Working paper

DYNAMIC PROCUREMENT UNDER UNCERTAINTY: OPTIMAL DESIGN AND IMPLICATIONS FOR INCOMPLETE CONTRACTS[1]

January 20, 2016

Malin Arve and David Martimort

Abstract. We characterize the optimal dynamic contract for a long-term basic service when an uncertain add-on is required later on. Introducing firm risk aversion has two impacts. Profits for the basic service can be backloaded to induce cheaper information revelation for this service: An *Income Effect* which reduces output distortions. The firm must bear some risk to induce information revelation for the add-on. This *Risk Effect* reduces the level of the add-on but hardens information revelation for the basic service. The interaction between these effects has important implications for the dynamics of distortions, contract renegotiation and the value of incomplete contracts.

Keywords. Procurement, dynamic mechanism design, asymmetric information, uncertainty, change orders, risk aversion.

1. INTRODUCTION

Contracts for public utilities such as delegated management in water, sanitation and transportation are long-term contracts that last for up to several decades. For these kinds of long lasting relationships, ample evidence suggests that project managers and public authorities generally expect a certain amount of *ex post* adaptations, regardless of how well the project was planned and executed. For instance, the National Audit Office (NAO) acknowledges that over time most deals that have been signed under the umbrella of the UK *Public Finance Initiative* (PFI) need to be modified to meet inevitable but

[1]We thank participants at the *Mannheim Workshop on Procurement and Contracts*, the *PSE Workshop on Financing Investments in Crisis Times*, the *MaCCI Competition and Regulation Day*, the *Fundação Getulio Vargas Workshop on Regulatory Environment and Institutions in Public Procurement*, the 18^{th} *SFB Meeting*, the *Padua Workshop on How Governance Complexity and Financial Constraints Affect Public-Private Contracts*, the *Barcelona GSE Summer Forum* and *the NHH-Telenor Workshop*, seminar participants at the University of Copenhagen, University of Mannheim, University of Tilburg, Humboldt-Universität zu Berlin, University of Vienna and NHH Norwegian School of Economics, Bernard Caillaud, Vinicius Carrasco, Elisabetta Iossa, Juan José Ganuza, Susanne Goldlücke, Georgia Kosmopoulou, Antonio Miralles, Andras Niedermayer and Frank Rosar for extremely valuable comments. We also thank four referees and an Editor of this journal for useful comments that helped us improve the paper. Financial support from the Deutsche Forschungsgemeinschaft (SFB/TR-15) and the program Investissements d'Avenir of the French government (ANR-10-LABX-93-01) is gratefully acknowledged. All errors are ours.

[a]NHH Norwegian School of Economics, malin.arve@nhh.no
[b]Paris School of Economics-EHESS, david.martimort@parisschoolofeconomics.eu

uncertain changes in the costs and demands for public services.[1] In standard project management, modifications to an initial contract come in the form of *change orders*.[2] These change orders may require modifications of the specifications of the basic good or service itself, but they may also correspond to additional work. In fact, the NAO points out that this second scenario is the most frequent one; the majority of changes to UK PFIs are additions rather than direct changes to the type or level of the service provided.[3][4]

Although much has been reported on the role that uncertain add-ons play on contractual hazards by practitioners and legal scholars,[5] the full consequences of such uncertainty and its timely resolution on the design of long-term procurement contracts have not received much attention in the theoretical literature. This paper aims at filling this gap. Our overall goal is to evaluate how standard lessons from the procurement and incentive regulation literatures must be amended to take into account risk sharing between the firm and the Agency in charge when the procurement process involves future *add-ons* that go beyond the basic requirement of the service. How much of the future risk that might be borne by contractors should already be incorporated into the price and scale of the basic provision at early contracting stages? What are the consequences of the timely resolution of such uncertainty on the intertemporal profile of payments received by the contractor? To what extent might the addition of such add-ons destabilize provision of basic services, trigger costly renegotiation and why? And, on a more theoretical ground, how does agency problems that may arise when uncertainty on the add-on is resolved impact on earlier periods' incentives and efficiency? These are questions that certainly fare high on the agenda of practitioners in the field. Our answers have implications for the dynamics of long-term procurement contracts, their robustness to renegotiation, but also for the value of incomplete contracting for long-term projects whose characteristics may change over time.

[1] NAO (2008).

[2] Meredith and Mantel (2009).

[3] See NAO (2008). Other prominent examples of procurement where additional work was required are the *"Big Dig"* highway project in Boston that led to changes in more than 150 contracts (Bajari et al. (2014), Cleland and Ireland (2008)) and the Getty Center Art Museum in Los Angeles that had to be redesigned (without much change in the scope of the project) due to site conditions that were hard to anticipate *ex ante* (Bajari and Tadelis (2001) and Chakravarty and MacLeod (2009)).

[4] Uncertain future contract adjustments are also a concern in *private* procurement. An example in order is the telecommunication sector. Telecommunication companies often obtain spectrum via auctions and the outcome of these auctions is clearly uncertain. It will nevertheless affect to what extent the company needs to rely on its network of base stations and will therefore affect contractual relations with outside contractors for the maintenance and development of these base stations. For instance, if the company obtains less spectrum then expected, it will have to rely on more base stations and also heavier maintenance of the existing ones. This example is just one variation of practices often referred to as *Managed Services*. In this case uncertainty is related to the auction outcome and its effect on costs, but uncertainty could also come from product or technology dimensions. We thank Terje Ambjørnsen (Telenor), Bjørn Hansen (Telenor) and Timothy Wyndham (NHH) for pointing out this example.

[5] NAO (2008), Cox (1997) and Callahan (2012).

Main elements of the model. We consider a long-term procurement contract that covers two periods. In the first period, the firm supplies a basic service whose cost is private information. This basic service is durable (for instance, an infrastructure of a given size, a commitment to serve a fixed fraction of demand, etc.) and its level remains fixed over the entire length of the relationship. In the second period, an additional project is needed. The cost of this add-on is *ex ante* unknown and, for simplicity, this cost remains independent of that for the basic service. In the spirit of *contingent clauses*, we assume that the necessity of such an add-on can be anticipated in the initial agreement. Thus contracting for the add-on takes place under symmetric but incomplete information, although later on the firm will privately learn the realization of its cost. As a shortcut for a more complete modeling of the impact of the financial constraints that the firm faces when looking for outside finance related to the add-on, we assume that the firm is risk averse in the second period. It cares about its own share of the risk associated with the uncertain cost of the add-on. We are interested in the impact of such second-period *background risk* on the design of the optimal long-term procurement contract. Of course, the level of this risk depends on the informational environment surrounding the contract and this design depends on fine details of the information structure.

Income and Risk Effects. In a first pass, suppose that the cost of the add-on is verifiable. The Agency can fully insure the firm against this shock. As a result, the firm has no reason to manipulate information on earlier procurement stages to reduce its own risk exposure. Yet, the concavity of the firm's utility function has important consequences on contracting distortions. Indeed, the marginal gains from manipulating information on the cost of the basic service are evaluated at the marginal utility of income. Backloading profits for the basic service is thus a way for the Agency to decrease the firm's second period marginal utility and thereby relax the first-period incentive compatibility constraint that pertains to the revelation of the cost of the basic service: An *Income Effect.* Turning to output distortions, the level of the basic service comes closer to the first best than had the firm been risk neutral. Incentives on the basic service are high-powered and the scale of this project increases as a result of this *Income Effect.*

Consider now the more realistic scenario where the firm privately learns the cost of the add-on. The firm will of course require a greater compensation to cover this cost. Contracts now have to induce truthful revelation of such information, which implies distortions in the regulation of the add-on. To satisfy a second-period incentive compatibility constraint, the risk averse firm must now bear some *endogenous risk.* Even if the cost of the basic service were common knowledge, this *Risk Effect* would suffice to justify downward distortions of the level of the add-on and an additional increase of second-period payments to compensate for this risk. Reducing output diminishes the risk borne by the firm and saves on this risk premium.

Importantly, this (endogenous) *background risk* affects first-period incentives when the cost of the basic service is also private information. The firm may indeed be tempted to exaggerate this cost earlier on, asking for more compensation for the basic service so as to affect its future risk exposure. The timing of payments and the scale for the basic service should thus respond to these novel strategic possibilities. In other words, even if costs for the basic service and the add-on are uncorrelated and technologies unrelated, the solutions to the agency problems at the different stages cannot be disentangled. There exists an intertemporal *contractual externality* across agency problems.

It is interesting to see in more details how those agency problems interact. In this respect, remember that satisfying second-period incentive compatibility requires the firm to bear some risk. This endogenous second-period risk increases the marginal utility at this date, at least when the firm's utility function satisfies a standard assumption of *prudence*. As a result, the fact that the firm bears some risk on the add-on makes it even more valuable for it to manipulate first-period costs to enjoy more rent in the second period. The *Income Effect* is now exacerbated. Under these circumstances, the Agency backloads payments and profits for the basic service even more than when the cost of the add-on is common knowledge. Now the Agency is also more concerned with rent extraction and the distortion on the level of the basic service is more pronounced. In other words, risk in the second-period is a justification for greater output distortions and low powered contracts not only for the add-on but also for the basic service.

But the contractual externality also goes the other way. With *Decreasing Absolute Risk Aversion*, reducing the share of second-period risk borne by the firm also decreases its marginal utility of income at this date. As a result, the first-period incentive constraint exacerbates the *Risk Effect*. The optimal contract also exhibits lower powered incentives in the second period and the Agency implements stronger distortions of the add-on than when there is symmetric information in the first period.

Renegotiation and the Value of Incomplete Contracts. Our starting assumption is that the procurement Agency can fully commit to a comprehensive long-term contract that covers both the basic service and the add-on. Unfortunately, under very general conditions, the optimal long-term contract is *not renegotiation-proof*. Because of the intertemporal contractual externality that was just stressed, the Agency commits to strongly reducing the level of the add-on to make first-period incentives cheaper. However, once first-period costs have been revealed, these extra distortions are no longer needed. The Agency would want to renegotiate the contract towards higher powered schemes, letting the firm bear more risk for the sole purpose of inducing cheaper information revelation in the second period. Interestingly, with Constant Absolute Risk Aversion ($CARA$), the optimal *complete* long-term contract under full commitment remains renegotiation-proof. Furthermore, with $CARA$ preferences, the optimal long-term contract can be implemented by simply postponing contracting for the add-on to the second-period. An incomplete contract that only governs the basic service together with the addition of a spot contract for the add-on later on can replicate what can be achieved with an optimal (complete) long-term contract. This result certainly has some appeal for real-world practices since administrative and transaction costs may prevent parties from drafting complete contracts. Beyond the case of $CARA$ preferences, there is nevertheless a positive value of writing a long-term contract for both the basic service and the add-on. As a result, renegotiation of long-term contracts should be common in these environments. This feature certainly echoes real-world practices and practitioners' concern for the destabilizing effect of additional clauses in long-term contracts.

Justifying risk aversion. A key feature of our analysis and a significant departure from most of the existing literature is the explicit modeling of the firm's risk attitude in a procurement context. Economic theory in general and, the literature on procurement and regulation more specifically, has mostly taken the short-cut of considering risk neutral firms. This common view is based on two implicit assumptions, one being related to the firm's relationship with its outside financiers and the other being concerned with its internal organization. The first implicit assumption is that firms have perfect access to financial markets. Relying on signaling arguments, Leland and Pyle (1977) argue that, when firms have private information about future returns but lack cash to finance outlay investments, a credible signal to convey the firm's value to outside investors is the amount of risk kept by existing owners. Firms thus remain imperfectly diversified. Issuing debt may also entail costly bankruptcy or auditing costs as in the financial contracting literature.[6] These costs may be convex in the firm's value so as to make the firm's net payoff concave.[7] As pointed out by Asplund (2002), there might also be other reasons that justify such risk attitude, for instance liquidity constraints, costly financial distress, imperfect risk management and nonlinear tax systems.[8]

The second implicit assumption justifying risk neutrality is that firms do not suffer from any internal agency problems. This is in sharp contradiction with a whole body of works on the Theory of the Firm which,[9] by stressing the separation between ownership and control, has indeed pointed out the existence of an important trade-off between risk and incentives. Imperfect insurance in a moral hazard context is an incentive device that forces firms (viewed as coalitions between shareholders and managers) to remain imperfectly diversified. A similar trade-off arises when firms subcontract with independent units and these relationships are themselves plagued by agency problems. For instance, Kawasaki

[6] Townsend (1979), Gale and Hellwig (1985) and Bolton and Sharfstein (1990) among others.

[7] See Che and Gale (1998) for the consequences of such convexity in an auction context.

[8] Holmström and Tirole (1997) study the financing options of firms which are risk neutral but liquidity constrained. Risk neutrality can be questioned in the case of small ventures which remain poorly diversified.

[9] Holmström (1979) and Prendergast (1999).

and McMillian (1987), Asanuma and Kikutani (1992) and Yun (1999) have applied a simple principal-agent framework to empirically study subcontracting and risk sharing in the relationship between manufacturers and contractors in Japan and Korea and found that contractors are indeed risk averse. There is thus little doubt that risk matters for the behavior of contractors and that this ingredient should be part of a more complete view of the procurement contexts which are of prime interest for our study.[10]

LITERATURE REVIEW. Our paper touches upon several trends in the literature that are now reviewed.

Risk aversion in static contracting models. While the literature on risk aversion in moral hazard environments is huge,[11] that on risk aversion in adverse selection settings, and especially regulation and procurement, remains surprisingly sparse.[12] More than by irrelevance, this scarcity is probably best explained by the technical complexity encountered when risk aversion and incentives interact. Salanié (1990) illustrates this complexity in his study of an adverse selection problem where contracting takes place *ex ante*, i.e., before the risk averse agent learns private information about his cost parameter. Laffont and Rochet (1998) instead focus on *ex post* participation constraints while Baron and Besanko (1987) introduce monitoring by the principal. The general take-away from this literature is that risk aversion shifts optimal contracts towards being low powered, with greater output distortions and lower informational rents than under risk neutrality. Bunching may possibly also arise in the limit of large degrees of risk aversion.[13]

Long-term procurement contracts are mostly allocated through tenders. Auction and bargaining mechanisms 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) and Campo et al. (2011)). These studies suggest that the assumption of risk neutrality is not always appropriate. Maskin and Riley (1984) and Matthews (1984) offer the first theoretical analysis of risk averse bidders while, on a more applied ground, McAfee and McMillan (1986) demonstrate how auctions of incentive contracts are tilted towards low-powered incentives. Eső and White (2004) show that bidders exhibiting decreasing absolute risk aversion may shade their bids for pure *"precaution-*

[10]We ignore the issue of risk aversion for the principals. Modeling risk averse principals, perhaps as a proxy for the budgetary pressures on public authorities, is an interesting extension that lies beyond the scope of this paper. We refer to Lewis and Sappington (1995) and Martimort and Sand-Zantman (2007) for examples of optimal regulatory designs by risk averse local governments.

[11]See Prendergast (1999) for a survey.

[12]This has been stressed in some important surveys in the field (Laffont (1994) and Armstrong and Sappington (2007)).

[13]Bunching might arise in these papers because the authors analyze continuous-type models. In most of the paper, we focus on a two-type model that generates some endogenous second-period risk in response to the incentive compatibility problem. It is well know that risk aversion in two-type models does not introduce bunching in canonical models (see for instance Laffont and Martimort (2002) (Chapter 2.11.2)). This simplifying assumption allows us to stress the role of such endogenous risk in its simplest form. The case of where second-period costs take values in a continuum is nevertheless studied in Section 8.

ary bidding" purposes.[14] In a dynamic bargaining context, White (2008) demonstrates how such precautionary behavior makes players more patient. Our procurement model differs in many respects from these papers but also shares some common concerns. In sharp contrast with Eső and White (2004) and White (2008), background risk is here endogenously determined by second-period incentive compatibility constraints.[15]

Optimal dynamic mechanism design. From a more theoretical viewpoint, our paper belongs to a quite active literature on dynamic mechanism design (see Baron and Besanko (1984) for a seminal contribution and Pavan, Segal and Toikka (2014) for a more recent venture) which is much too vast to be successfully summarized in the limited space of this review. Dynamic mechanism design stresses the value of history in long-term relationships, especially when types are serially correlated (Battaglini (2005), Zhang (2009), Battaglini and Lamba (2012), Eső and Szentes (2013), Kapicka (2013), Garrett and Pavan (2015)), when preferences are only unveiled over time (Courty and Li (2000), Krähmer and Strausz (2011)) and when current projects affect future technological frontier (Lewis and Yildirim (2002), Gärtner (2010), Auray et al. (2011)). In a nutshell and re-interpreting these models in the procurement environment of the present paper, this body of works shows how the Myersonian notion of *"virtual cost parameter"* must be amended to take into account dynamic considerations. This leads to output distortions that differ from the static optimal distortions. Because the bulk of this literature is developed for quasi-linear preferences, no attention has been devoted to the impact of current incentives on future marginal utility of income and thus on output distortions; something which is the focus of our procurement model. In other words, in this paper we revisit the notion of *"virtual cost parameter"* and its impact on output distortions in a dynamic and uncertain environment where concerns for income smoothing and insurance matter.

As pointed out by Bergemann and Pavan (2015) in their authoritative survey, moving beyond quasi-linearity in optimal dynamic design is of prime importance to understand key issues in intertemporal consumption smoothing, taxation, and insurance. That asymmetric information impedes income smoothing over time and insurance over states of nature is indeed at the core of an important literature on optimal dynamic insurance schemes in macroeconomic environments (Townsend (1982), Green (1987), Thomas and Worall (1988), Farinha Luz (2013)). In this literature, agents are subject to endowment shocks which are private information and are not productive as in our procurement context. Production is instead a key aspect of the *New Dynamic Public Finance* literature which

[14]There is a related strand of the auction literature that studies bidding behavior under uncertainty. Calvares et al. (2004) and Burguet et al. (2012) study the effect of cost uncertainty on firms' bidding behavior. Under limited liability, firms which are financially weaker tend to bid more aggressively. Arve (2014) characterizes the optimal procurement contract with risk neutral, but possibly financially constrained, firms.

[15]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 elected political principals in charge of designing regulatory policies.

is interested in finding formulae for the dynamics of the wedge between labor wages and productivity in Mirrleesian environments with serially correlated income shocks (Farhi and Werning (2013), Stantcheva (2014), Makris and Pavan (2015), Golosov et al. (2006)).

