*(\theta) \notin X^o(\theta)$, that is if $\theta \in \widehat{\Theta}$. Since from Remark **11.3.1**, for all principal $i$ in the contracting set $t_i^o(\mathbf{x}^o(\theta)) = g_i\left(\mathbf{x}^o(\theta)\right) - v_i^o$ for almost all $\theta$, it follows that

$$E_\theta\left[\sum_{i\in A^o}\left[g_i(\mathbf{x}^o(\theta)) - v_i^o\right] - r(A^o, \theta, \mathbf{x}^o(\theta))\right] \geq E_\theta\left[\sum_{i\in A^o}\left[g_i(\mathbf{x}^*(\theta)) - v_i^o\right] - r\left(A^o, \theta, \mathbf{x}^*(\theta)\right)\right]. \tag{114}$$

almost all $\theta$. Inequality (114) is strict for almost all $\theta \in \widehat{\Theta}$. Since $\widehat{\Theta}$ is a set of strictly positive measure, this says

$$E_\theta\left[\sum_{i\in A^o} g_i(\mathbf{x}^o(\theta)) - r(A^o, \theta, \mathbf{x}^o(\theta))\right] > E_\theta\left[\sum_{i\in A^o} g_i(\mathbf{x}^*(\theta)) - r\left(A^o, \theta, \mathbf{x}^*(\theta)\right)\right], \tag{115}$$

where the equilibrium payoffs $v_i^o$ have been cancelled on both sides. This later inequality is clearly not compatible with the definition of the efficient actions $\mathbf{x}^*(\theta)$. $\square$

### 11.3.4 Lemma 11.3.4 (Necessary and sufficient condition defining the Pareto frontier $\mathcal{V}_\Gamma^*$ of $V_\Gamma^* \equiv V_\Gamma(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta})$):

*The payoff vector $\mathbf{v}^o$ is in $\mathcal{V}_\Gamma^*$, the Pareto frontier of $V_\Gamma^*$, if and only if, for all $j$ there exists $S \subseteq A^o$, with $j \in S$, such that:*

$$\sum_{i\in S} v_i^o = E_\theta\left[W(\theta, A^o) - W(\theta, A^o \backslash S)\right]. \tag{116}$$

**Proof:** $(i)$ Assume that $\mathbf{v}^o = (v_1^o, v_2^o, .., v_n^o) \in \mathcal{V}_\Gamma(\theta, \mathbf{x})$ but for all $S \subseteq A^o$ with $j \in S$ we have:

$$\sum_{i\in S} v_i^o < E_\theta \left[W(\theta, A^o) - W(\theta, A^o \backslash S)\right]. \tag{117}$$

It is possible to slightly increase $v_j^o$ while remaining inside $V_\Gamma^* \equiv V_\Gamma(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta})$, a contradiction.

$(ii)$ Assume that for all $j$ there exists $S \subseteq A^o$, with $j \in S$, such that:

$$\sum_{i\in S} v_i^o = E_\theta \left[W(\theta, A^o) - W(\theta, A^o \backslash S)\right]. \tag{118}$$

and $\mathbf{v}^o = (v_1^o, v_2^o, .., v_n^o) \in V_\Gamma^*$, but $\mathbf{v}^o \notin \mathcal{V}_\Gamma^*$. Then there exists $\overline{\mathbf{v}} = (\overline{v}_1, \overline{v}_2, .., \overline{v}_n) \in \mathcal{V}_\Gamma^*$ with $(\overline{v}_1, \overline{v}_2, .., \overline{v}_n) \geq (v_1^o, v_2^o, .., v_n^o)$. Without loss of generality, assume that $\bar{v}_1 > v_1^o$. Consider $(\bar{v}_1, v_2^o, .., v_n^o)$. Since $(\overline{v}_1, \overline{v}_2, .., \overline{v}_n) \geq (\bar{v}_1, v_2^o, .., v_n^o)$, it is clear that $(\bar{v}_1, v_2^o, .., v_n^o) \in V_\Gamma^*$. But this cannot be since there exists $S \subseteq A^o$ with $j = 1 \in S$ such that:

$$E_\theta \left[W(\theta, A^o) - W(\theta, A^o \backslash S)\right] = \sum_{i\in S} v_i^o < \bar{v}_1 + \sum_{i\in S\backslash\{1\}} v_i^o, \tag{119}$$

a contradiction with Proposition **11.3.2**. $\square$

### 11.3.5 Proposition 11.3.5 (Pareto efficiency, or $\mathbf{x}^o(\theta) \in \mathcal{V}_\Gamma(\theta, \mathbf{x}^*(\theta))$):

*Any equilibrium is Pareto efficient, i.e. the equilibrium payoffs are in $\mathcal{V}_\Gamma^*$, the Pareto frontier of the set $V_\Gamma^* \equiv V_\Gamma(\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta})$.*

**Proof:** We consider the two cases evoqued in Lemma **11.2.3**, as follows.

- Case 1: $E_\theta \left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}^o, \mathbf{x})\right] = 0$

  By definition of $W(\theta, A^o) \equiv \max_{\mathbf{x}\in X} \left(\sum_{j\in A^o} g_j(\mathbf{x}) - r(A^o, \theta, \mathbf{x})\right)$, we know that:

$$E_\theta \left[W(\theta, A^o)\right] = E_\theta \left[\sum_{j\in A^o} g_j(\mathbf{x}^*(\theta)) - \sum_{j\in A^o} t_j^o(\mathbf{x}^*(\theta)) + \sum_{j\in A^o} t_j^o(\mathbf{x}^*(\theta)) - r\left(A^o, \theta, \mathbf{x}^*(\theta)\right)\right]. \tag{120}$$

Recalling from Proposition **11.3.3** that almost all elements $\mathbf{x}^*(\theta)$ of the efficient set

$$X^*(\theta) \equiv \arg\max_{\mathbf{x}\in X}\left(\sum_{j\in A^o} g_j(\mathbf{x}) - r(A^o,\theta,\mathbf{x})\right), \tag{121}$$

are also elements of the agent's profit-maximizing set

$$X^o(\theta) \equiv \arg\max_{\mathbf{x}\in X}\left(\sum_{j\in A^o} t_j^o(\mathbf{x}) - r(A^o,\theta,\mathbf{x})\right) \tag{122}$$

we obtain:

$$E_\theta\left[W(\theta,A^o)\right] = E_\theta\left[\sum_{j\in A^o}\left[g_j(\mathbf{x}^o(\theta)) - t_j^o(\mathbf{x}^o(\theta))\right]\right] + E_\theta\left[\sum_{j\in A^o} t_j^o(\mathbf{x}^o(\theta)) - r(A^o,\theta,\mathbf{x}^o(\theta))\right]. \tag{123}$$

Since $E_\theta\left[\max_{\mathbf{x}\in X}\left(\sum_{j\in A^o} t_j^o(\mathbf{x}) - r(A^o,\theta,\mathbf{x})\right)\right] = 0$, equation (123) becomes:

$$E_\theta\left[W(\theta,A^o)\right] = \sum_{j\in A^o} E_\theta\left[g_j(\mathbf{x}^o(\theta)) - t_j^o(\mathbf{x}^o(\theta))\right] = \sum_{j\in A^o} v_j^o \tag{124}$$

- Case 2: $E_\theta\left[\max_{\mathbf{x}\in X} v_{\mathcal{L}}^{A^o}(\theta,\mathbf{t}^o,\mathbf{x})\right] = \sup_{S\subseteq N\setminus\{i\}} E_\theta\left[\max_{\mathbf{x}\in X}\left(\sum_{j\in S} t_j^o(\mathbf{x}) - r(\theta,\mathbf{x})\right)\right] > 0$.

  Let $S_{-i} = \arg\sup_{S\subseteq N\setminus\{i\}} E_\theta\left[\max_{\mathbf{x}\in X}\left(\sum_{j\in S} t_j^o(\mathbf{x}) - r(\theta,\mathbf{x})\right)\right]$.

  Assume that $S_{-i} \subseteq A^o\setminus\{i\}$.

  From Lemma 1bis, we know that there exists a set of equilibrium actions such that a subset of the contracting principals gives the agent a zero expected contribution at equilibrium. More precisely, there exists a set of actions $\left\{\mathbf{x}_{-i}^o(\theta)\right\}$ in $X^o(\theta)$ such that the set of principals $T_{\{i\}} = A^o\setminus S_{-i}$ verifies

$$E_\theta\left[\sum_{j\in T_{\{i\}}} t_j^o\left(\mathbf{x}_{-i}^o(\theta)\right) - f_{T_{\{i\}}}^{A^o}\left(\theta,\mathbf{x}_{-i}^o(\theta)\right)\right] = 0, \tag{125}$$

  where $f_{T_{\{i\}}}^{A^o}\left(\theta,\mathbf{x}_{-i}^o(\theta)\right) = \left[r\left(A^o,\theta,\mathbf{x}_{-i}^o(\theta)\right) - r\left(A^o\setminus T_{\{i\}},\theta,\mathbf{x}_{-i}^o(\theta)\right)\right]$ is the incremental costs attached to the participation of the principals in $T_{\{i\}}$. Observe that, since $S_{-i} \subseteq$

$A^o \backslash \{i\}$, the set $T_{\{i\}}$ contains at least principal $i$. Since $\mathbf{x}^o_{-i}(\theta)$ and $\mathbf{x}^o(\theta)$ are in $X^o(\theta)$ by definition, we have:

$$E_\theta\left[\sum_{j\in A^o} t^o_j\left(\mathbf{x}^o(\theta)\right) - r\left(A^o,\theta,\mathbf{x}^o(\theta)\right)\right] = E_\theta\left[\sum_{j\in A^o} t^o_j\left(\mathbf{x}^o_{-i}(\theta)\right) - r\left(A^o,\theta,\mathbf{x}^o_{-i}(\theta)\right)\right]. \tag{126}$$

Recalling again from Proposition **11.3.3** that $\mathbf{x}^o(\theta)$ is in $X^*(\theta)$ for almost all $\theta$, we can substitute $\mathbf{x}^*(\theta)$ for $\mathbf{x}^o(\theta)$ in the latter expression to obtain:

$$E_\theta\left[\sum_{j\in A^o} t^o_j\left(\mathbf{x}^*(\theta)\right) - r\left(A^o,\theta,\mathbf{x}^*(\theta)\right)\right] = E_\theta\left[\sum_{j\in A^o} t^o_j\left(\mathbf{x}^o_{-i}(\theta)\right) - r\left(A^o,\theta,\mathbf{x}^o_{-i}(\theta)\right)\right]. \tag{127}$$

