Working paper

LONG-TERM PROCUREMENT UNDER UNCERTAINTY: OPTIMAL DESIGN AND IMPLICATIONS FOR RENEGOTIATION AND TENDER PROCEDURES[1]

May 23, 2014

Malin Arve and David Martimort

We study the effect of risk aversion on the optimal procurement of a basic service with an *uncertain* add-on. We identify two new effects that influence the optimal contract. First, a marginal utility of income effect reduces the cost of asymmetric information and increases the level of the basic service. However, we show that for incentive reasons some risk has to be transferred to the risk-averse firm and this leads to a risk effect that reduces the level of the add-on. We discuss the contractual externality between these two effects and its implication on renegotiation-proofness. We also show how these effects carry over to a competitive environment for a fixed level of the service.

Keywords: Procurement, asymmetric information, uncertainty, change orders, risk aversion.

[1]We thank participants at the *Mannheim Workshop on Procurement and Contracts*, the *PSE Workshop on Financing Investments in Times of Crisis*, the *MaCCI Competition and Regulation Day*, the *Fundação Getulio Vargas Workshop on Regulatory Environment and Institutions in Public Procurement, the* 18^{th} *SFB Meeting and the Padua Workshop on How Governance Complexity and Financial Constraints affect Public-Private Contracts", seminar participants at the University of Mannheim, University of Tilburg and Humboldt-Universität zu Berlin, Bernard Caillaud, Vinicius Carrasco, Elisabetta Iossa, Juan José Ganuza, Susanne Goldlücke, Andras Niedermayer and Frank Rosar for extremely valuable comments. Financial support by Deutsche Forschungsgemeinschaft (DFG), SFB/TR-15, is greatly acknowledged. All errors are ours.*

[a]University of Mannheim, marve@mail.uni-mannheim.de

[b]Paris School of Economics-EHESS, david.martimort@parisschoolofeconomics.eu

1. INTRODUCTION

Contracts for public utilities such as delegated management in water, sanitation and transportation are usually long-term contracts and can last for up to several decades. In these types of long-term contracts, project managers expect a certain amount of ex post adaptations. Regardless of how well the project is planned and executed some ex post adaptations are too be expected[1]. In a report, the National Audit Office acknowledges the fact that over time UK Public Finance Initiative (PFI) deals need to be changed to meet inevitable changes in public services (NAO (2008)).

In this paper we are interested in the effects of changes to the initial contract. In particular, we consider that a contract consists of both a basic service and additional future work where there is some uncertainty surrounding the additional work. We are interested in the effects of these *uncertain* add-ons on the design of the optimal procurement contract. Our analysis is undertaken in an environment with dynamic adverse selection. Ex post adverse selection regarding the add-on generates some risk and we are interested in understanding how this risk affects the firm's behavior and what the consequences for the design of the optimal procurement contract are.

In project management, any modifications to the initial contract comes is the form of a *change order*[2]. Roughly speaking, change orders can be classified into two categories; The first one is related to changes to the good or service itself. But changes can also be related to additional work that has to be provided and that was not clearly specified in the initial contract or that is related to characteristics that were unobservable at the initial time of contracting. The nature, but also the consequences of these two categories of change orders are very different. In this paper we focus on the latter. In fact, the NAO points out that the majority of changes to UK PFI seem to be additions rather than direct changes to the type or level of the service provided (NAO (2008)). Another example of contracts where additional work was required is the Getty Center Art Museum in Los Angeles which had to be redesigned due to site conditions that were hard to anticipate ex ante[3]. In fact, new geological information about the construction site and reassessment of design standards led to additional work but did not per se change the scope of the project[4]. In

[1] Kerzner (2013).

[2] Meredith and Mantel (2009).

[3] Bajari and Tadelis (2001) and Chakravarty and MacLeod (2009).

[4] This issue is also acknowledge in a study about relational contrast and subcontracting in highway procurement by Gil and Marion (2013). Legal issues related to change order and the issue of when a change can be part of the initial contract in public procurement are further discussed in Hartlev and Liljenbøl (2013).

fact, we consider that the basic service has some important features of a durable good or service and therefore its level will remain fixed over the entire contracting period.

An important feature of our analysis is that we explicitly model firm's risk attitude. Economic theory in general and, the literature on procurement and regulation more specifically, has mostly taken the modeling short-cut of considering firms to be risk-neutral. This view is based on two implicit assumptions; One is related to the firm's relationship with its outside investors while the other is concerned with its internal organization. The first implicit assumption is that firms have perfect access to financial markets. This means that equityholders have enough financial instruments to diversify away any risk they may bear by owning stocks in a firm. As recognized by Leland and Pyle (1977), this is not the case when firms have private information on the risk they must bear. In this case, outside investors who are solicited for financing new projects will then ask for credible signals on the quality of the venture. One such credible signal is the amount of risk kept by existing owners. The second implicit assumption is that firms are not subject to any agency problems. By stressing the separation between ownership and control, the agency literature[5] has pointed out the existence of an important trade-off between risk and insurance that leads the firm to being imperfectly diversified.

A similar trade-off arises when the firms itself subcontracts with independent contractors, and these relationships get plagued by agency problems. For instance, Kawasaki and McMillian (1987), Asanuma and Kikutani (1992) and Yun (1999) have applied a simple principal-agent framework to study subcontracting and risk-sharing in the relationship between manufacturers and contractors in Japan and Korea and empirically found that contractors are indeed risk averse.[6] There is thus little doubt that risk matters for firms' behavior and that this ingredient should be part of a more complete theory. This is especially true in the procurement contexts which are of prime interest for our study. Long-term relationships under changing conditions on demand and costs that require new developments and investments beyond existing services and assets and might also involve substantial subcontracting offer a perfect example for the two ingredients that justify introducing risk aversion as a modeling ingredient.

[5] Ross (1973), Holmström (1979) and Shavell (1979).

[6] Asplund (2002) points out that there might be other reasons that make firms act *as if* they were risk averse. These reasons include also factors such as liquidity constraints, costly financial distress, and non-linear tax-system. For instance the project might be so important that any loss or gain related to it has a huge impact on the firms overall profits and survival.

Overview. By introducing risk aversion and an uncertain add-on into a procurement environment with dynamic adverse selection, we identify two important effects that influence the optimal procurement contract.

First, we show how the *Marginal Utility of Income Effect* allows the procurement agency to relax first-period incentives by shifting payments to the second period. A higher second-period payment is evaluated at a lower marginal utility of income and therefore giving the firm a given level of utility is less costly to the procurement agency. This effect is present even when there is no risk transfer to the risk averse firm. By shifting payments to the second period, the procurement agency can exploit that the marginal utility of income is none constant under risk aversion. When the second-period marginal utility of income decreases, the firm's incentives to exaggerate first-period costs are reduced and truthtelling can be implemented at a lower cost for the procurement agency. This leads to less distortions on the basic service.

Second, we turn to the *Risk Effect*. Because of asymmetric information, the procurement agency has to make the risk averse firm bear some risk. This is costly both because it requires a risk premium to be paid to the risk averse firm and because it increases the marginal utility of income in the second period, thereby making first-period incentive compatibility more costly. The level of the risk can be reduced by reducing the level of the add-on. Finally, we analyze the interaction between these two effects and show that in this environment, even without any technological linkage, there is still a contractual externality between the two effects which influences level of the basic service and the add-on.

An important assumption in our environment is that the procurement agency can fully commit to a long-term contract. We argue that under very general conditions, the optimal long-term contract that we derive is not renegotiation-proof. Because of the contractual externality, ex ante the procurement agency wants to reduce the level of the add-on to influences the firm's first-period incentives. However, once in the second period, these extra distortions are no longer needed and if possible, the agency may want to renegotiation the contract. Therefore the optimal long-term contract might not be *sequentially optimal*. Hence exploiting the contractual externality that arises because of firms' risk aversion is only possible when there is a strong commitment not to renegotiate the original contract. Preventing interim renegotiation of long-term contract may lead to a reduced overall cost of long term contracting because it allows a better use of ex ante contractual spillovers. This is especially important in times of financial instability in which firms might exhibit a higher degree of risk aversion[7]. Notice however that when preferences exhibit constant absolute risk aversion (CARA), the optimal long term contract under full commitment is indeed renegotiation-proof. Furthermore, with CARA preferences, restricting attention to spot contracts comes at no loss of generality compared to the optimal long-term contract derived previously.

In an extension we allow for competition for the procurement contract and show how risk aversion affects the reserve price. We show that since the Marginal Utility of Income Effect eases the ex ante incentive problem, it also tends to make bids more acceptable. However, when introducing risk, the Risk Effect make bids less acceptable because introducing risk is costly both in terms of risk premium and information rents.

Literature review. The existing literature on risk aversion in adverse selection settings that have stimulated much of the procurement and regulation literatures is sparse.[8] This scarcity is probably best explained by the technical complexity of the analysis when risk aversion and incentive constraints interact. Salanié (1990) illustrates this complexity in his study of an adverse selection problem where contracting takes place ex ante, i.e., before the agent gets private information about his cost parameter. With CARA preferences, he shows how bunching can easily become an important issue. Laffont and Rochet (1998) instead focus on ex post participation and show that risk aversion induces greater distortions, lower informational rents and, again, possibly some bunching.[9] However, their assumptions differ from ours[10]. Furthermore. they only provide general results for the two-type case. In the continuous-type case, they restrict attention to CARA preferences.

Long-term procurement contracts are mostly allocated through tenders. Auctions and, more generally bargaining procedures, are competitive environments for which risk aversion has been widely documented both in experimental works (see Kagel (1995) for a survey) and econometrically (Athey and Levin (2001)). This suggests that the assumption of risk neutrality is not always appropriate. Eső and White (2004) take this issue seriously and show that decreasingly risk averse bidders may shade their bids for pure "precautionary bidding" reasons.[11] In a dynamic bargaining context, White (2008) shows

[7]Due to the increased costs and risks on the financial markets.

[8]Laffont (1994) and Armstrong and Sappington (2007).

[9]See also Maskin and Riley (1984) and Matthews (1984) for the case of auctions.

[10]For instance, they impose ex post individual rationality and this assumption considerably simplifies the modeling.

[11]There is a related strand of the auction literature that studies bidding behavior under uncertainty.

that a similar precautionary behavior may lead bidders to be more patient. Our procurement model differs in many respects from these papers but we share the idea that incentive constraints and thus optimal contracts depend on how much risk a privately informed agent will bear. In the current paper, this risk is endogenously determined as the solution to an agency problem in future periods of the relationship.[12]

From a more theoretical viewpoint, our paper also contributes to the dynamic mechanism design literature which goes beyond the study of long-term relationships in the specific procurement context under scrutiny in this paper. This literature stresses the value of history in long-term relationships, especially when types are serially correlated and/or when (for various reasons) current projects impact future technological frontiers. An important idea highlighted by this strand of literature is that commitment to future actions may help to screen current private information and it thereby reduces rents. As a by-product, contracts are not immune to renegotiation. However, following Baron and Besanko (1984) and Pavan, Segal and Toikka (2013), this literature mostly considers risk-neutral agents and has thus overlooked the impact that their current reporting strategies might have on future marginal utility of income. This is especially likely to be an important aspect when some production stages are long-lasting. In these case, the principal might find it attractive to distort future contracts to reduce the agent's marginal utility of income and improve rent extraction. We show that this effect creates a new value of commitment that holds even when types are independently drawn over time. With risk-neutrality, and more generally with CARA preferences, we demonstrate that the value of commitment disappears. The optimal contract is sequentially optimal (renegotiation-proof) and could even be implemented through some form of spot contracts.

Chakravarty and MacLeod (2009) show that the American standard form of construction contract can in fact be viewed an efficient mechanisms for implementing uncertain or risky building projects given existing US legal rules. However, they also point out that this might not be the case in other jurisdictions. We abstract from current (US or other) practices and study the optimal mechanism in a general contracting setting with a basic

Calvares et al. (2004) and Burguet et al. (2012) study the effect of cost uncertainty on firm's bidding behavior. They show that under limited liability, financially weak firms tend to bid more aggressively. We abstract from this effect and focus on the consequences of ex ante risky change orders when contracting with risk-averse firms. We unveil two new effects that are not present in Calvares et al. (2004) and Burguet et al. (2012).

[12] Faure-Grimaud and Martimort (2003) and Strausz (2011) study a very specific risk borne by regulated firms, the political risk coming from fluctuations in the preferences of political principals in charge of designing regulatory policies.

service and an uncertain add-on.

Organization. Our model is presented in Section 2. The set of incentive-feasible allocations in our dynamic context is described in Section 3. Section 4 provides some important partial results. These partial results prepare the stage for the more complete analysis of optimal contracts that is undertaken in Section 5. Section 6 discusses whether the optimal contract is sequentially optimal or not and provides condition under which it can be implemented through a sequence of contracts, one for the basic and long-lasting service and another one for the add-on. Section 7 shows how tender procedures must be modified to account for risk. Proofs are relegated to the Appendix.

2. THE MODEL

Consider the following procurement context: A government agency contracts with a firm for the provision of a public service. The costs and benefits from this service accumulates over two periods. To be more precise, this service includes a basic (or main) service provided in quantity q in each period but also some additional service (add-on) which is provided in quantity x but only needs to be delivered in the second period.[13] The idea is that the add-on represents long-term contractible variables associated with the service (refined specification, incremental services for new segments of demand, further developments of a prototype in defense procurement, etc.). The exact specifications required for these additional services are not be completely known by the two parties at the time of contracting. This uncertainty around the add-on may put the firm's returns at risk.[14] In the sequel, we will be particularly interested in the impact of this risk on contract design.

• Technology and preferences. The basic service yields a gross surplus $S(q)$ in each period. Motivated by the idea that this basic service is the choice of a capacity, the fixed size of a network, or the basic version of a long-term durable good, we assume

[13] To make the model more realistic, we could allow for the add-on only to occur with a certain probability. Since this does not qualitatively change our results and in order to keep the model as simple as possible, we assume that the add-on will be required for sure.

[14] We focus on the case of bundling of the basic service and the add-on. There are several reasons for this. First of all, in many long-term contracts such as Public-Private Partnerships, the bidding consortia have a rather ephemeral life and in later stages of the contract only the winning consortium is still available for providing the add-on. Another reason could be that the competitors have too high costs because of project specific developments during the first stage or simply because it is not physically possible to have two different providers of the basic service and the add-on.

that the quantity q is chosen once for all and does not vary over time. The firm provides this service at a constant marginal cost θ. The function S is increasing and concave ($S' > 0, S'' < 0$) with $S(0) = 0$ and satisfies the Inada condition $S'(0) = +\infty$.[15] The procurement contract also specifies the level x of the second-period add-on. That add-on is produced at constant marginal cost β. The principal enjoys a gross surplus $V(x)$ from consuming x units of this add-on. The function V is increasing, concave ($V' > 0, V'' < 0$) with $V(0) = 0$ and $V'(0) = +\infty$.

The payments for the basic service are denoted by $t(\theta)$ in period 1 and $t(\theta) + y(\theta)$ in period 2. $t(\theta)$ is the fixed per-period payment for the basic service $q(\theta)$ while the payment $y(\theta)$ represents an extra premium for the second-period. Indeed we show that payments for the basic service may not be stationary. On top of allowing for flexible payments for the basic service, we denote by $p(\theta, \beta)$ the second-period price for the add-on. Similar notations apply to the levels of the basic service and the add-on which are respectively denoted by $q(\theta)$ and $x(\theta, \beta)$.

Denoting by $1 - \delta$ and δ the relative weights of the first and second period, respectively, and normalizing intertemporal payoffs accordingly, the agency's expected gains from trading with a firm of type θ for the basic service and the add-on be written as:

$$S(q(\theta)) - t(\theta) - \delta y(\theta) + \delta E_\beta \left[V(x(\theta, \beta)) - p(\theta, \beta) \right].$$

Defining $u(\theta) = t(\theta) - \theta q(\theta)$ and $u_2(\theta, \beta) = p(\theta, \beta) - \beta x(\theta, \beta)$ as the firm's first- and second-period profit respectively, the agency's gains from trade can be rewritten as:

$$S(q(\theta)) - \theta q(\theta) - u(\theta) - \delta y(\theta) + \delta E_\beta \left[V(x(\theta, \beta)) - \beta x(\theta, \beta) - u_2(\theta, \beta) \right].$$

This expression already highlights the trade-off between efficiency and rent extraction that characterizes optimal contracting under informational asymmetries.

We are particularly interested in the consequences on contract design of introducing uncertainty on the add-on and therefore on the second-period returns. Analyzing this requires us to move away from the standard models of procurement that assume that the firm is risk neutral. We thus assume that, from an ex ante point of view, the firm is risk-averse and evaluates the second-period risky returns using a von Neuman-Morgenstern

[15] These latter two conditions ensure that "shutting-down" production even with the least efficient service provider is not a valuable option for the agency. This assumption simplifies our modeling and is without loss of economic insight. Section 7 relaxes this assumption in the specific case of a service of a fixed quantity.

utility function v which is increasing and concave, $v' > 0$, $v'' \leq 0$.[16] We also impose the following normalizations: $v(0) = 0$ and $v'(0) = 1$. Risk aversion should be viewed as a proxy for financial constraints that may for instance limit the firm's access to the capital market. The normalization $v'(0) = 1$ ensures that the firm's marginal utility of income at zero wealth is the same in both periods.[17] From the agency's viewpoint, the cost of increasing the firm's revenues by one extra dollar is thus the same in both periods.