Limited commitment. An important idea highlighted by the dynamic mechanism design literature is that commitment to future actions and payments helps screen current private information and reduces information rents. Yet, lessons from this literature are hardly robust when commitment is limited. Risk aversion is especially important in regulation and procurement when some production stages are long-lasting and firms may be reluctant to enter into extra rounds of contracting as they become available. In these cases, the public Agency in charge 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, this value of commitment disappears. In this case, the optimal contract is *renegotiation-proof* and, surprisingly, can even be implemented through a sequence of spot contracts. This finding is reminiscent of important work by Fudenberg, Holmström and Milgrom (1990), although the latter develop their theory in a pure moral hazard environment.

Organization. The model is presented in Section 2. Section 3 characterizes the set of incentive contracts that are feasible in our dynamic environment. Section 4 provides benchmark results for the case of risk neutrality. The general characterization of the optimal dynamic contract is undertaken in Section 5. This section shows how the *Income* and *Risk Effects* interact. Section 6 provides some partial results to illustrate these two effects separately so as to better understand their respective roles on contractual design. Section 7 discusses various limits on contracting and investigates to what extent contractual incompleteness is an obstacle to efficient incentives. In this section, we analyze the value of unbundling basic services and add-ons between several contractors, the cost of renegotiation and the value of spot contracting. Section 8 investigates the case where the second-period project is lumpy and second-period costs are drawn from a continuous distribution. Although details differ, we show that our main results remain by and large robust. Section 9 briefly concludes. Proofs are relegated to an Appendix while an online Appendix gathers material not of prime interest for the analysis in the main text.

2. THE MODEL

We consider the following model of procurement: An Agency (henceforth *the principal*) contracts with a firm for the provision of a service. The first component of this service is a durable component whose costs and benefits accumulate over two periods while an add-on is only required in the second period. To be more precise, the basic service is supplied in quantity q_1 in each period while the add-on is provided in quantity q_2 only in the second period.[16] The add-on represents long-term contractible variables associated with the service (a refined specification, some incremental services for new segments of demand, further stages of development of a prototype in defense procurement, etc.). The exact specifications required for the add-on are not completely known *ex ante* by the contracting parties. Uncertainty around the add-on puts the firm's returns at risk.[17] In the sequel, we will be interested in the impact of this risk on contract design.

Technology and preferences. The basic service generates a gross surplus $S_1(q_1)$ in each period. Motivated by the idea that this basic service is the choice of an infrastructure of a given capacity, the fixed size of a network, or the basic version of a long-term durable good, we assume that the quantity q_1 is chosen once and for all, and does not vary over time. The firm provides this service at a constant marginal cost θ_1. The function $S_1(\cdot)$ is increasing and strictly concave ($S_1' > 0, S_1'' < 0$) with the standard assumptions $S_1(0) = 0$ and the Inada condition $S_1'(0) = +\infty$.[18] The gross surplus from consuming q_2 units of the add-on is $S_2(q_2)$ where $S_2(\cdot)$ is increasing, strictly concave ($S_2' > 0, S_2'' < 0$) with $S_2(0) = 0$ and the Inada condition $S_2'(0) = +\infty$. The firm can provide the add-on at a constant marginal cost θ_2.

The fixed payments for $q_1(\theta_1)$ units of the basic service are denoted by $t_1(\theta_1)$ and $t_1(\theta_1) + y(\theta_1)$ for periods 1 and 2 respectively. The premium $y(\theta_1)$ captures the possible non-stationarity of payments for this service even though the quantity $q_1(\theta_1)$ remains constant. The second-period payment for the add-on is denoted by $t_2(\theta_1, \theta_2)$.

Denoting by $1-\beta$ and β the relative weights on the first and second period respectively,[19] the principal's expected gains from dealing with a firm of type θ_1 can be written as:

$$S_1(q_1(\theta_1)) - t_1(\theta_1) - \beta y(\theta_1) + \beta \mathbb{E}_{\theta_2} \left(S_2(q_2(\theta_1, \theta_2)) - t_2(\theta_1, \theta_2) \right).$$

Denoting by $u_1(\theta_1) = t_1(\theta_1) - \theta_1 q_1(\theta_1)$ the firm's first-period profit from the basic service and by $\mathcal{U}_2(\theta_1, \theta_2) = t_2(\theta_1, \theta_2) - \theta_2 q_2(\theta_1, \theta_2)$ its second-period profit from the add-on, the

[16]To make the model more realistic, we could allow for the add-on to only occur with some probability. This extra layer of uncertainty does not qualitatively change our results and is for simplicity ignored.

[17]We focus on the case where the basic service and the add-on are bundled. A first justification is that, in many long-term contracts such as Public-Private Partnerships, 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 would incur excessive costs of entry at later stages, maybe because of project specific developments during earlier periods or simply because it is not physically possible to find two different providers for the basic service and the add-on. The issue of unbundling is discussed in Section 7.1.

[18]These conditions ensure that *shutting-down* production even with the least efficient service provider is never optimal. This simplification allows us to concentrate on the impact at the intensive margin of second-period risk on first-period incentives. Section 8 investigates how our results might be modified when not providing the add-on is a relevant option.

[19]This formulation is equivalent to considering a discount factor δ and normalizing payoffs in two periods so that their discounted weights $\frac{1}{1+\delta}$ and $\frac{\delta}{1+\delta}$ sum up to one. The discount factor could be reinterpreted as the probability that an add-on is required.

principal's intertemporal payoff becomes:

$$S_1(q_1(\theta_1)) - \theta_1 q_1(\theta_1) - u_1(\theta_1) + \beta \mathbb{E}_{\theta_2}\left(S_2(q_2(\theta_1,\theta_2)) - \theta_2 q_2(\theta_1,\theta_2) - y(\theta_1) - \mathcal{U}_2(\theta_1,\theta_2)\right). \tag{2.1}$$

This expression highlights the rent-efficiency trade-off that characterizes optimal contracting under informational asymmetries. The principal cares about the social value of the project but would also like to minimize the share that accrues to the firm.

With these remarks in mind, we may also express the firm's intertemporal payoff as:

$$(1-\beta)u_1(\theta_1) + \beta \mathbb{E}_{\theta_2}\left(v\left(u_1(\theta_1) + y(\theta_1) + \mathcal{U}_2(\theta_1,\theta_2)\right)\right).$$

More on Preferences. We are interested in the consequences for optimal contract design of introducing uncertainty on the cost of the add-on. While we assume that the firm remains risk neutral w.r.t. first-period returns, studying how risk sharing impact on incentives requires to move away from more standard models of procurement and assume that the firm is risk averse w.r.t. second-period returns. Our assumption that risk aversion changes over time certainly deserves some comments. As suggested above, second-period risk aversion should be viewed as a proxy for the existing constraints that might limit the firm's access to the capital market when it raises outside funds to finance the necessary outlay investments associated with the add-on. That the firm remains risk neutral w.r.t. the first-period returns thus captures the idea that returns on the basic service are well-known and stable enough to limit these costs of outside finance.[20]

The firm's Bernoulli utility function $v(\cdot)$ is thus increasing and concave, ($v' > 0$, $v'' \leq 0$) with the normalizations $v(0) = 0$ and $v'(0) = 1$. We also assume that $v(\cdot)$ satisfies standard properties in the risk literature:[21]

Assumption 1 *Decreasing (resp. constant) absolute risk aversion (DARA) (resp. CARA):*

$$\frac{d}{dz}\left(-\frac{v''(z)}{v'(z)}\right) < 0 \ (\textit{resp.} = 0) \quad \forall z.$$

DARA can easily be motivated when risk aversion is viewed as a proxy for costly access to financial markets. Indeed, firms which already benefit from an activity (the basic service) that generates stable returns that can be used as pledgeable collateral also face less tight constraints and requirements on these markets.

To express the firm's intertemporal payoff in a more compact form, a first useful step

[20] We nevertheless discuss what happens when risk aversion applies to both periods at the end of Section 6.1.

[21] Holt and Laury (2002).

is to evaluate general properties of payoff functions once one adds some (zero-mean) background risk to a project that yields a fixed profit z. To fix ideas, this level of profit is what can be earned from running the basic service while the background risk comes from the add-on. As we will see below, returns on the add-on depend on the cost realization. They are thus random, can be normalized so as to have a zero mean and take values $(1-\nu)\varepsilon \geq 0$ with probability ν and $-\nu\varepsilon \leq 0$ with probability $1-\nu$, where ε can be viewed as a measure of risk. Consider thus a utility function $w(z,\varepsilon)$ defined over wealth z and risk levels $\varepsilon \geq 0$ as:

$$w(z,\varepsilon) \equiv \nu v(z+(1-\nu)\varepsilon) + (1-\nu)v(z-\nu\varepsilon). \tag{2.2}$$

The function $w(\cdot)$ inherits some important properties from $v(\cdot)$ as it is also increasing and concave in z. It is also decreasing in ε which captures the fact that more background risk reduces expected payoff. A last important property is that the cross derivative $w_{z\varepsilon}$ is non-negative as $v''' \geq 0$ and the firm exhibits *prudent* behavior when Assumption 1 holds. In other words, more background risk increases the firm's marginal value of income:[22]

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

For further references, let $\varphi(\zeta,\varepsilon)$ be the wealth level that guarantees ζ utils to the firm when the risk level is ε, i.e., $\zeta = w(\varphi(\zeta,\varepsilon),\varepsilon)$. The function $\varphi(\cdot)$ is increasing in ζ and ε.[23] For the sequel, it is useful to define the function $H(\cdot)$ as:

$$H(z,\varepsilon) \equiv w_{z\varepsilon}(z,\varepsilon) - \frac{w_{zz}(z,\varepsilon)w_{\varepsilon}(z,\varepsilon)}{w_z(z,\varepsilon)}. \tag{2.3}$$

That $H(\cdot)$ is non-negative follows from Assumption 1.[24] Importantly, $\frac{dw_z}{d\varepsilon}(\varphi(\zeta,\varepsilon),\varepsilon) = H(\varphi(\zeta,\varepsilon),\varepsilon)$. Hence, the fact that $H(\cdot)$ remains non-negative means that the marginal utility of income increases with ε if the firm's utility is left unchanged by raising z. The sign of this total derivative plays an important role in understanding agency distortions because it shows how a change in second-period risk sharing impacts on profits.

Information. At the time of contracting the firm has private information on the cost parameter θ_1. This variable is drawn from a (common knowledge) and atomless cumulative distribution $F(\cdot)$ with an everywhere positive density $f(\theta_1)$ whose support is $\Theta_1 = [\underline{\theta}_1, \overline{\theta}_1]$. Following a standard assumption in the screening literature,[25] the *monotone hazard rate property* holds:

[22] The concept of *prudence* goes back to Leland (1968) and Sandmo (1970). Experimental evidence (Deck and Schlesinger (2014), Noussair et al (2014)) is in line with this assumption.

[23] We have $\varphi_\zeta(\zeta,\varepsilon) = \frac{1}{w_z(\varphi(\zeta,\varepsilon),\varepsilon)} > 0$, and $\varphi_\varepsilon(\zeta,\varepsilon) = -\frac{w_\varepsilon(\varphi(\zeta,\varepsilon),\varepsilon)}{w_z(\varphi(\zeta,\varepsilon),\varepsilon)} > 0$.

[24] See Appendix B (Lemma B.1).

[25] Bagnoli and Bergstrom (2005).

Assumption 2 *Monotone hazard rate property:*

$$\frac{d}{d\theta_1}\left(\frac{F(\theta_1)}{f(\theta_1)}\right) \geq 0 \text{ for all } \theta_1 \in \Theta_1.$$

To capture the idea that the add-on is not yet completely defined at the time of contracting, we assume that its cost is uncertain at this stage. *Ex ante*, there is symmetric but incomplete information on the cost parameter θ_2. However, before producing the add-on, the firm learns θ_2. To maintain a tractable analysis throughout most of the paper, we consider the case where θ_2 is drawn from a common knowledge distribution on the discrete support $\Theta_2 = \{\underline{\theta}_2, \overline{\theta}_2\}$ (where $\Delta\theta_2 = \overline{\theta}_2 - \underline{\theta}_2 > 0$) with respective probabilities ν and $1-\nu$, where $\nu \in (0,1)$.[26,27]

First- and second-period cost parameters are independently drawn and, more generally, there is no technological linkage between periods. When deriving the optimal dynamic contract in this environment, any departure from the contract that would be optimal with a risk neutral firm comes from the fact that the firm is risk averse in the second period. Furthermore, our analysis unveils conditions under which a *contractual externality* between the first and the second period arises. Indeed, despite type independence and the absence of technological linkage, the design of first- and second-period incentives cannot be disentangled.

Contracts. The principal commits to a long-term contract that regulates the basic service over both periods and the add-on in the second period. There are several justifications for this commitment assumption. First, in some contexts 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. Changes that are outside these clauses might also be limited by law. For instance, 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. Second, the desire to build a reputation might force the principal to stick to her initial commitment even if renegotiation becomes attractive. Third, focusing on the full commitment scenario characterizes an upper bound on what long-term contracting can achieve. This allows us to unveil an important intertemporal linkage across periods that arises only because of asymmetric information.[28]

From the (dynamic version of the) Revelation Principle (Baron and Besanko

[26]Section 8 investigates the case where θ_2 is drawn from a continuous distribution.

[27]In some environments, the firm may already have some private signals about the cost of the add-on even at the *ex ante* stage. Existing models of sequential screening, for instance Courty and Li (2000) and Krähmer and Strausz (2011), suggest that such an extra layer of informational asymmetries might be a source of further information rents for the firm. As long as this early signal remains imperfect and future uncertainty puts the firm's profits at risk, the effects described in this paper will persist.

[28]The issues of renegotiation and incomplete contracts are discussed in Sections 7.2 and 7.3.

(1984), Myerson (1986)), there is no loss of generality in restricting the analysis to incentive-compatible direct revelation mechanisms. A mechanism, denoted by $\mathcal{C}$, thus stipulates payments and outputs for each period as a function of the firm's report of its current type and, possibly, the past history of reports, namely $\mathcal{C} = \left\{t_1(\hat{\theta}_1), y(\hat{\theta}_1), q_1(\hat{\theta}_1), t_2(\hat{\theta}_1, \hat{\theta}_2), q_2(\hat{\theta}_1, \hat{\theta}_2)\right\}_{\hat{\theta}_1 \in \Theta_1, \hat{\theta}_2 \in \Theta_2}$ where $\hat{\theta}_1$ and $\hat{\theta}_2$ are the firm's announcements of its cost parameters for the basic service and the add-on respectively.[29] These reports are of course truthful in equilibrium.

Timing. The contracting game unfolds as follows:

1. The firm privately learns its cost parameter θ_1 for the basic service.
2. The principal offers the long-term contract $\mathcal{C}$. If the firm refuses, parties get their reservation payoffs which are, without loss of generality, normalized to zero.
3. Following acceptance, the firm reports its cost for the basic service. This report $\hat{\theta}_1$ determines both the level of the basic service $q_1(\hat{\theta}_1)$ and the payments $t_1(\hat{\theta}_1)$ and $t_1(\hat{\theta}_1) + y(\hat{\theta}_1)$ in each period.
4. The firm learns the value of the cost of the add-on, θ_2. It then reports $\hat{\theta}_2$ and provides the corresponding level of the add-on, $q_2(\hat{\theta}_1, \hat{\theta}_2)$, at a price $t_2(\hat{\theta}_1, \hat{\theta}_2)$.

Complete Information Benchmark. As a first pass, suppose that θ_1 and θ_2 are both common knowledge, but recall that at the time of contracting the cost θ_2 is not yet realized. The solution to the contracting problem is obvious. First, because transferring risk to a risk averse firm is costly, the principal should keep all risk associated with the add-on so as to perfectly ensure the firm against second-period cost uncertainty. Second, the firm must keep the same marginal utility of income in both periods so as to smooth the cost of subsidies over time. Given the normalization $v(0) = 0$ and $v'(0) = 1$, this means that, for all realizations of its costs parameters, the firm should make zero profit in each period. This normalization provides a convenient benchmark that allows us to conclude that any non-stationarity in payments and profits follows from asymmetric information.[30] Lastly, the principal can request efficient production in both periods. Payments and output in period i are thus given by:[31]

$$t_i^{fb}(\theta_i) = \theta_i q_i^{fb}(\theta_i) \text{ and } S_i'(q_i^{fb}(\theta_i)) = \theta_i.$$

Example: *CARA* preferences. Suppose that $v(\cdot)$ is *CARA*. Given the normalizations $v(0) = 0$ and $v'(0) = 1$, $v(z) = \frac{1}{\tau}(1 - exp(-\tau z))$ and $w(z, \varepsilon) = \frac{1}{\tau}(1 - exp(-\tau z)\eta(\tau, \varepsilon))$ where $\eta(\tau, \varepsilon) = \nu exp(-\tau(1-\nu)\varepsilon) + (1-\nu)exp(\tau\nu\varepsilon)$. Finally, $H(z, \varepsilon) = 0$ for all (z, ε).

[29] For technical reasons, we will assume that all feasible q_1 and q_2 are respectively bounded above by some levels Q_1 and Q_2 (both large enough).

[30] Had the firm also had the same concave utility function in the first period, the same result would hold. Profits would be zero in each period.

[31] The superscript fb stands for *first-best* and it indexes optimal variables in the complete information benchmark.

3. INCENTIVE-FEASIBLE ALLOCATIONS

Since the Revelation Principle applies in this dynamic context with full commitment, we can define the firm's intertemporal payoff as

$$\mathcal{U}(\theta_1) = \max_{\hat{\theta}_1 \in \Theta_1, \hat{\theta}_2 \in \Theta_2} (1-\beta)(t_1(\hat{\theta}_1) - \theta_1 q_1(\hat{\theta}_1)) + \beta \mathbb{E}_{\theta_2}\left(v\left(t_1(\hat{\theta}_1) - \theta_1 q_1(\hat{\theta}_1) + y(\hat{\theta}_1) + t_2(\hat{\theta}_1, \hat{\theta}_2) - \theta_2 q_2(\hat{\theta}_1, \hat{\theta}_2)\right)\right) \quad \forall \theta_1 \in \Theta_1 \tag{3.1}$$

where the maximum above is achieved for truthful strategies.

