Contract as reference points: A new approach to contracting and its implication for relationships among levels of government

Eric Brousseau

Issue 2018/01
Contract as reference points: A new approach to contracting and its implication for relationships among levels of government

Eric Brousseau
University Paris-Dauphine, PSL Université Paris

This paper draws on recent developments in contract theory, initiated by Oliver Hart since 2008 and qualified as Contract as Reference Point Theory (CRPT). I explores how the CRPT can be used to better analyse the contractual relationships among levels of government (or between the government and a private operator). The theory focuses on the issues raised by the ex post adaptation, interpretation and renegotiation of these contracts and points out that a contract cannot control ex post the behaviours of the parties. Henceforth, it might be useless to attempt to design either a sophisticated contractual arrangement that would endeavour to foresee response to any contingency, or a highly flexible contract that would grant authority to one of the parties when future adjustments will be needed because these contracts could be subject to divergence of interpretations, yielding aggrievement (disappointment) and shading (underperformance). Indeed the theory establish a clear distinction between the spirit of a contract and its letter, and highlight that parties can deliver performance only along the letter, while the other party expect a cooperative spirit in ex-post adaption when needed. The theory therefore calls for contracts that are less subject to interpretation and manipulability than the usual recommendations drawn from past theories. Everything being equal, more rigid or shorter term contracts should be preferred. Moreover the conditions in which contracts are negotiated and renegotiated are crucial to guarantee the legitimacy of mutual commitments.
1. The theory and its essential takeaways for the understanding of contracting

A series of recent papers develops the notion that ex ante contracts serve as reference points for ex post trade (see Hart and Moore [2008]; Hart [2009]; Hart and Holmström [2010]). The idea is that an initial contract circumscribes what parties feel entitled to but may not pin down a unique “fair” outcome. In the simplest version, each trading party has a self-serving bias that leads him/her to hope for the most favourable outcome permitted by the contract. In flexible contracts that allow for multiple outcomes, parties to a contract typically have different favourite outcomes so that misaligned reference points emerge. Although the parties will end up compromising, if they feel that they have not received what they are entitled to, they will be aggrieved and engage in counterproductive ex post behaviour (qualified as shading). A rigid contract avoids this situation by limiting the number of outcomes and thereby aligning reference points. The downside of rigidity is that the terms cannot be adjusted to the realised state of the world, which can lead to ex post inefficiency. The theory implies that there is a trade-off between contractual rigidity and flexibility.

To allow the reader to better understand how this approach of contracting sheds new light on contract design and parties’ behaviours, we successively discuss to what extent this approach differs from previous ones: reviewing its main assumptions; discussing its concepts and line of reasoning; and highlighting four key takeaways from the CRPT.

a. The CRPT in the light of alternative contract theories

The theory departs from the previous literature on contracting in that it focuses on parties’ expectations about the outcome of a contract rather than on their incentives to provide performance.

- In the incentive/principal-agent theory, it is assumed that these incentives can be completely designed ex ante so that the (complete) contract monitors precisely both parties’ behaviours in all possible states of the world. The central issue is to deal with the information asymmetries between the parties and, indeed, between them and the entity responsible for the enforcement of the agreement. No ex post adaptation is needed, and the theory has nothing to say about it. If for various reasons a contract is no longer relevant, it should be voided and replaced by a completely new contract. This might lead to compensations in favour of the loser, which is not an issue since the third party responsible for enforcement can access all the relevant information to compute gains and losses as well as identify deviations from the contract and the party liable for deviating.

- The incomplete contract/property right literature focuses on the incentives to invest in non-contractible investments that impact on contractual performance. Parties can, however, design a contract ex ante that gives optimal incentives to both parties to invest (given the sensitivity of the collective result to their respective investments). Once the uncertainty about these investments is resolved, the parties can adapt the conditions of exchange to the situation they are facing.