We can now write the firm's intertemporal payoff as

$$(1-\delta)u(\theta) + \delta E_\beta \left(v\left(u(\theta) + y(\theta) + u_2(\theta,\beta)\right)\right).$$

• Information. The agent has private information on the efficiency parameter θ at the time of contracting. This variable is drawn from a (common knowledge) cumulative distribution $F(\cdot)$ with an atomless and everywhere positive density $f(\theta) = F'(\theta)$ on the support $\Theta = [\underline{\theta}, \bar{\theta}]$. As is standard in the screening literature[18], we a monotone hazard rate property:

Assumption 1 *Monotone hazard rate property:*

$$\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right) \geq 0 \textit{ for all } \theta \in \Theta.$$

To model ex ante uncertainty on the net gains from the add-on, we also assume that there is ex ante *symmetric but incomplete* information on the cost parameter β. In the second period, before providing the add-on, the firm becomes privately informed about the realized value of β. To maintain a tractable analysis, we consider the case where β is drawn from a discrete-support distribution on $\mathcal{B} = \left\{\underline{\beta}, \bar{\beta}\right\}$ (with $\Delta\beta = \bar{\beta} - \underline{\beta} > 0$) with respective probabilities ν and $1-\nu$, where $\nu \in (0,1)$. This distribution is common knowledge.

[16]There are many ways of modeling risk aversion and what affects risk aversion. In general, the overall environment in which a firm operates might affect the firm's risk attitude. Here we abstract from many of these factors and assume that risk attitudes are only affected by the return and risk of the given period. One justification for why second-period risk aversion is not affected by first-period returns could simply be that first-period returns are redistributed to the shareholders in the first period and therefore the firm is facing a tight budget constraint. Risk aversion can then be seen as a consequence of this tight budget constraint and the interaction between investors and the firm. For a more general model of this issue (without risk aversion) see Martimort, Pouyet and Sand-Zantman (2014).

[17]This implies that under complete information, it is not optimal to transfer payments between periods.

[18]See Bagnoli and Bergstrom (2005) and Chapter 3 in Laffont and Martimort (2001).

It is important to stress that first- and second-period cost parameters are independent draws so that there is no technological linkage across periods. When deriving the optimal dynamic contract in this environment, any departure from the optimal spot contracts can thus only be explained by the fact that externalities between first- and second-period contracting stages nevertheless arise and that these are due to the firm's risk-averse behavior. Much of our analysis below consists in unveiling conditions under which this is the case.[19] This is also the reason why we do not consider any correlation between the surplus of the basic service and the add-on.

• Incentive mechanisms. The agency commits to a long-term contract that regulates the basic service and the add-on over both periods. There are several justifications behind this assumption. First, in some contexts to which our model applies, like for instance PPP contracts, public officials commit over periods up to thirty years but include adaptation clauses to react to changes in the environment.[20] Changes that are outside these clauses might be limited by law.[21] Second, reputation-like arguments may force the agency to stick to her initial commitment even if renegotiating these long-term agreements could appear as an attractive option as time unfolds. Third, focusing on the full commitment scenario allows us to find an upper bound on what long-term contracting can achieve and to unveil intertemporal links across periods that arise under asymmetric information. We relax this assumption and discuss the issue of renegotiation in Section 6.

From the (dynamic version of the) Revelation Principle (Myerson (1986), Baron and Besanko (1984)), there is no loss of generality in restricting the analysis to incentive-compatible direct revelation mechanisms. These mechanisms stipulate payments and outputs in each period as a function of the firm's report of his type at this date and, possibly, the past history of reports. Such mechanisms are thus of the form $\left\{t(\hat{\theta}), y(\hat{\theta}), q(\hat{\theta}), p(\hat{\theta},\hat{\beta}), x(\hat{\theta},\hat{\beta})\right\}_{\hat{\theta}\in\Theta,\hat{\beta}\in\mathcal{B}}$ where $\hat{\theta}\in\Theta$ and $\hat{\beta}\in\mathcal{B}$ are the firm's announcements of his costs for the basic service and the add-on respectively.[22]

[19]Extensions of our framework could allow for such technological linkages, for instance, by introducing serial correlation across types in both periods, or learning by doing effects directly embedded into the expression of second-period costs. The strand of literature studying sequential screening with correlation across two types goes back to Baron and Besanko (1984) and recent work on this topic includes Courty and Li (2000).

[20]We assume that these clauses completely pin down the contract related to the add-on.

[21]In the European Union, add-ons that are outside the scope of the initial contract might be seen as a violation of Art. 101 of the Treaty on the Functioning of the European Union.

[22]For technical reasons, we will assume that all feasible q and x are bounded above by some levels Q and X, respectively. These upper bounds are large enough to contain the efficient domain for the

• Timing. The contracting game unfolds as follows:

1. The firm privately learns its cost parameter θ for the basic service.
2. The agency offers a long-term contract $\mathcal{C}$. If the firm refuses, the game ends with parties getting their reservation payoffs which are, without loss of generality, normalized to zero.
3. If the firm accepts the offer, it announces its first-period report $\hat{\theta}$. This report determines both the level of the basic service $q(\hat{\theta})$ and payments $t(\hat{\theta})$ and $t(\hat{\theta})+y(\hat{\theta})$ for this service in each period.
4. The second period cost β is realized and the firm privately learns its value. The firm then reports $\hat{\beta}$, provides the corresponding level of the add-on, $x(\hat{\theta},\hat{\beta})$, and receives the payment $p(\hat{\theta},\hat{\beta})$.

• Preliminaries. In the sequel, detailed (albeit standard) properties of v play a key role in the characterization of the optimal contract. Beyond the familiar assumptions of being increasing and concave, we also assume that v satisfies:

Assumption 2

$$\forall z, \quad v'''(z) \geq 0 \text{ and } v^{(4)}(z) \leq 0.$$

The first condition captures the firm's prudence[23] while the last one is generally referred to as temperance[24]. Since these requirements are less common, we point out whenever they are required for our results to hold.

For the remainder of the paper, it will be useful to express the firm's intertemporal payoff in a more compact form. To this end, a first useful step, is to add to a fixed-return project that yields a profit z, a second random project with zero mean and that yields $(1-\nu)\varepsilon$ with probability ν and $-\nu\varepsilon$ with probability $1-\nu$, and analyze the impact of such changes on the firm's expected utility. Let us thus introduce a second-period utility function $w(z,\epsilon)$ defined over wealth and risk levels (z,ε):

$$w(z,\epsilon) \equiv \nu v(z+(1-\nu)\varepsilon) + (1-\nu)v(z-\nu\varepsilon).$$

corresponding variable.

[23]The importance of prudence goes back to Leland (1968) and Sandmo (1970). The term prudence was coined in Kimball (1990). For a more general introduction to the literature on risk and uncertainty see Gollier (2004) and references therein. Prudence is sometimes referred to as downside risk aversion (Menezes et al (1980)).

[24]The importance of temperance, or outer risk aversion (Menezes and Wang (2005)), goes back to Kimball (1992).

The function w inherits some important properties from v. It is straightforward to check that w is increasing and concave in z while it remains decreasing in ε.[25] The firm's marginal utility of income in the second period is positive for a given level of second-period risk but decreases as the income level increases. Following Assumption 2, the firm exhibits a prudent behavior and its marginal utility of income in the second period instead increases in more uncertainty environments:

$$w_{z\varepsilon}(z,\varepsilon) = \nu(1-\nu)(v''(z+(1-\nu)\varepsilon) - v''(z-\nu\varepsilon)) \geq 0.$$

Intuitively, the firm is better able to cope with second-period risk at higher levels of wealth. This assumption captures the idea that at the margin agency costs in accessing financial markets may be lower for firms which are able to pledge more revenues from their basic service.[26]

For further references, let $\varphi(z,\varepsilon)$ be the second-period wealth level that guarantees to the firm z units of utility in the second period when risk at this date has variance $\nu(1-\nu)\varepsilon^2$:

$$z = w(\varphi(z,\varepsilon),\varepsilon).$$

In particular, observe that $\varphi(z,0) = h(z)$ where $h = v^{-1}$. It is also immediate that φ is increasing in z and ε.[27]

Let us denote by $z^*(\varepsilon)$ the solution to

$$w_z(\varphi(z^*(\varepsilon),\varepsilon),\varepsilon) = 1.$$

By definition, $z^*(\varepsilon)$ is the second-period utility level such that the marginal utility of income is just equal to its value in the first period (which is normalized to one). At this utility level and this level of risk, the firm is exactly indifferent between receiving one more dollar in the first or receiving it in the second period. Observe that:

$$\dot{z}^*(\varepsilon) = -\frac{w_z(\varphi(z^*(\varepsilon),\varepsilon),\varepsilon)}{w_{zz}(\varphi(z^*(\varepsilon),\varepsilon),\varepsilon)} H(\varphi(z^*(\varepsilon),\varepsilon),\varepsilon), \tag{2.1}$$

where the function H is defined for all pairs (z,ε) as:

$$H(z,\varepsilon) = w_{z\varepsilon}(z,\varepsilon) - \frac{w_{zz}(z,\varepsilon)w_\varepsilon(z,\varepsilon)}{w_z(z,\varepsilon)}.$$

[25] Indeed, we have: $w_z(z,\varepsilon) = \nu v'(z+(1-\nu)\varepsilon) + (1-\nu)v'(z-\nu\varepsilon) > 0$, $w_{zz}(z,\varepsilon) = \nu v''(z+(1-\nu)\varepsilon) + (1-\nu)v''(z-\nu\varepsilon) \leq 0$ and $w_\varepsilon(z,\varepsilon) = \nu(1-\nu)(v'(z+(1-\nu)\varepsilon) - v'(z-\nu\varepsilon)) \leq 0$ (since $v'' \leq 0$).

[26] Similarly, $v^{(4)} \leq 0$ also implies: $w_{zz\varepsilon}(z,\varepsilon) = \nu(1-\nu)(v'''(z+(1-\nu)\varepsilon) - v'''(z-\nu\varepsilon)) \leq 0$.

[27] Indeed, we have: $\varphi_z(z,\varepsilon) = \frac{1}{w_z(\varphi(z,\varepsilon),\varepsilon)} \geq 0$, and $\varphi_\varepsilon(z,\varepsilon) = -\frac{w_\varepsilon(\varphi(z,\varepsilon),\varepsilon)}{w_z(\varphi(z,\varepsilon),\varepsilon)} \geq 0$.

We will restrict attention to preferences such that $z^*(\varepsilon)$ increases with ε. This means that if second-period returns are more risky, the second-period utility increases in order to keep the marginal utility of income constant over time. To ensure that this property holds, we assume:

Assumption 3 *Generalized decreasing absolute risk aversion (GDARA):*

$H(z,\varepsilon) \geq 0$ *for all* (z,ε).

This condition which is imposed for all pairs (z,ε) implies that $z^*(\varepsilon)$ is a decreasing function. More intuition for Assumption 3 can be given by observing that, in the limit of small risk, the following approximation holds:

$$H(z,\varepsilon) \underset{\varepsilon\approx 0}{\approx} \nu(1-\nu)\varepsilon\left(v'''(z) - \frac{(v''(z))^2}{v'(z)}\right) = -\nu(1-\nu)\varepsilon v'(z)\frac{d}{dz}\left(-\frac{v''(z)}{v'(z)}\right).$$

Hence, in the limit of small risks, the fact $\dot{z}^*(\varepsilon) \geq 0$ simply follows from assuming decreasing absolute risk aversion, an assumption which is standard in the risk literature.[28]

More generally, H is the sum of two terms with opposite signs. The first one, $w_{z\varepsilon}(z,\varepsilon)$, is non-negative when the firm exhibits some prudent behavior. It simply means that more uncertainty increases the marginal utility of income, making it more valuable to transfer revenues towards the second period: an income effect. The second term, $-\frac{w_{zz}(z,\varepsilon)w_\varepsilon(z,\varepsilon)}{w_z(z,\varepsilon)}$, is negative, it captures the idea that more uncertain prospects decrease utility in the second period, making such transfers less attractive: a risk effect. Assumption 3 ensures that the income effect dominates. In the sequel, we will point out specifically when we rely on Assumption 3 for our results.

Example (CARA preferences). Suppose that v is CARA, which given our normalizations means $v(z) = \frac{1-exp(-rz)}{r}$. We can write $w(z,\varepsilon) = \frac{1-exp(-rz)\eta(r,\varepsilon)}{r}$ where $\eta(r,\varepsilon) = \nu exp(-r(1-\nu)\varepsilon) + (1-\nu)exp(r\nu\varepsilon)$. Finally, we have $H(z,\varepsilon) \equiv 0$ for all (z,ε) so that the income and risk effect exactly compensate each other.

[28] The literature on risk aversion provides evidence for individuals exhibiting decreasing absolute risk aversion (see for instance Holt and Laury (2002)). When interpreting risk aversion as a proxy for costly access to financial markets, then it seems reasonable that wealthy firms face less of these constraints simply because they can diversify and pledge more income on the financial market. Small or specialized firms do not have the possibility to diversify or pledge income on the financial markets in the same way as wealthier firms and therefore face tighter financial constraints. In other words, firms exhibit decreasing absolute risk aversion because an increase in wealth allows them to pledge more income on the financial markets.

3. INCENTIVE-FEASIBLE ALLOCATIONS

This section describes the set of incentive feasible allocations in this dynamic context with full commitment. From the Revelation Principle, we can define the firm's intertemporal payoff as:

(3.1)
$$\mathcal{U}(\theta) = \max_{\hat\theta\in\Theta,\hat\beta\in\mathcal{B}} (1-\delta)(t(\hat\theta)-\theta q(\hat\theta))+\delta E_\beta\left(v\left(t(\hat\theta)-\theta q(\hat\theta)+y(\hat\theta)+p(\hat\theta,\hat\beta)-\beta x(\hat\theta,\hat\beta)\right)\right).$$

Second-period incentive compatibility. The requirement of incentive compatibility can be applied recursively.[29] Observe first that the firm's second-period profit $u_2(\theta,\beta)$ must satisfy:

$$u_2(\theta,\beta) = \max_{\hat\beta\in\mathcal{B}} p(\theta,\hat\beta)-\beta x(\theta,\hat\beta), \quad \forall\theta\in\Theta. \tag{3.2}$$

Incentive compatibility in the second period requires in particular that a firm facing a low cost of producing the add-on prefers the requested option:

$$u_2(\theta,\underline{\beta}) \geq u_2(\theta,\bar\beta)+\Delta\beta x(\theta,\bar\beta), \quad \forall\theta\in\Theta. \tag{3.3}$$ [30]

This condition tells us that, when the firm has ex post private information on the cost of producing the add-on, the distribution of second-period profits must necessarily be risky in order to satisfy incentive compatibility.[31]

Furthermore, observe that there is no loss of generality in assuming that in expectations the firm makes zero profits on the add-on, i.e.,

$$E_\beta\left(u_2(\theta,\beta)\right) = 0. \tag{3.4}$$ [32]

[29] See Baron and Besanko (1984), Battaglini (2005) and Pavan, Segal and Toikka (2013).

[30] In this two-type model, it is routine to check that the second-period incentive constraint of a firm facing a high cost of producing the add-on, namely $u(\hat\theta,\bar\beta)\geq u(\hat\theta,\underline{\beta})-\Delta\beta x(\hat\theta,\underline{\beta})$, is automatically satisfied when (3.3) is binding and $x(\hat\theta,\underline{\beta})\geq x(\hat\theta,\bar\beta)$ as requested by the standard monotonicity condition. This monotonicity condition holds for the optimal contract that will be derived below. Therefore, we simplify presentation by focusing only on the low-cost incentive constraint (3.3).

[31] This is true at least as long as the add-on is produced in positive quantities, $x(\theta,\bar\beta)>0$. This non-negativity requirement is satisfied by the optimal levels when the Inada condition holds.

From this remark, second-period profits can be expressed as:

$$p(\theta,\underline{\beta})-\underline{\beta}x(\theta,\underline{\beta})=(1-\nu)\varepsilon(\theta) \text{ and } p(\theta,\bar{\beta})-\bar{\beta}x(\theta,\bar{\beta})=-\nu\varepsilon(\theta) \tag{3.5}$$

for some function $\varepsilon(\theta)$. It is thus equivalent to view a direct mechanism as a menu $\left\{t(\hat{\theta}),y(\hat{\theta}),q(\hat{\theta}),\varepsilon(\hat{\theta}),x(\hat{\theta},\hat{\beta})\right\}_{\hat{\theta}\in\Theta}$ where $\varepsilon(\hat{\theta})$ is the amount of risk borne by the firm in the second period and we can write

$$u_2(\theta,\underline{\beta})=(1-\nu)\varepsilon(\theta) \text{ and } u_2(\theta,\bar{\beta})=-\nu\varepsilon(\theta), \tag{3.6}$$

for some $\varepsilon(\theta)$. That amount of risk is endogenously determined by incentive compatibility in the second period and (3.3) indeed amounts to:

$$\varepsilon(\theta)\geq\Delta\beta x(\theta,\bar{\beta}), \quad \forall\theta\in\Theta. \tag{3.7}$$

In the sequel, we are particularly interested in tracing out the consequences of this endogenous risk on first-period incentives.

First-period incentive compatibility. Since is equivalent to characterize incentive compatibility conditions by means of a direct and truthful mechanism $\mathcal{C}$ or through the allocation $(U(\theta),q(\theta),u(\theta),y(\theta),\varepsilon(\theta),x(\theta,\beta))$ induced by that mechanism. We adopt the dual approach and can rewrite the firm's intertemporal payoff as

$$\mathcal{U}(\theta)=\max_{\hat{\theta}\in\Theta}(1-\delta)(t(\hat{\theta})-\theta q(\hat{\theta}))+\delta w(t(\hat{\theta})-\theta q(\hat{\theta})+y(\hat{\theta}),\varepsilon(\hat{\theta})). \tag{3.8}$$

Recalling that $u(\theta)=t(\theta)-\theta q(\theta)$, the following lemma provides necessary and sufficient conditions satisfied by any incentive compatible allocation.