Furthermore, the requirement of incentive compatibility can be applied recursively.[32] For any first-period report $\hat{\theta}_1$, the second-period report, which is truthful from the Revelation Principle, should maximize the firm's continuation payoff $t_2(\hat{\theta}_1, \hat{\theta}_2) - \theta_2 q_2(\hat{\theta}_1, \hat{\theta}_2)$. To get a compact characterization of incentive compatibility, we define the firm's second-period profit from the add-on $\mathcal{U}_2(\hat{\theta}_1, \theta_2)$ as:

$$\mathcal{U}_2(\hat{\theta}_1, \theta_2) = \max_{\hat{\theta}_2 \in \Theta_2} t_2(\hat{\theta}_1, \hat{\theta}_2) - \theta_2 q_2(\hat{\theta}_1, \hat{\theta}_2), \quad \forall \hat{\theta}_1 \in \Theta_1. \tag{3.2}$$

Second-period incentive compatibility requires that a firm facing a low cost of producing the add-on prefers the requested option:

$$\mathcal{U}_2(\hat{\theta}_1, \underline{\theta}_2) \geq \mathcal{U}_2(\hat{\theta}_1, \overline{\theta}_2) + \Delta\theta_2 q_2(\hat{\theta}_1, \overline{\theta}_2), \quad \forall \hat{\theta}_1 \in \Theta_1.\text{[33]} \tag{3.3}$$

This incentive compatibility constraint implies that, when the firm has private information on the cost of the add-on, second-period profits must remain risky.[34]

Furthermore, because any non-zero expected payment could, by a simple redefinition of payment variables, be incorporated into the second-period premium for the basic service, $y(\hat{\theta}_1)$, there is no loss of generality in assuming that the firm makes zero expected profit on the add-on:

$$\mathbb{E}_{\theta_2}\left(\mathcal{U}_2(\hat{\theta}_1, \theta_2)\right) = 0, \quad \forall \hat{\theta}_1 \in \Theta_1. \tag{3.4}$$

[32]See Baron and Besanko (1984), Battaglini (2005) and Pavan, Segal and Toikka (2014).

[33]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 $\mathcal{U}_2(\hat{\theta}_1, \overline{\theta}_2) \geq \mathcal{U}_2(\hat{\theta}_1, \underline{\theta}_2) - \Delta\theta_2 q_2(\hat{\theta}_1, \underline{\theta}_2)$, is automatically satisfied when (3.3) is binding and $q_2(\hat{\theta}_1, \underline{\theta}_2) \geq q_2(\hat{\theta}_1, \overline{\theta}_2)$ as required by the standard monotonicity condition. This monotonicity condition holds for the optimal contract that will be derived below. Therefore, we simplify the presentation by focusing only on the low-cost type's incentive constraint (3.3).

[34]This is true as long as the add-on is produced in positive quantities, $q_2(\theta_1, \overline{\theta}_2) > 0$. This non-negativity requirement is satisfied by the optimal levels when the Inada condition $S_2'(0) = +\infty$ holds. Section 8 investigates a model where the *shut-down* of the least efficient types in the second-period is used as a screening device.

Second-period profits can thus be expressed as a random variable with zero mean:

$$(3.5)\qquad \mathcal{U}_2(\hat{\theta}_1, \underline{\theta}_2) = (1-\nu)\varepsilon(\hat{\theta}_1) \text{ and } \mathcal{U}_2(\hat{\theta}_1, \overline{\theta}_2) = -\nu\varepsilon(\hat{\theta}_1),$$

for some function $\varepsilon(\hat{\theta}_1)$ which represents the amount of risk borne by the firm in the second period. This risk is endogenously determined by second-period incentive compatibility. Indeed, (3.3) amounts to:

$$(3.6)\qquad \varepsilon(\hat{\theta}_1) \geq \Delta\theta_2 q_2(\hat{\theta}_1, \overline{\theta}_2), \quad \forall \hat{\theta}_1 \in \Theta_1.$$

Notice that (3.6) is necessarily binding in the second period if one wants to provide the minimal amount of risk consistent with truthtelling at this date.[35] Of course, this second-period endogenous risk impacts on first-period incentives. To see how, we use our previous definition of payoffs in terms of $w(\cdot)$ and rewrite the firm's intertemporal payoff as:

$$(3.7)\qquad \mathcal{U}(\theta_1) = \max_{\hat{\theta}_1 \in \Theta_1} (1-\beta)(t_1(\hat{\theta}_1) - \theta_1 q_1(\hat{\theta}_1)) + \beta w(t_1(\hat{\theta}_1) - \theta_1 q_1(\hat{\theta}_1) + y(\hat{\theta}_1), \varepsilon(\hat{\theta}_1)).$$

Recall that $u_1(\theta_1) = t_1(\theta_1) - \theta_1 q_1(\theta_1)$ and $\mathcal{U}(\theta_1) = (1-\beta)u_1(\theta_1) + \beta w(u_1(\theta_1) + y(\theta_1), \varepsilon(\theta_1))$. This allows us to express the premium $y(\theta_1)$ in terms of other variables as:

$$(3.8)\qquad u_1(\theta_1) + y(\theta_1) = \varphi\left(\frac{\mathcal{U}(\theta_1) - (1-\beta)u_1(\theta_1)}{\beta}, \varepsilon(\theta_1)\right).$$

This condition tells us how the second-period profit $u_1(\theta_1) + y(\theta_1)$ on the basic service should be modified to keep the second-period utility $\frac{1}{\beta}(\mathcal{U}(\theta_1) - (1-\beta)u_1(\theta_1))$ constant if the second-period risk $\varepsilon(\theta_1)$ is modified.

With this change of variables, an incentive-compatible allocation appears as a quadruplet $(\mathcal{U}(\theta_1), u_1(\theta_1), q_1(\theta_1), \varepsilon(\theta_1))$ that stipulates an intertemporal rent, first-period profit and output and a level of risk for the second period in terms of the firm's first-period cost. Equipped with this dual specification of incentive-compatible allocations, we present a lemma which provides necessary and sufficient conditions satisfied by any such allocation.

LEMMA 1 NECESSARY CONDITION. *Any incentive-compatible allocation* $(\mathcal{U}(\theta_1), u_1(\theta_1), q_1(\theta_1), \varepsilon(\theta_1))$ *is such that* $\mathcal{U}(\theta_1)$ *is absolutely continuous in* θ_1 *(and thus almost everywhere differentiable) with at any point of differentiability:*

$$(3.9)\qquad \dot{\mathcal{U}}(\theta_1) = -q_1(\theta_1)\left(1-\beta+\beta w_z\left(\varphi\left(\frac{\mathcal{U}(\theta_1) - (1-\beta)u_1(\theta_1)}{\beta}, \varepsilon(\theta_1)\right), \varepsilon(\theta_1)\right)\right).$$

SUFFICIENT CONDITION. *An allocation* $(\mathcal{U}(\theta_1), u_1(\theta_1), q_1(\theta_1), \varepsilon(\theta_1))$ *is incentive compat-*

[35] This assertion that simplifies the presentation is proved in Appendix A.

ible if $\mathcal{U}(\theta_1)$ is absolutely continuous, satisfies (3.9) at any point of differentiability and is convex.

The sources of information rent. To understand the envelope condition (3.9), it is useful to consider the benefits that a firm with first-period cost θ_1 gets when pretending to have a marginally higher cost $\theta_1 + d\theta_1$. Doing so means that it can produce the requested basic service $q_1(\theta_1 + d\theta_1)$ at a slightly lower cost and thus save an amount $q_1(\theta_1 + d\theta_1)d\theta_1 \approx q_1(\theta_1)d\theta_1$. This gain is evaluated at the margins $1-\beta$ for the first period and at β multiplied by the marginal utility of income in the second period. This marginal utility itself depends on the second-period profit for the basic service which roughly amounts to $u_1(\theta_1 + d\theta_1) + y(\theta_1 + d\theta_1) \approx u_1(\theta_1) + y_1(\theta_1)$. It also depends on how much risk is borne by the firm for the provision of the add-on. This amount of risk is worth $\varepsilon(\theta_1 + d\theta_1) \approx \varepsilon(\theta_1)$. Putting these facts together, a firm with cost θ_1 is not tempted to mimic the behavior of a $\theta_1 + d\theta_1$ type if it receives an extra rent $\mathcal{U}(\theta_1) - \mathcal{U}(\theta_1 + d\theta_1) \approx -\dot{\mathcal{U}}(\theta_1)d\theta_1$ worth $q_1(\theta_1)\left(1-\beta+\beta w_z(u_1(\theta_1)+y(\theta_1), \varepsilon(\theta_1))\right)d\theta_1$. Simplifying yields (3.9).

The right-hand side of (3.9) already shows the basic forces at play in the optimal contract. First, reducing the output $q_1(\theta_1)$ of a type θ_1 helps relaxing the incentive compatibility constraint. A slightly more efficient firm with type $\theta_1 - d\theta_1$ would find it less tempting to exaggerate its cost when this output is reduced. Reducing output also reduces the slope of the rent, namely $\dot{\mathcal{U}}(\theta_1)$, and helps to save on rents and payments for all inframarginal types below θ_1. This is a familiar distortion in screening environments.

However, the right-hand side of (3.9) also unveils a less familiar effect. Indeed, among all intertemporal profiles of profits $(u_1(\theta_1), u_1(\theta_1) + y(\theta_1))$ that leaves the overall rent $\mathcal{U}(\theta_1)$ of a given type θ_1 unchanged, the principal benefits from shifting more of these profits towards the second period. Reducing the second-period marginal utility also diminishes the slope $\dot{\mathcal{U}}(\theta_1)$ and the principal again saves on the rents and payments for all inframarginal types below θ_1.

Lastly, and still stemming from the concavity of the utility function, playing on the risk $\varepsilon(\theta_1)$ borne by the firm in the second period is a useful tool to save on rents. The rest of our analysis will precisely show how.

Sufficiency. It should be clear from (3.9) that the convexity of $\mathcal{U}(\cdot)$ is guaranteed when the future does not matter much (i.e., β small enough) and $q_1(\theta_1)$ is decreasing. As we will see in the characterization of the optimal outputs under various scenarios, this monotonicity is always satisfied by the solution to the relaxed problem where such requirement is omitted provided that Assumption 2 holds and β is small enough. To simplify the exposition and rule out the uninteresting technicalities that would otherwise arise, we will make these provisos implicit in all statements below. The optimality conditions obtained in the sequel nevertheless apply for a broader set of values of β provided that the rent profile remains convex which is the case when $v(\cdot)$ is not too concave.

A REMARK ON SHORT-TERM PROJECTS. To better understand the analysis below, it is useful to consider the case where the first-period project entails surplus and costs that are only realized in the first period. The initial project and the add-on can thus be viewed as two independent projects or, when costs are identically distributed and surplus functions remain identical, as two independent realizations of the same project. In this case, the firm earns no second-period information rent from exaggerating its first-period costs. In other words, the envelope condition (3.9) is modified to:

$$\dot{\mathcal{U}}(\theta_1) = -(1-\beta)q_1(\theta_1),$$

where the right-hand side is just the short-run benefits of saving on costs when exaggerating the first-period type. This formula no longer depends on the firm's risk aversion and, as a result, first-period distortions are not affected by future risk attitude. Intuitively, all incentives for first-period information revelation are best given in the first period of the relationship and the same output distortion as in the optimal static contract (namely the standard Baron and Myerson (1982) allocation soon to be defined) is implemented. In other words, the optimal contract would not feature any history dependance.[36]

4. THE SPECIAL CASE OF RISK NEUTRALITY

The case of risk neutrality provides an interesting and simple benchmark for the rest of the analysis. This setting best illustrates the situation where the firm is a well-diversified venture that has perfect access to financial markets. Under these circumstances, the principal can easily structure incentives to induce efficient production of the add-on and extract all profits from this activity even if the firm has *ex post* private information on its cost. The only contracting friction then comes from the fact that the firm has private information on the cost of the basic service at the time of contracting. To see how, consider the following second-period payment:

$$\hat{t}_2^{fb}(\hat{\theta}_2) = S_2(q_2^{fb}(\hat{\theta}_2)) - S_2^{fb}, \quad \forall \hat{\theta}_2 \in \Theta_2,$$

where $S_2^{fb} = \mathbb{E}_{\theta_2}(S_2(q_2^{fb}(\theta_2)) - \theta_2 q_2^{fb}(\theta_2))$ is the expected (first-best) value of the add-on. Together with the output requirement $q_2^{fb}(\hat{\theta}_2)$, this forms a second-period direct revelation mechanism that makes the risk neutral firm residual claimant for the social value of the add-on and solves the second-period screening problem at no cost. Furthermore, this second-period mechanism, being independent of the first-period report, has no impact on

[36]Things would be different if the firm was also risk averse in the first period with the same degree of risk aversion as in the second period. In this case, any reward or punishment for incentivizing first-period revelation is best given by being smoothed over time; an argument which is familiar from the repeated moral hazard literature starting with Rogerson (1985).

first-period incentives. The solution to the first-period screening problem thus reduces to the standard Baron and Myerson (1982) allocation that would be optimal in a static context. Second-period risk has no impact on first-period incentives. Because at the time of contracting the firm is privately informed about θ_1, it must receive an information rent to reveal this parameter. The optimal contract thus exhibits the familiar trade-off between information rent and efficiency. The Baron and Myerson (1982) output $q_1^{bm}(\theta_1)$ is distorted downward below the first-best level $q_1^{fb}(\theta_1)$ for all but the most efficient type. The marginal benefit of the project must be equal to its *virtual marginal cost* which leads to the standard formula:

$$S_1'\left(q_1^{bm}(\theta_1)\right) = \theta_1 + \frac{F(\theta_1)}{f(\theta_1)}, \quad \forall \theta_1 \in \Theta_1. \tag{4.1}$$[37]

In the sequel, we are interested in unveiling how this virtual cost must be modified by uncertainty and income smoothing considerations.

5. DYNAMIC INCENTIVES: THE GENERAL CASE

This section characterizes the different effects that are at play when a risk averse firm faces second-period risky returns. We show how our main result can be decomposed into two elementary effects stemming from the concavity of the utility function. On the one hand, letting the risk averse firm bear the endogenous risk needed to ensure incentive compatibility is costly. This means that an optimal contract has to balance insurance concerns and incentives for truthtelling for the add-on; a *Risk Effect*. On the other hand, inducing information revelation of the first-period cost requires to pay up some information rent which has an impact on the second-period marginal utility of income; an *Income Effect*.

Positive Second-Period Profits. Recall that under complete information, the firm's rent can be fully extracted while income smoothing requires to keep the firm's marginal utility of income constant over time. Given our normalizations ($v(0) = 0$ and $v'(0) = 1$), this means that the firm should earn zero profit in each period. Had the firm been given a positive rent, the same smoothing argument would imply that all extra profits should be given in the first period. Of course, asymmetric information on the cost of the basic service provides a rationale for such positive rent. Yet, if the principal was giving all such rent for a given type θ_1 through first-period profits, a slightly more efficient type $\theta_1 - d\theta_1$ would find it very attractive to exaggerate its first-period cost. From the right-hand side of (3.9), it would then benefit from a cost saving approximatively worth $q_1(\theta)(1 - \beta + \beta \times v'(0))d\theta_1 = q_1(\theta)d\theta_1$. The principal can improve on this outcome and reduce the rent left to all inframarginal types by offering a positive profit for the basic

[37]Provided that Assumption 2 holds, the firm's information rent $\mathcal{U}^{bm}(\theta_1) = \int_{\theta_1}^{\overline{\theta}_1} q_1^{bm}(x)dx$ is always non-negative, decreasing and convex as requested by sufficiency in Lemma 1.

service in the second-period. From the concavity of the second-period utility function, the firm's marginal utility of income at this date diminishes. A slightly more efficient firm thus finds it less attractive to exaggerate its cost because the return of doing so is lower in the second period. This simple intuition also carries over when there is asymmetric information on the cost of the add-on and some endogenous risk is borne by the firm in the second period as it is confirmed in the next proposition.

PROPOSITION 1 *The firm's marginal utility of income decreases over time:*[38]

$$w_z(u_1^{sb}(\theta_1)+y^{sb}(\theta_1),\varepsilon^{sb}(\theta_1)) = 1+q_1^{sb}(\theta_1)\frac{F(\theta_1)}{f(\theta_1)}w_{zz}(u_1^{sb}(\theta_1)+y^{sb}(\theta_1),\varepsilon^{sb}(\theta_1)) \le 1. \tag{5.1}$$

Condition (5.1) means that profits are *backloaded* to relax first-period incentive compatibility. Furthermore, the stronger the risk aversion (as measured by $w_{zz}(\cdot)$ which in turn reflects the concavity of $v(\cdot)$) the more attractive backloading is.

To further illustrate this effect and stress the role of second-period uncertainty, we use (5.1) and the fact that $w_{z\varepsilon} \ge 0$ to get the following string of inequalities:

$$1 \ge w_z(u_1^{sb}(\theta_1)+y^{sb}(\theta_1),\varepsilon^{sb}(\theta_1)) \ge w_z(u_1^{sb}(\theta_1)+y^{sb}(\theta_1),0) = v'(u_1^{sb}(\theta_1)+y^{sb}(\theta_1)).$$

This in turn implies that the second-period profit from the basic service $u_1^{sb}(\theta_1)+y^{sb}(\theta_1)$ is always non-negative. Consider now a type θ_1 slightly lower than $\overline{\theta}_1$. Because the rent profile is decreasing, rent minimization calls for leaving this worst type just indifferent between participating or not, i.e., $\mathcal{U}(\overline{\theta}_1)=0$. Putting together this condition with the fact that second-period profits for the basic service are always positive gives

$$u_1^{sb}(\theta_1) \le 0 \le u_1^{sb}(\theta_1)+y^{sb}(\theta_1)$$

for such a type. To deter the most efficient firms from mimicking those with large first-period costs, the optimal contract stipulates a first-period loss if large costs are reported and this loss is only recouped later on.

OUTPUT DISTORTIONS. BASIC SERVICE. Distortions of the basic service reflect the firm's incentives to manipulate first-period costs. These distortions thus depend on the concavity of $v(\cdot)$. Yet, the fact that the firm bears some risk in the second period affects its second-period marginal utility and thus the magnitude of these distortions.

PROPOSITION 2 *The production of the basic service is always distorted downward below the first-best level but remains above the Baron and Myerson (1982) outcome,* $q_1^{bm}(\theta_1) \le$

[38] The superscript *sb* stands for *second-best* and it indexes variables in the optimal contract. Throughout, we assume that the optimization problem is concave. Conditions for that are provided in Appendix C. These conditions always hold provided that β is small enough and that the surplus functions are sufficiently concave.

$q_1^{sb}(\theta_1) \leq q_1^{fb}(\theta_1)$:

$$S_1'(q_1^{sb}(\theta_1)) = \theta_1 + \frac{F(\theta_1)}{f(\theta_1)}(1 - \beta + \beta w_z(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), \varepsilon^{sb}(\theta_1))). \tag{5.2}$$

From (5.1), we know that the firm's benefit from exaggerating its first-period cost is evaluated at a lower marginal utility of income in the second period. As a result, the principal does not need to distort production as much as under risk neutrality. Output distortions are thus lower. Contracts for the basic service are high powered.