---

1 An extended version of this paper, with a detailed survey of the literature was prepared for the OECD and is available on its website: Brousseau, E. (2017), “Contract as reference points: A new approach to contracting and its implication for relationships among levels of government”, Background paper prepared for the seminar “Multi-level governance for regional economic development: a focus on flexibility and adaptability”, OECD, Paris, 23 January.
However, there is a hold-up problem: as a consequence of a *de facto* renegotiation of their conditions of exchange, each party shares some of the fruits of its initial investments with the other party. Anticipating this, each party under-invests. In that approach, all the future adaptations are, *in fact*, anticipated at the contract design stage, which makes it of little relevance to analyse *ex post* contractual adaptation and renegotiation. The *ex ante* contract monitors *ex post* behaviours.

- In the transaction cost approach, behaviours are also monitored from the beginning, but either by writing a complete contract when it is feasible or by delegating authority to the parties or to a third party when it is too costly to write and implement a complete contract. Both the costs of contracting and of haggling (*ex ante and ex post*) over the distribution of surplus are the main drivers of the trade-off between complete and incomplete contracting. However, neither the distribution of the surplus, nor its impact on the behaviours of the parties is considered. Also, the appropriate choice of a governance structure *ex ante* (e.g. market vs. hierarchy) solves *a priori* the issue of *ex post* adaptation. Thus, the theory finds it difficult to argue for alternative design of the renegotiation mechanisms of a contract.

b. The main theoretical insights of the Contract as Reference Points Theory

*Retaliations when the outcomes of a contract do not meet the parties’ expectations*

The basis of the CRPT was proposed in Hart and Moore (2008). It explains in what sense a contract provides a reference point for the parties’ trading relationship: more precisely for their feelings of entitlement. Each party to a contract has the discretion to provide “perfunctory” (basic) or “consummate” (exemplary) performance. This is a significant departure from the standard approach of contracting in economics that usually assumes that trade is perfectly enforceable *ex post* (e.g. by a court of law). Hart and Moore assume that only perfunctory performance can be enforced: consummate performance is always discretionary.

What determines whether a party provides consummate performance? The authors appeal to the behavioural economics literature. It is assumed that a party is roughly indifferent between providing perfunctory and consummate performance; consummate performance costs only slightly more or may even be slightly more pleasurable. A party will provide consummate performance if he is “well treated” but not if he is “badly treated” (negative reciprocity). The precise assumption made in Hart and Moore (2008) is that each party feels entitled to the best outcome consistent with the contract and “shades” on consummate performance in proportion to the amount he/she feels aggrieved. A party who is shortchanged thus shades on performance, which causes a deadweight loss.

One way the parties can reduce this deadweight loss is for them to write an *ex ante* contract that pins down future outcomes very precisely and that therefore leaves little room for disagreement and aggrievement. The drawback of such a contract is that it does not allow the parties to adjust the outcome to the state of the world. Hart and Moore (2008) thus study the trade-off between rigidity and flexibility. This analysis provides a basis for long-term contracting in the absence of non-contractible relationship-specific investments.
Shading: Verifiable instead of non-verifiable performance when dissatisfaction occurs

To better understand the fundamental insights, it is useful to relate the ideas developed in Hart and Moore (2008) to incomplete contract theory. A typical model in that literature goes as follows. A buyer and seller meet initially. Because the future is hard to anticipate, they write an incomplete contract. As time passes and uncertainty is resolved, the parties can and do renegotiate their contract, in a Coasian fashion, to generate an ex post efficient outcome. However, as a consequence of this renegotiation, each party shares some of the benefits of prior (non-contractible) relationship-specific investments with the other party. Recognising this, each party underinvests ex ante, which results in a lower/second-best level of performance.

Hart and Moore (2008) depart from the existing literature in two key ways. First, they drop the assumption made in almost all of the literature that ex post trade is perfectly contractible. Instead, they suppose that trade is only partially contractible. Specifically, they distinguish between perfunctory performance and consummate performance, that is, performance within the letter of the contract and performance within the spirit of the contract. Perfunctory performance can be judicially enforced whereas consummate performance cannot.