Lemma 1 **Necessary condition.** *$\mathcal{U}(\theta)$ is absolutely continuous in θ and thus almost everywhere differentiable with at any point of differentiability:*

$$\dot{\mathcal{U}}(\theta)=-q(\theta)\left(1-\delta+\delta w_z(u(\theta)+y(\theta),\varepsilon(\theta))\right) \quad \forall\theta\in\Theta. \tag{3.9}$$

Sufficient conditions. *A rent profile $\mathcal{U}(\theta)$ absolutely continuous in θ and satisfying (3.9) corresponds to an incentive compatible allocation if it is convex.*

[32]Suppose on the contrary that the second period profit on the add-on has a non-zero mean $E_\beta\left(p(\theta,\beta)-\beta x(\theta,\beta)\right)\neq 0$ for some payments $(t(\theta),y(\theta),p(\theta,\beta))$. Keeping all outputs unchanged, we can redefine a set of new payments $(\tilde{t}(\theta),\tilde{y}(\theta),\tilde{p}(\theta,\beta))$ such that $\tilde{p}(\theta,\beta)=p(\theta,\beta)-E_\beta\left(p(\theta,\beta)-\beta x(\theta,\beta)\right)$, $\tilde{t}(\theta)=t(\theta)+E_\beta\left(p(\theta,\beta)-\beta x(\theta,\beta)\right)$ and $\tilde{y}(\theta)=y(\theta,\beta)-E_\beta\left(p(\theta,\beta)-\beta x(\theta,\beta)\right)$. The per-period total payment remains unchanged since $\tilde{t}(\theta)+\tilde{y}(\theta)=t(\theta,\beta)+y(\theta)$ and $\tilde{t}(\theta)+\tilde{p}(\theta,\beta)=t(\theta)+p(\theta,\beta)$ while, by construction, second-period profits have zero mean: $E_\beta\left(\tilde{p}(\theta,\beta)-\beta x(\theta,\beta)\right)=0$.

Because the firm is privately informed about its efficiency parameter θ at the time of contracting, a firm with type θ must receive an information rent to reveal its type. This informational rent corresponds to how much extra utility $\mathcal{U}(\theta)-\mathcal{U}(\theta+d\theta) \approx -\dot{\mathcal{U}}(\theta)d\theta$ this type θ gets from mimicking the behavior of a slightly less efficient firm with type $\theta+d\theta$. Exaggerating its first-period cost allows the firm to pocket some information rent because it is asked to produce the same amount as a slightly less efficient type but does so at lower marginal cost, thereby saving saving $q(\theta+d\theta)d\theta \approx q(\theta)d\theta$ in each period. Of course, these effects must be weighted in each period by the corresponding marginal utility of income. This explain the term $w_z(u(\theta)+y(\theta), \varepsilon(\theta))$ that appears on the right-hand side of (3.9).

It is clear from (3.9) that the convexity of $\mathcal{U}$ is guaranteed when δ is small enough and $q(\theta)$ is decreasing. As we will see below in the characterization of the optimal outputs under various scenarios, these conditions are always satisfied when Assumption 1 holds. This proviso on δ validates our approach that will consist in solving the relaxed problem obtained by omitting the convexity condition.

Observe also that (3.9) implies that $\mathcal{U}(\theta)$ is non-increasing and the participation constraint (4.1) holds everywhere if it holds for $\bar{\theta}$:

$$\mathcal{U}(\bar{\theta}) \geq 0. \tag{3.10}$$

We may also express the second-period profit $u(\theta)+y(\theta)$ in terms of $\mathcal{U}(\theta)$ as:

$$u(\theta)+y(\theta) = \varphi\left(\frac{\mathcal{U}(\theta)-(1-\delta)u(\theta)}{\delta}, \varepsilon(\theta)\right). \tag{3.11}$$

This allows us to get a more compact expression of incentive compatibility as:

$$\dot{\mathcal{U}}(\theta) = -q(\theta)\left(1-\delta+\delta w_z\left(\varphi\left(\frac{\mathcal{U}(\theta)-(1-\delta)u(\theta)}{\delta}, \varepsilon(\theta)\right), \varepsilon(\theta)\right)\right). \tag{3.12}$$

This expression shows that, when the firm's output is reduced, the incentive-compatibility constraint is relaxed and less rent is left to the firm. This is a familiar distortion in screening environments. However, here the firm's rent is also reduced if the second-period marginal utility of income decreases. This is a specific feature of our environment which comes from the concavity of the firm's utility function in the second period. As we will see below, playing both on the second-period profit $u(\theta)+y(\theta)$ and on the risk $\varepsilon(\theta)$ helps the agency extract more of the firm's rent. In particular, this risk effect will indeed be at the source of an important linkage between the different stages of the incentive problems.

4. PARTIAL RESULTS

This section introduces the different effects that are at play when the firm's is risk-averse and faces a risky second period. We focus in particular on the impact that concave payoffs have on first-period incentives. To help build intuition and before undertaking the full-fledged analysis, we decompose the analysis into several elementary steps that can be studied independently of each other.

First, we analyze the simple case where cost realizations in each period are common knowledge. This benchmark allows us to demonstrate the consequences on the agency problem of introducing a concave utility function. Second, we turn to the case of a risk-neutral firm. With risk neutrality we demonstrate that a simple solution to the agency problem can be obtained by making the firm residual claimant for any second-period profits that can be realized on the add-on. Finally, we move to the case of a concave second-period utility function and distinguish two elementary effects coming from the concavity of the firm's utility function. The first effect, the *Marginal Utility of Income Effect*, is present even when there is symmetric information in the second period. The second effete, the *Risk Effect*, occurs even when there is complete information in the first period. To illustrate these effects, the agency problem in each period is taken in isolation and there is no feedback from one period to the other.

4.1. *Complete Information*

Suppose that θ and β are both common knowledge, but recall that at the time of contracting the second-period cost β is not realized. The solution to the contracting problem is nevertheless obvious. It entails perfect insurance against second-period cost realizations for the add-on, the same marginal utility of income in both periods so that the marginal social cost of public subsidies remains constant over time and efficient provision levels for both the basic service and the add-on. To prepare for the rest of the analysis, we will nevertheless be more explicit about the agency's problem in this simple informational environment.

First, equipped with our more compact notations, notice that the firm's intertemporal payoff can be written as:

$$\mathcal{U}(\theta) = (1-\delta)u(\theta) + \delta w(u(\theta) + y(\theta), \varepsilon(\theta)).$$

A contract is thus accepted by the firm whenever it at least breaks even from accepting the contract:

$$(4.1) \qquad \mathcal{U}(\theta) \geq 0 \qquad \forall \theta \in \Theta.$$

Taking into account that second-period expected profits from the add-on are zero, the agency's expected payoff can be rewritten as:

$$S(q(\theta)) - \theta q(\theta) - (1-\delta)u(\theta) - \delta\varphi\left(\frac{\mathcal{U}(\theta) - (1-\delta)u(\theta)}{\delta}, \varepsilon(\theta)\right) + \delta E_\beta\left[V(x(\theta,\beta)) - \beta x(\theta,\beta)\right].$$

Maximizing this expression subject to the break-even constraint (4.1) immediately yields that the basic service and the add-on are both produced at efficient levels $q^*(\theta)$ and $x^*(\beta)$. These are given by:

$$(4.2) \qquad S'(q^*(\theta)) = \theta, \quad \forall \theta \in \Theta, \text{ and } V'(x^*(\beta)) = \beta, \quad \forall \beta \in \mathcal{B}.$$

The agency's expected payoff is maximized when the firm only gets its reservation value (i.e., (4.1) is binding since $\varphi_z \geq 0$):

$$(4.3) \qquad U^*(\theta) = 0, \quad \forall \theta \in \Theta.$$

Because letting the firm bears some risk is costly (i.e., $\varphi_\varepsilon \geq 0$), it should bet fully insured against variations in second-period costs:

$$(4.4) \qquad \varepsilon^*(\theta) = 0, \quad \forall \theta \in \Theta.$$

The last important property that follows from this optimization is related to the intertemporal distribution of profits. Differentiating the principal's objective with respect to $u(\theta)$ then yields the following first-order condition:

$$\varphi_z\left(-\frac{(1-\delta)}{\delta}u^*(\theta), 0\right) = 1 \Leftrightarrow v'\left(h\left(-\frac{(1-\delta)}{\delta}u^*(\theta)\right)\right) = 1, \quad \forall \theta \in \Theta.$$

Profits should be set so that the marginal utility of income remains constant over time. To keep a constant marginal utility of income and because $v'(0) = 1$ and $h(0) = 0$, the firm should actually make zero profit in each period:

$$(4.5) \qquad u^*(\theta) = y^*(\theta) = 0, \quad \forall \theta \in \Theta.$$

As a result, the optimal contract under complete information entails a stationary payment for the basic service. Second-period payments just provide insurance against second-period uncertainty.

4.2. *The Simple Case of Risk Neutrality*

A risk-neutral firm has no cost associated with bearing risk related to the costs of the add-on. This case can thus be viewed as a short-cut for a well-diversified venture that has perfect access to financial markets. Under these circumstances, by making the firm residual claimant for the add-on provision, the agency can easily structure incentives to induce efficient production of the add-on and extract all profits from this activity. To see more clearly how this can be done, consider the following simple mechanism which is designed to regulate the production of the add-on:

$$p(x) = V(x) - V^*, \quad \forall x.$$

The fee $V^* = E_\beta(V(x^*(\beta)) - \beta x^*(\beta))$ is the expected surplus generated by the add-on at the first-best. $p(x)$ is simply a nonlinear price independent of the first-period announcement of θ. Facing this nonlinear contract, the firm chooses to provide the add-on at the efficient level and the agency captures the entire expected net surplus.

With this continuation for the second period, the characterization of first-period incentive compatibility given in Lemma 1 applies. The profile $\mathcal{U}(\theta)$ is obtained as a maximum of linear functions of θ and as such is convex, absolutely continuous and thus almost everywhere differentiable with at any point of differentiability:

$$\dot{\mathcal{U}}(\theta) = -q(\theta). \tag{4.6}$$

The requirement of convexity then amounts to $q(\theta)$ being monotonically decreasing.[33]

We summarize the main features of the optimal contract in the next proposition.[34]

Proposition 1 *Assume that the firm is risk neutral. The optimal contract exhibits the following features.*

- *The add-on is always produced at the efficient level $x^*(\beta)$ for all $\beta \in \mathcal{B}$*
- *The production of the basic service is distorted below the first-best level, $q^{bm}(\theta) \leq q^*(\theta)$ for all $\theta \in \Theta$. This output is given by the standard Baron and Myerson (1982)*

---

[33] In the screening literature, it is common to only focus on the necessary condition (3.9) and check ex post that the solution to the so-called relaxed optimization problem satisfies this extra requirement. With risk aversion, this second step is much harder due to the fact that the marginal utility of income is not constant. Yet, as long as this marginal utility remains close to one (for instance, because $\varepsilon(\theta)$ is small enough), the monotonicity of q is enough to ensure sufficiency.

[34] The proof is omitted. Under risk neutrality, the analysis reduces to the standard framework à la Baron and Myerson (1982).

formula:

$$S'\left(q^{bm}(\theta)\right) = \theta + \frac{F(\theta)}{f(\theta)}, \quad \forall \theta \in \Theta. \tag{4.7}$$

When Assumption 1 holds, $q^{bm}(\theta)$ is monotonically decreasing in θ

- *Unless the firm is if the least efficient type it always gets a positive information rent:*

$$\mathcal{U}^{bm}(\theta) = \int_{\theta}^{\bar{\theta}} q^{bm}(x)dx, \quad \forall \theta \in \Theta.$$

Making the risk-neutral firm residual claimant for all second-period profits allows for efficient decisions in the second-period. Provided that the firm bears all risk ex post, the first-period screening problem reduces to the Baron and Myerson (1982) allocation. Because at the time of contracting the firm is privately informed about its efficiency parameter θ, it must receive an information rent to reveal its type. The optimal contract exhibits the usual trade-off between this information rent and efficiency. As a result, output is distorted downward below the first best for all but the most efficient type.

4.3. *Marginal Utility of Income*

Suppose now that the agency and the firm share ex post knowledge about the cost parameter β. Contracting still takes place ex ante, i.e. before the realization of this cost, and the firm remains privately informed about θ at the time of contracting. Of course, second-period incentive constraints do not matter in this environment and letting the firm bear some risk is costly for two reasons. First, it requires a risk premium to paid to the firm. Second, when Assumption 2 holds, the second-period marginal utility of income increases with more risk, $w_{z\varepsilon} \geq 0$. Increasing second-period risk thus renders information revelation in the first period more difficult. Both of these reasons lead to the same conclution that

$$\varepsilon(\theta) = 0, \quad \forall \theta \in \Theta.$$

Profit levels. The first consequence of having concave payoffs in the second-period is that it allows the agency to play on the firm's second-period marginal utility of income to reduce its information rent. We call this the *Marginal Utility of Income Effect.*

PROPOSITION 2 *Assume that β is common knowledge in the second-period. At the optimal contract, the firm's marginal utility of income is lower in the second than in the* *first period (with equality for $\underline{\theta}$ only):*

$$v'(u^{mu}(\theta)+y^{mu}(\theta)) = 1 + q^{mu}(\theta)\frac{F(\theta)}{f(\theta)}v''(u^{mu}(\theta)+y^{mu}(\theta)) \leq 1, \quad \forall\theta\in\Theta. \tag{4.8}$$

Unless the firm is of the least efficient type, it always gets a positive information rent:

$$\mathcal{U}^{mu}(\theta) \geq \mathcal{U}^{mu}(\bar{\theta}) = 0, \quad \forall\theta\in\Theta. \tag{4.9}$$

This "payment-smoothing" result is similar to the consumption smoothing results in the new dynamic public finance literature[35]. However, here this result is embedded in a model of procurement and will have further effects on service levels.

Starting from the full information scenario and increasing second-period profits from the basic service above zero, i.e., setting $u^{mu}(\theta)+y^{mu}(\theta) \geq 0$ as implied by (4.8), decreases the firm's second-period marginal utility of income. This makes lying about its cost less attractive for the firm because the corresponding benefits are not as much at this date. To reduce the information rent left to all infra-marginal types these distortions are more pronounced as the firm becomes less efficient.

Payments are no longer stationary as was the case under complete information, but instead backloaded. Indeed, putting together (4.8) and 4.9) immediately yields that, for inefficient types (θ close enough to $\bar{\theta}$):

$$u^{mu}(\theta) \leq 0 \leq u^{mu}(\theta) + y^{mu}(\theta).$$

To deter more efficient types from announcing a very high type, the first-period pay-of is negative for high enough θ. Since first-period pay-offs are evaluated at a higher marginal utility of income than second-period pay-offs, this is an efficient tool for incentive compatibility in this environment. Furthermore, the stronger the risk aversion[36], the more important backloading payments becomes and more payments are shifted to the second period.

Turning now to the lower tail of the types distribution, observe that (4.8) also implies that $u^{mu}(\underline{\theta}) + y^{mu}(\underline{\theta}) = 0$ (with $u^{mu}(\theta) + y^{mu}(\theta) = 0$ in the right-neighborhood of $\underline{\theta}$). Since a firm with the most efficient type gets a positive rent (i.e., $\mathcal{U}^{mu}(\underline{\theta}) > 0$), payments are front-loaded for types close enough to $\underline{\theta}$:

$$u^{mu}(\theta) > u^{mu}(\theta) + y^{mu}(\theta) \geq 0.$$

In general, firms do not have have an incentive to understate costs. For low enough types, the procurement agency can therefore shift payoffs to the first-period. At the optimal

[35] Kocherlakota (2010) and Golosov et al (2006)

[36] Measured by the importance of $v''(\cdot)$.

contract a given utility level is obtained by a lower payment in the first period compared to the required second-period payment (because of the difference in marginal utility of income). It is therefore less costly for the principal to pay out in the first period. However, as θ increases this is no longer optimal for incentive-compatibility reasons.

Outputs. Output distortions now depend on the concavity of the firm's utility function.

PROPOSITION 3 *Assume that β is common knowledge in the second-period. The optimal production of the basic service is distorted downward below the first best, $q^{mu}(\theta) \leq q^*(\theta)$ (with equality at $\underline{\theta}$ only):*

$$S'(q^{mu}(\theta)) = \theta + \frac{F(\theta)}{f(\theta)}(1 - \delta + \delta v'(u^{mu}(\theta) + y^{mu}(\theta))), \quad \forall\theta \in \Theta. \tag{4.10}$$

The add-on is always produced at the efficient level: $x^{mu}(\theta, \beta) = x^*(\beta)$ for all $\beta \in \mathcal{B}$.

From (4.8), we know that the firm's benefits from exaggerating its costs are evaluated at a marginal utility of income which is lower in the second period. As a result, the agency does not need to distort provision levels as much. In particular, output distortions are lower than had the firm remained risk neutral in the second period and evaluated the gains from lying in the same way in both periods:

$$q^{mu}(\theta) \geq q^{bm}(\theta), \quad \forall\theta \in \Theta.$$

We can thus conclude that the *Marginal Utility of Income Effect* favors high powered contracts for the basic service.

Furthermore the effects uncovered above can be shown to be more important the higher the weight on the second period. Intuitively, the more important the second period, the more the first-period incentive problem is eased when payoffs are shifted to the second period. This of course makes it more attractive to increase the second-period payoff which again makes distortions less desirable.[37]

Note also that the optimal mechanism remains deterministic in our environment. This can be seen as a consequence of Strausz (2006) who shows that when the optimal deterministic mechanism does not involve bunching, then it is also optimal with respect to stochastic mechanisms. Furthermore, this can be seen from the concavity of the Hamiltonian associated with our maximization problem.[38]

[37]This can be formally shown by differentiation (4.8) and (4.10) with respect to δ.

[38]See Appendix B for a discussion of this concavity. For the same reasons, the result of the optimal mechanism being deterministic also hold in the general case in Section 5.