The right-hand side of the optimality condition (5.2) generalizes the familiar Myersonian formula for *virtual costs*. As usual, the optimal output must trade off the marginal surplus against the overall marginal cost of production, including the information cost over both periods. To understand this optimality condition observe that, raising the quantity $q_1(\theta_1)$ produced by all types on an interval $[\theta_1, \theta_1 + d\theta_1]$ by a small amount dq_1 yields a marginal gain in net surplus worth

$$(S_1'(q_1(\theta_1)) - \theta_1)f(\theta_1)d\theta_1 dq_1.$$

However, it also increases the information rent of all inframarginal types less than θ_1, whose overall mass is $F(\theta_1)$, by an amount dq_1 times the firm's marginal utility of income (where marginal utilities at each date are weighted by the relative importance of those periods):

$$(1 - \beta + \beta w_z(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), \varepsilon^{sb}(\theta_1)))F(\theta_1)d\theta_1 dq_1.$$

The costs and benefits of the marginal change dq_1 are equal when (5.2) holds.

Backloaded profits imply a lower marginal utility of income in the second period and thus lower distortions than under risk neutrality. This effect is all the more pronounced the more the futur matters (higher values of β). It then becomes cheaper to relax first-period incentive compatibility by increasing second-period profits for the basic service. This makes output distortions less useful to induce information revelation.

Since $w_{z\varepsilon} \geq 0$, the fact that the firm bears some risk in the second period increases the marginal utility of income at this date. Thus the firm has stronger incentives to exaggerate its first-period costs for higher risk levels because the corresponding cost savings are evaluated at a higher marginal utility. This suggests that the principal should be more concerned with rent extraction than in the absence of risk. This points at an important complementarity between the first- and the second-period agency problems even when types are independent and there are no technological linkages across periods.

Output Distortions. Add-On. We have seen how the concavity of the second-period utility function impacts on first-period incentives. The interaction between the agency problems also goes the other way. Second-period profits from the basic service have an impact on the firm's risk attitude and thus on how much risk it should bear; an effect which is unveiled in the next proposition.

Proposition 3 *The add-on is always produced at the first-best level when second-period costs are low but below the first-best level otherwise:* $q_2^{sb}(\theta_1, \underline{\theta}_2) = q_2^{fb}(\underline{\theta}_2)$ *and* $0 < q_2^{sb}(\theta_1, \overline{\theta}_2) \leq q_2^{fb}(\underline{\theta}_2)$ *where*

$$(5.3) \quad (1-\nu)(S_2'\left(q_2^{sb}(\theta_1,\overline{\theta}_2)\right) - \overline{\theta}_2) = \Delta\theta_2 \left(\underbrace{\varphi_\varepsilon(w(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), \varepsilon^{sb}(\theta_1)), \varepsilon^{sb}(\theta_1))}_{\textit{Risk-Premium Distortion}} + \underbrace{q_1^{sb}(\theta_1)\frac{F(\theta_1)}{f(\theta_1)}H(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), \varepsilon^{sb}(\theta_1))}_{\textit{Direct and Substitution Effects}} \right).$$

The firm always bears some risk in the second period, namely $\varepsilon^{sb}(\theta_1) = \Delta\theta_2 q_2^{sb}(\theta_1, \overline{\theta}_2) > 0$ and must thus be compensated with a risk premium. Reducing this compensation means that the second-period output must also be reduced. This distortion, which is captured by the first term on the right-hand side of (5.3) relaxes the second-period incentive-compatibility constraint (3.6). However, there is also an interaction with the first-period agency problem as can be seen from the last term of (5.3). When choosing how much risk should be borne by the firm, the principal must anticipate all consequences of increasing risk on the firm's marginal utility of income.

There are two effects at play and they can best be viewed by noticing that (2.3) shows $H(z,\varepsilon)$ as the sum of two terms. First, keeping second-period profits from the basic service fixed, more risk increases the marginal utility of income ($w_{z\varepsilon}(z,\varepsilon) \geq 0$). This makes first-period incentive compatibility more costly. This *Direct Effect of Risk* on the marginal utility of income calls for reducing the share of second-period risk borne by the firm. However, letting the firm bear more risk in the second period also requires raising second-period profits on the basic service so as to keep the firm's utility constant ($-\frac{w_{zz}(z,\varepsilon)w_\varepsilon(z,\varepsilon)}{w_z(z,\varepsilon)} \leq 0$). This *Substitution Effect of Risk* reduces the second-period marginal utility of income, making first-period incentives less costly. Overall, Assumption 1 guarantees that the *Direct Effect* dominates. This justifies a stronger downward distortion of $q_2(\theta_1, \bar{\theta}_2)$ that is captured by the second term on the right-hand side of (5.3).

Example (*CARA* preferences - continued). Straightforward computations yield the following closed-form solutions:

$$(5.4)\qquad (1-\nu)\left(S_2'\left(\frac{\varepsilon^{sb}}{\Delta\theta_2}\right)-\overline{\theta}_2\right)=\Delta\theta_2\frac{\eta_\varepsilon(\tau,\varepsilon^{sb})}{\tau\eta(\tau,\varepsilon^{sb})}>0,$$

$$(5.5)\qquad S_1'(q_1^{sb}(\theta_1))=\theta_1+\frac{F(\theta_1)}{f(\theta_1)}\left(1-\beta+\frac{\beta}{1+\tau q_1^{sb}(\theta_1)\frac{F(\theta_1)}{f(\theta_1)}}\right),$$

$$(5.6)\qquad u_1^{sb}(\theta_1)+y^{sb}(\theta_1)=\underbrace{\frac{1}{\tau}ln\left(\eta(\tau,\varepsilon^{sb})\right)}_{Risk\ Effect}+\underbrace{\frac{1}{\tau}ln\left(1+\tau q_1^{sb}(\theta_1)\frac{F(\theta_1)}{f(\theta_1)}\right)}_{Income\ Effect}.$$

$CARA$ preferences offer a nice generalization of the risk neutral case; all previous findings being simply obtained by taking the limit when τ goes to 0 in the above expressions (5.4) to (5.6). Importantly for our discussion of the value of commitment and incomplete contracts that follows in Section 7, there is again a complete dichotomy between the agency problems for the basic service and for the add-on. The second-period distortion ε^{sb} defined in (5.4) is independent of the first-period announcement while the distortion of $q_1^{sb}(\theta_1)$ does not depend on how much risk the firm bears in the second period. Also, the second-period profit for the basic service is the sum of a risk premium related to second-period risk and an extra payment required to reduce the marginal utility of income. With more general preferences, condition (5.3) shows that this dichotomy fails.

When τ goes to infinity, the firm becomes infinitely risk averse and we obtain a more familiar screening distortion in the second-period which is the same as in the two-type screening models with *ex post* participation constraints:[39]

$$\lim_{\tau\to+\infty} S_2'(q_2^{sb})=\overline{\theta}_2+\frac{\nu}{1-\nu}\Delta\theta_2.$$

Indeed, an infinitively risk averse firm only cares about having the high-cost in the second period, i.e., the worst possible scenario. This requirement hardens the firm's participation constraint and calls for strong second-period distortions. When considering information manipulations in the first period, the firm now anticipates that the cost savings obtained by exaggerating those costs provide no utility gains in the second-period. As a result, incentives to exaggerate costs are solely driven by first-period considerations. This explains that incentive distortions are weighted by the relative importance of the first period in the expression of the optimal output. We retrieve here the same distortion as if the first-period project was only short-term:

$$\lim_{\tau\to+\infty} S_1'(q_1^{sb}(\theta_1))=\theta_1+(1-\beta)\frac{F(\theta_1)}{f(\theta_1)}.$$

[39]See for instance Chapter 2, Laffont and Martimort (2002).

6. DECOMPOSING THE IMPACTS OF THE INCOME AND THE RISK EFFECTS

This section isolates the respective impacts of the *Income* and the *Risk Effects.* To do so, we decompose the overall agency problem into two distinct scenarios. We show how the *Income Effect* arises even under symmetric information on the cost of the add-on. Instead, the *Risk Effect* follows from the second-period agency problem and occurs even under complete information on the cost of the basic service.

6.1. *Income Effect*

The impact of the *Income Effect* on contract design can be seen when the second-period risk can be fully insured. To see how, suppose that the second-period cost θ_2 is common knowledge and verifiable. Second-period incentive constraints do not matter in this environment. The principal should bear all the risk associated to the add-on so that $\varepsilon^i(\theta_1) = 0$ for all θ_1.[40] Because of symmetric information and perfect insurance, the add-on is still produced at the first-best level and $q_2^i(\theta_1, \theta_2) = q_2^{fb}(\theta_1)$ for all (θ_1, θ_2).

Yet, by a reasoning which is by now familiar from Proposition 1 and which is independent of whether the firm bears some risk or not on the add-on, the *Income Effect* implies that profits are backloaded so as to reduce the cost of first-period incentive compatibility.

This *Income Effect* remains even with full insurance. The firm's marginal utility of income still decreases over time:

$$v'(u_1^i(\theta_1) + y^i(\theta_1)) = 1 + q_1^i(\theta_1)\frac{F(\theta_1)}{f(\theta_1)}v''(u_1^i(\theta_1) + y^i(\theta_1)) \leq 1. \tag{6.1}$$

Second, the optimal level of the basic service is again distorted downwards below the first-best level, $q_1^i(\theta_1) \leq q_1^{fb}(\theta_1)$ with the new expression of the rent-efficiency trade-off given by:

$$S_1'(q_1^i(\theta_1)) = \theta_1 + \frac{F(\theta_1)}{f(\theta_1)}(1 - \beta + \beta v'(u_1^i(\theta_1) + y^i(\theta_1))). \tag{6.2}$$

Since the firm's benefit from exaggerating its first-period cost is evaluated at a lower marginal utility of income in the second period, the principal does not need to distort production as much as in the Baron and Myerson (1982) scenario. As before, output distortions are lower than with risk neutrality and contracts for the basic service are tilted towards high-powered incentives and lower output distortions.

IMPACT OF SECOND-PERIOD ASYMMETRIC INFORMATION. The fact that $w_{z\varepsilon} \geq 0$ ensures that the second-period marginal utility of income decreases when the firm is fully insured

[40] We use the notation z^i to indicate the optimal value of any variable z under asymmetric information on θ_1 only.

against risk. In other words, the following condition holds:

$$S_1'(q_1^{sb}(\theta_1)) \geq \theta_1 + \frac{F(\theta_1)}{f(\theta_1)}(1-\beta+\beta v'(u_1^{sb}(\theta_1)+y^{sb}(\theta_1))).$$

Starting from the output and second-period profit levels that satisfy (6.2), the principal further reduces the level of the basic service and further backloads profits towards the second period to cope with asymmetric information on the costs of the add-on.

Remark on first-period Risk Aversion. Still assuming no second-period risk, we could easily generalize the envelope condition for incentive compatibility to the case where the firm exhibits the same degree of risk aversion in both periods so as to obtain:

$$\dot{\mathcal{U}}(\theta_1) = -q_1(\theta_1)\left((1-\beta)v'(u_1(\theta_1)) + \beta v'\left(\varphi\left(\frac{\mathcal{U}(\theta_1)-(1-\beta)v(u_1(\theta_1))}{\beta}, 0\right)\right)\right). \tag{6.3}$$

A slightly more efficient type $\theta_1 - d\theta_1$ would thus gain from exaggerating its costs an amount worth approximatively $q_1(\theta_1)v'(u_1(\theta_1))d\theta_1$ if the principal were to choose to give to type θ_1 the same profit $u_1(\theta_1)$ in each period so that $\varphi\left(\frac{\mathcal{U}(\theta_1)-(1-\beta)v(u_1(\theta_1))}{\beta}, 0\right) = u_1(\theta_1)$.

Inspired by our previous findings, we may wonder whether the principal could benefit from modifying the profit profile so that it becomes increasing over time while still ensuring that type θ_1 gets a rent $\mathcal{U}(\theta_1)$. To fix ideas, consider reducing the first-period profit by $\beta\epsilon$ and increasing the second-period profit by $(1-\beta)\epsilon$. Up to terms of order two, $\mathcal{U}(\theta_1)$ remains almost unchanged while the principal overall payment remains exactly the same. A slightly more efficient type $\theta_1 - d\theta_1$ would now gain from exaggerating its costs by $d\theta_1$ an amount which is worth approximatively $q_1(\theta_1)(v'(u_1(\theta_1)) - v''(u_1(\theta_1))(\beta(1-\beta)-\beta(1-\beta))\epsilon d\theta_1 = q_1(\theta_1)v'(u_1(\theta_1))d\theta_1$. Hence, the right-hand side of (6.3) is left unchanged and the principal cannot relax incentive compatibility by backloading profits. The *Income Effect* disappears. As a result, the basic service is produced at its Baron and Myerson (1982) level.[41] When θ_2 is private information, things are obviously more complex because the second-period *Risk Effect* changes the marginal utility of income in the second-period and gives a motive for backloading profits even without private information on θ_1.[42]

6.2. *Risk Effect*

Suppose now that θ_2 is privately learned by the firm while θ_1 remains common knowledge. The first-period incentive constraint (3.9) now disappears from the principal's optimization problem and the basic service is always produced at the first-best level $q_1^r(\theta_1) = q_1^{fb}(\theta_1)$ for all θ_1.[43] The principal is only concerned with the conflicting objec-

[41] Proof in Appendix E (not for publication).

[42] The analysis of this more complex scenario is left for future research.

[43] We now use the notation z^r to indicate the optimal value of any variable z under asymmetric information on θ_2 only.

tives of providing insurance against uncertain second-period costs and inducing revelation of these costs. Inducing information revelation calls for letting the firm bear some risk and reducing the corresponding risk premium requires a by now familiar downward distortion of the level of the add-on. These objectives are independent of the cost of the basic service so that the *Direct* and *Substitution Effects of Risk* are no longer relevant. Therefore, the second-period risk $\varepsilon^r(\theta_1) = \varepsilon^r = \Delta\theta_2 q_2^r > 0$ and the second-period output q_2^r when a high second-period cost is revealed do not depend on θ_1:

$$(6.4) \qquad (1-\nu)(S_2'(q_2^r) - \overline{\theta}_2) = \Delta\theta_2 \varphi_\varepsilon(w(u_1^r + y^r, \varepsilon^r), \varepsilon^r).$$

Finally, and in contrast to previous scenarios, the principal adjusts second-period profits to ensure that the firm's marginal utility of income also remains constant, i.e., $u_1^r(\theta_1) + y^r(\theta_1) = u_1^r + y^r$ for all θ_1, with:

$$(6.5) \qquad w_z(u_1^r + y^r, \varepsilon^r) = 1.$$

Because $w_{z\varepsilon} \geq 0$, condition (6.5) implies that $w_z(u_1^r + y^r, 0) = v'(u_1^r + y^r) \leq 1$ and second-period profits are necessarily positive:

$$u_1^r + y^r = \varphi\left(\frac{-(1-\beta)u_1^r}{\beta}, \varepsilon^r\right) \geq 0,$$

where the first equality follows from the definition of $\varphi(\cdot)$ and the fact that the firm's intertemporal payoff can be fully extracted when θ_1 is common knowledge. Even in the absence of the first-period incentive problem, profits are backloaded for the sole reason that the *Risk Effect* requires the principal to pay a risk premium to the firm.

Impact of first-period asymmetric information. From (6.4) and the non-negativeness of $H(\cdot)$, it immediately follows that:

$$(1-\nu)(S_2'(q_2^r) - \overline{\theta}_2) \leq \Delta\theta_2\left(\varphi_\varepsilon(w(u_1^r + y^r, \varepsilon^r), \varepsilon^r) + q_1^{fb}(\theta_1)\frac{F(\theta_1)}{f(\theta_1)}H(u_1^r + y^r, \varepsilon^r)\right).$$

The comparison with (5.3) is straightforward. Starting from the output and second-period profit levels that satisfy (6.4), the principal further reduces the level of the add-on when there is asymmetric information on the costs of the basic service.

Example (*CARA* preferences - continued). The closed-form expressions for risk and profits are now as follows:

$$(6.6) \qquad \varepsilon^r = \varepsilon^{sb}, \quad y^r = \frac{1}{\tau} ln\left(\eta(\tau, \varepsilon^r)\right) > 0 = u_1^r, \quad \mathcal{U}_1(\theta_1) = 0, \quad \forall \theta_1 \in \Theta_1.$$

With *CARA*, the risk borne by the firm is independent of whether there is asymmetric information on first-period costs or not. However, dealing with this risk still requires to pay a risk premium $\frac{1}{\tau}ln\left(\eta(\tau,\varepsilon^r)\right)$, but it is independent of the first-period cost.

7. INCOMPLETE CONTRACTS

This section analyzes the possible costs that parties incur when they are not able to perfectly commit *ex ante* to a complete contract with a single firm in charge of providing both the basic service and the add-on. Such scenarios are meant to capture the highly incomplete contracting environments that may surround long-term contracts, a concern that has repeatedly been brought forward by practitioners in the PPPs sector.[44] Section 7.1 discusses the costs and benefits of bundling tasks, an issue of particular importance in the PPP literature. Section 7.2 analyzes the issue of renegotiation while Section 7.3 considers the scenario where the add-on is not even included in the initial contract.

7.1. *Unbundling Tasks*

Bundling the basic service and the add-on into a grand-contract signed with a single supplier was the sole possibility considered above. This might be the most appropriate assumption for many add-ons that build on the firm's expertise and investment in the basic service. When these add-ons are less specific, provision might be adequately undertaken by other suppliers. In such contexts, whether the two tasks should be bundled or not becomes an important issue. The extant literature on bundling versus unbundling in PPP contexts has focused on investments, cost externalities and budget constraints as a rationale for the economies of scope behind the bundling decision (Hart (2003), Bennett and Iossa (2006), Martimort and Pouyet (2008), Iossa and Martimort (2012), Schmitz (2013), Martimort and Straub (2016)). In our model there is no technological linkage across projects; and still agency costs at each stage of the production process are strongly linked. Of course, the $CARA$ case offers an interesting benchmark where bundling and unbundling reach similar welfare levels.

Beyond this specific case, contracting for the add-on with a separate firm would allow the principal to isolate second-period distortions from first-period concerns. To illustrate this, in the somewhat optimistic scenario where the degree of risk aversion of the specialized firm would be identical to that of an integrated supplier, the principal could still replicate the same second-period production as in Section 6.2. Because the firm in charge of the add-on also certainly faces a costly access to financial markets, its preferences certainly exhibit risk aversion. However, this degree of risk aversion might be of a higher magnitude than that of the integrated provider because the more specialized firm can no longer pledge the profit made on the basic service to facilitate access to financial markets. This points at a possible cost of unbundling. Turning now to the possible benefits of unbundling, notice that the firm in charge of the basic service does not have to finance

[44] Guasch et al (2007), Guasch et al (2008).

any new investment. It remains risk neutral for the second period. The *Income Effect* disappears under unbundling. The regulation of the basic service follows the familiar Baron and Myerson (1982) formula (4.1) and there is always less provision of the basic service and more rent extraction than with an integrated supplier.