The latter expression can be decomposed over the set $T_{\{i\}}$, as introduced above, and its complementary set in $A^o$, namely $S_{-i}$. This yields:

$$\begin{aligned}&E_\theta\left[\sum_{j\in A^o} t^o_j\left(\mathbf{x}^*(\theta)\right) - r\left(A^o,\theta,\mathbf{x}^*(\theta)\right)\right]\\ &= E_\theta\left[\sum_{j\in T_{\{i\}}} t^o_j\left(\mathbf{x}^o_{-i}(\theta)\right) - f^{A^o}_{T_{\{i\}}}\left(\theta,\mathbf{x}^o_{-i}(\theta)\right)\right] + E_\theta\left[\sum_{j\in S_{-i}} t^o_j\left(\mathbf{x}^o_{-i}(\theta)\right) - r\left(S_{-i},\theta,\mathbf{x}^o_{-i}(\theta)\right)\right].\end{aligned}$$

We now focus on the right-hand side of the latter displayed expression. Recall that $E_\theta\left[\sum_{j\in T_{\{i\}}} t^o_j\left(\mathbf{x}^o_{-i}(\theta)\right) - f^{A^o}_{T_{\{i\}}}\left(\theta,\mathbf{x}^o_{-i}(\theta)\right)\right] = 0$. Moreover, since $\mathbf{x}^o_{-i}(\theta)$ is an equilibrium action, from Remark **11.3.1** we know that for all principal $i$ in the contracting set $t^o_i(\mathbf{x}^o_{-i}(\theta)) = g_i\left(\mathbf{x}^o_{-i}(\theta)\right) - v^o_i$ for almost all $\theta$. Recalling from Proposition **11.3.3** that $\mathbf{x}^o_{-i}(\theta) \in X^*(\theta)$, almost all $\theta$, we thus have $t^o_j\left(\mathbf{x}^o_{-i}(\theta)\right) = g_j\left(\mathbf{x}^o_{-i}(\theta)\right) - E_\theta\left[g_j(\mathbf{x}^*(\theta)) - t^o_j(\mathbf{x}^*(\theta))\right]$, almost all $\theta$. This leads us to rewrite (128) as:

$$\begin{aligned}&E_\theta\left[\sum_{j\in A^o} t^o_j\left(\mathbf{x}^*(\theta)\right) - r\left(A^o,\theta,\mathbf{x}^*(\theta)\right)\right]\\ &\qquad = E_\theta\left[\sum_{j\in S_{-i}}\left[g_j\left(\mathbf{x}^o_{-i}(\theta)\right) - g_j\left(\mathbf{x}^*(\theta)\right) + t^o_j\left(\mathbf{x}^*(\theta)\right)\right] - r\left(S_{-i},\theta,\mathbf{x}^o_{-i}(\theta)\right)\right].\end{aligned}$$

Now we add $E_\theta\left[\sum_{j\in T_{\{i\}}} g_j\left(\mathbf{x}^*\left(\theta\right)\right)\right]$ on both sides of the equality sign and simplify the terms in $t_j^o\left(\mathbf{x}^*\left(\theta\right)\right)$, all $j\in S_{-i}$, to obtain:

$$E_\theta\left[\sum_{j\in T_{\{i\}}} g_j\left(\mathbf{x}^*\left(\theta\right)\right)+\sum_{j\in T_{\{i\}}} t_j^o\left(\mathbf{x}^*\left(\theta\right)\right)-r\left(A^o,\theta,\mathbf{x}^*\left(\theta\right)\right)\right]$$
$$=E_\theta\left[\sum_{j\in T_{\{i\}}} g_j\left(\mathbf{x}^*\left(\theta\right)\right)+\sum_{j\in S_{-i}}\left[g_j\left(\mathbf{x}_{-i}^o\left(\theta\right)\right)-g_j\left(\mathbf{x}^*\left(\theta\right)\right)\right]-r\left(S_{-i},\theta,\mathbf{x}_{-i}^o\left(\theta\right)\right)\right],$$

which can be reorganized to obtain:

$$E_\theta\left[\sum_{j\in T_{\{i\}}} g_j\left(\mathbf{x}^*\left(\theta\right)\right)-\sum_{j\in T_{\{i\}}} t_j^o\left(\mathbf{x}^*\left(\theta\right)\right)\right]$$
$$=E_\theta\left[\left(\sum_{j\in A^o} g_j\left(\mathbf{x}^*\left(\theta\right)\right)-r\left(A^o,\theta,\mathbf{x}^*\left(\theta\right)\right)\right)-\left(\sum_{j\in S_{-i}} g_j\left(\mathbf{x}_{-i}^o\left(\theta\right)\right)-r\left(S_{-i},\theta,\mathbf{x}_{-i}^o\left(\theta\right)\right)\right)\right].$$

By definition of $x_{S_{-i}}^*\left(\theta\right)$,

$$\sum_{j\in S_{-i}} g_j\left(\mathbf{x}_{-i}^o\left(\theta\right)\right)-r\left(S_{-i},\theta,\mathbf{x}_{-i}^o\left(\theta\right)\right)\leq\sum_{j\in S_{-i}} g_j\left(x_{S_{-i}}^*\left(\theta\right)\right)-r\left(S_{-i},\theta,x_{S_{-i}}^*\left(\theta\right)\right)$$

and recalling that for all principals in the contracting set $v_j^o=E_\theta\left[g_j\left(\mathbf{x}^o\left(\theta\right)\right)-t_j^o\left(\mathbf{x}^o\left(\theta\right)\right)\right]=E_\theta\left[g_j\left(\mathbf{x}^*\left(\theta\right)\right)-t_j^o\left(\mathbf{x}^*\left(\theta\right)\right)\right]$ from Proposition **11.3.3,** we obtain:

$$\sum_{j\in T_{\{i\}}} v_j^o\geq E_\theta\left[W(\theta,A^o)-W(\theta,A^o\backslash T_{\{i\}})\right]. \tag{132}$$

Since we know from Proposition **11.3.3** that $\sum_{j\in S} v_j^o\leq E_\theta\left[W(\theta,A^o)-W(\theta,A^o\backslash S)\right]$ all $S\subseteq A^o$, it follows that there is equality. As a result of Lemma **11.3.4**, $\mathbf{v}^o\in\mathcal{V}_\Gamma^*$.$\square$

### 11.3.6 Lemma 11.3.6

*Assume that, for all $\mathbf{x}\in X$ and all $i\in N$, the transfers $t_i(\mathbf{x})$ are defined by:*

$$t_i(\mathbf{x})=\sup\left\{g_i(\mathbf{x})-\overline{v}_i,0\right\} \tag{133}$$

*and verify*

$$\sum_{i\in S} t_i(\mathbf{x}) - f_S^{A^o}(\theta, \mathbf{x}) \geq 0, \tag{134}$$

*all* $S \subseteq A^o$ *and almost all* $\theta$.

*Assume furthermore that for some* $S \subseteq A^o$:

$$\sum_{i\in S} \overline{v}_i = E_\theta \left[ W(\theta, A^o) - W(\theta, A^o \backslash S) \right]. \tag{135}$$

*Then, if* $S \subsetneq A^o$, *the following three properties hold:*

a) $\sum_{i\in A^o} t_i\left(\mathbf{x}^*_{A^o\backslash S}(\theta)\right) - r\left(A^o\backslash S, \theta, \mathbf{x}^*_{A^o\backslash S}(\theta)\right) = \sum_{i\in A^o} t_i\left(\mathbf{x}^*_{A^o}(\theta)\right) - r\left(A^o, \theta, \mathbf{x}^*_{A^o}(\theta)\right)$ *almost all* $\theta$;

b) $\sum_{i\in S} t_i\left(\mathbf{x}^*_{A^o\backslash S}(\theta)\right) - f_S^{A^o}\left(\theta, \mathbf{x}^*_{A^o\backslash S}(\theta)\right) = 0$ *almost all* $\theta$;

c) *For all* $i \in A^o\backslash S$, $t_i(\mathbf{x}^*_{A^o\backslash S}(\theta)) = g_i(\mathbf{x}^*_{A^o\backslash S}(\theta)) - \overline{v}_i$ *almost all* $\theta$.

**Proof of a)** Let $\mathbf{x}^o(\theta)$ denote an equilibrium action. As shown in Remark **11.3.1**, for all principals in the contracting set $A^o$, it must be the case that $g_i(\mathbf{x}^o(\theta)) \geq E_\theta\left[g_i(\mathbf{x}^o(\theta)) - t_i(\mathbf{x}^o(\theta))\right] = v_i^o$, almost all $\theta$. Thus, if strategies are defined according to (133), $v_i^o = \overline{v}_i$ and $t_i(\mathbf{x}^o(\theta)) = g_i(\mathbf{x}^o(\theta)) - \overline{v}_i$, almost all $\theta$. From Proposition **11.3.3**, any equilibrium action is an efficient action hence, we can substitute $\mathbf{x}^*(\theta)$ for $\mathbf{x}^o(\theta)$ in the later expression to obtain $t_i(\mathbf{x}^*(\theta)) = g_i(\mathbf{x}^*(\theta)) - \overline{v}_i$ almost all $\theta$. It follows that equation (135) rewrites:

$$\sum_{i\in S} \left[g_i(\mathbf{x}^*(\theta)) - t_i(\mathbf{x}^*(\theta))\right]$$
$$= E_\theta\left[\left(\sum_{i\in A^o} g_i(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta))\right) - \left(\sum_{j\in A^o\backslash S} g_i(\mathbf{x}^*_{A^o\backslash S}(\theta)) - r(A^o\backslash S, \theta, \mathbf{x}^*_{A^o\backslash S}(\theta))\right)\right],$$

almost all $\theta$.Since the later equality holds true for almost all $\theta$, it also holds true in expectation. We simplify it by decomposing the set $A^o$ into $S \cup (A^o \backslash S)$ and reorganizing terms we obtain:

$$E_\theta\left[\sum_{j\in A^o\backslash S} g_i(\mathbf{x}^*_{A^o\backslash S}(\theta)) - r(A^o\backslash S, \theta, \mathbf{x}^*_{A^o\backslash S}(\theta))\right] - E_\theta\left[\sum_{i\in A^o\backslash S} g_i(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta))\right]$$
$$= E_\theta\left[\sum_{i\in S} t_i(\mathbf{x}^*(\theta))\right].$$

We also know by assumption, *i.e.* from equation (134), that

$$E_\theta\left[\sum_{i\in S} t_i\left(\mathbf{x}^*_{A^o\backslash S}(\theta)\right) - f_S^{A^o}\left(\theta, \mathbf{x}^*_{A^o\backslash S}(\theta)\right)\right] \geq 0.$$