Example (CARA preferences - continued). Here, we can use (4.8) and (4.10) to obtain the following closed-form expressions for outputs, per-period profits and rent:

$$S'(q^{mu}(\theta)) = \theta + \frac{F(\theta)}{f(\theta)}\left(1-\delta+\frac{\delta}{1+rq^{mu}(\theta)\frac{F(\theta)}{f(\theta)}}\right), \quad \forall\theta\in\Theta,$$

$$\mathcal{U}^{mu}(\theta) = \int_{\theta}^{\bar{\theta}} q^{mu}(s)\left(1-\delta+\frac{\delta}{1+rq^{mu}(s)\frac{F(s)}{f(s)}}\right)ds, \quad \forall\theta\in\Theta,$$

$$u^{mu}(\theta)+y^{mu}(\theta) = \frac{1}{r}ln\left(1+rq^{mu}(\theta)\frac{F(\theta)}{f(\theta)}\right), \quad \forall\theta\in\Theta,$$

$$u^{mu}(\theta) = \frac{1}{1-\delta}\left(\mathcal{U}^{mu}(\theta)-\frac{\delta q^{mu}(\theta)\frac{F(\theta)}{f(\theta)}}{1+rq^{mu}(\theta)\frac{F(\theta)}{f(\theta)}}\right), \quad \forall\theta\in\Theta.$$

4.4. *Risk*

Suppose now that in the second period β is privately learned by the firm while θ is common knowledge. In this environment, the only concern is how to draft a contract that simultaneously gives the firm insurance against uncertain second-period costs and provides incentives to reveal these costs truthfully. In this context, the first-period incentive constraint (3.9) disappears from the agency's optimization problem.

Profit levels. The consequence of having concave payoffs in the second period is now to introduce risk behavior for a firm that faces random returns. This is what we call the *Risk Effect*. Of course, as the next proposition unveils, such risk behavior affects the second-period incentive problem.

PROPOSITION 4 *Assume that β is privately observed by the firm in the second period, but that θ is common knowledge. At the optimal contract,*

- *The firm always bears some risk but this risk is independent of the first-period cost, $\varepsilon^{ri}(\theta) = \varepsilon^{ri} > 0$ for all $\theta \in \Theta$.*
- *The firm's marginal utility of income remains constant over time and second-period profits are independent of the first-period cost, $u^{ri}(\theta) + y^{ri}(\theta) = u^{ri} + y^{ri}$ for all $\theta \in \Theta$ where:*

$$w_z(u^{ri} + y^{ri}, \varepsilon^{ri}) = 1. \tag{4.11}$$

- *There is always full extraction of the firm's rent:*

$$\mathcal{U}^{ri}(\theta) = 0, \quad \forall \theta \in \Theta. \tag{4.12}$$

When θ is common knowledge there is no reason to condition second-period payments and outputs on the first-period cost. Indeed, payments are used to provide insurance and incentives in the second period and to ensure that the firm's marginal utility of income remains constant over time. These two objectives are independent of the firm's first-period production.

From (3.7), satisfying the second-period incentive-compatibility constraint always requires to let the firm to bear some minimal amount of risk. This is obviously costly for the agency since a risk premium must be paid to ensure the firm's participation. Yet, the agency's marginal cost of funds must remain the same over time which requires adjusting the second-period subsidies to ensure that the firm's marginal utility of income remains constant over time. Because $w_{z\varepsilon} \geq 0$, (4.11) also implies that $w_z(u^{ri} + y^{ri}, 0) = v'(u^{ri} + y^{ri}) < 1$ and second-period profits are necessarily positive. This is a striking difference with the case where costs in both periods remain common knowledge. Indeed, we have:

$$u^{ri} + y^{ri} = \varphi\left(\frac{-(1-\delta)u^{ri}}{\delta}, \varepsilon^{ri}\right) > 0,$$

where the first equality follows from the definition of φ and the fact that $\mathcal{U}(\theta) = 0$. Yet, more precise results on the intertemporal profile of payments can be obtained.

Corollary 1 *Assume that θ is common knowledge in the first period, that β is privately observed by the firm in the second period and that Assumptions 2 and 3 hold. Payments are always backloaded:*

$$y^{ri} \geq u^{ri} + y^{ri} > 0 \geq u^{ri}, \quad \forall \theta \in \Theta.$$

There are two effects at play in determining profit levels in each period. Transferring some wealth to the second period and choosing $y^{ri} \geq 0$ of course acts as precautionary savings against second-period uncertainty; an optimal response when the firm exhibits some prudent behavior ($v''' \geq 0$).[39] Yet, risk aversion also calls for giving the firm a risk-premium which requires that the overall intertemporal profit $u^{ri} + \delta y^{ri}$ remains positive.

[39]For more details on the interaction between risk aversion and precautionary savings, see Kimball and Weil (2009).

Taken in tandem, Assumptions 2 and 3 ensures that the second effect remains strong enough even when there are positive profit in the first period.

Outputs. Because the cost parameter θ is common knowledge, there is no reason to distort the level of the basic service. Downward distortions of output only arise in the second period and concern the level of the add-on. Relaxing the second-period incentive compatibility constraint (3.7) calls for a downward distortion of its provision. This distortion is summarized in the next proposition.

PROPOSITION 5 *Assume that θ is common knowledge in the first period and β is privately observed by the firm in the second period.*

- *The basic service is always produced at the efficient level: $q^{ri}(\theta) = q^*(\theta)$ for all $\theta \in \Theta$.*
- *The add-on is always produced at an efficient level when second-period costs are low, below the first best otherwise but at a level x^{ri} which is independent of the first-period cost: $x^{ri}(\theta, \underline{\beta}) = x^*(\underline{\beta})$ and $0 < x^{ri}(\theta, \bar{\beta}) = x^{ri} = \frac{\varepsilon^{ri}}{\Delta\beta} \leq x^*(\bar{\beta})$ where*

$$(4.13) \quad (1-\nu)(V'\left(x^{ri}\right) - \bar{\beta}) = \Delta\beta\varphi_{\varepsilon}(w(u^{ri} + y^{ri}, \varepsilon^{ri}), \varepsilon^{ri}).$$

When there is no first-period asymmetric information, the optimal contract is the same regardless of the importance of the second period (as measured by δ). Because the first-period incentive problem can be ignored, there is no reason to distort the marginal utility of income away from its socially optimal level. Furthermore all the distortions related to the Risk Effect only play a role in the second period and therefore they are not affected by the value of δ.[40]

Example (CARA preferences - continued). We now use (4.11), (4.12) and (4.13) to obtain closed-form expressions for per-period profits and output as follows:

$$(4.14) \quad (1-\nu)\left(V'\left(\frac{\varepsilon^{ri}}{\Delta\beta}\right) - \bar{\beta}\right) = \Delta\beta\frac{\eta_{\varepsilon}(r, \varepsilon^{ri})}{r\eta(r, \varepsilon^{ri})} > 0,$$

$$y^{ri} = \frac{1}{r} ln\left(\eta(r, \varepsilon^{ri})\right) > 0 = u^{ri}.$$

With CARA preferences, coping with second-period uncertainty requires a risk-premium to be paid to the firm. Second-period profits are thus positive. At the same time, the firm's marginal utility of income remains unchanged when this compensation for the second-period risk has been taken into account and there is no need to transfer wealth from the first to the second period.

[40]Formally, this can be viewed in (4.11) and (4.13) which are independent of δ.

5. OPTIMAL CONTRACT: THE GENERAL CASE

We now suppose that both θ and β are privately known by the firm. The analysis in this section thus merges the two specific contexts analyzed in Section 4 and highlights how the Marginal Utility of Income and Risk Effects interact. In fact we show that the optimal contract exhibits a contractual externality despite of there being no technological linkage. We find that output distortions on the basic service are more pronounced than when there is no agency problem in the second period and that the firm must bear less risk in the second-period compared to when there is no incentive problem in the first period. Compared to the partial results presented previously, the introduction of risk leads to lower powered incentives in both periods. Finally, profits are backloaded, because this relaxes the incentive problems in both periods.

The optimal contract illustrates how the incentive problems in each period feed back into each other.

Profit levels. A first obvious finding is that, at the optimal contract, there is full rent extraction only for the least efficient firm:

$$\mathcal{U}^{as}(\theta) \geq \mathcal{U}^{as}(\bar{\theta}) = 0, \quad \forall \theta \in \Theta. \tag{5.1}$$

More specific to the present context are changes in the firm's marginal utility of income both over time and by comparison with the case where there is no private information in the second period. This is the purpose of next proposition.

PROPOSITION 6 *Assume that both θ and β are private information. At the optimal contract, the firm's marginal utility of income is lower in the second period than in the first one (with equality for $\underline{\theta}$ only):*

$$w_z(u^{as}(\theta)+y^{as}(\theta), \varepsilon^{as}(\theta)) = 1+q^{as}(\theta)\frac{F(\theta)}{f(\theta)}w_{zz}(u^{as}(\theta)+y^{as}(\theta), \varepsilon^{as}(\theta)) \leq 1, \quad \forall \theta \in \Theta. \tag{5.2}$$

When Assumption 2 holds, we also have:

$$v'(u^{as}(\theta)+y^{as}(\theta)) \leq 1+q^{as}(\theta)\frac{F(\theta)}{f(\theta)}v''(u^{as}(\theta)+y^{as}(\theta)), \quad \forall \theta \in \Theta. \tag{5.3}$$

In a similar way to what we found in Section 4.3, the Marginal Utility of Income effect means that payments are backloaded so as to reduce the cost of first-period incentive compatibility. But with an incentive problem in the second period, the firm must bear some risk and $\varepsilon^{as}(\theta) > 0$. Assumption 2 guarantees that the firm's marginal utility of income is higher when second-period returns are riskier. To compensate for this effect of the second-period agency problem on the first period, the cost of first-period incentive compatibility is reduced by increasing even further second-period profits; a mechanism similar to thr one highlighted in Section 4.3 when only the *Risk Effect* was at play. Because of the complementarily between the *Marginal Utility of Income* and the *Risk Effects*, payments are even more backloaded as can be seen by comparing (4.8) and (5.3).[41]

Outputs. Screening considerations in the first period again call for output distortions and the magnitude of these distortions now depends on how much risk is borne by the firm in the second period.

PROPOSITION 7 *Assume that both θ and β are private information. At the optimal contract, the production of the basic service is distorted downwards below the first best but remains above the Baron and Myerson (1982) outcome, $q^{bm}(\theta) \le q^{as}(\theta) \le q^*(\theta)$ (with equality at $\underline{\theta}$ only):*

$$S'(q^{as}(\theta)) = \theta + \frac{F(\theta)}{f(\theta)}(1 - \delta + \delta w_z(u^{as}(\theta) + y^{as}(\theta), \varepsilon^{as}(\theta))), \quad \forall \theta \in \Theta. \tag{5.4}$$

When Assumption 2 holds, we also have:

$$S'(q^{as}(\theta)) \ge \theta + \frac{F(\theta)}{f(\theta)}(1 - \delta + \delta v'(u^{as}(\theta) + y^{as}(\theta))), \quad \forall \theta \in \Theta. \tag{5.5}$$

Since $w_{z\varepsilon} \ge 0$, the fact that the firm now bears some risk in the second period implies a higher marginal utility of income at this date. The firm therefore has more incentives to lie about its costs because the corresponding cost saving is evaluated at a higher rate of return for the second-period. This suggests that the agency should be more concerned with rent extraction than in the absence of such risk (condition (5.5) below) and output distortions are stronger. Condition (5.2) further ensures that distortions remain less pronounced than if the firm was risk neutral in the second period.

The interaction between the agency problems in each period also goes the other way. The rent left to the firm for the production of the add-on has an impact on the amount of risk it has to bear. This impact is unveiled in the next proposition.

[41] The attentive reader will of course have noticed that the outputs $q^{mu}(\theta)$ and $q^{as}(\theta)$ change across the two scenarios. Yet our reasoning remains indicative of the direction in which payments move.

PROPOSITION 8 *Assume that both θ and β are private information. At the optimal contract,*

- *The firms always bear some risk in the second period:* $\varepsilon^{as}(\theta) = \Delta\beta x^{as}(\theta, \bar{\beta}) > 0$.
- *The add-on is always produced at the efficient level when second-period costs are low but below the first best otherwise:* $x^{as}(\theta, \underline{\beta}) = x^*(\underline{\beta})$ *and* $0 < x^{as}(\theta, \bar{\beta}) \leq x^*(\underline{\beta})$ *where*

$$(5.6) \qquad (1-\nu)(V'\left(x^{as}(\theta,\bar{\beta})\right) - \bar{\beta})$$
$$= \Delta\beta\left(\varphi_\varepsilon(w(u^{as}(\theta) + y^{as}(\theta), \varepsilon^{as}(\theta)), \varepsilon^{as}(\theta)) + q^{as}(\theta)\frac{F(\theta)}{f(\theta)}H(u^{as}(\theta) + y^{as}(\theta), \varepsilon^{as}(\theta))\right), \quad \forall \theta \in \Theta.$$

When Assumption 3 holds, we have:

$$(5.7) \quad (1-\nu)(V'\left(x^{as}(\theta,\bar{\beta})\right)-\bar{\beta}) \geq \Delta\beta\varphi_\varepsilon(w(u^{as}(\theta)+y^{as}(\theta), \varepsilon^{as}(\theta)), \varepsilon^{as}(\theta)), \quad \forall \theta \in \Theta.$$

The findings in Proposition 8 are closely related to those highlighted in Section 4.4 when only second-period incentive compatibility was a concern. The first difference is illustrated in condition (5.6). When choosing how much risk should be borne by the firm, the agency anticipates the consequences of that risk on the firm's marginal utility of income at that date. There are two effects at play here. First, keeping second-period profits fixed, more second-period risk increases the marginal utility of income at that date which makes first-period incentive compatibility more costly. This first effect calls for reducing risk in the second-period. However, leaving the firm with more risk in the second period also requires an increase in second-period profits if one wants to keep the firm's intertemporal utility constant. Thus this second effect reduces the marginal utility of income in the second period, making first-period incentives less costly. Overall, Assumption 3 guarantees that the first effect dominates. This means that the agency chooses to give the firm low powered incentives and more insurance against second-period uncertainty. As illustrated by (5.7), the optimal contract exhibits lower powered incentives in the second period and a lower distortion on the add-on than when there is no asymmetric information in the first period.

Example (CARA preferences - continued). Straightforward computations immediately lead to:

$$(5.8) \qquad \varepsilon^{as}(\theta) = \varepsilon^{ri} \text{ and } q^{as}(\theta) = q^{mu}(\theta), \quad \forall \theta \in \Theta.$$

Outputs distortions in both periods are now the same as when the incentive problem at each dates is taken in isolation. The second-period endogenous risk is independent of whether there is asymmetric information on first-period costs or not. Reciprocally, distortions on the basic service are independent on whether there is symmetric or asymmetric information on the second-period cost.

However, second-period uncertainty requires payments to be backloaded and a second-period risk premium to be paid for the endogenous risk that the privately informed firm must bear. Indeed, with CARA preferences, the second-period risk premium paid to the firm, namely $\frac{1}{r}ln\left(\eta(r,\varepsilon^{ri})\right)$, is independent of how the profits from the basic service. Therefore, there is no reason to modify the firm's information rent and first-period output in view of reducing the second-period agency costs.

An immediate consequence of these findings is that the firm's information rent is also the same as when only the *Marginal Utility of Income Effect* is at play:

$$\mathcal{U}^{as}(\theta) = \mathcal{U}^{mu}(\theta), \quad \forall \theta \in \Theta.$$

Second-period profits are simply the sum of the values obtained when considering the *Marginal Utility of Income* and the *Risk Effects* separately:

$$u^{as}(\theta) + y^{as}(\theta) = \frac{1}{r}ln\left(\eta(r,\varepsilon^{ri})\right) + \frac{1}{r}ln\left(1 + rq^{mu}(\theta)\frac{F(\theta)}{f(\theta)}\right), \quad \forall \theta \in \Theta.$$

As a result, the marginal utility of income is unchanged by the addition of some endogenous risk for the second-period. Namely

$$w_z(u^{as}(\theta) + y^{as}(\theta), \varepsilon^{ri}) \equiv v'(u^{mu}(\theta) + y^{mu}(\theta)) \quad \forall \theta \in \Theta.$$

This equality has an important implication: The second-period agency problem has no impact on first-period incentives.

Overall, and similarly to the case of a risk-neutral firm analyzed in Section 4.2, there is now a dichotomy between solving incentives problem in the each period. The case of risk neutrality is simply obtained by taking the limit when r goes to 0 and in this way we can recover the findings of Proposition 1. When r goes to infinity, the firm becomes infinitely risk averse and we obtain a screening distortion (à la Baron-Myerson) which as in the two-type screening models with ex post participation constraints:[42]

$$\lim_{r\to+\infty} V'(x^{ri}) = V'(x^{ri}_{\infty}) = \bar{\beta} + \frac{\nu}{1-\nu}\Delta\beta.$$

[42]See Chapter 2, Laffont and Martimort (2001) for instance.

Indeed, an infinitively risk-averse firm only cares about the worst possible returns for the second-period. This requirement hardens the participation constraint and calls for strong distortions in the second period. When considering information manipulation in the first period, the firm now anticipates that cost savings provide no utility gains in the second-period. This reduces how much rent the firm can make from the basic service. As a result, output distortions are also weaker and information rents are lower:

$$\lim_{r\to+\infty} S'(q^{as}(\theta)) = S'(q^{as}_\infty(\theta)) = \theta+(1-\delta)\frac{F(\theta)}{f(\theta)} \text{ and } \lim_{r\to+\infty} \mathcal{U}^{as}(\theta) = (1-\delta)\int_\theta^{\bar{\theta}} q^{as}_\infty(s)ds, \quad \forall\theta\in\Theta.$$

■

6. RENEGOTIATION

In this section, we argue that under very general conditions, the dynamics of the optimal contract in our environment leaves open the door to renegotiation. The aim of this section is not to characterize the optimal renegotiation-proof contract but to give some insights into when renegotiation may or may not hinder the performances of a long-term agreement. The results of this section highlights the difficult trade-off between flexibility and rigidity of long-term contracts: In general, a long-term contract needs to be be flexible enough to include the add-on, but rigid enough to avoid (costly) renegotiation.

Under full commitment, the agency can fully anticipate the impact of private information in the second period on first-period incentive compatibility. Therefore, the agency might find it attractive to play on how much second-period risk is borne by the firm in order to ease information revelation earlier on. This important feature of long-term contracting is apparent from comparing the second-period distortions with and without informational asymmetry in the first period, (4.13) and (5.6) respectively. It is immediate that as soon as $H(z,\varepsilon)>0$ these expressions differ.