7.2. *Renegotiation*

The dynamics of the optimal contract features a strong linkage between first- and second-period distortions, at least beyond the specific case of $CARA$ preferences where agency costs in each period can be fully disentangled. This suggests that this long-term contract might not be immune to renegotiation when the decision on the level of the basic service has been sunk and new opportunities for risk-sharing arise before uncertainty on the cost of the add-on realizes. Our aim here is not to characterize the optimal renegotiation-proof contract, but to give some insights into when and why renegotiation may hinder the performance of long-term agreements.

Consider a long-term contract $\mathcal{C} = \left\{t_1(\hat{\theta}_1), y(\hat{\theta}_1), q_1(\hat{\theta}_1), \varepsilon(\hat{\theta}_1)\right\}_{\hat{\theta}_1 \in \Theta_1}$.[45] With this contract, the firm of type θ_1 which reported a first-period cost $\hat{\theta}_1$ gets a second-period expected utility worth $w(t_1(\hat{\theta}_1) + y(\hat{\theta}_1) - \theta_1 q_1(\hat{\theta}_1), \varepsilon(\hat{\theta}_1))$. With limited commitment, the Revelation Principle may fail and 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. When the second-period comes along but before second-period costs are observed by the firm, the principal may nevertheless propose a new contract for the provision of the add-on. Indeed, Proposition 3 showed how the principal reduces the second-period risk borne by the firm to decrease the second-period marginal utility of income and facilitate information revelation earlier on. When information has already been revealed, this motive for extra insurance is no longer needed. When renegotiating the second-period agreement, the principal wants the firm to bear more risk and, by incentive compatibility, wants to increase the production of the add-on. Such a change is accompanied by an increase $\tilde{y}(\hat{\theta}_1)$ in the second-period premium to give an expected utility of $w(t_1(\hat{\theta}_1)+y(\hat{\theta}_1)+\tilde{y}(\hat{\theta}_1)-\theta_1 q_1(\hat{\theta}_1), \tilde{\varepsilon}(\hat{\theta}_1))$ where $\tilde{\varepsilon}(\hat{\theta}_1) = \Delta\theta_2 q_2(\hat{\theta}_1, \overline{\theta})$ is the second-period risk borne by the firm.

Among all possible contracts that can be accepted by the firm at the renegotiation stage, i.e., contracts that improve the firm's payoff, there is always the null contract that consists of offering no extra payment to the firm, $\tilde{y}(\hat{\theta}_1) \equiv 0$, and leaving the levels of the add-on and the overall risk unchanged, i.e., $\tilde{\varepsilon}(\hat{\theta}_1) \equiv \varepsilon(\hat{\theta}_1)$. Borrowing a definition of *renegotiation-*

[45]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 the condition of zero expected profits in (3.4) and by an amount of risk in the second-period that satisfies (3.6). Second, we restrict attention to the case of direct mechanisms. To motivate this approach, recall that Bester and Strausz (2001) show that, with discrete types, this restriction is without loss of generality even in an environment with limited commitment if one is ready to entertain the possibility that reports are no longer truthful.

proofness due to Dewatripont (1988), a long-term contract $\mathcal{C}$ is *renegotiation-proof* if, given the principal's posterior beliefs about the firm's type θ_1 at the beginning of the second period, the principal finds it optimal to offer this null contract at the renegotiation stage, leaving the initial offer unchanged.

We are now ready to check whether the optimal contract under full commitment $\mathcal{C}^{sb}$ is robust to renegotiation and under which circumstances.

PROPOSITION 4 *The optimal contract under full commitment $\mathcal{C}^{sb}$ is not (resp. always) renegotiation-proof when $v(\cdot)$ satisfies DARA (resp. CARA).*

With *CARA* preferences, the amount of risk borne by the firm in the second period is, even in the optimal full commitment contract, independent of how much profit has been promised for the delivery of the basic service. Incentive problems in each period are not linked and the optimal contract under full commitment remains renegotiation-proof. A particularly interesting case arises when the firm is risk neutral. To the extent that risk neutrality captures the idea that the firm has perfect access to financial markets, our model predicts that long-term contracts with firms having perfect access to financial markets are stable and robust to further rounds of negotiations. Instead, costly access to financial markets might destabilize long-term contracts when add-ons become necessary.

7.3. *Spot Contracts*

In practice, parties might face unforeseen contingencies that could not be anticipated and written into the initial contract, especially if this contract covers the provision of a basic service over many years. To model such settings, we now suppose that *ex ante* parties can only agree on a highly incomplete long-term contract which does not even specify payments and output requirements for the add-on. Of course and in accordance with the incomplete contracting literature,[46] the mere opportunity of such additional projects can be anticipated. An important issue is thus to assess whether, in this context, contracting on the add-on *on the spot* entails any loss to the parties.

In fact, even if parties can only contract on the add-on at the interim stage, the same allocation as in the optimal long-term contract $\mathcal{C}^{sb}$ can still be implemented in the case of *CARA* preferences. To see how, consider a long-term agreement $\{t^{sb}(\hat{\theta}_1), y^{sb}(\hat{\theta}_1) - y^r, q_1^{sb}(\hat{\theta}_1)\}_{\hat{\theta}_1\in\Theta_1}$ that regulates the basic service over the whole relationship and, as such, does not specify any risk premium nor any add-on specification. When the second period comes, parties agree on a spot contract to regulate this add-on. This spot contract specifies a risk premium y^r, the levels of the add-on in the different states of nature $q_2^{sb}(\theta_1, \theta_2)$, and the risk $\varepsilon^r = \Delta\theta_2 q_2^{sb}(\theta_1, \overline{\theta}_2)$ borne by the firm. This spot contract does not modify the firm's risk attitude and, even though it is anticipated by parties, it has

[46] Grossman and Hart (1986).

no impact on first-period incentives. Compounding the impact of this spot contract with the initial contract for the basic service replicates the optimal long-term contract.

PROPOSITION 5 *Suppose that the firm has CARA preferences. There is no loss of generality in contracting for the add-on only at the interim stage.*

This proposition bears some strong resemblance with findings in Fudenberg, Holmström and Milgrom (1990). These authors also discuss conditions under which an optimal long-term contract can be implemented with a sequence of short-term contracts. Beyond the fact that one of the contract regulates a long-term service and thus covers both periods, our setting also differs from their more general inquiry in several respects. First, those authors allow the principal and the firm to borrow from the financial markets on equal terms. While this assumption is certainly relevant for some employment relationships, it is less so 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 friction in the contract, Fudenberg, Holmström and Milgrom (1990) study a repeated moral hazard contracting model under the assumption that technology is common knowledge (in the sense that future contract outcomes is completely determined by current history). These conditions (plus other more technical assumptions) are then shown to be sufficient to obtain the irrelevance of long-term contracting. Our model illustrates that such sequential optimality can be found in other, admittedly more informationally constrained, environments as well.

8. ROBUSTNESS: LUMPY ADD-ON WITH CONTINUOUS COSTS

So far our analysis has been simplified by assuming that the cost of the uncertain add-on was drawn from a binary distribution. Although this assumption allows us to consider the consequences of an endogenous background risk on earlier incentives in a stripped down manner, a more symmetric treatment requires the cost of the add-on to take a continuum of values. We thus assume that θ_2 is distributed according to a continuous and atomless cumulative distribution $F_2(\theta_2)$ (with a positive density $f_2(\theta_2)$) on $\Theta_2 = \left[\underline{\theta}_2, \overline{\theta}_2\right]$. The technical difficulty pointed out by both Salanié (1990) and Laffont and Rochet (1998) for such models is that, even in simpler static settings, complicated areas of bunching might arise for the optimal level of add-on when the degree of risk aversion is sufficiently large.

One way to extend our analysis without falling into such technicalities is to consider a setting where the second-period project is lumpy. Possible examples would be the expansion of an existing infrastructure, or the addition of services into new geographical areas or new segments of demand. This add-on, whose fixed value is denoted by S_2, is only pursued when the principal pays a price that covers the cost θ_2. Bunching thus takes a simpler form: The project is only done for costs below a threshold. We also assume that $\underline{\theta}_2 < S_2 < \overline{\theta}_2$, meaning that implementing the add-on is not always efficient even under complete information. This assumption stands in contrast to our previous analysis where an Inada condition imposed on the second-period surplus implied that the add-on was always valuable and thus always provided even in the second-best scenario. It nevertheless still implies that the firm's second-period returns remain risky with part of the risk coming from the possibility to give up the project if it turns out to be too costly.

The firm's expected second-period payoff with an arbitrary price $p \in \Theta_2$ for the add-on can now be written as:

$$w(z,p) = \int_{\underline{\theta}_2}^{p} v(z + p - \theta_2) f_2(\theta_2) d\theta_2 + v(z)(1 - F_2(p)). \tag{8.1}$$

Observe also that an increase in p makes it more likely to implement the add-on. It thus shifts the distribution of second-period profits in the sense of first-order stochastic dominance and this reduces the firm's second-period marginal utility of income since:

$$w_{zp}(z,p) = \int_{\underline{\theta}_2}^{p} v''(z + p - \theta_2) f_2(\theta_2) d\theta_2 \leq 0. \tag{8.2}$$

Following the same procedure as previously, we may also re-define two new functions $H(z,p) = w_{zp}(z,p) - \frac{w_{zz}(z,p)}{w_z(z,p)} w_p(z,p)$ and $\varphi(\zeta,p)$ such that $\zeta = w(\varphi(\zeta,p),p)$.[47] Assumption 1 then ensures that $H(\cdot)$ remains non-negative.[48] Thus, $\frac{dw_z}{dp}(\varphi(\zeta,p),p)$ is also non-negative. Contrary to our main scenario, this condition implies that the *Direct Effect* of increasing p is now dominated by the *Substitution Effect*. Increasing p reduces the marginal utility of income but is also requires to decrease the second-period profit made on the basic service to maintain second-period utility constant, which in turn increases the marginal utility of income more than the direct decrease.

Although details of the model differ, the analysis bears some resemblance to our previous findings. A first common feature is that the principal can reduce the cost of information rent by decreasing the firm's marginal utility of income in the second period. Indeed, at the optimal contract, we have:

$$w_z(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), p^{sb}(\theta_1)) = 1 + q_1^{sb}(\theta_1) \frac{F(\theta_1)}{f(\theta_1)} w_{zz}(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), p^{sb}(\theta_1)) \leq 1. \tag{8.3}$$

As a result and by a mechanism which is now familiar, output distortions for the basic service are also less pronounced than in the Baron and Myerson (1982) outcome:

$$S_1'(q_1^{sb}(\theta_1)) = \theta_1 + \frac{F(\theta_1)}{f(\theta_1)}(1 - \beta + \beta w_z(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), p^{sb}(\theta_1))). \tag{8.4}$$

[47] In particular, we have $\varphi_p(\zeta,p) = -\frac{w_p(\varphi(\zeta,p),p)}{w_z(\varphi(\zeta,p),p)} < 0$ and $\varphi_\zeta(\zeta,p) = \frac{1}{w_z(\varphi(\zeta,p),p)} > 0$.

[48] See Appendix D, Lemma D.1.

Because now the *Substitution Effect* dominates, relaxing the firm's first-period incentive constraint calls for decreasing the second-period price below its level in the absence of a first-period incentive problem. Indeed, the following condition holds:

$$(8.5)\qquad (S_2 - p^{sb}(\theta_1))f_2(p^{sb}(\theta_1)) - F_2(p^{sb}(\theta_1)) \geq \varphi_p(w(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), p^{sb}(\theta_1)), p^{sb}(\theta_1)).$$

Even though details differ, there is a common thread to this setting and our previous model. To isolate the first-period agency problem from the second-period one, the principal makes the second-period project less relevant either by reducing its size (in our main model) or by reducing the likelihood of its implementation in the present setting.

9. CONCLUDING REMARKS

We have studied optimal procurement contracts in the context of long-lasting services which are later on augmented by add-ons whose costs are unknown at the time of contracting. As a reduced form for the firm's limited ability to raise outside finance to cover new investments required by these add-ons, we have assumed that the firm is risk averse at the date of producing the add-on. Risk aversion has intricate impacts on both second- and first-period incentives. On the one hand, backloading payments for the basic service reduces the firm's marginal utility of income at the time of producing the add-on and makes strategic manipulations of information on the costs of basic service cheaper from the principal's viewpoint. This *Income Effect* calls for smaller distortions in the provision of this basic service and high powered incentives at the inception of the relationship. On the other hand, risk aversion also implies that the firm must keep part of the risk associated with the add-on to reveal the corresponding costs. This *Risk Effect* calls for downscaling the add-on below its efficient level and implementing low powered incentives on those additional clauses. Those two effects reinforce each other when the firm's preferences exhibit $DARA$. Backloading profits makes it easier for the firm to support risk associated with the add-on; while keeping more of that risk also increases the firm's marginal utility of income and makes backloading more attractive. Except in the case of $CARA$ preferences, there is a value of writing a long-term contract specifying also how add-ons should be supplied although such comprehensive long-term contracts are prone to renegotiation. Incomplete contracts, with unspecified clauses on add-ons, may nevertheless fare well in that $CARA$ case.

Our analysis could be extended along several lines. The first obvious extension would consist in offering a more detailed modeling of the financial constraints faced by the firm with a view on how those constraints trigger the renegotiation of long-term contracts. We are quite confident on the robustness of our results to such micro-foundations even though a more detailed analysis of the role that financiers play, especially when modifying financial contracts in response to the renegotiation of procurement contracts, may unveil some specific features.

Second, the renegotiation of long-term contracts may be a particular acute concern when it is led by corrupt public officials with significant discretion in drafting additional clauses. A basic tenet of the collusion literature[49] is that preventing collusion certainly hardens the trade-off between risk and insurance and thus calls for implementing low powered incentives on the add-on at the renegotiation stage. In turn, those low powered incentives reduce the second-period marginal utility of income and makes backloading payments less effective. A consequence is that low-powered incentives might also be preferred for the provision of the basic service. The threat of corruption may thus contaminate contractual clauses which *a priori* remain beyond the responsibility of corrupt officials.

Third, we have cast our model in terms of cost uncertainty. However, uncertainty at the contracting stage may bear on which add-on is the most appropriate one or maybe on whether any add-on will be needed at all.[50] We conjecture that such extra risk dimensions won't change the bulk of our arguments although details may matter.

Finally, our analysis has shown that risk aversion was an important ingredient to link agency problems at different stages of the production process. Other features like cost correlation or technological economies of scope between the two stages, whose effects are better known, could easily be added to our analysis and would superimpose their more traditional impacts to the effects unveiled by our analysis.

REFERENCES

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

Arve, M. (2014), "Procurement and Predation: Dynamic Sourcing from Financially Constrained Suppliers," *Journal of Public Economics*, 120: 157-168.

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*, 6: 1-29.

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

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

Auray, S., T. Mariotti and F. Moizeau (2011), "Dynamic Regulation of Quality," *The RAND Journal of Economics,* 42: 246-265.

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

[49] Tirole (1986).

[50] We are grateful to Georgia Kosmopoulou for pointing out this possible extension.

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

Bajari, P., S. Houghton and S. Tadelis (2014), "Bidding for Incomplete Contracts: An Empirical Analysis of Adaptation Costs" *The American Economic Review*, 104: 1288-1319.

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

Baron, D. P. and D. Besanko (1987), "Monitoring, Moral Hazard, Asymmetric Information, and Risk Sharing in Procurement Contracting," *The RAND Journal of Economics*, 18: 509-532.

Baron, D. and R. Myerson (1982), "Regulating a Monopolist with Unknown Costs," *Econometrica*, 50: 911-930.

Battaglini, M. (2005), "Long-Term Contracting with Markovian Consumers," *American Economic Review*, 95: 637-658.

Battaglini, M. and R. Lamba (2012), "Optimal Dynamic Contracting," *Economic Theory Center Working Paper*, (46-2012).

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

Bergemann, D. and A. Pavan (2015), "Introduction to Symposium on Dynamic Contracts and Mechanism Design," *Journal of Economic Theory* 159: 679-701.

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

Bolton, P. and D. Scharfstein (1990), "A Theory of Predation Based on Agency Problems in Financial Contracting," *The American Economic Review,* 80: 93-106.

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

Callahan, M. (2012), "Change Order Basics - Part 1," *Contractor Magazine*, January: 30.

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

Campo, S., E. Guerre, E., I. Perrigne and Q. Vuong (2011), "Semi-Parametric Estimation of First-Price Auctions with Risk-Averse Bidders," *The Review of Economic Studies*, 78: 112-147.

Chakravarty, S. and W.B. MacLeod (2009), "Contracting in the Shadow of the Law," *The RAND Journal of Economics*, 40: 533-557.

Che, Y. K. and I Gale (1998), "Standard Auctions with Financially Constrained Bidders," *The Review of Economic Studies,* 65: 1-21.

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*, 67: 697-717.

Cox, R.K. (1997), "Managing Change Orders and Claims," *Journal of Management in Engineering*, 13: 24-29.

Deck, C. and H. Schlesinger (2014), "Consistency of Higher Order Risk Preferences," *Econometrica*, 82: 1913-1943.

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

Eső, P. and L. White (2004), "Precautionary Bidding in Auctions," *Econometrica*, 72: 77-92.

Eső, P. and B. Szentes (2013), "Dynamic Contracting with Adverse Selection: An Irrelevance Result," Working paper Oxford University.

Farhi, E. and I. Werning (2013), "Insurance and Taxation over the Life Cycle," *The Review of Economic Studies*, 80: 596-635.

Farinha Luz, V. (2013), "Dynamic Competitive Insurance," Working paper Yale Economics Department.

Faure-Grimaud, A. and D. Martimort (2003), "Regulatory Inertia," *The RAND Journal of Economics*, 34: 413-437.

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

Gale, D. and M. Hellwig (1985), "Incentive-Compatible Debt Contracts: The One-Period Problem," *The Review of Economic Studies,* 52: 647-663.

Gärtner, D. (2010), "Monopolistic Screening Under Learning By Doing," *The RAND Journal of Economics,* 41: 574-597.

Garrett, D. and A. Pavan (2015), "Dynamic Managerial Compensation: A Variational Approach," *Journal of Economic Theory*, 159: 775-818.

Golosov, M., Tsyvinski, A. and Werning, I. (2006), *New Dynamic Public Finance: a User's Guide*, in *NBER Macroeconomics Annual* 2006, Vol. 21.