Adding this inequality to the previous equation gives:

$$E_\theta\left[\sum_{j\in A^o\backslash S} g_i(\mathbf{x}^*_{A^o\backslash S}(\theta)) - r\left(A^o\backslash S, \theta, \mathbf{x}^*_{A^o\backslash S}(\theta)\right)\right] - E_\theta\left[\sum_{i\in A^o\backslash S} g_i(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta))\right]$$
$$+ E_\theta\left[\sum_{i\in S} t_i\left(\mathbf{x}^*_{A^o\backslash S}(\theta)\right) - f_S^{A^o}\left(\theta, \mathbf{x}^*_{A^o\backslash S}(\theta)\right)\right] \geq E_\theta\left[\sum_{i\in S} t_i(\mathbf{x}^*(\theta))\right].$$

From assumption (133) (and Proposition **11.3.3**) we also get that for all $i \in A^o$, all $\mathbf{x} \in X$:

$$t_i(\mathbf{x}) \geq g_i(\mathbf{x}) - \overline{v}_i = g_i(\mathbf{x}) - E_\theta\left[g_i(\mathbf{x}^*(\theta)) - t_i(\mathbf{x}^*(\theta))\right],$$

hence by considering $\mathbf{x} = \mathbf{x}^*_{A^o\backslash S}(\theta)$ and taking the expectation in $\theta$ we have:

$$E_\theta\left[\sum_{i\in A^o\backslash S} t_i\left(\mathbf{x}^*_{A^o\backslash S}(\theta)\right)\right] \geq \sum_{i\in A^o\backslash S} E_\theta\left[g_i\left(\mathbf{x}^*_{A^o\backslash S}(\theta)\right) - g_i(\mathbf{x}^*(\theta))\right] + E_\theta\left[\sum_{i\in A^o\backslash S} t_i(\mathbf{x}^*(\theta))\right].$$

Add the latter inequality to equation (136) to obtain:

$$E_\theta\left[\sum_{i\in A^o} t_i\left(\mathbf{x}^*_{A^o\backslash S}(\theta)\right) - r\left(A^o, \theta, \mathbf{x}^*_{A^o\backslash S}(\theta)\right)\right] \geq E_\theta\left[\sum_{i\in A^o} t_i(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta))\right], \tag{137}$$

where we used $r(A^o, \theta, \mathbf{x}) = f_S^{A^o}(\theta, \mathbf{x}) + r(A^o\backslash S, \theta, \mathbf{x})$.

Inequality (137) rewrites

$$E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta, \mathbf{t}, \mathbf{x}^*_{A^o\backslash S}(\theta)\right)\right] \geq E_\theta\left[v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}, \mathbf{x}^*(\theta))\right]. \tag{138}$$

From Proposition **11.3.3,** we know that $\mathbf{x}^*(\theta) \in X^o(\theta)$, almost all $\theta$. Hence, for all $\mathbf{x} \in X$ and for almost all $\theta$

$$\sum_{i\in A^o} t_i(\mathbf{x}) - r(A^o,\theta,\mathbf{x}) \leq \sum_{i\in A^o} t_i(\mathbf{x}^*(\theta)) - r(A^o,\theta,\mathbf{x}^*(\theta)) \equiv \max_{\mathbf{x}\in X}\left\{v_{\mathcal{L}}^{A^o}(\theta,\mathbf{t},\mathbf{x})\right\}.$$

It follows that inequality (138) cannot be strict, so that

$$E_\theta\left[\sum_{i\in A^o} t_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right) - r\left(A^o,\theta,\mathbf{x}^*_{A^o\setminus S}(\theta)\right)\right] = E_\theta\left[\sum_{i\in A^o} t_i(\mathbf{x}^*(\theta)) - r(A^o,\theta,\mathbf{x}^*(\theta))\right]. \tag{139}$$

Moreover, since $\sum_{i\in A^o} t_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right) - r\left(A^o,\theta,\mathbf{x}^*_{A^o\setminus S}(\theta)\right) \leq \sum_{i\in A^o} t_i(\mathbf{x}^*(\theta)) - r(A^o,\theta,\mathbf{x}^*(\theta))$, almost all $\theta$, the equality

$$\sum_{i\in A^o} t_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right) - r\left(A^o,\theta,\mathbf{x}^*_{A^o\setminus S}(\theta)\right) = \sum_{i\in A^o} t_i(\mathbf{x}^*(\theta)) - r(A^o,\theta,\mathbf{x}^*(\theta))$$

holds true almost all $\theta$.This proves **a)**

**Proof of b) and c)** Assume that $E_\theta\left[\sum_{i\in S} t_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right) - f_S^{A^o}\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right)\right] > 0$. Then inequality (136) is strict which implies that inequality (138) is also strict, contradicting assertion **a)** of the lemma. This proves that $E_\theta\left[\sum_{i\in S} t_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right) - f_S^{A^o}\left(\theta,\mathbf{x}^*_{A^o\setminus S}(\theta)\right)\right] = 0$. Since from (134), $\sum_{i\in S} t_i(\mathbf{x}) - f_S^{A^o}(\theta,\mathbf{x}) \geq 0$, all $\mathbf{x} \in X$, almost all $\theta$, this implies in turn that $\sum_{i\in S} t_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right) - f_S^{A^o}\left(\theta,\mathbf{x}^*_{A^o\setminus S}(\theta)\right) = 0$, almost all $\theta$, namely **b)**.

Similarly assume that for some $i \in A^o\setminus S$, $E_\theta\left[t_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right)\right] > E_\theta\left[g_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right)\right] - \overline{v}_i$. This would imply again that inequality (137) is also strict, contradicting assertion **a)** of the lemma. This proves that $E_\theta\left[t_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right)\right] = E_\theta\left[g_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right)\right] - \overline{v}_i$ all $i \in A^o\setminus S$. Since $t_i(\mathbf{x}) \geq g_i(\mathbf{x}) - \overline{v}_i$, all $\mathbf{x} \in X$, it implies that $t_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right) = g_i\left(\mathbf{x}^*_{A^o\setminus S}(\theta)\right) - \overline{v}_i$ all $i \in A^o\setminus S$, and almost all $\theta$ namely **c)**. □

### 11.3.7 Proposition 11.3.7 (Any vector in $\mathcal{V}_\Gamma^*$ can be supported by a Truthful Nash-Equilibrium):

*Any vector* $\overline{\mathbf{v}}$ *in* $I_\Gamma^* \cap \mathcal{V}_\Gamma^*$, that is in *the intersection of the implementable payoff set* $I_\Gamma^*$ and *the Pareto frontier* $\mathcal{V}_\Gamma^*$ *of* $V_\Gamma^*$, *can be obtained as an equilibrium distribution of payoffs.*

**Proof:** Consider a vector $\overline{\mathbf{v}} \in I_\Gamma^* \cap \mathcal{V}_\Gamma^*$. We shall prove that there exists a $\mathbf{t}^o$ such that the triplet $\left(\mathbf{t}^o, A^o, \{\mathbf{x}^*(\theta)\}_{\theta\in\Theta}\right)$ is a Truthful Nash Equilibrium that generates the equilibrium payoffs $\overline{\mathbf{v}}$. Toward this aim, consider the truthful strategies:

$$t_i^o(\mathbf{x}) = \sup\{g_i(\mathbf{x}) - \overline{v}_i, 0\} \qquad \text{all } \mathbf{x} \in X, \text{ all } i \in N. \tag{140}$$

As $\overline{\mathbf{v}} \in I_\Gamma^*$, we know that there exists $\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta}$ in $\{\mathbf{X}^*(\theta)\}_{\theta\in\Theta}$ such that, for all $i \in N$,

$$\overline{v}_i \leq g_i(\mathbf{x}^*(\theta)),$$

almost all $\theta$. It follows that

$$t_i^o(\mathbf{x}^*(\theta)) = g_i(\mathbf{x}^*(\theta)) - \overline{v}_i$$

almost all $\theta$.As a result, if the triplet $(A^o, \mathbf{t}^o, \mathbf{x}^*(\theta))$ constitutes an equilibrium, for any principal $i$ in $A^o$, the equilibrium payoff writes

$$E_{\theta\in\Theta}\left[g_i(\mathbf{x}^*(\theta)) - t_i^o(\mathbf{x}^*(\theta))\right] = \overline{v}_i.$$

We now show that the triplet $(A^o, \mathbf{t}^o, \mathbf{x}^*(\theta))$ indeed constitutes an equilibrium. To do this, we have to show that the four conditions of Theorem **3** are satisfied.

- Condition (1) says that the equilibrium action almost always maximizes the agent's benefits. We thus have to show that $\mathbf{x}^*(\theta) \in X^o(\theta)$ almost all $\theta$. Assume it is not the case *i.e.* there exist $\widehat{\Theta}$ of strictly positive measure and $\hat{\mathbf{x}}(\boldsymbol{\theta}) \neq \mathbf{x}^*(\theta)$ such that:

$$\sum_{j\in A^o} t_j^o(\hat{\mathbf{x}}(\theta)) - r(A^o, \theta, \hat{\mathbf{x}}(\theta)) > \sum_{j\in A^o} t_j^o(\mathbf{x}^*(\theta)) - r(A^o, \theta, \mathbf{x}^*(\theta)), \tag{141}$$

for all $\theta \in \widehat{\Theta}$. Note that, by definition:

$$\begin{aligned}\sum_{j\in A^o} t_j^o(\mathbf{x}^*(\theta)) - r(A^o,\theta,\mathbf{x}^*(\theta)) &= \sum_{j\in A^o} g_j(\mathbf{x}^*(\theta)) - \left(\sum_{j\in A^o}\left[g_j(\mathbf{x}^*(\theta)) - t_j^o(\mathbf{x}^*(\theta))\right]\right) \\ &\quad -r(A^o,\theta,\mathbf{x}^*(\theta)) \\ &= \sum_{j\in A^o} g_j(\mathbf{x}^*(\theta)) - \sum_{j\in A^o}\overline{v}_j - r(A^o,\theta,\mathbf{x}^*(\theta)), \qquad (142)\end{aligned}$$

almost all $\theta$, which allows us to rewrite the right hand-side of the previous inequality. Let $S(\theta) \subseteq A^o$ be the set of principals for which the expected transfers are zero when the chosen action is $\hat{\mathbf{x}}(\theta)$. By definition $S(\theta) = \left\{j \in A^o \mid t_j^o(\hat{\mathbf{x}}(\theta)) = 0\right\}$, and from truthfulness, $t_j^o(\hat{\mathbf{x}}(\theta)) = g_j(\hat{\mathbf{x}}(\theta)) - \overline{v}_j$ for all $j \in A^o\backslash S(\theta)$. This allows us to rewrite the left hand-side of inequality (141) to obtain:

$$\sum_{j\in A^o\backslash S(\theta)}[g_j(\hat{\mathbf{x}}(\theta)) - \overline{v}_j] - r(A^o,\theta,\hat{\mathbf{x}}(\theta)) > \sum_{j\in A^o} g_j(\mathbf{x}^*(\theta)) - \sum_{j\in A^o}\overline{v}_j - r(A^o,\theta,\mathbf{x}^*(\theta)), \qquad (143)$$

for almost all $\theta \in \widehat{\Theta}$. It follows that:

$$\sum_{j\in S(\theta)}\overline{v}_j > \left(\sum_{j\in A^o} g_j(\mathbf{x}^*(\theta)) - r(A^o,\theta,\mathbf{x}^*(\theta))\right) - \left(\sum_{j\in A^o\backslash S(\theta)} g_j(\hat{\mathbf{x}}(\theta)) - r(A^o,\theta,\hat{\mathbf{x}}(\theta))\right), \qquad (144)$$

for almost all $\theta \in \widehat{\Theta}$. Since, by definition of $\mathbf{x}^*_{A^o\backslash S(\theta)}(\theta)$, we have:

$$\sum_{j\in A^o\backslash S(\theta)} g_j(\hat{\mathbf{x}}(\theta)) - r(A^o,\theta,\hat{\mathbf{x}}(\theta)) \le \sum_{j\in A^o\backslash S(\theta)} g_j\left(\mathbf{x}^*_{A^o\backslash S(\theta)}(\theta)\right) - r(A^o,\theta,\mathbf{x}^*_{A^o\backslash S(\theta)}(\theta)), \qquad (145)$$

for almost all $\theta \in \widehat{\Theta}$. This rewrites as:

$$\sum_{j\in S(\theta)}\overline{v}_j > W(\theta,N) - W(\theta,N\backslash S(\theta)), \qquad (146)$$

for almost all $\theta \in \widehat{\Theta}$, which contradicts the fact that $\overline{\mathbf{v}} \in \mathcal{V}_\Gamma(\theta,\mathbf{x}^*(\theta)) \subset V_\Gamma(\theta,\mathbf{x}^*(\theta))$, almost all $\theta$.

- Condition (2) states that the agent's partipation constraint is binding for all the contractual relationships it entertains with principals in in the contracting set. Formally, for all principals $i$ in $A^o$:

$$E_\theta\left[v_{\mathcal{L}}^{A^o}\left(\theta,\mathbf{t}^o,\mathbf{x}^o\left(\theta\right)\right)\right]=\sup\left\{0,\sup_{S\subseteq N\backslash\{i\}}E_\theta\left[\max_{\mathbf{x}\in X}v_{\mathcal{L}}^{S}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right]\right\}. \tag{147}$$

To see that this holds is true, observe by Lemma **11.3.4** that $\overline{\mathbf{v}}\in\mathcal{V}_\Gamma(\theta,\mathbf{x}^*\left(\theta\right))$ implies that, for all $i\in A^o$, there exists $S\subseteq A^o$, with $i\in S$, such that:

$$\sum_{j\in S}\overline{v}_j=E_\theta\left[\sum_{j\in A^o}g_j\left(\mathbf{x}^*\left(\theta\right)\right)-r\left(A^o,\theta,\mathbf{x}^*\left(\theta\right)\right)-W(\theta,A^o\backslash S)\right]. \tag{148}$$

If $S=A^o$, observe that, from condition (1), $\mathbf{x}^*\left(\theta\right)\in X^\circ\left(\theta\right)$ almost all $\theta$, hence

$$E_\theta\left[\max_{\mathbf{x}\in X}\left\{\sum_{j\in A^o}t_j^o(\mathbf{x})-r(A^o,\theta,\mathbf{x})\right\}\right]=E_\theta\left[\sum_{j\in A^o}t_j^o\left(\mathbf{x}^*\left(\theta\right)\right)-r\left(A^o,\theta,\mathbf{x}^*\left(\theta\right)\right)\right]. \tag{149}$$

Thus a simple decomposition gives

$$\begin{aligned}E_\theta\left[\max_{\mathbf{x}\in X}\left\{\sum_{j\in A^o}t_j^o(\mathbf{x})-r(A^o,\theta,\mathbf{x})\right\}\right] &= E_\theta\left[\sum_{j\in A^o}g_j(\mathbf{x}^*\left(\theta\right))-r\left(A^o,\theta,\mathbf{x}^*\left(\theta\right)\right)\right]\\ &\quad-\sum_{j\in A^o}E_\theta\left[g_j(\mathbf{x}^*\left(\theta\right))-t_j^o(\mathbf{x}^*\left(\theta\right))\right]\\ &= E_\theta\left[\sum_{j\in A^o}g_j(\mathbf{x}^*\left(\theta\right))-r\left(A^o,\theta,\mathbf{x}^*\left(\theta\right)\right)-W(\theta,\emptyset)\right]\\ -\sum_{j\in A^o}\overline{v}_j &= 0\end{aligned} \tag{150}$$

from equation (148).

If $S\subsetneq A^o$, we know by Lemma **11.3.6(a)** that, for this particular subset $S$, we have:

$$\sum_{j\in A^o}t_j\left(\mathbf{x}_{A^o\backslash S}^*\left(\theta\right)\right)-r\left(A^o,\theta,\mathbf{x}_{A^o\backslash S}^*\left(\theta\right)\right)=\sum_{j\in A^o}t_j\left(\mathbf{x}^*\left(\theta\right)\right)-r\left(A^o,\theta,\mathbf{x}^*\left(\theta\right)\right) \tag{151}$$

almost all $\theta$. In other words, $\mathbf{x}^*\left(\theta\right)$ leads to the same benefits as $\mathbf{x}_{A^o\backslash S}^*\left(\theta\right)$ for the agent. Therefore $\mathbf{x}^*\left(\theta\right)\in X^o\left(\theta\right)$, almost all $\theta$, as obtained from Condition (1), implies that

$\mathbf{x}^*_{A^o\backslash S}(\theta) \in X^o(\theta)$, almost all $\theta$. And from Lemma **11.3.6(b)**,

$$\sum_{j\in S} t_j\left(\mathbf{x}^*_{A^o\backslash S}(\theta)\right) - f_S^{A^o}\left(\theta, \mathbf{x}^*_{A^o\backslash S}(\theta)\right) = 0, \tag{152}$$

almost all $\theta$. Therefore, substracting equation (152) from (151) we obtain:

$$\begin{aligned} v_{\mathcal{L}}^{A^o}(\theta, \mathbf{t}^o, \mathbf{x}^*(\theta)) &= \sum_{i\in A^o\backslash S} t_j\left(\mathbf{x}^*_{A^o\backslash S}(\theta)\right) + f_S^{A^o}\left(\theta, \mathbf{x}^*_{A^o\backslash S}(\theta)\right) - r\left(A^o, \theta, \mathbf{x}^*_{A^o\backslash S}(\theta)\right) \\ &= \sum_{j\in A^o\backslash S} t_j\left(\mathbf{x}^*_{A^o\backslash S}(\theta)\right) - r\left(A^o\backslash S, \theta, \mathbf{x}^*_{A^o\backslash S}(\theta)\right), \end{aligned} \tag{153}$$

almost all $\theta$. Remind that principal $i \in S$. Thus $A^o\backslash S \subseteq N\backslash\{i\}$. Moreover, from Lemma **11.3.6(c)**, for all $j \in A^o\backslash S$, $t_j(x^*_{A^o\backslash S}(\theta)) = g_j(x^*_{A^o\backslash S}(\theta)) - \overline{v}_j$, almost all $\theta$. It follows that $\mathbf{x}^*_{A^o\backslash S}(\theta) \in \arg\max_{\mathbf{x}}\left\{\sum_{j\in A^o\backslash S} t_j^o(\mathbf{x}) - r(A^o\backslash S, \theta, \mathbf{x})\right\}$, almost all $\theta$. As a result

$$E_\theta\left[\max_{\mathbf{x}\in X}\left\{\sum_{j\in A^o} t_j^o(\mathbf{x}) - r(A^o, \theta, \mathbf{x})\right\}\right]$$

later equation states that refore we have shown that there exists an equilibrium strategy $\mathbf{x}^o_{-i}(\theta)$, which here is given by $\mathbf{x}^*_{N\backslash S}(\theta)$, such that the expected transfer from principal $i$ is zero.

- Condition (3) states that, given the strategies $t_j^o(\mathbf{x})$, all $j \neq i$, of all other principals, the equilibrium action maximizes the joint payoffs of the agent and principal $i$, any $i \in A^o$. To see that, observe that, as $\overline{\mathbf{v}} \in I_\Gamma^*$, there exists $\{\mathbf{x}^*(\theta)\}_{\theta\in\Theta}$ in $\{\mathbf{X}^*(\theta)\}_{\theta\in\Theta}$ such that, for all $i \in A^o$,

$$\overline{v}_i \leq g_i(\mathbf{x}^*(\theta)), \tag{154}$$

almost all $\theta$. By definition:

$$t_i^o(\mathbf{x}) = \sup\{0, g_i(\mathbf{x}) - \overline{v}_i\} \tag{155}$$

all $\mathbf{x}$, hence

$$t_i^o(\mathbf{x}^*(\theta)) = g_i(\mathbf{x}^*(\theta)) - \overline{v}_i, \tag{156}$$

almost all $\theta$. Now, as truthfulness implies

$$t_i^o(\mathbf{x}) \geq g_i(\mathbf{x}) - \overline{v}_i, \tag{157}$$

all $\mathbf{x}$, and considering from (156) that $\overline{v}_i = g_i(\mathbf{x}^*(\theta)) - t_i^o(\mathbf{x}^*(\theta))$, almost all $\theta$, we obtain:

$$g_i(\mathbf{x}^*(\theta)) - g_i(\mathbf{x}) \geq t_i^o(\mathbf{x}^*(\theta)) - t_i^o(\mathbf{x}), \tag{158}$$

all $i \in A^o$, all $\mathbf{x}$, almost all $\theta$.