Whenever these two conditions differ, history plays an active role in the contract and whatever first-period type has been revealed has an impact on risk-sharing related to the add-on. More precisely, the agency cares about reducing that risk to save on the firm's information rent and this concern is greater the greater the level of the basic service.[43] Yet, once in the second period, these extra distortions are no longer needed. In other words, the optimal long-term contract might not be *sequentially optimal* in a sense that we now define more precisely.

[43]Formally, the term $\Delta\beta q^{as}(\theta)\frac{F(\theta)}{f(\theta)}H(u^{as}(\theta)+y^{as}(\theta),\varepsilon^{as}(\theta))$ depends on how much of the basic service is being provided.

Let us consider a long-term contract $\mathcal{C} = \left\{t(\hat{\theta}), y(\hat{\theta}), q(\hat{\theta}), \varepsilon(\hat{\theta})\right\}_{\hat{\theta}\in\Theta}$.[44] In the second period, such a contract yields an expected utility of $w(t(\hat{\theta})+y(\hat{\theta})-\theta q(\hat{\theta}), \varepsilon(\hat{\theta}))$ to the firm of type θ when it chooses to report a first-period cost type $\hat{\theta}$. In a context with limited commitment, there is no reason to expect that the firm reports truthfully in the first place. In fact, by misreporting its type early on, the firm might secure a more attractive renegotiation later on. We denote by $M(\hat{\theta}|\theta)$ the distribution of reports that type θ may choose.

In the second period, but before second-period costs are observed by the firm, the agency might want to propose a new contract for the provision of the add-on. Such a contract may let the firm bear more or less risk than the initial contract. For the firm to accept this, the new contract must stipulate an average extra profit level $\tilde{y}(\hat{\theta})$, production levels of the add-on $\tilde{x}(\hat{\theta}, \hat{\beta})$ and an amount of risk $\tilde{\varepsilon}(\hat{\theta})$[45] that are more profitable to the firm than the initial contract: For any type θ such that the report $\hat{\theta}$ is in the support of $M(\cdot|\theta)$, we must have:

$$(6.1) \qquad w(t(\hat{\theta}) + y(\hat{\theta}) + \tilde{y}(\hat{\theta}) - \theta q(\hat{\theta}), \tilde{\varepsilon}(\hat{\theta})) \geq w(t(\hat{\theta}) + y(\hat{\theta}) - \theta q(\hat{\theta}), \varepsilon(\hat{\theta})).$$

The new contract must also remain incentive compatible:

$$(6.2) \qquad \tilde{\varepsilon}(\hat{\theta}) \geq \Delta\beta\tilde{x}(\hat{\theta}, \bar{\beta}).$$

Among all contracts that are acceptable at the renegotiation stage, there is always the null contract that consists of offering no extra payments to the firm, $\tilde{y}(\hat{\theta}) \equiv 0$, leaving levels of the add-on and overall risk unchanged, so that $\tilde{\varepsilon}(\hat{\theta}) \equiv \varepsilon(\hat{\theta})$.

A long-term contract $\mathcal{C}$ (with an associated first-period reporting strategy $M(\cdot|\theta)$) is *renegotiation-proof* if, given her posterior beliefs on the firm's type θ, the agency still finds it optimal to offer the null contract (i.e., $\tilde{y}(\hat{\theta}) = 0$ and $\tilde{\varepsilon}(\hat{\theta}) = \varepsilon(\hat{\theta})$) at the renegotiation stage.[46]

We are now ready to check whether the optimal contract under full commitment $\mathcal{C}^{as}$ is actually robust to further rounds of contracting.

[44]For the sake of simplifying notations and the presentation, we first take the short-cut of considering that the second-period contract is fully determined by a condition of zero expected profits from (3.4) and an amount of risk in the second-period that satisfies (3.7). Second, we restrict attention to the case of direct mechanisms. The analysis of Bester and Strausz (2001) shows that even in an environment with limited commitment this is without loss of generality in the case of a finite set of types.

[45]In other words the new prices for each levels of the add-ons are now $\tilde{p}(\hat{\theta}, \underline{\beta}) - \underline{\theta}\tilde{x}(\hat{\theta}, \underline{\beta}) = \tilde{y}(\hat{\theta}) + (1 - \nu)\tilde{\varepsilon}(\hat{\theta})$ and $\tilde{p}(\hat{\theta}, \bar{\beta}) - \bar{\theta}\tilde{x}(\hat{\theta}, \bar{\beta}) = \tilde{y}(\hat{\theta}) - \nu\tilde{\varepsilon}(\hat{\theta})$.

[46]Even if the informational contacts slightly different, we use this definition of renegotiation-proofness from Dewatripont (1988).

Proposition 9 *If there exists $\theta \in \Theta$ such that*

$$H(u^{as}(\theta) + y^{as}(\theta), \varepsilon^{as}(\theta)) > 0, \tag{6.3}$$

then the optimal contract under full commitment $\mathcal{C}^{as}$ is not renegotiation-proof.

The optimal contract under full commitment $\mathcal{C}^{as}$ is always renegotiation-proof if the firm has CARA preferences.

With general preferences, in Section 4.4 we showed that the agency may want to reduce the second-period risk as a mean of facilitating information revelation in the first period. When information has been revealed, there is no longer any reason for providing the firm with insurance against the realization of second-period uncertainty. Renegotiation may push the agency to increase the level of the add-on and thus to shift more risk to the firm.[47]

However, with CARA preferences, the amount of risk borne by the firm in the second period is, even at the optimal contract under full commitment, independent of how much profit has been promised to the firm for delivering the basic service and thus independent of its costs for the basic service. In the absence of such income effects, the two incentive problems are not linked. The optimal contract is thus renegotiation-proof.

A particular case of interest arises when the firm is risk neutral. To the extent that risk neutrality captures the idea that the firm has a perfect access to financial markets, our model predicts that long-term contracts are then stable and robust to further rounds of negotiations. Instead, a costly access to the financial markets may be a reason for destabilizing renegotiations of long-term contracts.

Spot contracting. This robustness to recontracting in the case of CARA preferences also holds when the long-term contract is highly incomplete and cannot even specify payments and output requirements for the add-on at the time it is signed. Indeed, with CARA preferences, it can easily be checked that even if parties are somewhat restricted to trade with a sequence of contracts, the same allocation as when an optimal long-term contract $\mathcal{C}^{as}$ can be implemented arise. By "somewhat restricted trade" we mean that the initial long-term agreement $\{t^{as}(\hat{\theta}), y^{as}(\hat{\theta}) - y^{ri}, q^{as}(\hat{\theta})\}_{\hat{\theta}\in\Theta}$ which is signed ex ante, regulates the basic service over the whole relationship. However it does not include any specifications regarding the add-on and when the second period arrives, parties agree on a

[47]If the add-on only becomes contractible at the interim stage, Proposition 9 also tells us that the outcome of $\mathcal{C}^{as}$ is only attainable when $H(u^{as}(\theta) + y^{as}(\theta), \varepsilon^{as}(\theta)) = 0$ (for instance in the CARA case).

spot contract for the add-on. This spot contract only involves an extra fixed payment y^{ri}, the levels of the add-on in the different states of nature $x^{as}(\cdot, \beta)$ and the associated risk ε^{ri}. Indeed, when anticipating this continuation for the second period, the firm chooses to always accept the spot contract because for any report $\hat{\theta}$ it has made earlier on:

$$w(t^{as}(\hat{\theta})+y^{as}(\hat{\theta})-\theta q^{as}(\hat{\theta}), \varepsilon^{ri}) = w(t^{as}(\hat{\theta})+y^{as}(\hat{\theta})-y^{ri}-\theta q^{as}(\hat{\theta}), 0) \Leftrightarrow w(y^{ri}, \varepsilon^{ri}) = 0 \tag{6.4}$$

where we have used the value of y^{ri} and the definition of w for CARA functions. Moreover, anticipating acceptance of such spot contract, the firm thus chooses to report θ truthfully since from (5.8) the following equality holds:

$$\max_{\hat{\theta}\in\Theta}(1-\delta)(t^{as}(\hat{\theta}) - \theta q^{as}(\hat{\theta})) + \delta w(t^{as}(\hat{\theta}) - \theta q^{as}(\hat{\theta}) + y^{as}(\hat{\theta}) - y^{ri} + y^{ri}, \varepsilon^{ri})$$

$$= \max_{\hat{\theta}\in\Theta}(1-\delta)(t^{as}(\hat{\theta}) - \theta q^{as}(\hat{\theta})) + \delta w(t^{as}(\hat{\theta}) - \theta q^{as}(\hat{\theta}) + y^{as}(\hat{\theta}), \varepsilon^{as}(\theta)) = \mathcal{U}^{as}(\theta).$$

We state this important finding in the next proposition.

PROPOSITION 10 *Suppose that the firm has CARA preferences. Then there is no loss of generality in delaying contracting on the add-on to the second period.*

This proposition bears some resemblance to Fudenberg, Holmström and Milgrom (1990). There are two main difference between the present paper and this more general inquiry. First, these authors allow the principal and the agent to borrow from the financial markets on equal terms. While this assumption is certainly relevant for some employment relationships, it is less relevant in our procurement context. Indeed our modeling of second-period risk aversion for the firm is precisely meant to capture such frictions. Second, while we insist on (repeated) adverse selection as a fundamental frictions in contracts, Fudenberg, Holmström and Milgrom (1990) study a repeated contracting model under the assumption that technology is common knowledge (in the sense that future contract outcomes is completely determined by current history). Those two conditions (plus other more technical requirements) are then shown to be sufficient to obtain the implementation of the optimal long-term contracts by means of spot contracts. Our model illustrates that such sequential optimality can be found in other more informationally constrained environments as well.

7. AUCTIONS AND RESERVE PRICES

Service provision in most infrastructure sectors (water, transportation, waste disposal, energy,...) is allocated among competing firms through competitive bidding. Yet, at the time of tender, both the bidders and the agency may not be aware of specific needs and costs from future stages of the project. Taken in isolation, the behavior of each individual bidder at the tendering stage may reflect how they perceive that risk. In response, the agency may have to modify reserve prices. In this section we investigate the direction in which it has to so.

Because our focus will be on competition, we simplify a bit the modeling on the demand side and assume that the agency wants to procure only one unit of good or service and the social value from this procurement is S. In the second period there will also only be one unit of the add-on to procure. Because of switching costs, or strong complementarities between provision at each stage, the winning firm in the first stage of the tender is necessarily in charge of producing the add-on later on.[48] We will assume that its value V is large enough (typically $V > \bar{\beta} + \frac{\nu}{1-\nu}\Delta\beta$ gives a sufficient condition) to ensure that this additional stage of the project is always valuable even under asymmetric information. Intuitively, this means that the firm gets paid a fixed amount for the provision of the add-on so that no screening in terms of quantities can help reduce the agency costs in the second period. From our earlier and more general analysis made in Section 5, asymmetric information on the second-period cost then requires the firm to bear all risk associated with these second-period returns. Formally, we have $\varepsilon(\theta) = \Delta\beta$ for all θ.

In the first period, the agency runs a first-price auction with a reserve price. We assume that there are $n + 1$ bidders whose first-period costs are independently drawn from the same distribution F. Similarly, these firms are symmetric in terms of the distribution of their second-period costs.

Given the ex ante symmetry across firms, we look for a symmetric equilibrium bidding strategy that determines the price $\{b(\hat{\theta})\}_{\hat{\theta}\in\Theta}$ for the basic service that is a function of the firm's announcement of its costs. As before, we denote by $\{y(\hat{\theta})\}_{\hat{\theta}\in\Theta}$ a bonus paid to the firm when it delivers the add-on. For the agency, fixing a reserve price amounts to defining a cutoff for the winning bidder's type above which it is preferable not to engage in the long-term project. We denote by $\tilde{\theta}$ such cut-off.[49]

[48] To motivate this setting, remember that tenders for long-lived projects like PPPs might involve consortia which are short-term ventures and are not even around later on, leaving thereby the winning firm and the agency in a situation of bilateral lock-in.

[49] In Section 5, we assumed Inada conditions on the surplus functions for the basic service and the add-on. In fact, those conditions implied that it was always optimal to procure the service even when the firm had the highest possible cost $\bar{\theta}$. With unit projects (which technically amounts to assuming linear surplus cum a capacity constraint, endogenizing the set of active types becomes a true concern. In this auction setting, it is done by fixing a reserve price.

Using symmetry and assuming that the bidding strategy $b(\hat{\theta})$ is increasing, the probability that any bidder claiming a first-period cost $\hat{\theta}$ wins is $(1-F(\hat{\theta}))^n$. This remark allows us to rewrite the requirement of incentive compatibility for a bidder with type θ as:

$$\mathcal{U}(\theta) = \max_{\hat{\theta}\in\Theta}(1-F(\hat{\theta}))^n \left((1-\delta)(b(\hat{\theta})-\theta) + \delta w(b(\hat{\theta})-\theta+y(\hat{\theta}),\Delta\beta)\right).$$

Following our earlier notations, we denote first-period profit conditionally on winning as $u(\theta) = b(\theta) - \theta$. Using the Envelope Theorem and the equilibrium value of $\mathcal{U}$, we can then rewrite a necessary condition for incentive compatibility as:[50]

$$(7.1) \qquad \dot{\mathcal{U}}(\theta) = -(1-F(\theta))^n \left(1-\delta+\delta w_z\left(\varphi\left(\frac{\frac{\mathcal{U}(\theta)}{(1-F(\theta))^n}-(1-\delta)u(\theta)}{\delta},\Delta\beta\right),\Delta\beta\right)\right).$$

Once a bidding strategy $b(\theta)$ (or alternatively a first-period profit $u(\theta)$) is given the rent profile $\mathcal{U}$ is fully determined by the differential equation (7.1) for the most efficient types $\theta \leq \tilde{\theta}$ altogether with the boundary condition:

$$(7.2) \qquad \mathcal{U}(\tilde{\theta}) = 0.$$

To compute the agency's expected payoff, observe that, thanks to equilibrium symmetry, everything happens as if the agency was dealing with a single firm but having a first-period cost drawn from the distribution of the minimum of $n+1$ independent variables. The corresponding distribution function is $G(\theta) = 1-(1-F(\theta))^{n+1}$ (with corresponding density $g(\theta) = (n+1)f(\theta)(1-F(\theta))^n$.

Next propositions summarize the main features of the solution to the agency's problem in this competitive bidding environment.

Proposition 11 *The second-period profits of the winning firm satisfy:*

$$(7.3) \qquad w_z(u^{as}(\theta)+y^{as}(\theta),\Delta\beta) = 1+\frac{F(\theta)}{f(\theta)}w_{zz}(u^{as}(\theta)+y^{as}(\theta),\Delta\beta) \leq 1, \quad \forall\theta\leq\tilde{\theta}.$$

When Assumptions 2 holds, these profits are higher than when second-period costs are common knowledge.

Proposition 11 captures the distortions needed to ease information revelation in the first period. Formally, (7.3) looks very similar to (5.2) with the only proviso that there is

[50]In the sequel, we will content ourselves with the analysis of a relaxed optimization problem for the principal where we neglect the second-order incentive compatibility constraint. Following Lemma 1, this condition now amounts to $w_z\left(u(\theta)+y(\theta),\Delta\beta\right)$ being decreasing in θ.

only one unit of procurement in each period (at least for all firms whose bids are below the reserve price). The agency commits to a high price for the delivery of the add-on as a means to reduce the marginal utility of income in the second-period. Doing so implies that inducing information revelation of the cost of the basic service becomes easier. This marginal utility of income effect is stronger when there is also asymmetric information in the second period. Such asymmetry makes it more valuable to backload payments for precautionary purposes.

Let us now turn to the determination of the reserve price below which the agency prefers not to ask for delivery. For the sake of comparison, it is useful to recall the value of the cutoff $\tilde{\theta}^{rn}$ that would be achieved with risk neutrality. Of course, when fixing the reserve price the agency trades off the overall value of the project (including the expected benefits from the add-on) and its cost taking into account information rents left to the winning firm. It is routine to determine that cutoff as the solution to:

$$S + \delta(V - E_\beta(\beta)) = \tilde{\theta}^{rn} + \frac{F(\tilde{\theta}^{rn})}{f(\tilde{\theta}^{rn})}. \tag{7.4}$$

To ensure an interior solution, we will from now on assume that

$$\underline{\theta} < S + \delta(V - E_\beta(\beta)) < \bar{\theta} + \frac{1}{f(\bar{\theta})}.$$

Similarly, we may define $\tilde{\theta}^{mu}$ and $\tilde{\theta}^{as}$ as the cutoffs when second-period costs are common knowledge and when they are privately known, respectively. We can now turn to the more general characterization of those cut-offs.

Proposition 12 *When β is common knowledge in the second period, the Marginal Utility of Income Effect increases the optimal cutoff $\tilde{\theta}^{mu}$:*

$$\tilde{\theta}^{mu} \geq \tilde{\theta}^{rn}. \tag{7.5}$$

When β is private information and Assumptions 2 holds, the Risk Effect decreases the optimal cutoff $\tilde{\theta}^{as}$:

$$\tilde{\theta}^{as} \leq \tilde{\theta}^{mu}. \tag{7.6}$$

Following the intuition built in Section 4.3, we expect that the *Marginal Utility of Income Effect* makes it less attractive for the firm to exaggerate costs in the first-period. This implies that the agency can raise the optimal reserve price beyond its value under risk neutrality. However, the impact of second-period uncertainty on that reserve price goes the other way around. Uncertainty has two impact. First, the *Risk Effect* mechanically requires to pay an extra risk premium to be paid to the firm. This reduces the value of contracting and decreases the reserve price. Second, second-period risk increases the marginal utility of income when $w_{z\varepsilon} \geq 0$ and makes the firm's first-period rent more costly. This also pushes towards lower a reserve price.