Green, E. (1987), "Lending and the Smoothing of Uninsurable Income," in Prescott, E. and Wallace, N. (eds.), *Contractual Arrangements for Intertemporal Trade*, 3-25. Minneapolis: University of Minnesota Press.

Grossman, S. and O. Hart (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy,* 94: 691-719.

Guasch, J.L., J.J. Laffont and S. Straub (2007), " Concessions of infrastructure in Latin America: Government-Led renegotiations," *Journal of Applied Econometrics*, 22, 1267-1294.

Guasch, J.L., J.J. Laffont and S. Straub (2008), "Renegotiation of Concession Contracts in Latin America: Evidence from the Water and Transport Sectors," *International Journal of Industrial Organization*, 26, 421-442.

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

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

Holmström, B. and J. Tirole (1997), "Financial Intermediation, Loanable Funds, and the Real Sector ," *The Quarterly Journal of Economics*, 112: 663-691.

Holt, C.A. and S.K. Laury (2002), "Risk Aversion and Incentive Effects," *The American Economic Review*, 92: 1644-1655.

Iossa, E. and D. Martimort (2012), "Risk Allocation and the Costs and Benefits of Public-Private Partnerships," *The RAND Journal of Economics*, 43: 442-474.

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

Kapicka, M. (2013), "Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-Order Approach," *The Review of Economic Studies,* 80: 1027-1054.

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

Krähmer, D. and R. Strausz (2011), "The Benefits of Sequential Screening," http://epub.ub.uni-muenchen.de/13191/1/363.pdf.

Laffont, J.J. (1994), "The New Economics of Regulation Ten Years After," *Econometrica*, 62: 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*, 25: 149-173.

LELAND, H.E. (1968), "Saving and Uncertainty: The Precautionary Demand for Saving," *Quarterly Journal of Economics*, 82: 465-473.

LELAND, H.E. AND D.H. PYLE (1977), "Informational Asymmetries, Financial Structure, and Financial Intermediation," *The Journal of Finance*, 32: 371-387.

LEWIS, T. AND D. SAPPINGTON (1995), "Optimal Capital Structure in Agency Relationships," *The RAND Journal of Economics*, 26: 343-361.

LEWIS, T. AND H. YILDIRIM (2002), "Learning By Doing and Dynamic Regulation," *The RAND Journal of Economics,* 33: 22-36.

MCAFEE, P. AND J. MCMILLAN (1986), "Bidding for Contracts: A Principal-Agent Analysis," *The RAND Journal of Economics,* 17: 326-338.

MAKRIS, M. AND A. PAVAN (2015), "Taxation under Learning By Doing: Incentives for Endogenous Types," Discussion paper. Northwestern University.

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*, 26: 393-411.

MARTIMORT, D. AND W. SAND-ZANTMAN (2007), "Signaling and the Design of Delegated Management Contracts for Public Utilities," *The RAND Journal of Economics*, 37: 763-782.

MARTIMORT, D AND S. STRAUB (2016), "How to Design Infrastructure Contracts in a Warming World; A Critical Appraisal of Public-Private Partnerships," forthcoming *International Economic Review.*

MASKIN E. AND J. RILEY (1984), "Optimal Auctions with Risk Averse Buyers," *Econometrica*, 52: 1473-1518.

MATTHEWS, S. (1984), "On the Implementability of Reduced Form Auctions," *Econometrica*, 52: 1519-1522.

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*, 70: 583-601.

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

MYERSON, R. (1986), "Multistage Games with Communication," *Econometrica*, 54: 323-358.

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

NOUSSAIR, C., S. TRAUTMANN, G. VAN DE KUILEN (2014), "Higher Order Risk Attitudes, Demographics, and Financial Decisions," *The Review of Economic Studies*, 81: 325-355.

PAVAN, A., I. SEGAL AND J. TOIKKA (2014), "Dynamic Mechanism Design: A Myersonian Approach," *Econometrica*, 82: 601-653.

PRENDERGAST, C. (1999), "The Provision of Incentives in Firms," *Journal of Economic Literature*, 37: 7-63.

ROGERSON, W. (1985), "Repeated Moral Hazard," *Econometrica*, 53: 69-76.

SALANIÉ, B. (1990), "Sélection Adverse et Aversion pour le Risque," *Annales d'Economie et de Statistiques*, 18: 131-149.

SANDMO, A. (1970), "The Effect of Uncertainty on Saving Decisions," *The Review of Economic Studies,* 37: 353-360.

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

SEIERSTAD, A. AND K. SYDSAETER (1987), *Optimal Control Theory with Economic Applications*, North-Holland, Amsterdam.

STANTCHEVA, S. (2014), "Optimal Taxation and Human Capital Policies over the Life Cycle," *Mimeo.*

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

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

TIROLE, J. (1986), "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations," *Journal of Law Economics and Organization,* 2: 181-214.

THOMAS, J. AND T. WORRALL (1990), "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem," *Journal of Economic Theory*, 51: 367-390.

TOWNSEND, R. (1979), "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory,* 21: 265-293.

TOWNSEND, R. (1982), "Optimal Multiperiod Contracts and the Gain from Enduring Relationships under Private Information," *The Journal of Political Economy,* 90: 1166-1186.

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

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

ZHANG, Y. (2009), "Dynamic Contracting with Persistent Shocks," *Journal of Economic Theory,* 144: 635-675.

APPENDIX A: MAIN PROOFS

PROOF OF LEMMA 1: NECESSITY. From Theorem 2 and Corollary 1 in Milgrom and Segal (2002) and the fact that q_2 and q_1 are positive and bounded from above, it immediately follows that $\mathcal{U}(\theta_1)$ is absolutely continuous and thus almost everywhere differentiable with:

$$\dot{\mathcal{U}}(\theta_1) = -q_1(\theta_1)\left(1-\beta+\beta w_z(u_1(\theta_1)+y(\theta_1),\varepsilon(\theta_1))\right) \tag{A.1}$$

at any point of differentiability. Using (3.8), we obtain (3.9).

SUFFICIENCY. We rewrite (3.7) as:

$$\begin{aligned}\mathcal{U}(\theta_1)-\mathcal{U}(\hat{\theta}_1) &\geq (1-\beta)(\hat{\theta}_1-\theta_1)q_1(\hat{\theta}_1) \\ &+\beta\left(w(u_1(\hat{\theta}_1)+y(\hat{\theta}_1)+(\hat{\theta}_1-\theta_1)q_1(\hat{\theta}_1),\varepsilon(\hat{\theta}_1))-w(u_1(\hat{\theta}_1)+y(\hat{\theta}_1),\varepsilon(\hat{\theta}_1))\right) \quad \forall(\theta_1,\hat{\theta}_1)\in\Theta_1^2\end{aligned} \tag{A.2}$$

where the right-hand side can also be expressed in integral form as:

$$q_1(\hat{\theta}_1)\int_{\theta_1}^{\hat{\theta}_1}\left(1-\beta+\beta w_z\left(u_1(\hat{\theta}_1)+y(\hat{\theta}_1)+(\tilde{\theta}_1-\theta_1)q_1(\hat{\theta}_1),\varepsilon(\hat{\theta}_1)\right)\right)d\tilde{\theta}_1.$$

Using (A.1) and absolute continuity, the rent profile $\mathcal{U}(\theta_1)$ satisfies:

$$\mathcal{U}(\theta_1)-\mathcal{U}(\hat{\theta}_1)=\int_{\theta_1}^{\hat{\theta}_1}q_1(\tilde{\theta}_1)\left(1-\beta+\beta w_z(u_1(\tilde{\theta}_1)+y(\tilde{\theta}_1),\varepsilon(\tilde{\theta}_1))\right)d\tilde{\theta}_1, \quad \forall(\theta_1,\hat{\theta}_1)\in\Theta_1^2.$$

Condition (A.2) thus holds when:

$$\int_{\theta_1}^{\hat{\theta}_1}q_1(\tilde{\theta}_1)\left(1-\beta+\beta w_z(u_1(\tilde{\theta}_1)+y(\tilde{\theta}_1),\varepsilon(\tilde{\theta}_1))\right)d\tilde{\theta}_1 \geq \tag{A.3}$$

$$q_1(\hat{\theta}_1)\int_{\theta_1}^{\hat{\theta}_1}\left(1-\beta+\beta w_z\left(u_1(\hat{\theta}_1)+y(\hat{\theta}_1)+(\tilde{\theta}_1-\theta_1)q_1(\hat{\theta}_1),\varepsilon(\hat{\theta}_1)\right)\right)d\tilde{\theta}_1 \quad \forall(\theta_1,\hat{\theta}_1)\in\Theta_1^2.$$

Because $w(\cdot)$ is concave w.r.t. its first argument, we have:

$$w_z\left(u_1(\hat{\theta}_1)+y(\hat{\theta}_1)+(\tilde{\theta}_1-\theta_1)q_1(\hat{\theta}_1),\varepsilon(\hat{\theta}_1)\right)\le w_z\left(u_1(\hat{\theta}_1)+y(\hat{\theta}_1),\varepsilon(\hat{\theta}_1)\right) \quad \forall\tilde{\theta}_1\ge\theta_1.$$

Therefore, a sufficient condition for condition (A.3) to hold is that for all $(\theta_1,\hat{\theta}_1)\in\Theta_1^2$:

$$\int_{\theta_1}^{\hat{\theta}_1} q_1(\tilde{\theta}_1)\left(1-\beta+\beta w_z(u_1(\tilde{\theta}_1)+y(\tilde{\theta}_1),\varepsilon(\tilde{\theta}_1))\right)d\tilde{\theta}_1\ge(\hat{\theta}_1-\theta_1)q_1(\hat{\theta}_1)\left(1-\beta+\beta w_z\left(u_1(\hat{\theta}_1)+y(\hat{\theta}_1),\varepsilon(\hat{\theta}_1)\right)\right),$$

which amounts to the convexity of $\mathcal{U}(\cdot)$. *Q.E.D.*

PROOFS OF PROPOSITIONS 1, 2 AND 3 : The principal's expected payoff in (2.1) becomes:

$$\mathbb{E}_{\theta_1}\left(S_1(q_1(\theta_1))-\theta_1q_1(\theta_1)-u_1(\theta_1)+\beta\mathbb{E}_{\theta_2}\left(S_2(q_2(\theta_1,\theta_2))-\theta_2q_2(\theta_1,\theta_2)-y(\theta_1)-\mathcal{U}_2(\theta_1,\theta_2)\right)\right).$$

Taking into account the expression of $y_1(\theta_1)$ given by (3.8) and the expressions of $\mathcal{U}_2(\theta_1,\theta_2)$ in terms of the second-period risk $\varepsilon(\theta_1)$ given by (3.5), we may rewrite this objective as:

$$\text{(A.4)}\quad \mathbb{E}_{\theta_1}\Bigg(S_1(q_1(\theta_1))-\theta_1q_1(\theta_1)-(1-\beta)u_1(\theta_1)+\beta\mathbb{E}_{\theta_2}\left(S_2(q_2(\theta_1,\theta_2))-\theta_2q_2(\theta_1,\theta_2)\right)$$
$$-\beta\varphi\left(\frac{\mathcal{U}(\theta_1)-(1-\beta)u_1(\theta_1)}{\beta},\varepsilon(\theta_1)\right)\Bigg).$$

Omitting the sufficiency condition for incentive compatibility given by (A.3) and focusing on a so called relaxed optimization problem, the principal's problem is to maximize (A.4) among all possible allocations $(\mathcal{U}(\theta_1),u_1(\theta_1),q_1(\theta_1),\varepsilon(\theta_1))$ subject to the necessary condition for first-period incentive compatibility (3.9), the second-period incentive compatibility condition (3.6) and the firm's participation constraint.[51] A few preliminary remarks are worth to be made.

• First, observe that (3.9) implies that $\mathcal{U}(\theta_1)$ is non-increasing so that the participation constraint:

$$\text{(A.5)}\quad \mathcal{U}(\theta_1)\ge 0 \quad \forall\theta_1\in\Theta_1$$

reduces to

$$\text{(A.6)}\quad \mathcal{U}(\overline{\theta}_1)\ge 0.$$

Of course, this constraint is binding at the optimum. Suppose it is not, then one could keep outputs the same in both periods, modify payments in those periods so that $u(\theta_1)$ is reduced by a small uniform amount $\frac{\epsilon}{1-\beta}$ (while increasing $y(\theta_1)$ by a small amount so as to keep constant

[51]The optimal contract is deterministic. The argument is similar to Strausz (2006) who shows that, when the optimal deterministic mechanism does not involve bunching, stochastic mechanisms cannot help.

$\frac{\mathcal{U}(\theta_1)-(1-\beta)u_1(\theta_1)}{\beta}$ and thus $\dot{\mathcal{U}}(\theta_1)$). This transformation thus reduces $\mathcal{U}(\theta_1)$ by ϵ. If ϵ is small enough, then (A.6) still holds and (A.5) also holds everywhere. Yet, such a modification of the profile would improve the principal's expected payoff by an amount ϵ.

• Second, the second-period incentive constraint (3.6) is necessarily binding at the optimum since $\varphi_\varepsilon > 0$. We therefore get $\varepsilon(\theta_1) = \Delta\theta_2 q_2(\theta_1, \overline{\theta}_2)$.

• Third, optimizing w.r.t. $q_2(\theta_1, \underline{\theta}_2)$ gives $q_2^{sb}(\theta_1, \underline{\theta}_2) = q_2^{fb}(\underline{\theta}_2)$ for all θ_1.

Therefore, we may simplify the expression of the principal's payoff from the add-on to:

$$\nu(S_2(q_2^{fb}(\underline{\theta}_2)) - \underline{\theta}_2 q_2^{fb}(\underline{\theta}_2)) + (1-\nu)\left(S_2\left(\frac{\varepsilon(\theta_1)}{\Delta\theta_2}\right) - \overline{\theta}_2 \frac{\varepsilon(\theta_1)}{\Delta\theta_2}\right).$$

Equipped with this expression, and denoting by λ the costate variable for (3.9) we can write the Hamiltonian for the principal's problem as:

(A.7) $\mathcal{H}(\mathcal{U}, q_1, u_1, \varepsilon, \lambda, \theta_1) =$

$$f(\theta_1)\Bigg(S_1(q_1) - \theta_1 q_1 - (1-\beta)u_1 - \beta\varphi\left(\frac{\mathcal{U}-(1-\beta)u_1}{\beta}, \varepsilon\right) + \nu\beta\left(S_2(q_2^{fb}(\underline{\theta}_2)) - \underline{\theta}_2 q_2^{fb}(\underline{\theta}_2)\right) +$$
$$(1-\nu)\beta\left(S_2\left(\frac{\varepsilon}{\Delta\theta_2}\right) - \overline{\theta}_2\frac{\varepsilon}{\Delta\theta_2}\right)\Bigg) - \lambda q_1\left(1-\beta+\beta w_z\left(\varphi\left(\frac{\mathcal{U}-(1-\beta)u_1}{\beta}, \varepsilon\right), \varepsilon\right)\right).$$

We shall assume that $\mathcal{H}(\mathcal{U}, q_1, u_1, \varepsilon, \lambda, \theta_1)$ is concave in $(\mathcal{U}, q_1, u_1, \varepsilon)$ and use the Pontryagin Principle to get optimality conditions satisfied by an extremal arc $(\mathcal{U}^{sb}(\theta_1), u_1^{sb}(\theta_1), q_1^{sb}(\theta_1), \varepsilon^{sb}(\theta_1))$. Sufficient conditions for concavity of $\mathcal{H}(\mathcal{U}, q_1, u_1, \varepsilon, \lambda, \theta_1)$ are listed in Appendix C. Concavity is ensured provided that β is small enough and $S_2(\cdot)$ is sufficiently concave.

• *Costate variable.* $\lambda(\theta_1)$ is continuous, piecewise continuously differentiable and such that:

$$\dot{\lambda}(\theta_1) = -\frac{\partial\mathcal{H}}{\partial\mathcal{U}}(\mathcal{U}^{sb}(\theta_1), q_1^{sb}(\theta_1), u_1^{sb}(\theta_1), \varepsilon^{sb}(\theta_1), \theta_1),$$

which amounts to

(A.8)
$$\frac{\dot{\lambda}(\theta_1)}{\varphi_\zeta\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta}, \varepsilon^{sb}(\theta_1)\right)} = f(\theta_1) + \lambda(\theta_1)q_1^{sb}(\theta_1)w_{zz}\left(\varphi\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta}, \varepsilon^{sb}(\theta_1)\right), \varepsilon^{sb}(\theta_1)\right).$$

• *Transversality condition.* Because (A.6) is binding at the optimum, this condition is:

(A.9) $\lambda(\underline{\theta}_1) = 0.$

• *First-order optimality condition w.r.t. u_1:*

(A.10)
$$\frac{f(\theta_1)}{\varphi_\zeta\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta}, \varepsilon^{sb}(\theta_1)\right)} = f(\theta_1) + \lambda(\theta_1)q_1^{sb}(\theta_1)w_{zz}\left(\varphi\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta}, \varepsilon^{sb}(\theta_1)\right), \varepsilon^{sb}(\theta_1)\right).$$

- *First-order optimality condition w.r.t. q_1:*

(A.11)
$$S_1'(q_1^{sb}(\theta_1)) = \theta_1 + \frac{\lambda(\theta_1)}{f(\theta_1)}\left(1-\beta+\beta w_z\left(\varphi\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta},\varepsilon^{sb}(\theta_1)\right),\varepsilon^{sb}(\theta_1)\right)\right).$$

- *First-order optimality condition w.r.t. ε:*

(A.12)
$$\frac{1-\nu}{\Delta\theta_2}\left(S_2'\left(\frac{\varepsilon^{sb}(\theta_1)}{\Delta\theta_2}\right)-\overline{\theta}_2\right) = \varphi_\varepsilon\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta},\varepsilon^{sb}(\theta_1)\right)$$
$$+q_1^{sb}(\theta_1)\frac{\lambda(\theta_1)}{f(\theta_1)}\Bigg(w_{zz}\left(\varphi\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta},\varepsilon^{sb}(\theta_1)\right),\varepsilon^{sb}(\theta_1)\right)\varphi_\varepsilon\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta},\varepsilon^{sb}(\theta_1)\right)$$
$$+w_{z\varepsilon}\left(\varphi\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta},\varepsilon^{sb}(\theta_1)\right),\varepsilon^{sb}(\theta_1)\right)\Bigg).$$

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

- Proposition 1. Inserting (A.10) into (A.8) and simplifying yields $\dot{\lambda}(\theta_1) = f(\theta_1)$, which together with (A.9) gives us:

(A.13)
$$\lambda(\theta_1) = F(\theta_1), \quad \forall \theta_1 \in \Theta_1.$$

Inserting this expression into (A.10), taking into account that $\varphi_\zeta(z,\varepsilon) = \frac{1}{w_z(\varphi(\zeta,\varepsilon),\varepsilon)}$, using (3.8) and simplifying yields (5.1).