Condition (1) has established that, for almost all $\theta$, $\mathbf{x}^*(\theta)$ is an equilibrium choice of the agent when strategies are $\mathbf{t}^o$. It follows that:

$$\left(\sum_{j\in A^o} t_j^o(\mathbf{x}) - r(A^o,\theta,\mathbf{x})\right) - \left(\sum_{j\in A^o} t_j^o(\mathbf{x}^*(\theta)) - r(A^o,\theta,\mathbf{x}^*(\theta))\right) \leq 0, \tag{159}$$

almost all $\theta$. Sum (158) and (159) to obtain:

$$\begin{aligned}
&g_i(\mathbf{x}^*(\theta)) - g_i(\mathbf{x}) \\
&\geq t_i^o(\mathbf{x}^*(\theta)) - t_i^o(\mathbf{x}) + \left(\sum_{j\in A^o} t_j^o(\mathbf{x}) - r(A^o,\theta,\mathbf{x})\right) - \left(\sum_{j\in A^o} t_j^o(\mathbf{x}^*(\theta)) - r(A^o,\theta,\mathbf{x}^*(\theta))\right), \\
&\qquad = \left(\sum_{j\in A^o\setminus\{i\}} t_j^o(\mathbf{x}) - r(A^o,\theta,\mathbf{x})\right) - \left(\sum_{j\in A^o\setminus\{i\}} t_j^o(\mathbf{x}^*(\theta)) - r(A^o,\theta,\mathbf{x}^*(\theta))\right),
\end{aligned}$$

all $i \in A^o$, all $\mathbf{x}$, almost all $\theta$.Since from Proposition **11.3.3,** $\mathbf{x}^*(\theta) \in X^o(\theta)$, almost all $\theta$, we have

$$g_i(\mathbf{x}^o(\theta)) - f_i^{A^o}(\theta,\mathbf{x}^o(\theta)) + v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}^o(\theta)\right) \geq g_i(\mathbf{x}) - f_i^{A^o}(\theta,\mathbf{x}) + v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right),$$

all $i \in A^o$, all $\mathbf{x}$, almost all $\theta$. The later rewrites

$$X^o(\theta) \subseteq \arg\max_{\mathbf{x}\in X}\left[g_i(\mathbf{x}) - f_i^{A^o}(\theta,\mathbf{x}) + v_{\mathcal{L}}^{A^o\setminus\{i\}}\left(\theta,\mathbf{t}_{-i}^o,\mathbf{x}\right)\right]$$

all $i \in A^o$, almost all $\theta$.

Since all four conditions of Theorem **3** are verified, the triplet $(A^o, \mathbf{t}^o, \mathbf{x}^*(\theta))$ appears to be an equilibrium. We thus exhibited a TNE that leads to the desired vector of payoffs. □

**Remark 3 (Set of principals with nul expected-transfer and efficiency of the corresponding action for common agency game restricted to the complementary set of principals):** *This Remark use to follow proposition 11.3.5*

*Assume that the expected profit of the agent is strictly positive and let* $x^{\circ}_{\{i\}}(\theta)$ *in* $X^{\circ}(\theta)$ *denote the actions such that* $t^{\circ}_i\left(\mathbf{x}^{\circ}_{\{i\}}(\theta)\right) = 0$, *almost all* $\theta$. *Let* $S_{\{i\}}$ *be the set of principals for which the expected transfer is zero for this set of actions:*

$$S_{\{i\}} = \left\{ j \in N \mid E_\theta \left[ t^{\circ}_j \left( \mathbf{x}^{\circ}_{\{i\}}(\theta) \right) \right] = 0 \right\}.$$

*The actions* $x^{\circ}_{\{i\}}(\theta)$ *are efficient actions of the common agency game restricted to the complementary set of principals* $N \backslash S_{\{i\}}$ *Formally:*

$$\mathbf{x}^{\circ}_{\{i\}}(\theta) \in \arg\max_{\mathbf{x} \in X} \left\{ \sum_{j \in N \backslash S_{\{i\}}} g_j(\mathbf{x}) - r(\theta, \mathbf{x}) \right\} \equiv X^*_{N \backslash S_{\{i\}}}(\theta)$$

*almost all* $\theta$.

**Proof:** A consequence of Proposition 4 is that inequality (**??**) is actually an equality. Hence

$$E_\theta \left[ \sum_{j \in N \backslash S_{\{i\}}} g_j \left( \mathbf{x}^{\circ}_{\{i\}}(\theta) \right) - r(\theta, \mathbf{x}^{\circ}_{\{i\}}(\theta)) \right] = E_\theta \left[ \sum_{j \in N \backslash S_{\{i\}}} g_j \left( x^*_{N \backslash S_{\{i\}}}(\theta) \right) - r(\theta, x^*_{N \backslash S_{\{i\}}}(\theta)) \right],$$

and since by definition of $x^*_{N \backslash S_{\{i\}}}(\theta)$,

$$\sum_{j \in N \backslash S_{\{i\}}} g_j \left( \mathbf{x}^{\circ}_{\{i\}}(\theta) \right) - r(\theta, \mathbf{x}^{\circ}_{\{i\}}(\theta)) \le \sum_{j \in N \backslash S_{\{i\}}} g_j \left( x^*_{N \backslash S_{\{i\}}}(\theta) \right) - r(\theta, x^*_{N \backslash S_{\{i\}}}(\theta)),$$

it follows that there is equality for almost all $\theta$.$\square$

### 11.4 Truthfullness and Coalition-proofness3

NB: *Notations, Marker et references to be updated*

**Proposition** 6 **(Any TNE is an efficient action for the restricted game $\Gamma^S$):** *Consider a common agency game $\Gamma$ and its restriction $\Gamma^S$. Any element of the agent's profit-maximizing set $X^{\circ}(\theta)$ for $\Gamma$ is an element the efficiency set $X_S^*(\theta)$ for $\Gamma^S$.*

**Proof:** The equilibrium action maximizes the joint profits of the agent and the principals in $S$, all $S$, so that no coalition can ever contemplate to be (jointly) better off. We have already demonstrated that, in all truthful equilibria, and for all realizations of the random variable $\theta$, the agent selects an efficient action. In other words, a maximizer of the agent's benefit function $\pi_{\mathcal{L}}(\theta, \mathbf{t}, \mathbf{x})$ is also a maximizer of $W^N(\theta, \mathbf{x})$. We now want to show that for all $S \subseteq N$, this maximizer is also a joint maximizing action for the agent and the coalition $S$ of principals, given the other principals' strategies $\left\{\hat{t}_i(\mathbf{x})\right\}_{i \in N \backslash S}$. More precisely, we want to show that the maximizer of $\pi_{\mathcal{L}}(\theta, \mathbf{t}, \mathbf{x})$ is also a maximizer of the joint-payoff function $\Phi^S(\theta, \mathbf{x})$ defined as follows:

$$\Phi^S(\theta, \mathbf{x}) = \sum_{i \in S} g_i(\mathbf{x}) + \sum_{j \in N \backslash S} t_j(\mathbf{x}) - r(\theta, \mathbf{x}). \tag{161}$$

Note that $\Phi^S(\theta, \mathbf{x}) = W^S(\theta, \mathbf{x}) + \sum_{j \in N \backslash S} t_j(\mathbf{x})$. Denote by $\hat{\mathbf{x}}(\theta) \equiv (\hat{x}_1(\theta), \hat{x}_2(\theta), .., \hat{x}_n(\theta))\hat{x}$ the action produced by the agent at equilibrium of the restricted agency game, when the principals offer the transfers $\hat{\mathbf{t}}(\mathbf{x}) \equiv (\hat{t}_1(\mathbf{x}), \hat{t}_2(\mathbf{x}), .., \hat{t}_n(\mathbf{x}))$ and the realization of the random variable is $\theta$. Clearly $\hat{\mathbf{x}}(\theta)$ is the maximizer of $\pi_{\mathcal{L}}(\theta, \mathbf{t}, \mathbf{x})$. Given the principals strategies $\hat{\mathbf{t}}(\mathbf{x})$, the equilibrium payoff of principal $i \in N$ is $\hat{k}_i \equiv g_i(\hat{\mathbf{x}}(\theta)) - \hat{t}_i(\hat{\mathbf{x}}(\theta))$ which, as a result of truthfulness, does not depend on $\theta$ as long as $\hat{t}_i(\hat{\mathbf{x}}(\theta)) > 0$. In what follows, we will assume this is the case.

Assume that $\hat{\mathbf{x}}(\theta)$ does not maximize $\Phi^S(\theta, \mathbf{x})$. More precisely, let $\tilde{\mathbf{x}}(\theta)$ be the maximizer of $\Phi^S(\theta, \mathbf{x})$ as defined in (161), and assume that there exists a set $\Theta$ of strictly positive Lebesgue measure such that, for all $\theta \in \Theta$:

$$\Phi^S(\theta, \tilde{\mathbf{x}}(\theta)) > \Phi^S(\theta, \hat{\mathbf{x}}(\theta)), \tag{162}$$

that is:

$$\sum_{i \in S} g_i(\tilde{\mathbf{x}}(\theta)) - r(\theta, \tilde{\mathbf{x}}(\theta)) + \sum_{j \in N \backslash S} \hat{t}_j(\tilde{\mathbf{x}}(\theta)) > \sum_{i \in S} g_i(\hat{\mathbf{x}}(\theta)) - r(\theta, \hat{\mathbf{x}}(\theta)) + \sum_{j \in N \backslash S} \hat{t}_j(\hat{\mathbf{x}}(\theta)). \tag{163}$$

The objective is now to demonstrate that $(\hat{\mathbf{t}}, \hat{\mathbf{x}})$ cannot be a (Nash) equilibrium of the restricted common agency game.

Let $\varepsilon$ be the expected difference between $\max \Phi^S(\theta, \mathbf{x}) \equiv \Phi^S(\theta, \tilde{\mathbf{x}}(\theta))$ and $\Phi^S(\theta, \hat{\mathbf{x}}(\theta))$, that is:

$$\varepsilon = E_\theta\left[\Phi^S(\theta, \tilde{\mathbf{x}}(\theta)) - \Phi^S(\theta, \hat{\mathbf{x}}(\theta))\right], \tag{164}$$

which is positive from equation (162), by assumption. Now define the truthful strategies $\left\{\tilde{t}_i(\mathbf{x})\right\}_{i \in S}$, such that:

$$\tilde{t}_i(\mathbf{x}) = g_i(\mathbf{x}) - \tilde{k}_i, \tag{165}$$

where the equilibrium payoff of the principal $i$ is:

$$\tilde{k}_i = \hat{k}_i + \varepsilon / (1 + |S|), \tag{166}$$

and $|S|$ is the size of the set $S$. Clearly, if the agent accepts the set of offers $\mathbf{t}(\mathbf{x}) \equiv (\tilde{\mathbf{t}}_{i \in S}(\mathbf{x}), \hat{\mathbf{t}}_{j \in N \backslash S}(\mathbf{x}))$, all principals $i$ in $S$ are better off (for all realizations of $\theta$) by proposing the truthful strategy $\tilde{t}_i(\mathbf{x})$ rather than $\hat{t}_i(\mathbf{x})$. We now show that the *expected* benefits of the agent are higher in the latter case, so that, given the payment strategies $\hat{\mathbf{t}}_{j \in N \backslash S}(\mathbf{x})$ as proposed by the other principals, the agent does accept to contract when the principals $i$ in $S$ propose the payments $\tilde{t}_i(\mathbf{x})$ in lieu of $\hat{t}_i(\mathbf{x})$.