Example (CARA preferences - continued). Our example allows us to quantify the relative impact of both effects and show that their overall impact may be ambiguous. First, observe that (7.3) now gives us the following expression of second-period profits:

$$u^{as}(\theta) + y^{as}(\theta) = \frac{1}{r} ln\left(\eta(r, \Delta\beta)\right) + \frac{1}{r} ln\left(1 + r\frac{F(\theta)}{f(\theta)}\right), \quad \forall \theta \in \Theta.$$

Inserting this into (7.1) and taking into account (7.2), we obtain:

$$\mathcal{U}^{as}(\theta) = \int_{\theta}^{\tilde{\theta}^{as}} (1 - F(\theta))^n \left(1 - \delta + \frac{\delta}{1 + r\frac{F(s)}{f(s)}}\right) ds, \quad \forall \theta \in \Theta$$

where, after simplifications, $\tilde{\theta}^{as}$ solves

$$S + \delta(V - E_\beta(\beta)) = \tilde{\theta}^{as} + \frac{F(\tilde{\theta}^{as})}{f(\tilde{\theta}^{as})} + \delta\left(\frac{1}{r} ln\left(\eta(r, \Delta\beta)\right) + \frac{1}{r} ln\left(1 + r\frac{F(\tilde{\theta}^{as})}{f(\tilde{\theta}^{as})}\right) - \frac{F(\tilde{\theta}^{as})}{f(\tilde{\theta}^{as})}\right).$$

When r is sufficiently small, a Taylor expansion of the last bracket on the right-hand side shows that it is close to:

$$\frac{\delta r}{2}\left(\nu(1-\nu)(\Delta\beta)^2 - \frac{F^2(\tilde{\theta}^{as})}{f^2(\tilde{\theta}^{as})}\right).$$

From this, it immediately follows that for the *Risk Effect* dominates when $\tilde{\theta}^{as}$ is high enough while the *Marginal Utility of Income Effect* dominates otherwise.

8. DISCUSSION AND CONCLUSION

In this paper we unveil two important effects that occur when firms are risk averse and that influence the optimal procurement contract for a basic service and an add-on. First we show that by transferring payoffs to the second period, the second-period marginal utility of income is below the socially optimal marginal utility of income. This allows the procurement agency to ease the first-period incentive problem and reduce rents. Because incentive compatibility becomes less costly under this Marginal Utility of Income Effect, the level of the basic service is less distorted than under risk neutrality where this effect does not occur. However, there is also a Risk Effect. Because of asymmetric information, some risk has to be transferred to the risk-averse firm. This is first of all costly because it requires a risk premium to be paid to the firm. But it also increases the marginal utility of income and makes the first-period incentive problem more costly. This risk is somewhat reduced by increasing the distortions on the add-on. However, this highlights that even in the absence of technological linkage, there is still a contractual externality that is at work through the Marginal Utility of Income Effect. In general this externality leads to the optimal long-term contract not being renegotiation-proof or sequentially optimal.

We have on purpose kept the model as simple as possible and kept externalities at a minimum in order to show the pure effect of risk aversion. We will now discuss some of our simplifying assumptions.

First of all, we have focused on a basic service which level does not change over time. Although we believe that this is a natural assumption when we consider *durable* goods and services, there might also be cases where the level of the basic service can be adjusted over time. In this case, the final (second-period) level of the basic service would be less distorted as a result of the effects described in this paper and the initial level of the basic service would be given by the standard Baron and Myerson (1982) outcome.

The cost of the uncertain add-on is a binary random variable. In full generality, one could consider that this cost can take a continuum of values within a given support. Both Salanié (1990) and Laffont and Rochet (1998) show how even in *simple* settings bunching becomes an important issue with a continuum of types and risk aversion. In order to avoid dealing with this issue and to be able to focus completely on the effect of risky additional works on the optimal contract, we abstract from this issue. In our specific setting, only the mean and the variance of the pay-off on the risky add-on are important. In a more general setting other moments and the entire distribution itself might also matter.

We consider a complete contracting framework where ex ante the additional work can be included in the contract. In reality, contractors are often faced with unforeseen contingencies which could not be anticipated and written into the initial contract. In our analysis of renegotiation-proofness we show that the optimal full commitment contract is in general not renegotiation-proof. This also implies that whether the additional work can be included in the initial contract or not matters. Only in the CARA case do we obtain the same level of services and overall expected payments regardless of whether the additional work was included in the initial contract or added at a later stage. However, most contracts involve clauses that describe what happens when unforeseen risks materialize, in the UK standardized PFI contracts this is the change mechanism clauses. So to some extent, even unforeseen events are included and governed by the initial contract. To some extent this justifies considering complete contracts.[51]

Bundling of the basic service and the add-on into one contract is not a choice in our model. In fact, we assume that because of switching costs or strong complementarities between the provision at each stage, the two goods or services have to be bundled. This might be the most appropriate assumption for many add-ons, but not for all. When the firms exhibit CARA, the choice of bundling versus unbundling becomes irrelevant as expected gains for the procurement agency is the same under the two regimes.[52] However, in the more general case, this result does not hold. In fact, when contracting for the add-on with a separate firm (in a separate contract at the ex ante or interim stage), the contract for the basic service leads to higher welfare under unbundling compared to the welfare obtained from the part of the bundling contract that relates to the basic service. This is because the unbundled contract for the basic service involves no risk and the negative effect that risk has one the Marginal Utility of Income in the second period vanishes. However, the (unbundled) contract for the add-on is more costly than the part of the bundling contract that relates to the add-on. This is because unbundling leads to higher risk and because the firm does not benefit from the income effect from the basic service.

We have focused our analysis on service provision. However, our model can straightforwardly be reinterpreted in terms of the provision of a public good (for instance a building or some other kind of infrastructure) and our results still hold. In that case the discount factor should be reinterpreted as the probability that an add-on to the initial good is required and period i becomes state of the world i. Furthermore what was previously interpreted as backloaded payments should now be reinterpreted as payments being shifted to states in which the add-on is required. Alternatively the model presented in Section 2 can be rewritten to exclude surplus and production costs in the first period, i.e. both the production of the basic good and production of the add-on take place in the second period but the payment for the basic service can be made both upfront before the pro-

[51]Moreover, the Maskin-Tirole critique (Maskin and Tirole (1999)) applies to our setting and further justifies focusing on complete contracts.

[52]The literature on bundling versus unbundling has focused on investment and cost externalities (Hart (2003), Bennett and Iossa (2006), Martimort and Pouyet (2008) and Iossa and Martimort (forthcoming)). In our model there is no technological linkage and in the CARA case our model is in line with the results of this literature. A notable exception from the aforementioned literature is Schmitz (2013) who studies the effect of the government's budget constraint on the bundling decision.

duction phase starts or in the production phase. Our results still hold (qualitatively) in this modified setting.

REFERENCES

Alnuaimi, A., Taha, R., Al Mohsin, M., and A. Al-Harthi (2010), "Causes, Effects, Benefits, and Remedies of Change Orders on Public Construction Projects in Oman", *Journal of Construction Engineering and Management*, vol. 136 n^0 5, 615-622.

Armstrong, M. and D. E. M. Sappington (2007), *Recent Developments in the Theory of Regulation*, in Handbook of Industrial Organization, Volume 3, ed. Mark Armstrong and Robert H. Porter, 1557-1700. Amsterdam: Elsevier, North-Holland.

Arnott, R. and J.E. Stiglitz (1988), "Randomization with Asymmetric Information", *RAND Journal of Economics*, vol. 19, n^0 3, 344-362.

Asanuma, B. and T. Kikutani (1992), "Risk Absorption in Japanese Subcontracting: A Microeconometric Study of the Automobile Industry", *Journal of the Japanese and International Economies*, vol. 6, 1-29.

Asplund (2002). "Risk Averse Firms in Oligopoly", *International Journal of Industrial Organization*, vol. 20, 995-1012.

Athey S. and J. Levin (2001), "Information and Competition in US Forest Service Timber Auctions", *Journal of Political Economy*, n^0 109, 375-417.

Bagnoli, C. and T. Bergstrom (2005), "Log-Concave Probability and its Applications", *Economic Theory*, vol. 26, 445-469.

Bajari, P. and S. Tadelis (2001), Incentives Versus Transaction Costs: A Theory of Procurement Contracts", *RAND Journal of Economics*, vol. 32, n^0 3, 287-307.

Bajari, P., S. Houghton and S. Tadelis (2007), "Bidding for Incomplete Contracts: An Empirical Analysis of Adaptation Costs" *mimeo.*

Baron, D. and D. Besanko (1984), "Regulation and Information in a Continuing Relationship", *Information Economics and Policy 1*, 267-302.

Baron, D. and B. Myerson (1982), "Regulating a Monopolist with Unknown Costs", *Econometrica*, vol. 50, n^0 4, 911-930.

Battaglini (2005), "Long-Term Contracting with Markovian Consumers", *American Economic Review*, vol. 95, n^0 3, 637-658.

Bennett, J. and E. Iossa (2006), "Bundling and Managing Facilities for Public Services", *Journal of Public Economics*, vol. 90, 2143-2160.

Bester, H. and R. Strausz (2001), "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case", *Econometrica*, vol. 69, 1077-1098.

Burguet, R., J.-J. Ganuza and E. Hauk (2012), "Limited Liability and Mechanism Design in Procurement", *Games and Economic Behavior*, vol. 76, 15-25.

Calveras, A., J.-J. Ganuza and E. Hauk (2004), "Wild Bids. Gambling for Resurrection in Procurement Contracts", *Journal of Regulatory Economics*, vol. 26, n^0 1, 41-68.

Chakravarty, S. and W.B. MacLeod (2009), "Contracting in the Shadow of the Law", *RAND Journal of Economics*, vol. 40, n^0 3, 533-557.

Cleland, D.I. and L.R. Ireland (2008), "Project Manager's Handbook: Applying Best Practices across Global Industries", *McGraw-Hill*.

Courty, P. and H. Li (2000), "Sequential Screening", *The Review of Economic Studies*, vol. 67, 697-717.

Dewatripont, M. (1988), "Commitment Through Renegotiation-Proof Contracts with Third Parties", *Review of Economic Studies*, vol. 55, 377-390.

Eső, P. and L. White (2004), "Precautionary Bidding in Auctions", *Econometrica*, vol. 72, n^0 1, 77-92.

Faure-Grimaud, A. and D. Martimort (2003), "Regulatory Inertia", *RAND Journal of Economics*, vol. 34, n^0 3, 413-437.

Fudenberg, D., B. Holmström and P. Milgrom (1989), "Short-Term Contracts and Long-Term Agency Relationship", *Journal of Economic Theory*, vol. 51, 1-31.

Gil, R. and J. Marion (2013), "Self-Enforcing Agreements and Relational Contracting: Evidence from California Highway Procurement", *Journal of Law, Economics & Organizations*, vol. 29, n^0 2, 239-277.

Gollier, C. (2004), *The Economics of Risk and Time*, MIT Press.

Golosov, M., Tsyvinski, A. and Werning, I. (2006), *New dynamic public finance: a users guide*, in NBER Macroeconomics Annual 2006, Volume 21.

Guthrie, G. (2006), "Regulating Infrastructure: The Impact on Risk and Investment", *Journal of Economic Literature*, vol. 64, 925-972.

Hart, O. (2003), "Incomplete Contracts and Public Ownership: Remarks, and an Application to Public-Private Partnerships", *Economic Journal*, vol. 113, n^0 486, C69-C76.

Hartlev, K. and M. W. Liljenbøl (2013), " Changes to Existing Contracts Under the EU Public Procurement Rules and the Drafting of Review Clauses to Avoid the Need for a New Tender", *Public Procurement Law Review*, vol. 22, n^0 2, 51-73.

Holmström, B. (1979), "Moral hazard and Observability", *Bell Journal of Economics*, vol. 10, 74-91.

Holmström, B. and P. Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics, & Organization*, vol. 7, 24-52.

Holmström, B. and P. Milgrom (1994), "The Firm as an Incentive System", *American Economic Review*, vol. 84, n^0 4, 972-991.

Holt, C.A. and S.K. Laury (2002), "Risk Aversion and Incentive Effects," *American Economic Review*, vol. 92, n^0 5, 1644-1655.

Ibbs, W., Nguyen, L., and S. Lee (2007), "Quantified Impacts of Project Change", *Journal of Professional Issues in Engineering Education and Practice*, vol. 133, n^0 1, 45-52.

Ibbs, W. (2012), "Construction Change: Likelihood, Severity, and Impact on Productivity," *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, vol. 4 n^0 3, 67-73.

Iossa, E. and D. Martimort (forthcoming), "The Simple Micro-economics of Public Private Partnerships," *JJournal of Public Economic Theory.*

Kagel, J. (1995), "Auctions: A Survey of Experimental Research", *Handbook of Experimental Economics*, Princeton University Press.

Kawasaki, S. and J. McMillan (1987), "The Design of Contracts: Evidence from Japanese Subcontracting," *Journal of the Japanese and International Economies*, vol. 1, 327-349.

Kerzner H.R. (2013), *Project Management: Case Studies*, John Wiley & Sons.

Kimball, M.S. (1990), "Precautionary Saving in the Small and in the Large", *Econometrica*, vol. 58, n^0 1, 53-73.

Kimball, M.S. (1992), "Precautionary Motives for Holding Assets", in *The New Pelgrave Dictionary of Money and Finance*, vol. 3, 158-161.

Kimball, M.S. and P. Weil (2009), "Precautionary Saving and Consumption Smoothing across Time and Possibilities", *Journal of Money, Credit and Banking*, vol. 41, n^0 2-3, 245-284.

Kimball, M.S. (1993), "Standard Risk Aversion", *Econometrica*, vol. 61, n^0 3, 589-611.

Kocherlakota N.R. (2010), *The New Dynamic Public Finance*, Princeton University Press.

Kovac, E. and T. Mylovanov (2009), "Stochastic Mechanisms in Settings without Monetary Transfers: The Regular Case", *Journal of Economic Theory*, vol. 144, 1373-1395.

Laffont, J.J. (1994), *The New Economics of Regulation Ten Years After*, *Econometrica*, vol. 62, n^0 3, 507-537.

Laffont, J.J. and D. Martimort (2002), *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press.

Laffont, J.J. and J.C. Rochet (1998), "Regulation of a Risk Averse Firm", *Games and Economic Behavior*, vol. 25, n^0 2, 149-173.

Leland, H.E. (1968), "Saving and Uncertainty: The Precautionary Demand for Saving", *Quarterly Journal of Economics*, vol. 82, n^0 3, 465-473.

Leland, H.E. and D.H. Pyle (1977), "Informational Asymmetries, Financial Structure, and Financial Intermediation", *the Journal of Finance*, vol. 32, n^0 2, 371-387.

Martimort, D. and E. Iossa (2012), "Risk Allocation and the Costs and Benefits of Public–Private Partnerships", *RAND Journal of Economics*, vol. 43, n^0 3, 442-474.

Martimort, D. and J. Pouyet (2008), "To Build or Not to Build: Normative and Positive Theories of Public-Private Partnerships", *International Journal of Industrial Organization*, vol. 26, 393-411.

Martimort, D., J. Pouyet and W. Sand-Zantman (2014), "How to Regulate a Firm under a Hard Budget Constraint", *mimeo*.

Maskin E. and J. Riley (1984), "Optimal Auctions with Risk Averse Buyers", *Econometrica*, vol. 52, n^0 6, 1473-1518.

Maskin E. and J. Tirole (1999), "Unforeseen Contingencies and Incomplete Contracts", *Review of Economic Studies*, vol. 66, n^0 1, 83-114.

Matthews, S. (1984), "On the Implementability of Reduced Form Auctions", *Econometrica*, vol. 52, n^0 6, 1519-1522.

Menezes, C.F., C.H. Geiss and J.F. Tressler (1980), "Increasing Downside Risk", *American Economic Review*, vol. 70, n^0 5, 921-932.

Menezes, C.F. and X.H. Wang (2005), "Increasing Outer Risk", *Journal of Mathematical Economics*, vol. 41, n^0 7, 875-886.

Meredith, J.R. and S.J. Mantel (2009), *Project Management: A Managerial Approach*, John Wiley & Sons.

Milgrom, P. and Segal I. (2002), "Envelope Theorems for Arbitrary Choice Sets", *Econometrica*, vol. 70, n^0 2, 583-601.

Myerson, R. (1982), "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems", *Journal of Mathematical Economics*, vol. 10, 67-81.

Myerson, R. (1986), "Multistage Games with Communication", *Econometrica*, vol. 54, n^0 2, 323-358.

NAO, National Audit Office (2007), *Improving the PFI Tendering Process*, National Audit Office, HC 149.

NAO, National Audit Office (2008), *Making Changes in Operational PFI Projects*, National Audit Office, HC 205.

Pavan, A., I. Segal and J. Toikka (2013), "Dynamic Mechanism Design: A Myersonian Approach", mimeo.

Ross, S.A. (1973), "The Economic Theory of Agency: The Principal's Problem", *American Economic Review*, vol. 63, n^0 2, 134-139.

Salanié, B. (1990), "Sélection adverse et aversion pour le risque", *Annales d'Economie et de Statistiques*, vol. 18, 131-149.

Sandmo, A. (1970), "The Effect of Uncertainty on Saving Decisions," *The Review of Economic Studies,*, vol. 37, n^0 3, 353-360.

Schmitz, P.W.. (2013), "Public Procurement in Times of Crisis: The Bundling Decision Reconsidered", *Economics Letters*, vol. 121, 533-536.

Seierstad, A. and K. Sydsaeter (1987), *Optimal Control Theory with Economic Applications*, North-Holland, Amsterdam.

Serag, E., and A. Oloufa (2007), "Change Orders Impact on Project Cost, AC2007-3039", *American Society for Engineering Education*, http://www.icee.usm.edu/ICEE/conferences/asee2007/papers/3039_CHANGE_ORDERS_IMPACT_ON_pROJECT_COST.pdf.

Serag, E., Oloufa, A., Malone, L., and E. Radwan (2010), "Model for Quantifying the Impact of Change Orders on Project Cost for U.S. Roadwork Construction," *Journal of Construction Engineering and Management*, vol. 136, n^0 9, 1015-1027.