- Proposition 2. Inserting (A.13) into (A.10) and simplifying using (3.8) gives us (5.2).

From (5.1), we know that $w_z(u_1^{sb}(\theta_1)+y^{sb}(\theta_1),\varepsilon^{sb}(\theta_1)) \leq 1$. Therefore, (5.2) implies that:
$$S_1'(q_1^{sb}(\theta_1)) \leq \theta_1 + \frac{F(\theta_1)}{f(\theta_1)}, \quad \forall \theta_1 \in \Theta_1,$$
and thus $q_1^{sb}(\theta_1) \geq q_1^{bm}(\theta_1)$.

- Proposition 3. Inserting (A.13) into (A.12) and simplifying using (3.8) gives us (5.3). *Q.E.D.*

Proofs of Results in Section 6.1: When θ_2 is common knowledge and the firm can be fully insured against risk so that $\varepsilon^i(\theta_1)=0$, $q_2(\theta_1,\theta_2)$ is set at the first-best level: $q_2^i(\theta_1,\theta_2) = q_2^{fb}(\theta_2)$, for all (θ_1,θ_2). We can now write the Hamiltonian for the corresponding optimization problem as:
$$\mathcal{H}^i(\mathcal{U},q_1,u,\lambda,\theta_1) = f(\theta_1)\left(S_1(q_1)-\theta_1 q_1-(1-\beta)u_1-\beta\varphi\left(\frac{\mathcal{U}-(1-\beta)u_1}{\beta},0\right)+\beta\mathbb{E}_{\theta_2}\left(S_2(q_2^{fb}(\theta_2))-\theta_2 q_2^{fb}(\theta_2)\right)\right)$$
$$-\lambda q_1\left(1-\beta+\beta v'\left(\varphi\left(\frac{\mathcal{U}-(1-\beta)u_1}{\beta},0\right)\right)\right).$$

The optimization then follows the same steps as the proofs for Proposition 1 and 2. Details are thus omitted. *Q.E.D.*

PROOFS OF RESULTS IN SECTION 6.2: When θ_1 is common knowledge, the principal maximizes (A.4) subject to (3.6) and (A.5). Because $\varphi_\zeta \geq 0$, (A.5) is necessarily binding. Inserting $\mathcal{U}(\theta_1) = 0$ into the maximand, this maximand is decreasing in ε and thus (3.6) is also binding. From there, pointwise optimization yields the results. In particular, it is trivial to check that second-period distortions and the fixed per-period profits are independent of the first-period cost. Moreover, the Inada condition $S_2'(0) = +\infty$ ensures that q_2^r and thus ε^r are both positive. *Q.E.D.*

PROOF OF PROPOSITION 4: With limited commitment, the Revelation Principle *a priori* fails and there is no reason to expect that the firm reports truthfully. By misreporting its first-period type, the firm might secure a more attractive renegotiation later on. Let $M(\hat{\theta}_1|\theta_1)$ be the distribution of optimal first-period reports for the type θ_1. The renegotiated offer may entail lower output distortions $\tilde{q}_2(\hat{\theta}_1, \overline{\theta}_2)$ and thus have the firm bear more risk $\tilde{\varepsilon}(\hat{\theta}_1)$ than the initial contract. For the firm to accept this new offer, it must stipulate an extra premium $\tilde{y}(\hat{\theta}_1)$, production levels of the add-on $\tilde{q}_2(\hat{\theta}_1, \hat{\theta}_2)$ and an amount of risk $\tilde{\varepsilon}(\hat{\theta}_1)$[52] that are found more profitable by the principal than the initial contract and leave the firm with at least the same payoff as in the initial offer. For any θ_1 and $\hat{\theta}_1$ that lies in the support of $M(\cdot|\theta_1)$, we must have:

$$w(t_1(\hat{\theta}_1) + y(\hat{\theta}_1) + \tilde{y}(\hat{\theta}_1) - \theta_1 q_1(\hat{\theta}_1), \tilde{\varepsilon}(\hat{\theta}_1)) \geq w(t_1(\hat{\theta}_1) + y(\hat{\theta}_1) - \theta_1 q_1(\hat{\theta}_1), \varepsilon(\hat{\theta}_1)). \tag{A.14}$$

The new contract must also remain incentive compatible for the second period:

$$\tilde{\varepsilon}(\hat{\theta}_1) \geq \Delta\theta_2 \tilde{q}_2(\hat{\theta}_1, \overline{\theta}_2) \quad \forall \hat{\theta}_1 \in \Theta_1. \tag{A.15}$$

To check whether a long-term contract $\mathcal{C}$ (and in particular the optimal contract under full commitment, $\mathcal{C}^{sb}$) is renegotiation-proof when the firm truthfully reveals its type in the first period (i.e., the first-period strategy $M(\cdot|\theta_1)$ puts unit mass on $\hat{\theta}_1 = \theta_1$) we first look for the optimal continuation for the second period following such truthful strategy. The first step is to observe that (A.14) becomes:

$$w(t_1(\theta_1)+y(\theta_1)+\tilde{y}(\theta_1)-\theta_1 q_1(\theta_1), \tilde{\varepsilon}(\theta_1)) \geq w(t_1(\theta_1)+y(\theta_1)-\theta_1 q_1(\theta_1), \varepsilon(\theta_1)), \quad \forall \theta_1 \in \Theta_1. \tag{A.16}$$

The principal's problem at the renegotiation stage is thus to maximize his second-period expected payoff from offering the new contract, namely $\mathbb{E}_{\theta_2}\left(S_2(\tilde{q}_2(\theta_1, \theta_2)) - \theta_2 \tilde{q}_2(\theta_1, \theta_2)\right) - \tilde{y}(\theta_1)$ subject to (A.15) and (A.16).

Observe that the acceptance condition (A.16) is necessarily binding because, otherwise decreasing $\tilde{y}(\theta_1)$ would increase the principal's expected payoff. Denoting again the fixed per-period profit in the long-term contract as $u_1(\theta_1) = t_1(\theta_1) - \theta_1 q_1(\theta_1)$ and the firm's reservation

[52] In other words the new prices for each levels of the add-ons are now $\tilde{t}_2(\hat{\theta}_1, \underline{\theta}_2) - \underline{\theta}_1 \tilde{q}_2(\hat{\theta}_1, \underline{\theta}_2) = \tilde{y}(\hat{\theta}_1) + (1-\nu)\tilde{\varepsilon}(\hat{\theta}_1)$ and $\tilde{t}_2(\hat{\theta}_1, \overline{\theta}_2) - \overline{\theta}_1 \tilde{q}_2(\hat{\theta}_1, \overline{\theta}_2) = \tilde{y}(\hat{\theta}_1) - \nu\tilde{\varepsilon}(\hat{\theta}_1)$.

payoff for the second period as $w_0(\theta_1) = w(u_1(\theta_1) + y(\theta_1), \varepsilon(\theta_1))$, we can thus write:

$$\text{(A.17)} \quad u_1(\theta_1) + y(\theta_1) + \tilde{y}(\theta_1) = \varphi(w_0(\theta_1), \tilde{\varepsilon}(\theta_1)), \quad \forall \theta_1 \in \Theta_1.$$

Inserting this expression of $\tilde{y}(\theta_1)$ into the maximand of the principal's problem and optimizing w.r.t. $\tilde{\varepsilon}(\theta_1)$ shows that (A.14) is necessarily binding. The last step of the optimization gives us $q_2^{re}(\theta_1, \underline{\theta}_2) = q_2^{fb}(\underline{\theta}_2)$ for all θ_1. When the second-period cost is $\overline{\theta}_2$, the distortion is:

$$\text{(A.18)} \quad (1-\nu)(S_2'(q_2^{re}(\theta_1, \overline{\theta}_2)) - \overline{\theta}_2) = \Delta\theta_2 \varphi_\varepsilon(w_0(\theta_1), \varepsilon^{re}(\theta_1)), \quad \forall \theta_1 \in \Theta_1.$$

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

$$w_0(\theta_1) = w(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), \varepsilon^{sb}(\theta_1)), \varepsilon^{re}(\theta_1) = \varepsilon^{sb}(\theta_1), q_2^{re}(\theta_1, \overline{\theta}_2) = q_2^{sb}(\theta_1, \overline{\theta}_2), \text{ and } \tilde{y}(\theta_1) = 0, \forall \theta_1 \in \Theta_1.$$

Inserting these conditions into (A.18) yields:

$$(1-\nu)(S_2'(q_2^{sb}(\theta_1, \overline{\theta}_2)) - \overline{\theta}_2) = \Delta\theta_2 \varphi_\varepsilon(w(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), \varepsilon^{sb}(\theta_1)), \varepsilon^{sb}(\theta_1)), \quad \forall \theta_1 \in \Theta_1.$$

Comparing this condition to (5.3) immediately tells us that, whenever $H(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), \varepsilon^{sb}(\theta_1)) > 0$, $\mathcal{C}^{sb}$ is not renegotiation-proof. On the other hand, for $CARA$ preferences, we have $H(z, \varepsilon) \equiv 0$ and thus the long-term contract $\mathcal{C}^{sb}$ is always renegotiation-proof. *Q.E.D.*

PROOF OF PROPOSITION 5: Suppose that the contract $(t_1^{sb}(\hat{\theta}_1), y^{sb}(\hat{\theta}_1), q_1^{sb}(\hat{\theta}_1))$ regulates the basic service over the two periods of the relationship. The firm accepts a spot contract $(\tilde{y}(\hat{\theta}_1), \tilde{\varepsilon}(\hat{\theta}_1))$ for the add-on if, for any report $\hat{\theta}_1$, its payoff weakly increases:

$$\text{(A.19)} \quad w(t_1^{sb}(\hat{\theta}_1) + y^{sb}(\hat{\theta}_1) - \theta_1 q_1^{sb}(\hat{\theta}_1) - y^r + \tilde{y}(\hat{\theta}_1), \tilde{\varepsilon}(\hat{\theta}_1)) \geq w(t_1^{sb}(\hat{\theta}_1) + y^{sb}(\hat{\theta}_1) - y^r - \theta_1 q_1^{sb}(\hat{\theta}_1), 0).$$

For $CARA$ preferences, this condition amounts to:

$$\text{(A.20)} \quad w(\tilde{y}(\hat{\theta}_1), \tilde{\varepsilon}(\hat{\theta}_1)) \geq 0.$$

When proposing the spot contract (y, ε), the principal's wants to maximize his second-period expected payoff, namely:

$$\mathbb{E}_{\theta_2}\left(S_2(\tilde{q}_2(\theta_1, \theta_2)) - \theta_2 \tilde{q}_2(\theta_1, \theta_2)\right) - \tilde{y}(\theta_1),$$

subject to second-period incentive compatibility expressed in (A.15) and the acceptance condition (A.20). Clearly, both constraints above are binding and the solution under $CARA$ preference yields second-period levels of the add-on given by $q_2^{sb}(\theta_1, \theta_2)$ in (5.4), $\tilde{y}(\theta_1) = y^r$ and $\tilde{\varepsilon}(\hat{\theta}_1) = \varepsilon^r$. From these findings, we get that the following equality holds:

$$w(t_1^{sb}(\hat{\theta}_1) - \theta_1 q_1^{sb}(\hat{\theta}_1) + y^{sb}(\hat{\theta}_1) - y^r + y^r, \varepsilon^r) = w(t_1^{sb}(\hat{\theta}_1) - \theta_1 q_1^{sb}(\hat{\theta}_1) + y^{sb}(\hat{\theta}_1), \varepsilon^{sb}(\theta_1)).$$

This condition means that the firm, anticipating acceptance of the spot contract (y^r, ε^r), also truthfully reveals its type to get the same payoff as in the second-best contract $\mathcal{C}^{sb}$, namely:

$$\mathcal{U}^{sb}(\theta_1) = \max_{\hat{\theta}_1 \in \Theta_1} (1-\beta)(t_1^{sb}(\hat{\theta}_1) - \theta_1 q_1^{sb}(\hat{\theta}_1)) + \beta w(t_1^{sb}(\hat{\theta}_1) - \theta_1 q_1^{sb}(\hat{\theta}_1) + y^{sb}(\hat{\theta}_1) - y^r + y^r, \varepsilon^r).$$

This shows that the same outcome as with $\mathcal{C}^{sb}$ can be obtained. *Q.E.D.*

APPENDIX B: PROPERTIES OF THE UTILITY FUNCTION

From (2.2), $w(\cdot)$ is increasing and concave in z while it remains decreasing in ε:

$$w_z(z,\varepsilon) = \nu v'(z+(1-\nu)\varepsilon)+(1-\nu)v'(z-\nu\varepsilon) > 0, w_{zz}(z,\varepsilon) = \nu v''(z+(1-\nu)\varepsilon)+(1-\nu)v''(z-\nu\varepsilon) \leq 0,$$

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

Lemma B.1 *Suppose that $v(\cdot)$ is DARA (resp. CARA). Then,*

$$H(z,\varepsilon) \geq 0 \ (resp. = 0) \quad \forall (z,\varepsilon) \in \mathbb{R} \times \mathbb{R}_+. \tag{B.1}$$

Proof of Lemma B.1: The inequality in (B.1) can be rewritten as:

$$\begin{aligned}&\left(v''(z+(1-\nu)\varepsilon) - v''(z-\nu\varepsilon)\right)\left(\nu v'(z+(1-\nu)\varepsilon) + (1-\nu)v'(z-\nu\varepsilon)\right)\\ &\geq \left(v'(z+(1-\nu)\varepsilon) - v'(z-\nu\varepsilon)\right)\left(\nu v''(z+(1-\nu)\varepsilon) + (1-\nu)v''(z-\nu\varepsilon)\right) \quad \forall (z,\varepsilon) \in \mathbb{R} \times \mathbb{R}_+.\end{aligned}$$

Simplifying this condition yields:

$$v'(z-\nu\varepsilon)v''(z+(1-\nu)\varepsilon) - v'(z+(1-\nu)\varepsilon)v''(z-\nu\varepsilon) \geq 0 \quad \forall (z,\varepsilon) \in \mathbb{R} \times \mathbb{R}_+.$$

This amounts to:

$$-\frac{v''(z+(1-\nu)\varepsilon)}{v'(z+(1-\nu)\varepsilon)} \leq -\frac{v''(z-\nu\varepsilon)}{v'(z-\nu\varepsilon)} \quad \forall (z,\varepsilon) \in \mathbb{R} \times \mathbb{R}_+;$$

a condition which holds if preferences are *DARA* or *CARA*. *Q.E.D.*

APPENDIX C: CONCAVITY CONDITIONS (NOT FOR PUBLICATION)

In this Appendix, we present conditions of the Mangasarian type that ensure that the allocation $(\mathcal{U}^{sb}, q_1^{sb}, u_1^{sb}, \varepsilon^{sb})$ characterized in the Proofs of Propositions 1, 2 and 3 through necessary conditions is indeed the solution. We follow Seierstad and Sydsaeter (1987) (Chapter 2, Theorem 4) and check that the Hamiltonian $\mathcal{H}(\mathcal{U}, q_1, u_1, \varepsilon, \lambda(\theta_1), \theta_1)$ defined in (A.7) is indeed concave in $(\mathcal{U}, q_1, u_1, \varepsilon)$. This yields conditions on the principal minors of the Hessian matrix for $\mathcal{H}$. For notational brevity $\mathcal{H}_{xy}$ denotes the cross derivative of $\mathcal{H}$ w.r.t. x and y and where we omit variables for simplifying notations.

In the general case where both θ_1 and θ_2 are private information, we require the following four conditions to hold:

(C.1) $f(\theta_1)S_1''(q_1(\theta_1))\mathcal{H}_{\mathcal{U}\mathcal{U}} - (\lambda(\theta_1)w_{zz}(\varphi(u_1(\theta_1)+y(\theta_1)),\varepsilon)\varphi_\zeta(u_1(\theta_1)+y(\theta_1),\varepsilon))^2 \geq 0,$

(C.2) $f(\theta_1)S_1''(q_1(\theta_1))\mathcal{H}_{\varepsilon\varepsilon}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1)$
$- (\beta\lambda(\theta_1)\,(w_{z\varepsilon}(\varphi(u_1(\theta_1)+y(\theta_1),\varepsilon),\varepsilon) + w_{zz}(\varphi(u_1(\theta_1)+y(\theta_1),\varepsilon),\varepsilon)\varphi_\varepsilon(u_1(\theta_1)+y(\theta_1),\varepsilon)))^2 \geq 0,$

(C.3) $\mathcal{H}_{\mathcal{U}\mathcal{U}}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1)\mathcal{H}_{\varepsilon\varepsilon}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1) - (\mathcal{H}_{\mathcal{U}\varepsilon}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1))^2 \geq 0,$

(C.4) $f(\theta_1)S_1''(q_1(\theta_1))\left(\mathcal{H}_{\mathcal{U}\mathcal{U}}\mathcal{H}_{\varepsilon\varepsilon} - (\mathcal{H}_{\mathcal{U}\varepsilon}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1))^2\right)$
$+ 2\mathcal{H}_{q_1\mathcal{U}}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1)\mathcal{H}_{q_1\varepsilon}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1)\mathcal{H}_{\mathcal{U}\varepsilon}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1)$
$- (\mathcal{H}_{q_1\varepsilon}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1))^2\,\mathcal{H}_{\mathcal{U}\mathcal{U}}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1)$
$- (\mathcal{H}_{q_1\mathcal{U}}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1))^2\,\mathcal{H}_{\varepsilon\varepsilon}(q_1,u_1,\varepsilon,\mathcal{U},\theta_1) \leq 0,$

For sufficiently concave surplus functions S_1 and S_2 compared to the degree of risk aversion, these conditions hold as this insures sufficiently high negative values of the diagonal of the Hessian matrix compared to the off-diagonal elements.

Furthermore, it can be checked (through straightforward but tedious computations) that in the case where the firm exhibits $CARA$ preferences and τ is small, these conditions are always satisfied (regardless of the degree of concavity of the surplus functions).