If $\mathbf{t} = \left(\hat{\mathbf{t}}_{i\in S}, \hat{\mathbf{t}}_{i\in N\setminus S}\right) = \hat{\mathbf{t}}$ and $\mathbf{x} = \hat{\mathbf{x}}(\theta)$ then:

$$\begin{aligned}
\pi_{\mathcal{L}}(\theta, \hat{\mathbf{t}}, \hat{\mathbf{x}}) &= \sum_{i\in S} \hat{t}_i(\hat{\mathbf{x}}(\theta)) - r(\theta, \hat{\mathbf{x}}(\theta)) + \sum_{j\in N\setminus S} \hat{t}_j(\hat{\mathbf{x}}(\theta)), \\
&= \sum_{i\in S} \left[g_i(\hat{\mathbf{x}}(\theta)) - \hat{k}_i\right] - r(\theta, \hat{\mathbf{x}}(\theta)) + \sum_{j\in N\setminus S} \hat{t}_j(\hat{\mathbf{x}}(\theta)), \qquad (167)\\
&= \Phi^S\left(\theta, \hat{\mathbf{x}}(\theta)\right) - \sum_{i\in S} \hat{k}_i. \qquad (168)
\end{aligned}$$

If $\mathbf{t} = \left(\tilde{\mathbf{t}}_{i\in S}, \hat{\mathbf{t}}_{j\in N\setminus S}\right)$ and $\mathbf{x} = \tilde{\mathbf{x}}(\theta)$ then:

$$\begin{aligned}
\pi_{\mathcal{L}}(\theta, \mathbf{t}, \tilde{\mathbf{x}}) &= \sum_{i\in S} \hat{t}_i(\tilde{\mathbf{x}}(\theta)) - r(\theta, \tilde{\mathbf{x}}(\theta)) + \sum_{j\in N\setminus S} \hat{t}_j(\tilde{\mathbf{x}}(\theta)), \qquad (169)\\
&= \sum_{i\in S} \left[g_i(\tilde{\mathbf{x}}(\theta)) - \tilde{k}_i\right] - r(\theta, \tilde{\mathbf{x}}(\theta)) + \sum_{j\in N\setminus S} \hat{t}_j(\tilde{\mathbf{x}}(\theta)), \qquad (170)\\
&= \Phi^S\left(\theta, \tilde{\mathbf{x}}(\theta)\right) - \sum_{i\in S} \tilde{k}_i. \qquad (171)
\end{aligned}$$

Clearly the *expected* benefits of the agent for $\mathbf{t} = \left(\hat{\mathbf{t}}_{-j}, \tilde{t}_j\right)$ and $\mathbf{x} = \tilde{\mathbf{x}}(\theta)$ are higher than for $\mathbf{t} = \hat{\mathbf{t}}$ and $\mathbf{x} = \hat{\mathbf{x}}(\theta)$, that is:

$$\begin{aligned}
&E_\theta\left[\pi_{\mathcal{L}}(\theta, \left(\tilde{\mathbf{t}}_{i\in S}, \hat{\mathbf{t}}_{j\in N\setminus S}\right), \tilde{\mathbf{x}})\right] - E_\theta\left[\pi_{\mathcal{L}}(\theta, \hat{\mathbf{t}}, \hat{\mathbf{x}})\right] \qquad (172)\\
&= E_\theta\left[\left(\Phi^S\left(\theta, \tilde{\mathbf{x}}(\theta)\right) - \sum_{i\in S} \tilde{k}_i\right) - \left(\Phi^S\left(\theta, \hat{\mathbf{x}}(\theta)\right) - \sum_{i\in S} \hat{k}_i\right)\right], \qquad (173)\\
&= \varepsilon - \sum_{i\in S}\left(\tilde{k}_i - \hat{k}_i\right) = \varepsilon / \left(1 + |S|\right). \qquad (174)
\end{aligned}$$

That is, if $(\mathbf{t}, \mathbf{x}) = \left(\tilde{\mathbf{t}}, \tilde{\mathbf{x}}(\theta)\right)$ the agent's expected benefits are greater than $E_\theta\left[\max_{\mathbf{x}} \pi_{\mathcal{L}}(\theta, \hat{\mathbf{t}}, \mathbf{x})\right]$, a contradiction. □

**Proposition 7 (Any TNE induces payoffs in the Pareto frontier of the set all coalition-proof payoff distributions for the restricted game $\Gamma^S$):** *Consider a common agency game $\Gamma$ and its restriction $\Gamma^S$. Any TNE yields a vector of principals' payoffs in $\mathcal{V}_\Gamma^S(\theta, \mathbf{x}^*(\theta))$, that is the Pareto frontier of $V_\Gamma^S(\theta, \mathbf{x}^*(\theta))$ for $\Gamma^S$.*

**Proof:** We argue that if $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$ is a truthful equilibrium of the agency game $\Gamma$ for the realization $\theta$ of the random variable, then for all $S \subseteq N$, the equilibrium payoff vector $\left(g_i\left(\mathbf{x}^{\circ}(\theta)\right) - t_i^{\circ}\left(\hat{\mathbf{x}}(\theta)\right)\right)_{i\in N}$ is in $\mathcal{V}_{\Gamma}^{S}(\theta, \mathbf{x}^*(\theta))$, that is the Pareto frontier of the set:

$$V_{\Gamma}^{S}(\theta, \mathbf{x}^*) = \left\{ \overline{\mathbf{v}} \in \Re_+^n : \text{for all } T \subseteq S,\ \sum_{i\in T} \overline{v}_i \leq \sum_{i\in S} g_i(\mathbf{x}^*) - r(\theta, \mathbf{x}^*) + \sum_{i\in N\backslash S} t_j^{\circ}(\mathbf{x}^*) - W(\theta, S\backslash T) \right\},$$

where $\mathbf{x}^*(\theta) \in X^*(\theta)$, the strategies $t_j^{\circ}(.)$ of the principal in $N\backslash S$ are considered as given, and where:

$$W^S(\theta, T) = \max_{\mathbf{x}\in X} \left\{ \sum_{i\in T} g_i(\mathbf{x}) - r(\theta, \mathbf{x}) + \sum_{j\in N\backslash S} t_j(\mathbf{x}) \right\}. \tag{175}$$

If this is true, it does mean that no subgroup can make a credible counterproposal that weakly benefits all of its members, and thus the equilibrium is self-enforcing.

From Proposition 4, we know that the vector of equilibrium payoffs is in $\mathcal{V}_{\Gamma}(\theta, \mathbf{x}^*(\theta))$, that is the Pareto frontier of $V_{\Gamma}(\theta, \mathbf{x}^*(\theta))$, where $\mathbf{x}^*(\theta) \in X^*(\theta)$.

From Lemma 2 we know that $\overline{\mathbf{v}} \in \mathcal{V}_{\Gamma}(\theta, \mathbf{x})$ if and only if $\overline{\mathbf{v}} \in V_{\Gamma}(\theta, \mathbf{x})$ and for all $j$ there exists $T \subseteq N$, with $i \in T$, such that:

$$\sum_{i\in T} \overline{v}_i = \sum_{i\in N} g_i(\mathbf{x}) - r(\theta, \mathbf{x}) - W(\theta, N\backslash T). \tag{176}$$

For any $j \in S$ it is thus possible to find a set $T \subseteq N$ with $j \in T$, such that:

$$\sum_{i\in T} \overline{v}_i = \left[ \sum_{i\in N} g_i(\mathbf{x}_N^*(\theta)) - r(\theta, \mathbf{x}_N^*(\theta)) \right] - \left[ \sum_{i\in N\backslash T} g_i(\mathbf{x}_{N\backslash T}^*(\theta)) - r(\theta, \mathbf{x}_{N\backslash T}^*(\theta)) \right]. \tag{177}$$

Noting that:

$$\sum_{i\in T} \overline{v}_i = \sum_{i\in T\cap S} \overline{v}_i + \sum_{i\in T\cap N\backslash S} \overline{v}_i, \tag{178}$$

and also that:

$$\sum_{i\in T\cap N\backslash S} \overline{v}_i = \sum_{i\in T\cap N\backslash S} g_i(\mathbf{x}_N^*(\theta)) - \sum_{i\in T\cap N\backslash S} t_i(\mathbf{x}_N^*(\theta)), \tag{179}$$

we see that:

$$\sum_{i\in T\cap S}\overline{v}_i=\sum_{i\in T}\overline{v}_i-\sum_{i\in T\cap N\setminus S}g_i(\mathbf{x}^*_N(\theta))+\sum_{i\in T\cap N\setminus S}t_i(\mathbf{x}^*_N(\theta)). \tag{180}$$

Again, by using (177), we obtain:

$$\begin{aligned}\sum_{i\in N}g_i(\mathbf{x}^*_N(\theta))-\sum_{i\in T\cap N\setminus S}g_i(\mathbf{x}^*_N(\theta)) &= \sum_{i\in S}g_i(\mathbf{x}^*_N(\theta))+\sum_{i\in N\setminus S}g_i(\mathbf{x}^*_N(\theta))-\sum_{i\in T\cap N\setminus S}g_i(\mathbf{x}^*_N(\theta)),\\ &= \sum_{i\in S}g_i(\mathbf{x}^*_N(\theta))+\sum_{i\in N\setminus S\cap N\setminus T}g_i(\mathbf{x}^*_N(\theta)), \qquad (181)\end{aligned}$$

and:

$$\sum_{i\in N\setminus T}g_i(\mathbf{x}^*_{N\setminus T}(\theta))=\sum_{i\in N\setminus T\cap S}g_i(\mathbf{x}^*_{N\setminus T}(\theta))+\sum_{i\in N\setminus T\cap N\setminus S}g_i(\mathbf{x}^*_{N\setminus T}(\theta)), \tag{182}$$

as well as:

$$\begin{aligned}\sum_{i\in T\cap N\setminus S}t_i(\mathbf{x}^*_N(\theta)) &= \sum_{i\in N\setminus S}t_i(\mathbf{x}^*_N(\theta))-\sum_{i\in N\setminus T\cap N\setminus S}t_i(\mathbf{x}^*_N(\theta)),\\ 0 &= \sum_{i\in N\setminus T\cap N\setminus S}t_i(\mathbf{x}^*_{N\setminus T}(\theta))+\sum_{i\in T\cap N\setminus S}t_i(\mathbf{x}^*_{N\setminus T}(\theta))-\sum_{i\in N\setminus S}t_i(\mathbf{x}^*_{N\setminus T}(\theta)). \qquad (183)\end{aligned}$$