Second, they introduce some important behavioural elements. They suppose that a party is happy to provide consummate performance if it feels that it is getting what it is entitled to, but will withhold some part of consummate performance if it is shortchanged; they refer to this as “shading.” An important assumption they are making is that a party’s sense of entitlement is determined by the contract they had written. This is the sense in which a contract is a “reference point.”

The benefits (and costs) of contractual rigidity

A flexible contract has the advantage that parties can adjust the outcome to the state of the world, but the disadvantage is that any outcome selected will typically cause at least one party to feel aggrieved and shortchanged, which leads to a loss of surplus from shading. An optimal contract trades off these two effects. This theory explains not only why parties will write somewhat rigid contracts, but also the nature of the rigidity. In short, a contract that sets the most sensitive dimensions of the exchange in advance eliminates ex post arguments (about interpretation of the situation in which the parties are) and aggrievement, and hence both parties will be willing to provide consummate performance, which reduces the related deadweight losses to zero, and may allow achievement of the first-best in some circumstances. Generally speaking, the parties are more likely to put restrictions on variables over which there is an extreme conflict of interest, such as price – i.e. a variable on which there is clearly a zero-sum game – than on variables over which conflict is less extreme, such as the nature or characteristics of the good to be traded.

Allocating authority/decision rights

Hart (2009) relies on the same model to analyse how decision rights should be allocated in the context of a relationship between a buyer (e.g. a party that expects the provision of a given service) and a seller (e.g. a party that would provide the service). He highlights that vertical integration by the buyer B might be good if the production method matters more to B than to the seller S (while outsourcing is preferable if the production method matters
more to S than to B). Indeed, if the production method matters more to B than to S, it is efficient that B chooses it. Moreover, and related, S’s aggrievement will be low because S does not care that much. The model also suggests that “outsourcing” is likely to be efficient when a detailed contract can be written on the nature of the good to be delivered, since in this case B’s value will be pretty insensitive to the choice of production method while S’s cost may not be. In contrast, if a detailed contract is hard to write and B’s value is very sensitive to the details of production, then in-house production may be better.

c. Four essential takeaways of the theory

i. Centrality of the “fundamental transformation”

Contracts as a reference of fairness for the parties

Hart and Moore (2008) argue that it is the combination of ex ante competition and ex post lock-in – what Oliver E. Williamson (1985) calls the “fundamental transformation” – that makes an initial contract a useful and salient reference point. The point is that a contract actually ties the hands of parties ex post as long as it is considered relevant and legitimate. Indeed, the parties can always shade on performance/act non-cooperatively ex post. The relevancy and legitimacy are based on the conditions under which the contract was negotiated. Hart and his co-authors insist in their models and experiments on the idea that a competitive market provides the parties with an anonymous and efficient mechanism to reveal the true values and costs; hence the “legitimacy” of the initial agreement. It is established on the basis of objective, relevant and non-manipulable information; and reflects balances between needs expressed by the demand and capabilities/capacities on the supply side. It is relevant in the sense— that if provides both parties with security of trade in the future.

Contractual commitments as a source of security, with opportunity costs

In addition, the initial contract is useful since it allows the parties to “reserve” capacities/capabilities for the transaction, allowing parties to efficiently perform at the time it is needed. However, this “reservation” aspect of the contract has a cost since ex post the parties are less flexible both between each other, and vis-à-vis all external options in terms of trade. This ex post mutual dependence is the raison d’être of the initial commitment and the cause of the limitations it imposes in ex post adjustment to the new state of the world. This explains why the parties rationally support the ex post expectations on trade within the contract, and care about the mechanisms, if any, to adapt them to new conditions if necessary.