APPENDIX D: PROOFS FOR SECTION 8 (NOT FOR PUBLICATION)

LEMMA D.1 *Suppose that $v(\cdot)$ is DARA (resp. CARA). Then,*

(D.1) $H(z,p) \geq 0$ (*resp.* $=0$) $\quad \forall (z,p) \in \mathbb{R}\times\Theta_2.$

PROOF OF LEMMA D.1: Simple properties of $w(\cdot)$ immediately follow from its definition (8.1):

$$w_p(z,p) = \int_{\underline{\theta}_2}^{p} v'(z+p-\theta_2)f_2(\theta_2)d\theta_2 > 0, w_{zp}(z,p) = \int_{\underline{\theta}_2}^{p} v''(z+p-\theta_2)f_2(\theta_2)d\theta_2 \leq 0,$$
$$w_z(z,p) = \int_{\underline{\theta}_2}^{p} v'(z+p-\theta_2)f_2(\theta_2)d\theta_2 + v'(z)(1-F_2(p)) > 0,$$
$$w_{zz}(z,p) = \int_{\underline{\theta}_2}^{p} v''(z+p-\theta_2)f_2(\theta_2)d\theta_2 + v''(z)(1-F_2(p)) \leq 0.$$

The inequality in (D.1) can be rewritten as:

$$\left(\int_{\underline{\theta}_2}^{p} v''(z+p-\theta_2)f_2(\theta_2)d\theta_2\right)\left(\int_{\underline{\theta}_2}^{p} v'(z+p-\theta_2)f_2(\theta_2)d\theta_2 + v'(z)(1-F_2(p))\right)$$
$$\geq \left(\int_{\underline{\theta}_2}^{p} v''(z+p-\theta_2)f_2(\theta_2)d\theta_2 + v''(z)(1-F_2(p))\right)\left(\int_{\underline{\theta}_2}^{p} v'(z+p-\theta_2)f_2(\theta_2)d\theta_2\right), \forall (z,p)\in\mathbb{R}\times\Theta_2$$

Developing and rearranging, this amounts to demonstrating that:

(D.2)

$$v'(z)\left(\int_{\underline{\theta}_2}^{p} v''(z+p-\theta_2)f_2(\theta_2)d\theta_2\right) \geq v''(z)\left(\int_{\underline{\theta}_2}^{p} v'(z+p-\theta_2)f_2(\theta_2)d\theta_2\right) \quad \forall (z,p)\in\mathbb{R}\times\Theta_2.$$

Now, observe that $v(\cdot)$ *DARA* (resp. *CARA*) implies:

$$-\frac{v''(z+p-\theta_2)}{v'(z+p-\theta_2)} \leq -\frac{v''(z)}{v'(z)} \quad \forall \theta_2 \leq p, \quad \forall z.$$

Multiplying both terms by $v'(z+p-\theta_2)$ and $v'(z)$ and integrating over $[\underline{\theta}_2, p]$ yields (D.2) and proves the Lemma. *Q.E.D.*

PROOF OF OTHER RESULTS: Trade in the second period occurs only when $\theta_2 \leq p(\theta_1)$. Adapting the general expression (2.1) to the present context, the principal's expected payoff becomes:

$$\mathbb{E}_{\theta_1}\left(S_1(q_1(\theta_1)) - \theta_1 q_1(\theta_1) - u_1(\theta_1) + \beta\mathbb{E}_{\theta_2}\left((S_2 - p(\theta_1))F_2(p(\theta_1)) - y(\theta_1)\right)\right),$$

which can again be re-expressed as:

$$\text{(D.3)} \quad \mathbb{E}_{\theta_1}\Bigg(S_1(q_1(\theta_1)) - \theta_1 q_1(\theta_1) - (1-\beta)u_1(\theta_1) + \beta\mathbb{E}_{\theta_2}\left((S_2 - p(\theta_1))F_2(p(\theta_1))\right) - \beta\varphi\left(\frac{\mathcal{U}(\theta_1) - (1-\beta)u_1(\theta_1)}{\beta}, p(\theta_1)\right)\Bigg).$$

In terms of first-period incentive compatibility, (3.9) is readily replaced with:

$$\text{(D.4)} \quad \dot{\mathcal{U}}(\theta_1) = -q_1(\theta_1)\left(1-\beta+\beta w_z\left(\varphi\left(\frac{\mathcal{U}(\theta_1)-(1-\beta)u_1(\theta_1)}{\beta}, \varepsilon(\theta_1)\right), p(\theta_1)\right)\right).$$

We now proceed as in the Proof of Propositions 1, 2 and 3 by relying on necessary conditions for optimality (details are omitted). Denoting again by λ the costate variable for (D.4), we now write the Hamiltonian for the principal's problem as:

$$\mathcal{H}(\mathcal{U}, q_1, u_1, p, \lambda, \theta_1) = f(\theta_1)\Bigg(S_1(q_1) - \theta_1 q_1 - (1-\beta)u_1 - \beta\varphi\left(\frac{\mathcal{U}-(1-\beta)u_1}{\beta}, p\right) + \beta(S_2-p)F_2(p)\Bigg) - \lambda q_1\left(1-\beta+\beta w_z\left(\varphi\left(\frac{\mathcal{U}-(1-\beta)u_1}{\beta}, p\right), p\right)\right).$$

Relying on the Pontryagin Principle to write the necessary conditions for an optimum $(\mathcal{U}^{sb}(\theta_1), u_1^{sb}(\theta_1), q_1^{sb}(\theta_1), p^{sb}(\theta_1))$, we obtain:

$$\text{(D.5)} \quad \frac{\dot{\lambda}(\theta_1)}{\varphi_\zeta\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta}, p^{sb}(\theta_1)\right)} = f(\theta_1)+\lambda(\theta_1)q_1^{sb}(\theta_1)w_{zz}\left(\varphi\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta}, p^{sb}(\theta_1)\right), p^{sb}(\theta_1)\right).$$

The transversality condition is still given by (A.9) and optimality w.r.t. u_1, q_1 and p yields:

$$\text{(D.6)}\qquad \frac{f(\theta_1)}{\varphi_\zeta\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta}, p^{sb}(\theta_1)\right)} = f(\theta_1) + \lambda(\theta_1)q_1^{sb}(\theta_1)w_{zz}\left(\varphi\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta}, p^{sb}(\theta_1)\right), p^{sb}(\theta_1)\right),$$

$$\text{(D.7)}\qquad S_1'(q_1^{sb}(\theta_1)) = \theta_1 + \frac{\lambda(\theta_1)}{f(\theta_1)}\left(1-\beta+\beta w_z\left(\varphi\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta}, p^{sb}(\theta_1)\right), p^{sb}(\theta_1)\right)\right),$$

$$\text{(D.8)}\qquad (S_2 - p^{sb}(\theta_1))f_2(p^{sb}(\theta_1)) - F_2(p^{sb}(\theta_1)) = \varphi_p\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta}, p^{sb}(\theta_1)\right) + q_1^{sb}(\theta_1)\frac{\lambda(\theta_1)}{f(\theta_1)}H\left(\varphi\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)u_1^{sb}(\theta_1)}{\beta}, p^{sb}(\theta_1)\right), p^{sb}(\theta_1)\right).$$

Notice that (D.5), (D.6) and (A.9) still imply (A.13). Condition (8.3) immediately follows from inserting (A.13) into (D.6). From (8.3), we know that $w_z(u_1^{sb}(\theta_1) + y^{sb}(\theta_1), p^{sb}(\theta_1)) \leq 1$. Therefore, (8.4) implies that $q_1^{sb}(\theta_1) \geq q_1^{bm}(\theta_1)$. Finally, inserting (A.13) into (D.8), simplifying and taking into account the result of Lemma D.1 gives us (8.5). *Q.E.D.*

APPENDIX E: FIRST-PERIOD RISK AVERSION (NOT FOR PUBLICATION)

Suppose that the agent also evaluates the first-period returns according to the same utility function $v(\cdot)$ as in the second period. We first analyze the case of a durable project. Then, and for the sake of completeness, we also report on the case of a non-durable, i.e., q_1 only arises in the first period. For simplicity and under both scenarios, we suppose that θ_2 remains common knowledge.

THE CASE OF A DURABLE FIRST-PERIOD PROJECT: The next proposition shows that the *Income Effect* disappears as suggested in the text. The principal finds no value in shifting payments towards the second period. As a result, the basic service is produced at its Baron-Myerson level.

PROPOSITION E.1 *Suppose that θ_2 remains common knowledge and that the first-period project is durable. The optimal contract has the following features:*

- *Constant profit over time for the durable:*

$$\text{(E.1)}\qquad y_2^{sb}(\theta_1) = 0.$$

- *The durable is produced at its Baron-Myerson level:*

$$\text{(E.2)}\qquad q_1^{sb}(\theta_1) = q_1^{bm}(\theta_1).$$

To show these results, observe that the principal's expected payoff can now be written as:

$$\text{(E.3)}\qquad \mathbb{E}_{\theta_1}\Big(S_1(q_1(\theta_1)) - \theta_1 q_1(\theta_1) - (1-\beta)u_1(\theta_1) + \beta\mathbb{E}_{\theta_2}\left(S_2(q_2(\theta_1,\theta_2)) - \theta_2 q_2(\theta_1,\theta_2)\right)$$

$$-\beta\varphi\left(\frac{\mathcal{U}(\theta_1)-(1-\beta)v(u_1(\theta_1))}{\beta},0\right)\Bigg).$$

Omitting the sufficiency condition for incentive compatibility given by (A.3) and focusing on a so called relaxed optimization problem, the principal's problem is to maximize (E.3) among all possible allocations $(\mathcal{U}(\theta_1), u_1(\theta_1), q_1(\theta_1))$ subject to the necessary condition for first-period incentive compatibility (6.3) and the firm's participation constraint (A.6) that again turns out to be binding at the optimum.

• Optimizing w.r.t. $q_2(\theta_1,\theta_2)$ gives $q_2^{sb}(\theta_1,\theta_2) = q_2^{fb}(\theta_2)$ for all (θ_1,θ_2). Therefore, we may simplify the expression of the principal's payoff from the add-on to:

$$\mathbb{E}_{\theta_2}(S_2(q_2^{fb}(\theta_2)) - \theta_2 q_2^{fb}(\theta_2)).$$

Equipped with this expression, and denoting by λ the costate variable for (6.3) we can write the Hamiltonian for the principal's problem as:

$$\text{(E.4)}\qquad \mathcal{H}(\mathcal{U}, q_1, u_1, \lambda, \theta_1) =$$

$$f(\theta_1)\left(S_1(q_1)-\theta_1 q_1-(1-\beta)u_1-\beta\varphi\left(\frac{\mathcal{U}-(1-\beta)v(u_1)}{\beta},0\right)+\mathbb{E}_{\theta_2}(S_2(q_2^{fb}(\theta_2))-\theta_2 q_2^{fb}(\theta_2))\right)$$

$$-\lambda q_1\left((1-\beta)v'(u_1)+\beta v'\left(\varphi\left(\frac{\mathcal{U}-(1-\beta)v(u_1)}{\beta},0\right)\right)\right).$$

We shall assume that $\mathcal{H}(\mathcal{U}, q_1, u_1, \lambda, \theta_1)$ is concave in $(\mathcal{U}, q_1, u_1)$ and use the Pontryagin Principle to get optimality conditions satisfied by an extremal arc $(\mathcal{U}^{sb}(\theta_1), u_1^{sb}(\theta_1), q_1^{sb}(\theta_1))$.

• *Costate variable.* $\lambda(\theta_1)$ is continuous, piecewise continuously differentiable and such that:

$$\text{(E.5)}\qquad \dot{\lambda}(\theta_1)v'(u_2^{sb}(\theta_1)) = f(\theta_1) + \lambda(\theta_1)q_1^{sb}(\theta_1)v''\left(u_2^{sb}(\theta_1)\right)$$

where the second-period profit is

$$\text{(E.6)}\qquad u_2^{sb}(\theta_1) = u_1^{sb}(\theta_1) + y^{sb}(\theta_1) = \varphi\left(\frac{\mathcal{U}^{sb}(\theta_1)-(1-\beta)v(u_1^{sb}(\theta_1))}{\beta},0\right).$$

• *Transversality condition.* Because (A.6) is binding at the optimum, this condition is:

$$\text{(E.7)}\qquad \lambda(\underline{\theta}_1) = 0.$$

• *First-order optimality condition w.r.t. u_1:*

$$\text{(E.8)}\qquad f(\theta_1)\frac{v'(u_1^{sb}(\theta_1))}{v'(u_2^{sb}(\theta_1))} = f(\theta_1) + \lambda(\theta_1)q_1^{sb}(\theta_1)\left(v''(u_1^{sb}(\theta_1)) - \frac{v''(u_2^{sb}(\theta_1))}{v'(u_2^{sb}(\theta_1))}v'(u_1^{sb}(\theta_1))\right).$$

• *First-order optimality condition w.r.t. q_1:*

$$\text{(E.9)}\qquad S_1'(q_1^{sb}(\theta_1)) = \theta_1 + \frac{\lambda(\theta_1)}{f(\theta_1)}\left((1-\beta)v'(u_1^{sb}(\theta_1)) + \beta v'(u_2^{sb}(\theta_1))\right).$$

A solution to (E.8) is given by:

(E.10) $u_1^{sb}(\theta_1) = u_2^{sb}(\theta_1).$

Inserting into (6.3) and (E.6) yields respectively:

(E.11) $\dot{\mathcal{U}}^{sb}(\theta_1) = -q_1^{sb}(\theta_1)v'(u_1^{sb}(\theta_1))$ where $\mathcal{U}^{sb}(\theta_1) = v(u_1^{sb}(\theta_1))$

which implies

(E.12) $\dot{u}_1^{sb}(\theta_1) = -q_1^{sb}(\theta_1).$

Inserting into (E.5) and using again (E.11) gives:

$$\frac{d}{d\theta}(\lambda(\theta_1)v'(u_1^{sb}(\theta))) = f(\theta_1).$$

Integrating and using (E.7) we obtain:

$$\lambda(\theta_1)v'(u_1^{sb}(\theta)) = F(\theta_1).$$

Inserting into (E.9) and again taking into account (E.11) gives (E.2).

Q.E.D.

THE CASE OF A NON-DURABLE FIRST-PERIOD PROJECT: The next proposition shows that the principal wants to push profits for the first-period project into the second period even if the first-period project is not a durable one. This project is produced below the first-best level.

PROPOSITION E.2 *Suppose that θ_2 remains common knowledge and that the first-period project's surplus and costs only arise in the first period. The optimal contract has the following features.*

- *The first-period project is rewarded in both periods but with declining profits:*

  (E.13) $u_1^{sb}(\theta_1) \geq u_2^{sb}(\theta_1)$

  with an equality only in the case of risk neutrality.

- *The first-period production is:*

  (E.14) $S_1'(q_1^{sb}(\theta_1)) = \theta_1 + \frac{v'(u_1^{sb}(\theta_1))}{f(\theta_1)} \int_{\underline{\theta}_1}^{\theta_1} \frac{f(\tilde{\theta})}{v'(u_2^{sb}(\tilde{\theta}))} d\tilde{\theta}.$

We first notice that, with a short-term project, the envelope condition for incentive compatibility becomes:

(E.15) $\dot{\mathcal{U}}(\theta_1) = -(1-\beta)q_1(\theta_1)v'(u_1(\theta_1)).$

The principal's expected payoff also takes into account that surplus and cost for q_1 only arise in the first period and have to be weighted accordingly:

$$\text{(E.16)}\quad \mathbb{E}_{\theta_1}\Bigg((1-\beta)(S_1(q_1(\theta_1))-\theta_1 q_1(\theta_1)-u_1(\theta_1))+\beta\mathbb{E}_{\theta_2}\left(S_2(q_2^{fb}(\theta_2))-\theta_2 q_2^{fb}(\theta_2)\right)$$

$$-\beta\varphi\left(\frac{\mathcal{U}(\theta_1)-(1-\beta)v(u_1(\theta_1))}{\beta},0\right)\Bigg).$$

Omitting the sufficiency condition for incentive compatibility given by (A.3) and focusing on a so called relaxed optimization problem, the principal's problem is to maximize (E.16) among all possible allocations $(\mathcal{U}(\theta_1), u_1(\theta_1), q_1(\theta_1))$ subject to the necessary condition for first-period incentive compatibility (E.15) and the firm's participation constraint (A.6) that again turns out to be binding at the optimum.

Denoting by λ the costate variable for (E.15) we can write the Hamiltonian for the principal's problem as:

$$\text{(E.17)}\quad \mathcal{H}(\mathcal{U},q_1,u_1,\lambda,\theta_1)=$$

$$f(\theta_1)\Bigg((1-\beta)(S_1(q_1)-\theta_1 q_1-u_1)-\beta\varphi\left(\frac{\mathcal{U}-(1-\beta)v(u_1)}{\beta},0\right)+\mathbb{E}_{\theta_2}(S_2(q_2^{fb}(\theta_2))-\theta_2 q_2^{fb}(\theta_2))\Bigg)$$

$$-\lambda(1-\beta)q_1 v'(u_1).$$

We shall assume that $\mathcal{H}(\mathcal{U},q_1,u_1,\lambda,\theta_1)$ is concave in $(\mathcal{U},q_1,u_1)$ and use the Pontryagin Principle to get optimality conditions satisfied by an extremal arc $(\mathcal{U}^{sb}(\theta_1), u_1^{sb}(\theta_1), q_1^{sb}(\theta_1))$.

- *Costate variable.* $\lambda(\theta_1)$ is continuous, piecewise continuously differentiable and such that:

$$\text{(E.18)}\quad \dot{\lambda}(\theta_1)v'(u_2^{sb}(\theta_1))=f(\theta_1).$$

- *Transversality condition.* Because (A.6) is binding at the optimum, this condition is:

$$\text{(E.19)}\quad \lambda(\underline{\theta}_1)=0.$$

- *First-order optimality condition w.r.t. u_1:*

$$\text{(E.20)}\quad f(\theta_1)\frac{v'(u_1^{sb}(\theta_1))}{v'(u_2^{sb}(\theta_1))}=f(\theta_1)+\lambda(\theta_1)q_1^{sb}(\theta_1)v''(u_1^{sb}(\theta_1)).$$

- *First-order optimality condition w.r.t. q_1:*

$$\text{(E.21)}\quad S_1'(q_1^{sb}(\theta_1))=\theta_1+\frac{\lambda(\theta_1)}{f(\theta_1)}v'(u_1^{sb}(\theta_1)).$$

From (E.18) and (E.19), $\lambda(\theta_1)$ satisfies:

$$\text{(E.22)}\quad \lambda(\theta_1)=\int_{\underline{\theta}_1}^{\theta_1}\frac{f(\tilde{\theta})}{v'(u_2^{sb}(\tilde{\theta}))}d\tilde{\theta}\geq 0.$$

Inserting into (E.20) immediately gives (E.13). Finally, inserting (E.22) into (E.21) yields (E.14). *Q.E.D.*