Shavell, S. (1979), "On Moral Hazard and Insurance", *Quarterly Journal of Economics*, vol. 93, 541-562.

Strausz, R. (2006), "Deterministic versus Stochastic Mechanisms in Principal-Agent Models", *Journal of Economic Theory*, vol. 128, 306-314.

Strausz, R. (2011), "Regulatory Risk under Optimal Monopoly Regulation" ,*Economic Journal*, vol. 121, 740-762.

White, L. (2008), "Prudence in Bargaining: The Effect of Uncertainty on Bargaining Outcomes", *Games and Economic Behavior*, vol. 62, 211-231.

Yun, M. (1999), "Subcontracting Relations in the Korean Automotive Industry: Risk Sharing and Technological Capability," *International Journal of Industrial Organization*, vol. 17, 81-108.

APPENDIX A: PROOFS

Proof of Lemma 1: *Necessity.* From Theorem 2 and Corollary 1 in Milgrom and Segal (2002) and the fact that x and q are bounded above, it follows immediately that $\mathcal{U}(\theta)$ is absolutely continuous and thus almost everywhere differentiable with (3.9) holding at any point of differentiability.

Sufficiency. We rewrite (3.8) as:

$$\text{(A.1)}\qquad \mathcal{U}(\theta) \geq \mathcal{U}(\hat{\theta}) + (1-\delta)(\hat{\theta}-\theta)q(\hat{\theta})$$
$$+\delta\left(w(u(\hat{\theta})+y(\hat{\theta})+(\hat{\theta}-\theta)q(\hat{\theta}),\varepsilon(\hat{\theta})) - w(u(\hat{\theta})+y(\hat{\theta}),\varepsilon(\hat{\theta}))\right) \quad \forall(\theta,\hat{\theta}).$$

Using (3.9) and absolute continuity, the rent profile $\mathcal{U}(\theta)$ satisfies:

$$\mathcal{U}(\theta) - \mathcal{U}(\hat{\theta}) = \int_{\theta}^{\hat{\theta}} q(\theta')\left(1-\delta+\delta w_z(u(\theta')+y(\theta'),\varepsilon(\theta'))\right) d\theta' \quad \forall(\theta,\hat{\theta}) \in \Theta^2$$

Condition (A.1) thus holds when:

$$\int_{\theta}^{\hat{\theta}} q(\theta')\left(1-\delta+\delta w_z(u(\theta')+y(\theta'),\varepsilon(\theta'))\right) d\theta' \geq$$
$$(1-\delta)(\hat{\theta}-\theta)q(\hat{\theta}) + \delta\left(w(u(\hat{\theta})+y(\hat{\theta})+(\hat{\theta}-\theta)q(\hat{\theta}),\varepsilon(\hat{\theta})) - w(u(\hat{\theta})+y(\hat{\theta}),\varepsilon(\hat{\theta}))\right).$$

Because w is concave in its first argument, we have:

$$w(u(\hat{\theta})+y(\hat{\theta})+(\hat{\theta}-\theta)q(\hat{\theta}),\varepsilon(\hat{\theta})) - w(u(\hat{\theta})+y(\hat{\theta}),\varepsilon(\hat{\theta})) \leq (\hat{\theta}-\theta)q(\hat{\theta})w_z(u(\hat{\theta})+y(\hat{\theta}),\varepsilon(\hat{\theta})).$$

A sufficient condition so that (A.1) holds is thus:

$$\int_{\theta}^{\hat{\theta}} q(\theta')\left(1-\delta+\delta w_z(u(\theta')+y(\theta'),\varepsilon(\theta'))\right) d\theta' \geq (\hat{\theta}-\theta)q(\hat{\theta})(1-\delta+\delta w_z(u(\hat{\theta})+y(\hat{\theta}),\varepsilon(\hat{\theta})),$$

or equivalently

$$\text{(A.2)}\qquad \int_{\theta}^{\hat{\theta}} \left((1-\delta)(q(\theta')-q(\hat{\theta})) + \delta(q(\theta')w_z(u(\theta')+y(\theta'),\varepsilon(\theta')) - q(\hat{\theta})w_z(u(\hat{\theta})+y(\hat{\theta}),\varepsilon(\hat{\theta})))\right) d\theta' \geq 0.$$

Observe now that $q(\theta)(1-\delta+\delta w_z(u(\theta)+y(\theta),\varepsilon(\theta)))$ weakly decreasing implies:

$$\int_{\theta}^{\hat{\theta}} q(\theta')(1-\delta+\delta w_z(u(\theta)+y(\theta),\varepsilon(\theta)))d\theta' \geq (\hat{\theta}-\theta)q(\hat{\theta})(1-\delta+\delta w_z(u(\hat{\theta})+y(\hat{\theta}),\varepsilon(\hat{\theta}))).$$

Hence, a sufficient condition to get (A.2) and thus (A.1) is given by

$$q(\theta)(1-\delta+\delta w_z(u(\theta)+y(\theta),\varepsilon(\theta))) \text{ weakly decreasing.}$$

Inserting into (3.9), this condition amounts to $\mathcal{U}$ convex. *Q.E.D.*

Proofs of Propositions 2 and 3 : Taking into account that $\varepsilon(\theta) \equiv 0$ and simplifying the principal's and the agent's objectives accordingly (and in particular taking into account that $w(z,0) = v(z)$), the agency's (relaxed) problem can be stated as follows:[53]

$$(\mathcal{P}_{mu}) : \quad \max_{(q(\theta),x(\theta,\beta),u(\theta),\mathcal{U}(\theta))} \int_{\underline{\theta}}^{\bar{\theta}} E_\beta \left(\mathcal{W}(q(\theta), x(\theta,\beta), u(\theta), 0, \mathcal{U}(\theta))\right) f(\theta) d\theta$$

subject to (3.10)-(3.12).

First, observe that at the optimum of $(\mathcal{P}_{mu})$ $x(\theta,\beta)$ is chosen at the efficient level $x^{mu}(\theta,\beta) = x^*(\underline{\beta})$ for all (θ,β). Accordingly, we may thus simplify the expression of the principal's second-period payoff as:

$$E_\beta(V(x^*(\beta)) - \beta x^*(\beta)) - \varphi\left(\frac{\mathcal{U}(\theta) - (1-\delta)u(\theta)}{\delta}, 0\right).$$

We now write the Hamiltonian $(\mathcal{P}_1)$ as:

$$\begin{aligned}
&\mathcal{H}_0(q,u,\mathcal{U},\theta) = \\
&f(\theta)\left(S(q) - \theta q - (1-\delta)u - \delta\varphi\left(\frac{\mathcal{U}-(1-\delta)u}{\delta}, 0\right) + \delta E_\beta[V(x^*(\beta)) - \beta x^*(\beta)]\right) \\
&-\lambda q\left(1 - \delta + \delta v'\left(\varphi\left(\frac{\mathcal{U}-(1-\delta)u}{\delta}, 0\right)\right)\right).
\end{aligned}$$

The optimization then follows the same steps (except for the optimization with respect to $\varepsilon(\theta)$ that is no longer here) as in the Proof of Propositions 6, 7 and 8. Details are left to the reader. *Q.E.D.*

Proofs of Propositions 4 and 5 : The agency's (relaxed) problem can now be written as follows:

$$(\mathcal{P}_2) : \quad \max_{(q(\theta),x(\theta,\beta),u(\theta),\varepsilon(\theta),\mathcal{U}(\theta))} E_\beta \left(\mathcal{W}(q(\theta), x(\theta,\beta), u(\theta), \varepsilon(\theta), \mathcal{U}(\theta))\right)$$

subject to (4.1) and (3.7).

Because $\varphi_z \geq 0$, (4.1) is necessarily binding. Inserting $\mathcal{U}(\theta) \equiv 0$ into the maximand, the maximand is decreasing in ε and thus obtained when (3.7) is also binding. Optimizing yields the result. In particular, it is trivial to check that second-distortions and profits in each periods are independent of the first-period cost. Moreover, the Inada condition $V'(0) = +\infty$ ensures that x^{ri} and thus ε^{ri} are positif. *Q.E.D.*

[53] We index the solution to this relaxed problem (indexed with a subscript mu). Similar notations are used below.

PROOF OF COROLLARY 1: From (4.11), (4.12) and the definition of z^*, we first get:

$$-\frac{1-\delta}{\delta}u^{ri} = z^*(\varepsilon^{ri}) \geq 0, \quad \forall\theta \in \Theta.$$

where the last inequality follows from $\varepsilon^{ri} > 0$, $z^*(0) = 0$ and the fact Assumption 3 holds so that z^* is non-decreasing.

Second, observe also that (4.12) implies:

$$y^{ri} = \varphi\left(-\frac{1-\delta}{\delta}u^{ri}, \varepsilon^{ri}\right) - u^{ri}.$$

Using again (4.11), we may as well rewrite:

$$y^{ri} = \varphi\left(z^*(\varepsilon^{ri}), \varepsilon^{ri}\right) + \frac{\delta}{1-\delta}z^*(\varepsilon^{ri}).$$

Let us now define $\psi(\varepsilon)$ as:

$$\psi(\varepsilon) = \varphi\left(z^*(\varepsilon), \varepsilon\right) + \frac{\delta}{1-\delta}z^*(\varepsilon).$$

Differentiating, we get:

$$\psi'(\varepsilon) = \left(\varphi_z\left(z^*(\varepsilon), \varepsilon\right) + \frac{\delta}{1-\delta}\right)\dot{z}^*(\varepsilon) + \varphi_\varepsilon\left(z^*(\varepsilon), \varepsilon\right).$$

Taking into account (2.1), we obtain:

$$\psi'(\varepsilon) = -\frac{w_{z\varepsilon}(\varphi(z^*(\varepsilon), \varepsilon), \varepsilon)}{w_{zz}(\varphi(z^*(\varepsilon), \varepsilon), \varepsilon)} + \frac{\delta}{1-\delta}\dot{z}^*(\varepsilon).$$

From Assumption 2, the first term is positive while Assumption 3 ensures that the second term is also positive. Because $\psi(0) = 0$ and $\varepsilon^{ri} > 0$, we deduce that:

$$y^{ri} = \psi(\varepsilon^{ri}) > 0.$$

Q.E.D.

PROOFS OF PROPOSITIONS 6, 7 AND 8 : We start with proving Proposition 6 before turning our attention to Propositions 7 and 8.

When both θ and β are private information, the agency 's (relaxed) problem can be written as follows:

$$(\mathcal{P}^{as}): \quad \max_{(q(\theta), x(\theta,\beta), u(\theta), \varepsilon(\theta), \mathcal{U}(\theta))} \int_{\underline{\theta}}^{\bar{\theta}} E_\beta\left(\mathcal{W}(q(\theta), x(\theta,\beta), u(\theta), \varepsilon(\theta), \mathcal{U}(\theta))\right) f(\theta)d\theta$$

subject to (3.7)-(3.10)-(3.12).

First, observe that the second-period incentive constraint (3.7) is necessarily binding at the optimum of $(\mathcal{P}^{as})$ so that $\varepsilon(\theta) = \Delta\beta x(\theta, \bar{\beta})$. Moreover, it should also be clear that optimizing with respect to $x(\theta, \underline{\beta})$ immediately gives us $x^{as}(\theta, \underline{\beta}) = x^*(\underline{\beta})$ for all θ. Accordingly, we may thus simplify the expression of the principal's second-period payoff:

$$\nu(V(x^*(\underline{\beta})) - \underline{\beta}x^*(\underline{\beta})) + (1-\nu)\left(V\left(\frac{\varepsilon(\theta)}{\Delta\beta}\right) - \bar{\beta}\frac{\varepsilon(\theta)}{\Delta\beta}\right) - \varphi\left(\frac{\mathcal{U}(\theta) - (1-\delta)u(\theta)}{\delta}, \varepsilon(\theta)\right).$$

Equipped with this expression, and denoting by λ the costate variable for (3.12) we now write the Hamiltonian for the problem ($\mathcal{P}^{as}$) as:

(A.3) $$\mathcal{H}(q, u, \varepsilon, \mathcal{U}, \theta) =$$
$$f(\theta)\Bigg(S(q) - \theta q - (1-\delta)u - \delta\varphi\left(\frac{U - (1-\delta)u}{\delta}, \varepsilon\right) + \nu\delta(V(x^*(\underline{\beta})) - \underline{\beta}x^*(\underline{\beta})) +$$
$$(1-\nu)\delta\left(V\left(\frac{\varepsilon}{\Delta\beta}\right) - \bar{\beta}\frac{\varepsilon}{\Delta\beta}\right)\Bigg) - \lambda q\left(1 - \delta + \delta w_z\left(\varphi\left(\frac{\mathcal{U} - (1-\delta)u}{\delta}, \varepsilon\right), \varepsilon\right)\right).$$

We shall assume that $\mathcal{H}(q, u, \varepsilon, \mathcal{U}, \theta)$ is concave in $(q, u, \varepsilon, \mathcal{U})$[54] and use the Pontryagyn Principle to get necessary and sufficient conditions for the optimum (see Chapter 2, Theorems 2 and 4 in Seierstad and Sydsaeter (1987)). These necessary and sufficient conditions are listed below.

• *Costate variable.* There exists λ which is is continuous and differentiable such that:

$$\dot{\lambda}(\theta) = -\frac{\partial \mathcal{H}}{\partial \mathcal{U}}(q(\theta), u(\theta), \varepsilon(\theta), \mathcal{U}(\theta), \theta)$$

or

(A.4)
$$\dot{\lambda}(\theta) = \left(f(\theta) + \lambda(\theta)q(\theta)w_{zz}\left(\varphi\left(\frac{\mathcal{U}(\theta) - (1-\delta)u(\theta)}{\delta}, \varepsilon(\theta)\right), \varepsilon(\theta)\right)\right)\varphi_z\left(\frac{\mathcal{U}(\theta) - (1-\delta)u(\theta)}{\delta}, \varepsilon(\theta)\right).$$

• *Transversality conditions.* Because (3.10) is necessarily binding at the optimum (otherwise, the principal's payoff could be improved by reducing $\mathcal{U}(\bar{\theta})$ by a small amount while still respecting the participation constraint for all types), the transversality conditions writes as:

(A.5) $$\lambda(\underline{\theta}) = 0.$$

• *Optimality condition with respect to u.* Using a first-order condition, we find:

(A.6)
$$f(\theta) = \varphi_z\left(\frac{\mathcal{U}(\theta) - (1-\delta)u(\theta)}{\delta}, \varepsilon(\theta)\right)\left(f(\theta) + \lambda(\theta)q(\theta)w_{zz}\left(\varphi\left(\frac{\mathcal{U}(\theta) - (1-\delta)u(\theta)}{\delta}, \varepsilon(\theta)\right), \varepsilon(\theta)\right)\right).$$

[54] See 8 for conditions under which this is the case.

• *Optimality condition with respect to q.* Using a first-order condition, we find:

$$S'(q(\theta)) = \theta + \frac{\lambda(\theta)}{f(\theta)}\left(1-\delta+\delta w_z\left(\varphi\left(\frac{\mathcal{U}(\theta)-(1-\delta)u(\theta)}{\delta},\varepsilon(\theta)\right),\varepsilon(\theta)\right)\right). \tag{A.7}$$

• *Optimality condition with respect to ε.* Using again a first-order condition, we find:

$$\frac{1-\nu}{\Delta\beta}\left(V'\left(\frac{\varepsilon(\theta)}{\Delta\beta}\right)-\bar{\beta}\right) = \varphi_\varepsilon\left(\frac{\mathcal{U}(\theta)-(1-\delta)u(\theta)}{\delta},\varepsilon(\theta)\right) \tag{A.8}$$
$$+q(\theta)\frac{\lambda(\theta)}{f(\theta)}\left(w_{zz}\left(\varphi\left(\frac{\mathcal{U}(\theta)-(1-\delta)u(\theta)}{\delta},\varepsilon(\theta)\right),\varepsilon(\theta)\right)\varphi_\varepsilon\left(\frac{\mathcal{U}(\theta)-(1-\delta)u(\theta)}{\delta},\varepsilon(\theta)\right)\right.$$
$$\left.+w_{z\varepsilon}\left(\varphi\left(\frac{\mathcal{U}(\theta)-(1-\delta)u(\theta)}{\delta},\varepsilon(\theta)\right),\varepsilon(\theta)\right)\right).$$

We now use these optimality conditions to derive more specific results.

• *Proposition 6.* Inserting (A.6) into (A.4) and simplifying yields:

$$\dot{\lambda}(\theta) = f(\theta)$$

which together with (A.5) yields

$$\lambda(\theta) = F(\theta). \tag{A.9}$$

Inserting this expression into (A.6), taking into account that $\varphi_z(z,\varepsilon) = \frac{1}{w_z(\varphi(z,\varepsilon),\varepsilon)}$ and the definition

$$u^{as}(\theta)+y^{as}(\theta) = \varphi\left(\frac{\mathcal{U}^{as}(\theta)-(1-\delta)u^{as}(\theta)}{\delta},\varepsilon^{as}(\theta)\right) \tag{A.10}$$

and simplifying yields (5.2). The inequality (5.3) then immediately follows from Assumption 2.

• *Proposition 7.* Inserting (A.9) into (A.6) and again simplifying using (A.10) gives us (5.4). The inequality (5.5) then immediately follows from Assumption 2.

From (5.2), we know that $w_z(u^{as}(\theta)+y^{as}(\theta),\varepsilon^{as}(\theta)) \le 1$. Therefore, (5.4) implies that

$$S'(q^{as}(\theta)) \le \theta + \frac{F(\theta)}{f(\theta)}, \quad \forall\theta\in\Theta.$$

and thus $q^{as}(\theta) \le q^{bm}(\theta)$.

• *Proposition 8.* Inserting (A.9) into (A.8) and simplifying using again (A.10) gives us (5.6). The inequality (5.7) then immediately follows from Assumption 3. *Q.E.D.*

Proof of Proposition 9: To check whether a long-term contract $\mathcal{C}$ (and in particular $\mathcal{C}^{as}$) is renegotiation-proof when the firm truthfully reveals its type in the first period, i.e., the first-period strategy $M(\cdot|\theta)$ puts unit mass on $\hat{\theta} = \theta$ and is thus fully revealing, we first look for the optimal continuation for the second period.