Eventually we get:

$$\begin{aligned}\sum_{i\in T\cap S}\overline{v}_i &= \left[\sum_{i\in S}g_i(\mathbf{x}^*_N(\theta))+\sum_{i\in N\setminus S}t_i(\mathbf{x}^*_N(\theta))-r(\theta,\mathbf{x}^*_N(\theta))\right] &(184)\\ &\quad-\left[\sum_{i\in N\setminus T\cap S}g_i(\mathbf{x}^*_{N\setminus T}(\theta))+\sum_{i\in N\setminus S}t_i(\mathbf{x}^*_{N\setminus T}(\theta))-r(\theta,\mathbf{x}^*_{N\setminus T}(\theta))\right] &(185)\\ &\quad+\left\{\left[\sum_{i\in N\setminus T\cap N\setminus S}g_i(\mathbf{x}^*_N(\theta))-\sum_{i\in N\setminus T\cap N\setminus S}t_i(\mathbf{x}^*_N(\theta))\right]\right. &(186)\\ &\quad\left.-\left[\sum_{i\in N\setminus T\cap N\setminus S}g_i(\mathbf{x}^*_{N\setminus T}(\theta))-\sum_{i\in N\setminus T\cap N\setminus S}t_i(\mathbf{x}^*_{N\setminus T}(\theta))\right]\right\} &(187)\\ &\quad+\sum_{i\in T\cap N\setminus S}t_i(\mathbf{x}^*_{N\setminus T}(\theta)). &(188)\end{aligned}$$

By Lemma 3(c), we know that a consequence of the equality (177) is that for all $i\in N\setminus T$, $t_i(\mathbf{x}^*_{N\setminus T}(\theta))=g_i(\mathbf{x}^*_{N\setminus T}(\theta))-\overline{v}_i$ where $\overline{v}_i=g_i(\mathbf{x}^*_N(\theta))-t_i(\mathbf{x}^*_N(\theta))$ is the equilibrium payoff.

Since $(N\backslash T \cap N\backslash S) \subseteq N\backslash T$, it follows that the term in curly brackets is identically zero. Now by Lemma 3(b), we know that another consequence of the equality (177) is that for all $i \in T$, $t_i(\mathbf{x}^*_{N\backslash T}(\theta)) = 0$. Since $(T \cap N\backslash S) \subseteq T$, we obtain that the last term of the latter displayed expression is also zero. As a result, we get:

$$\sum_{i\in T\cap S} \overline{v}_i = \left[\sum_{i\in S} g_i(\mathbf{x}^*_N(\theta)) + \sum_{i\in N\backslash S} t_i(\mathbf{x}^*_N(\theta)) - r(\theta, \mathbf{x}^*_N(\theta))\right] \\ - \left[\sum_{i\in N\backslash T\cap S} g_i(\mathbf{x}^*_{N\backslash T}(\theta)) + \sum_{i\in N\backslash S} t_i(\mathbf{x}^*_{N\backslash T}(\theta)) - r(\theta, \mathbf{x}^*_{N\backslash T}(\theta))\right]. \quad (189)$$

Now we argue that $\mathbf{x}^*_{N\backslash T}(\theta)$ maximizes $\sum_{i\in N\backslash T\cap S} g_i(\mathbf{x}) + \sum_{i\in N\backslash S} t_i(\mathbf{x}) - r(\theta, \mathbf{x})$. To see that, notice first that:

$$\begin{aligned} \sum_{i\in N\backslash S} t_i(\mathbf{x}^*_{N\backslash T}(\theta)) &= \sum_{i\in N} t_i(\mathbf{x}^*_{N\backslash T}(\theta)) - \sum_{i\in S} t_i(\mathbf{x}^*_{N\backslash T}(\theta)), \\ &= \sum_{i\in N} t_i(\mathbf{x}^*_{N\backslash T}(\theta)) - \sum_{i\in N\backslash T\cap S} t_i(\mathbf{x}^*_{N\backslash T}(\theta)) - \sum_{i\in T\cap S} t_i(\mathbf{x}^*_{N\backslash T}(\theta)), \quad (190) \\ &= \sum_{i\in N} t_i(\mathbf{x}^*_{N\backslash T}(\theta)) - \sum_{i\in N\backslash T\cap S} g_i(\mathbf{x}^{N\backslash T}(\theta)) + \sum_{i\in N\backslash T\cap S} \overline{v}_i, \quad (191) \end{aligned}$$

where the lowest line is again a consequence of Lemma 3(b) - 3(c) and of equation (177), which implies that $t_i(\mathbf{x}^*_{N\backslash T}(\theta)) = g_i(\mathbf{x}^*_{N\backslash T}(\theta)) - \overline{v}_i$, for all $i \in N\backslash T$, and $t_i(\mathbf{x}^*_{N\backslash T}(\theta)) = 0$ for all $i \in T$. Now we proceed *ad absurdum*, by supposing that there exists $\mathbf{x}$ such that:

$$\sum_{i\in N\backslash T\cap S} g_i(\mathbf{x}) + \sum_{i\in N\backslash S} t_i(\mathbf{x}) - r(\theta, \mathbf{x}) > \sum_{i\in N\backslash T\cap S} g_i(\mathbf{x}^*_{N\backslash T}(\theta)) + \sum_{i\in N\backslash S} t_i(\mathbf{x}^*_{N\backslash T}(\theta)) - r(\theta, \mathbf{x}^*_{N\backslash T}(\theta)). \quad (192)$$

Focus on the right-hand side of the latter expression. Substituting again the relation $g_i(\mathbf{x}^*_{N\backslash T}(\theta)) = \overline{v}_i + t_i(\mathbf{x}^*_{N\backslash T}(\theta))$ for all $i \in N\backslash T$, and adding $t_i(\mathbf{x}^*_{N\backslash T}(\theta)) = 0$ for all $i \in T \cap S$, we obtain that:

$$\begin{aligned} &\sum_{i\in N\backslash T\cap S} g_i(\mathbf{x}^*_{N\backslash T}(\theta)) + \sum_{i\in N\backslash S} t_i(\mathbf{x}^*_{N\backslash T}(\theta)) - r(\theta, \mathbf{x}^*_{N\backslash T}(\theta)) \quad (193) \\ &= \sum_{i\in N\backslash T\cap S} \overline{v}_i + \left[\sum_{i\in N} t_i(\mathbf{x}^*_{N\backslash T}(\theta)) - r(\theta, \mathbf{x}^*_{N\backslash T}(\theta))\right], \quad (194) \\ &= \sum_{i\in N\backslash T\cap S} \overline{v}_i + \left[\sum_{i\in N} t_i(\mathbf{x}^*_N(\theta)) - r(\theta, \mathbf{x}^*_N(\theta))\right], \quad (195) \end{aligned}$$

by Lemma 3(a) and equation (177). Moreover, from truthfulness with respect to the equilibrium values we know that:

$$\overline{v}_i \geq g_i(\mathbf{x}) - t_i(\mathbf{x}), \tag{196}$$

all $\mathbf{x}$, and all $i \in N \backslash T \cap S$. We also know that $t_i(\mathbf{x}) \geq 0$, all $i \in T \cap S$. Therefore assumption (192) can be rewritten as:

$$\begin{aligned} &\sum_{i \in N \backslash T \cap S} g_i(\mathbf{x}) + \sum_{i \in N \backslash S} t_i(\mathbf{x}) - r(\theta, \mathbf{x}) \\ &\qquad > \sum_{i \in N \backslash T \cap S} (g_i(\mathbf{x}) - t_i(\mathbf{x})) - \sum_{i \in T \cap S} t_i(\mathbf{x}) + \left[ \sum_{i \in N} t_i(\mathbf{x}_N^*(\theta)) - r(\theta, \mathbf{x}_N^*(\theta)) \right]. \end{aligned}$$

Collecting terms, we obtain:

$$\sum_{i \in N} t_i(\mathbf{x}) - r(\theta, \mathbf{x}) > \sum_{i \in N} t_i(\mathbf{x}_N^*(\theta)) - r(\theta, \mathbf{x}_N^*(\theta)), \tag{198}$$

a contradiction. □

*Proof by induction* We can now proceed by induction on the number of principals $n$, to demonstrate that for every common agency game $\Gamma$:

1. All (strictly) coalition-proof Nash equilibria yields an efficient action;
2. All (strictly) coalition-proof Nash equilibria yield a vector of principals' payoffs in $\mathcal{V}_\Gamma(\theta, \mathbf{x}^*(\theta))$, the Pareto frontier of $V_\Gamma(\theta, \mathbf{x}^*(\theta))$;
3. All truthful Nash equilibria are (strictly) coalition-proof.

**Step 0:** For $n = 1$, the three assertions are true.

By definition, for $n = 1$ a (strictly) coalition-proof Nash equilibrium is a Nash equilibrium. It is well known that, when the agent is risk-neutral and contracting takes place *ex-ante*, the optimal incentive contract implements the first-best. For any realization of the stochastic variable $\theta$, there is efficiency that is $\mathbf{x}^{\circ}(\theta) \in X^{*}(\theta)$ (point 1) and no rent is left to the agent that is $\overline{v}_1 = \max_{\mathbf{x}} \{g_1(\mathbf{x}) - r(\theta, \mathbf{x})\}$ (point 2). Moreover, a truthful Nash equilibrium is a Nash equilibrium, hence for $n = 1$ it is a (strictly) coalition-proof equilibrium (point 3).

**Step 1:** Suppose that all three assertions are true for every common agency game with no more than $n-1$ principals and consider the case of $n$ principals. Consider a coalition-proof Nash equilibrium as represented by the pair $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$. By definition, the pair $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$ is a strictly self-enforcing profile of strategies. This explains the relevance of the following claim.