The first step is to observe that (6.1) becomes:

$$(A.11) \quad w(t(\theta) + y(\theta) + \tilde{y}(\theta) - \theta q(\hat{\theta}), \tilde{\varepsilon}(\theta)) \geq w(t(\theta) + y(\theta) - \theta q(\theta), \varepsilon(\theta)) \quad \forall \theta \in \Theta.$$

The agency's problem at the renegotiation stage can be can be written as follows:

$$(\mathcal{P}^{re}) : \quad \max_{(\tilde{x}(\theta,\beta),\tilde{\varepsilon}(\theta),\tilde{y}(\theta))} E_\beta \left[V(\tilde{x}(\theta,\beta)) - \beta \tilde{x}(\theta,\beta) \right] - \tilde{y}(\theta)$$

subject to (3.7)-(A.11).

Observe that for the usual reasons the participation constraint (A.11) is necessarily binding. Denoting again the second-period profit in the long-term contract as $u(\theta) = t(\theta) - \theta q(\hat{\theta})$ and the firm's reservation payoff for the second period as $w_0(\theta) = w(u(\theta) + y(\theta), \varepsilon(\theta))$, we can thus write:

$$(A.12) \quad u(\theta) + y(\theta) + \tilde{y}(\theta) = \varphi(w_0(\theta), \tilde{\varepsilon}(\theta)), \quad \forall \theta \in \Theta.$$

Inserting this expression of $\tilde{y}(\theta)$ into the maximand of $(\mathcal{P}^{re})$ and optimizing with respect to $\tilde{\varepsilon}(\theta)$, we get that (3.7) is necessarily binding. The last step of the optimization gives us $x^{re}(\theta, \underline{\beta}) = x^*(\underline{\beta})$ for all θ. When the realized second period cost is $\bar{\beta}$, the distortion of the add-on is given by:

$$(A.13) \quad (1-\nu)(V'\left(x^{re}(\theta,\bar{\beta})\right) - \bar{\beta}) = \Delta\beta \varphi_\varepsilon(w_0(\theta), \varepsilon^{re}(\theta)), \quad \forall \theta \in \Theta.$$

If $\mathcal{C}^{as}$ is renegotiation-proof, we should have:

$$w_0(\theta) = w(u^{as}(\theta) + y^{as}(\theta), \varepsilon^{as}(\theta)), \quad x^{re}(\theta,\bar{\beta}) = x^{as}(\theta,\bar{\beta}) \text{ and } \tilde{y}(\theta) = 0, \quad \forall \theta \in \Theta.$$

Inserting this into (A.13) yields:

$$(A.14) \quad (1-\nu)(V'\left(x^{as}(\theta,\bar{\beta})\right) - \bar{\beta}) = \Delta\beta \varphi_\varepsilon(w(u^{as}(\theta) + y^{as}(\theta), \varepsilon^{as}(\theta)), \varepsilon^{as}(\theta)), \quad \forall \theta \in \Theta.$$

Comparing with (5.6), we can conclude that whenever (6.3) holds $\mathcal{C}^{as}$ is not renegotiation-proof.

On the other hand, for CARA preferences $H(z, \varepsilon) = 0$ and thus the long-term contract $\mathcal{C}^{as}$ is renegotiation-proof. *Q.E.D.*

Proof of Proposition 10: The proof follows directly from Section 6. First Proposition 9 show that spot contracting under CARA yields $x^{as}(\theta, \beta)$ as the effective level of the add-on. In the text we also show that the first-period incentive compatibility constraint when anticipating the continuation contrat for the add-on is the same incentive constraint as in the full commitment longterm contract and (6.4) tells us that no additional rent is given to the firm. *Q.E.D.*

Proof of Propositions 11 and 12 : Given the type distribution G of the winning firm's bid, the agency's intertemporal payoff when dealing with this firm can be written as:

$$\mathcal{W}(u(\theta),\mathcal{U}(\theta)) = S-\theta+\delta(V-E_\beta(\beta))-(1-\delta)u(\theta)-\delta\varphi\left(\frac{\frac{\mathcal{U}(\theta)}{(1-F(\theta))^n}-(1-\delta)u(\theta)}{\delta},\Delta\beta\right).$$

The (relaxed) problem with this representative firm can now be written as follows:

$$(\mathcal{P}^{as}): \quad \max_{(u(\theta),\mathcal{U}(\theta),\tilde{\theta})}\int_{\underline{\theta}}^{\tilde{\theta}}\mathcal{W}(u(\theta),\mathcal{U}(\theta))g(\theta)d\theta \text{ subject to (7.1)-(7.2)}.$$

Equipped with this expression, and denoting by λ the costate variable for (7.1) we can now write the Hamiltonian for the problem $(\mathcal{P}^{as})$ as:

$$\mathcal{H}(u,\mathcal{U},\theta) = g(\theta)\mathcal{W}(u(\theta),\mathcal{U}(\theta)) -\lambda(1-F(\theta))^n\left(1-\delta+\delta w_z\left(\varphi\left(\frac{\frac{\mathcal{U}(\theta)}{(1-F(\theta))^n}-(1-\delta)u(\theta)}{\delta},\Delta\beta\right),\Delta\beta\right)\right).$$

Since $\mathcal{H}(u,\mathcal{U},\theta)$ is concave in $(u,\mathcal{U})$[55], we can use the Pontryagyn Principle to get necessary and sufficient conditions for the optimum (see Chapter 2, Theorems 2 and 4 in Seierstad and Sydsaeter (1987)). These necessary and sufficient conditions are listed below.

• *Costate variable.* There exists λ which is is continuous and differentiable such that:

(A.15)

$$\dot{\lambda}(\theta) = -\frac{\partial\mathcal{H}}{\partial\mathcal{U}}(u(\theta),\mathcal{U}(\theta),\theta) = \left((n+1)f(\theta)+\lambda(\theta)w_{zz}\left(\varphi\left(\frac{\frac{\mathcal{U}(\theta)}{(1-F(\theta))^n}-(1-\delta)u(\theta)}{\delta},\Delta\beta\right),\Delta\beta\right)\right)$$
$$\times\varphi_z\left(\frac{\frac{\mathcal{U}(\theta)}{(1-F(\theta))^n}-(1-\delta)u(\theta)}{\delta},\Delta\beta\right).$$

• *Transversality conditions.* Because (3.10) is necessarily binding at the optimum (otherwise, the principal's payoff could be improved by slightly reducing $\mathcal{U}(\tilde{\theta})$ by a small amount while still respecting participation constraint for all types), the transversality conditions writes again as (A.5).

• *Optimality condition with respect to u.* Using the first-order condition, we find:

(A.16)

[55]It is straightforward to show that $\mathcal{H}(u,\mathcal{U},\theta)$ is concave in $(u,\mathcal{U})$ iff it is concave in $\mathcal{U}$. The condition for concavity in $\mathcal{U}$ simplifies to $1\geq\frac{\lambda(\theta)}{(n+1)f(\theta)}w_{zz}\left(\left(\varphi\left(\frac{\frac{\mathcal{U}(\theta)}{(1-F(\theta))^n}-(1-\delta)u(\theta)}{\delta},\Delta\beta\right),\Delta\beta\right)\right)$, which always holds.

$$1 = \varphi_z \left(\frac{\frac{\mathcal{U}(\theta)}{(1-F(\theta))^n} - (1-\delta)u(\theta)}{\delta}, \Delta\beta \right) \left(1 + \frac{\lambda(\theta)}{(n+1)f(\theta)} w_{zz} \left(\varphi \left(\frac{\frac{\mathcal{U}(\theta)}{(1-F(\theta))^n} - (1-\delta)u(\theta)}{\delta}, \Delta\beta \right), \Delta\beta \right) \right).$$

We now use these optimality conditions to derive more specific results.

• *Proposition 11.* Inserting (A.16) into (A.15) yields $\dot{\lambda}(\theta) = (n+1)f(\theta)$.Taking into account (A.5) yields $\lambda(\theta) = (n+1)F(\theta)$. Inserting this expression into (A.16), and simplifying yields (7.3).

• *Proposition 12.* We follow Seierstad and Sydsaeter (1985, p.145) to get the following optimality condition with respect to $\tilde{\theta}$:

$$\text{(A.17)} \quad \mathcal{H}(u(\tilde{\theta}), \mathcal{U}(\tilde{\theta}), \tilde{\theta}) = 0.$$

Taking into account (7.2), this optimality condition can be written as:

$$S + \delta(V - E_\beta(\beta)) = \tilde{\theta}$$

$$+(1-\delta)u(\tilde{\theta}) + \delta\varphi\left(\frac{-(1-\delta)}{\delta}u(\tilde{\theta}), \Delta\beta\right) + \frac{F(\tilde{\theta})}{f(\tilde{\theta})}(1 - \delta + \delta w_z(u(\tilde{\theta}) + y(\tilde{\theta}), \Delta\beta)).$$

Using again (7.2) and the definition of $\varphi(\cdot)$, this condition can be simplified to:

$$\text{(A.18)} \quad S + \delta(V - E_\beta(\beta)) = \tilde{\theta} + \frac{F(\tilde{\theta})}{f(\tilde{\theta})}$$

$$+\delta\left(u(\tilde{\theta}) + y(\tilde{\theta}) - w\left(u(\tilde{\theta}) + y(\tilde{\theta}), \Delta\beta\right) + \frac{F(\tilde{\theta})}{f(\tilde{\theta})}(w_z(u(\tilde{\theta}) + y(\tilde{\theta}), \Delta\beta) - 1)\right).$$

Mutatid mutants, we can use those conditions in the case where β is common knowledge in the second period. It is enough to replace $\Delta\beta$ by 0 into (A.18) to get the following expression for θ^{mu}:

$$\text{(A.19)} \quad S + \delta(V - E_\beta(\beta)) = \tilde{\theta}^{mu} + \frac{F(\tilde{\theta}^{mu})}{f(\tilde{\theta}^{mu})}$$

$$+\delta\left(u(\tilde{\theta}^{mu}) + y(\tilde{\theta}^{mu}) - v(u(\tilde{\theta}^{mu}) + y(\tilde{\theta}^{mu})) + \frac{F(\tilde{\theta}^{mu})}{f(\tilde{\theta}^{mu})}(v'(u(\tilde{\theta}^{mu}) + y(\tilde{\theta}^{mu})) - 1)\right).$$

Define $\zeta(\theta) \equiv u^{mu}(\theta) + y^{mu}(\theta)$. From (7.3) we know that $\zeta(\theta)$ is the implicit solution to:

$$\text{(A.20)} \quad v'(\zeta(\theta)) = 1 + \frac{F(\theta)}{f(\theta)}v''(\zeta(\theta)).$$

We will use the function $\zeta(\theta)$ to define $J(\theta)$ as:

$$\text{(A.21)} \quad J(\theta) = \zeta(\theta) - v(\zeta(\theta)) + \frac{F(\theta)}{f(\theta)}(v'(\zeta(\theta)) - 1).$$

First observe that from (7.3) and the normalizations of $v(\cdot)$, $J(\underline{\theta}) = 0$. Second, differentiating and taking into account (A.20) yields:

$$\text{(A.22)} \quad \dot{J}(\theta) = \frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right)\frac{F(\theta)}{f(\theta)}v''(\zeta(\theta)) \leq 0$$

where the last inequality follows from Assumption 1. From this, it follows that $J(\tilde{\theta}^{mu}) < 0$ when $\tilde{\theta}^{mu} > \underline{\theta}$. Inserting into (A.18), we deduce that:

$$S + \delta(V - E_\beta(\beta)) < \tilde{\theta}^{mu} + \frac{F(\tilde{\theta}^{mu})}{f(\tilde{\theta}^{mu})}$$

which gives us (7.5).

From Assumption 2, we have:

$$-w\left(u(\theta)+y(\theta),\Delta\beta\right)+\frac{F(\theta)}{f(\theta)}w_z(u(\theta)+y(\theta),\Delta\beta) \geq -v(u(\theta)+y(\theta))+\frac{F(\theta)}{f(\theta)}v'(u(\theta)+y(\theta)).$$

Using this for $\theta = \tilde{\theta}^{as}$ and Inserting it into (A.18) yields:

$$\text{(A.23)} \quad S + \delta(V - E_\beta(\beta)) \geq \tilde{\theta}_1 + \frac{F(\tilde{\theta})}{f(\tilde{\theta})}$$

$$+\delta\left(u(\tilde{\theta}) + y(\tilde{\theta}) - v(u(\tilde{\theta}) + y(\tilde{\theta})) + \frac{F(\tilde{\theta})}{f(\tilde{\theta})}(v'(u(\tilde{\theta}) + y(\tilde{\theta})) - 1)\right)$$

which implies (7.6).

Q.E.D.

APPENDIX B: CONCAVITY CONDITIONS

In this appendix we present condition under which the Hamiltonian in (A.3) is indeed concave in its arguments. Before going to the general conditions, we present conditions under which the partial results of Marginal Utility of Income and Risk effects are valid.

Concavity conditions when θ is private information and β is ex post public information

In the case of only the Marginal Utility of Income Effect, there is no private information on β and therefore no risk has to be transferred to the agent ($\varepsilon = 0$). Straightforward (but tedious) computations show that the problem is concave iff

$$\text{(B.1)} \qquad f(\theta)S''(q(\theta))\mathcal{H}_{\mathcal{U}\mathcal{U}} - (\lambda(\theta)w_{zz}(\varphi(u(\theta)+y(\theta)),\varepsilon))\varphi_z(u(\theta)+y(\theta),\varepsilon)^2 \geq 0,$$

where $\mathcal{H}_{xy}$ denotes the derivative of $\mathcal{H}$ with respect to x and y.

Concavity conditions when θ is public information and β is ex post private information

In the case where β is private information at the interim stage but θ is public information, the condition requires that the concavity of $V(\cdot)$ is sufficiently strong relative to the curvature of $v(\cdot)$ (and $\varphi(\cdot)$). Formally,

$$\text{(B.2)}\qquad \varphi_{zz}(u(\theta)+y(\theta),\varepsilon)\frac{1-\nu}{(\Delta\beta)^2}V''(\frac{\varepsilon}{\Delta\beta})$$
$$+\varphi_{zz}(u(\theta)+y(\theta),\varepsilon)\varphi_{\varepsilon\varepsilon}(u(\theta)+y(\theta),\varepsilon)-(\varphi_{z\varepsilon}(u(\theta)+y(\theta),\varepsilon))^2\leq 0,$$

where the second term is not necessarily negative.

Concavity conditions when both θ and β are private information

In the general case where both θ and β are private information, condition (B.2) becomes

$$\text{(B.3)}\qquad f(\theta)S''(q(\theta))\left[\mathcal{H}_{\mathcal{U}\mathcal{U}}(q,u,\varepsilon,\mathcal{U},\theta)\mathcal{H}_{\varepsilon\varepsilon}(q,u,\varepsilon,\mathcal{U},\theta)-(\mathcal{H}_{\mathcal{U}\varepsilon}(q,u,\varepsilon,\mathcal{U},\theta))^2\right]$$
$$+2\mathcal{H}_{q\mathcal{U}}(q,u,\varepsilon,\mathcal{U},\theta)\mathcal{H}_{q\varepsilon}(q,u,\varepsilon,\mathcal{U},\theta)\mathcal{H}_{\mathcal{U}\varepsilon}(q,u,\varepsilon,\mathcal{U},\theta)-(\mathcal{H}_{q\varepsilon}(q,u,\varepsilon,\mathcal{U},\theta))^2\,\mathcal{H}_{\mathcal{U}\mathcal{U}}(q,u,\varepsilon,\mathcal{U},\theta)$$
$$-(\mathcal{H}_{q\mathcal{U}}(q,u,\varepsilon,\mathcal{U},\theta))^2\,\mathcal{H}_{\varepsilon\varepsilon}(q,u,\varepsilon,\mathcal{U},\theta)\leq 0,$$

while condition (B.1) remains unchanged.

In addition the following two conditions are also required:

$$\text{(B.4)}\qquad f(\theta)S''(q(\theta))\mathcal{H}_{\varepsilon\varepsilon}(q,u,\varepsilon,\mathcal{U},\theta)$$
$$-(\delta\lambda(\theta)\left[w_{z\varepsilon}(\varphi(u(\theta)+y(\theta),\varepsilon),\varepsilon)+w_{zz}(\varphi(u(\theta)+y(\theta),\varepsilon),\varepsilon)\varphi_{\varepsilon}(u(\theta)+y(\theta),\varepsilon)\right])^2\geq 0,$$

and

$$\text{(B.5)}\qquad \mathcal{H}_{\mathcal{U}\mathcal{U}}(q,u,\varepsilon,\mathcal{U},\theta)\mathcal{H}_{\varepsilon\varepsilon}(q,u,\varepsilon,\mathcal{U},\theta)-(\mathcal{H}_{\mathcal{U}\varepsilon}(q,u,\varepsilon,\mathcal{U},\theta))^2\geq 0.$$

It can be checked (again through straightforward but tedious computations) that in the case where the firm exhibits CARA preferences and r is small, these conditions are always satisfied. With CARA preferences, this is also the case when the risk is sufficiently small (regardless of the value of r).

Concavity conditions when there is competition for a fixed size project

In the case of fixed project size and competition, the Hamiltionian needs to be concave in $(u,\mathcal{U})$. It can be shown that this is the case whenever

$$\mathcal{H}_{\mathcal{U}\mathcal{U}}\equiv-\left[\frac{g(\theta)}{(1-F(\theta))^n}-\lambda(\theta)w_{zz}(\varphi(u(\theta)+y(\theta),\varepsilon),\varepsilon)\right]\frac{\varphi_{zz}(u(\theta)+y(\theta),\varepsilon)}{\delta(1-F(\theta))^n}\leq 0.$$

In fact, this condition always holds.