**Claim 1:** *Let $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$ be a strictly self-enforcing strategy profile. Then for all $S \subseteq N$, and for all realizations of the stochastic variable $\theta$, the vector of principals' payoffs satisfies the inequalities:*

$$\sum_{i \in S} \overline{v}_i \leq \sum_{j \in N} g_j(\mathbf{x}^{\circ}(\theta)) - r(\theta, \mathbf{x}^{\circ}(\theta)) - W(\theta, N \backslash S). \tag{199}$$

**Proof:** Let $S \subset N$. By definition of a strictly self-enforcing strategy profile, the pair $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$ is a strictly coalition-proof Nash equilibrium of the restricted agency game:

$$\Gamma^{N \backslash S} = \left[\Theta, X, F, r(.) - \sum_{i \in S} t_i(.), \{g_i(.)\}_{i \in N \backslash S}\right].$$

By the induction hypothesis (1), $\mathbf{x}^{\circ}(\theta)$ is an efficient action of the restricted game $\Gamma_{N \backslash S}$, that is:

$$\mathbf{x}^{\circ}(\theta) \in \arg\max_{\mathbf{x} \in X} \left\{ \sum_{j \in N \backslash S} g_j(\mathbf{x}) - \left[ r(\theta, \mathbf{x}) - \sum_{j \in S} t_j^{\circ}(\mathbf{x}) \right] \right\}. \tag{200}$$

It follows that, for all $S \subset N$, all $\mathbf{x} \in X$, and all realizations of the stochastic variable $\theta$, we have:

$$\sum_{j \in S} t_j^{\circ}(\mathbf{x}^{\circ}(\theta)) - \sum_{j \in S} t_j^{\circ}(\mathbf{x}) \geq \left[ \sum_{j \in N \backslash S} g_j(\mathbf{x}) - r(\theta, \mathbf{x}) \right] - \left[ \sum_{j \in N \backslash S} g_j(\mathbf{x}^{\circ}(\theta)) - r(\theta, \mathbf{x}^{\circ}(\theta)) \right]. \tag{201}$$

Since $t_j^\circ(\mathbf{x}) \geq 0$ (by assumption), it can be dropped from the previous formula. Multiplying the inequality by $(-1)$ and adding $\sum_{j\in S} g_j(\mathbf{x}^\circ(\theta))$ on both sides, one gets:

$$\sum_{j\in S} g_j(\mathbf{x}^\circ(\theta)) - \sum_{j\in S} t_j^\circ(\mathbf{x}^\circ(\theta)) \leq \left[\sum_{j\in N} g_j(\mathbf{x}^\circ(\theta)) - r(\theta, \mathbf{x}^\circ(\theta))\right] - \left[\sum_{j\in N\setminus S} g_j(\mathbf{x}) - r(\theta, \mathbf{x})\right]. \tag{202}$$

This holds true in particular for $\mathbf{x} = \mathbf{x}^*_{N\setminus S}(\theta)$, so that:

$$\begin{aligned}\sum_{j\in S} \overline{v}_j &= \sum_{j\in S} \left[g_j(\mathbf{x}^\circ(\theta)) - t_j^\circ(\mathbf{x}^\circ(\theta))\right] \\ &\leq \sum_{j\in N} g_j(\mathbf{x}^\circ(\theta)) - r(\theta, \mathbf{x}^\circ(\theta)) - W(\theta, N\setminus S),\end{aligned} \tag{203}$$

for all $S \subset N$. Now for the case $S = N$, it is clear that:

$$\sum_{j\in N} t_j^\circ(\mathbf{x}^\circ(\theta)) - r(\theta, \mathbf{x}^\circ(\theta)) \geq -\min_{\mathbf{x}\in X} r(\theta, \mathbf{x}), \tag{204}$$

hence:

$$\sum_{j\in N} \left[g_j(\mathbf{x}^\circ(\theta)) - t_j^\circ(\mathbf{x}^\circ(\theta))\right] \tag{205}$$

$$\leq \sum_{j\in N} \left[g_j(\mathbf{x}^\circ(\theta)) - t_j^\circ(\mathbf{x}^\circ(\theta))\right] + \sum_{j\in N} t_j^\circ(\mathbf{x}^\circ(\theta)) - r(\theta, \mathbf{x}^\circ(\theta)) + \min_{\mathbf{x}\in X} r(\theta, \mathbf{x}), \tag{206}$$

$$\leq \sum_{j\in N} g_j(\mathbf{x}^\circ(\theta)) - r(\theta, \mathbf{x}^\circ(\theta)) + \min_{\mathbf{x}\in X} r(\theta, \mathbf{x}), \tag{207}$$

$$\leq \sum_{j\in N} g_j(\mathbf{x}^\circ(\theta)) - r(\theta, \mathbf{x}^\circ(\theta)) - W(\theta, \{\varnothing\}). \tag{208}$$

It follows that all coalition-proof Nash equilibrium as characterized by the pair $(\mathbf{t}^\circ, \mathbf{x}^\circ(\theta))$ lead to equilibrium payoffs in $V_\Gamma(\theta, \mathbf{x}^*(\theta))$. $\square$

The proof of Theorem 3 then continues by showing that the induction assumptions (3), (2), and (1) hold true with $n$ principals. To prove the induction assumption (3), namely that each truthful Nash equilibrium is a strictly coalition-proof Nash equilibrium, we need to demonstrate that (a) the pair $(\mathbf{t}^\circ, \mathbf{x}^\circ(\theta))$ is strictly self-enforcing, and (b) there is no other strictly self-enforcing profile that yields a higher payoff to all players.

**(a)** To demonstrate that the truthful Nash equilibrium as characterized by the pair $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$ is strictly self-enforcing, we must show that, for any $S \subset N$, the pair $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$ is a strictly coalition-proof Nash equilibrium of the restricted game:

$$\Gamma^{S} = \left[\Theta, X, F, r(.) - \sum_{i \in N \backslash S} t_i(.), \{g_i(.)\}_{i \in S}\right].$$

By the induction assumption (3), we know that, when there are up to $n-1$ principals, a truthful Nash equilibrium is strictly coalition-proof. It is thus sufficient to prove that, for $S \subset N$, the pair $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$ is a truthful Nash-equilibrium of the restricted game $\Gamma^{S}$. From Theorem 2, we know that any vector of payoffs in $\mathcal{V}_{\Gamma}(\theta, \mathbf{x}^{*}(\theta))$ can be supported by a truthful Nash equilibrium of the game $\Gamma$. We know also that $\mathbf{x}^{\circ}(\theta)$ is in the set $X^{S}(\theta)$ of efficient actions of the reduced agency game (from Proposition 6), and that the vector of payoffs is in the Pareto-frontier of the restricted game $\mathcal{V}_{\Gamma}^{S}(\theta, \mathbf{x}^{*}(\theta))$ (from Proposition 7). To summarize, the induction assumption (3), Theorem 2 and Propositions 6-7 allow us to show that a truthful Nash-equilibrium of the agency game gives rise to a pair $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$ which is strictly self-enforcing.

**(b)** To demonstrate that the pair $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$ is not Pareto dominated by any other self-enforcing pair, consider a self-enforcing pair $(\mathbf{t}', \mathbf{x}'(\theta))$ that yields to the vector of payoffs $\{\overline{v}_i'\}_{i \in N}$. From Claim 1, we know that, for any $S \subseteq N$, we have:

$$\sum_{i \in S} \overline{v}_i' \leq \sum_{j \in N} g_j\left(\mathbf{x}'(\theta)\right) - r\left(\theta, \mathbf{x}'(\theta)\right) - W(\theta, N \backslash S), \tag{209}$$

hence:

$$\sum_{i \in S} \overline{v}_i' \leq W(\theta, N) - W(\theta, N \backslash S), \tag{210}$$

which says that the payoff vector $(\overline{v}_i')_{i \in N}$ is in $V_{\Gamma}(\theta, \mathbf{x}^{*}(\theta))$. From Theorem 2, the equilibrium payoffs $(\overline{v}_i)_{i \in N}$ of the truthful Nash equilibrium are in $\mathcal{V}_{\Gamma}(\theta, \mathbf{x}^{*}(\theta))$, the Pareto frontier of $V_{\Gamma}(\theta, \mathbf{x}^{*}(\theta))$. It follows that $(\overline{v}_i')_{i \in N}$ cannot Pareto dominate the vector of equilibrium payoffs $(\overline{v}_i)_{i \in N}$. This concludes the proof of the induction assumption (3) for the case of $n$ principals. Any truthful Nash-equilibrium of the agency game gives rise to a pair $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$ which is strictly self-enforcing and which cannot be Pareto dominated, *i.e.* to a strictly coalition-proof Nash equilibrium.

We now turn to the proof of the induction assumption (2) for the case of $n$ principals. Previous argument led us to show that, for any self-enforcing pair $(\mathbf{t}', \mathbf{x}'(\theta))$ (hence any strictly coalition proof Nash equilibrium).give rise to vector of payoffs $(\overline{v}_i')_{i\in N}$ that belongs to $V_{\Gamma}(\theta, \mathbf{x}^*(\theta))$. Assume it does not belong to $\mathcal{V}_{\Gamma}(\theta, \mathbf{x}^*(\theta))$, the Pareto frontier $V_{\Gamma}(\theta, \mathbf{x}^*(\theta))$ that is it belongs to the interior of $V_{\Gamma}(\theta, \mathbf{x}^*(\theta))$. This does mean that there exist a vector $(\overline{v}_i)_{i\in N}$ in $\mathcal{V}_{\Gamma}(\theta, \mathbf{x}^*(\theta))$, the Pareto frontier $V_{\Gamma}(\theta, \mathbf{x}^*(\theta))$ that Pareto dominates the vector $(\overline{v}_i')_{i\in N}$. By Theorem 2, we know that this vector of payoffs $(\overline{v}_i)_{i\in N}$ can be obtained by a truthful Nash equilibrium, say the pair $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$. From induction assumption (3) we just proved, it is also a strictly coalition-proof Nash equilibrium. Hence the pair $(\mathbf{t}^{\circ}, \mathbf{x}^{\circ}(\theta))$ is self-enforcing. A contradiction since if $(\mathbf{t}', \mathbf{x}'(\theta))$ is a strictly coalition-proof Nash equilibrium, it cannot be dominated by any other one. This proves the induction assumption (2) for the case of $n$ principals.

To prove the induction assumption (1), consider a strictly coalition-proof Nash equilibrium as given by the pair $(\mathbf{t}', \mathbf{x}'(\theta))$, and suppose that the action $\mathbf{x}'(\theta)$ is not efficient, that is $\sum_{j\in N} g_j(\mathbf{x}'(\theta)) - r(\theta, \mathbf{x}'(\theta)) < W(\theta, N)$. Then recall that Claim 1 implies that, for all $S \subseteq N$ :

$$\begin{aligned}
\sum_{i\in S} \overline{v}_i' &\leq \sum_{j\in N} g_j\left(\mathbf{x}'(\theta)\right) - r\left(\theta, \mathbf{x}'(\theta)\right) - W\left(\theta, N\backslash S\right), \\
&< W(\theta, N) - W\left(\theta, N\backslash S\right), \qquad (211)
\end{aligned}$$

which means that $(\overline{v}_i')_{i\in N}$ belongs to the interior of $V_{\Gamma}(\theta, \mathbf{x}^*(\theta))$. Again, this is a contradiction, since there exists a vector of payoff $(\overline{v}_i)_{i\in N}$ that Pareto dominates $(\overline{v}_i')_{i\in N}$ and belongs to $\mathcal{V}_{\Gamma}(\theta, \mathbf{x}^*(\theta))$. By Theorem 2, this vector can be obtained by a truthful Nash equilibrium, and from the induction assumption (3), it is also a strictly coalition-proof Nash equilibrium. This constitutes a contradiction, since if $(\mathbf{t}', \mathbf{x}')$ is a strictly coalition-proof Nash equilibrium it cannot be dominated by any other one. This proves the induction assumption (1) for the case of $n$ principals. $\square$