In this analytical framework, therefore, the fundamental transformation is not linked, as in Williamson, to specific investments. It is intrinsic to any contract, and it is the essence of contracting. This leads to the second set of takeaways: the centrality of the legitimacy of processes to establish and revise contracts.

ii. Legitimacy of initial design and ex post adaptations

What is put forward in the analysis is the sensitivity of the parties to the conditions in which a contract is both established and revised.
Fairness at the negotiation stage

As stated above, negotiation among equals, i.e. parties benefiting from outside options and not submitted to liquidity/survival constraints, on a competitive market is one of the way to establish the legitimacy – that is, the relevancy, usefulness and fairness – of the initial agreement. However, there are other potential vectors of legitimacy. Indeed, the issue is to avoid the abuses of dominant position and the biases in matters of information asymmetries, even if for various reasons the conditions of a competitive market are not met. Thus, the procedure of negotiation might compensate. For instance, it is central to guarantee equity among parties. In the case of monopoly or monopsony, the counterparts should be backed-up and supported, for instance, by independent arbitrators supervising the negotiation, by the ability to join forces to rebalance the relationship (while it might lead to a bilateral monopoly triggering collusion against third parties as the final users or citizens), by providing the weak part with assistance and expertise (such as those that could be provided by independent agencies or by consulting firms). In the same spirit, information should be as shared and as complete as possible at the negotiation stage. Typically, for instance, contracting arrangements in similar situations should be made available to the negotiating parties to allow benchmarking. Of course, all possible relevant information to better grasp costs and values, and trends and risks for their future evolutions, should also be collected and shared among the parties. Again, institutional arrangements and procedure can help at this point. Typically, “independent observatories” and systematic mandatory reporting would be useful.

Neutrality of the revision/adaptation process

The second condition for the legitimacy and usefulness of an initial agreement is the condition under which it is revisable. The less manipulable it is, the less open it is to ex post haggling and aggrievement, which both result in deadweight losses and reduced credibility. This is the core of the flexibility-rigidity trade-off. Rigidity makes contractual protection credible, at the cost of poor performance when the environment is volatile. In case of repeated interactions, rigidity is desirable since parties would benefit from middle- to long-term mutual commitment to avoid permanent and costly negotiations. There is also assurance of the availability of resources to perform the transaction (either to secure the reservation of production capabilities or to develop them; i.e. invest). Thus, the implementation of mechanisms aiming at allowing adaptation is requested. However, both the theory and experimental evidences highlight that the more it is based on external non-manipulable factors or independent actors, the less aggrievement there is, and therefore the more legitimacy. The important point here is that it is not the adjustment which matters, but the way it is decided. To put it another way, one party can accept losing value and even a share of the surplus without feeling aggrieved if it is neutrally managed either by an automatic mechanism (as indexation will do) or by an external and neutral arbitrator (as an English trustee). Thus, while ex ante the distribution of cost and benefits should be agreed upon by the parties, the latter should not have a say ex post since any decision may be considered unfair because biased by the lock-in generated by the contract, even when made within the contractual boundaries. Non-manipulability of the mechanism and externality of both the information and of the agents involved matter more than anything else. It is important to point out that experimental results show that when a contract allows one of the parties to decide, aggrievement occurs even if this party adopts co-operative behaviour and takes decisions favourable to the other party.
to send signals of fairness. Fairness is thus not sufficient and independence seems to matter more.

**iii. Monetary vs. qualitative uncertainty**

The theory also puts forward the idea that the best response to the challenges of *ex post* adaptation depends on whether contracting parties face uncertainty about the monetary value of their contribution, i.e. the level of costs and values, or about its quality. In the first case, parties are facing a distribution challenge, in the sense that the way they decide how to split costs and benefits among them impacts upon the generated wealth. In the second case, they face an allocative challenge since the way they decide on economic value and quantity impacts upon the generated welfare. In the second case they manipulate two variables, price and quality, whereas in the first they manipulate only one: price.

**Rigidity as the best response to distribution challenges**

In cases of uncertainty about value and cost only (which means that agreement on quality can be established *ex ante*, not that this is a transaction on a simple good or service), contingent contracts that would allow *ex post* adaptation to uncertainty are implementable if and only if the variables on which contingency is established are not manipulable by the parties and cannot be subjectively interpreted (i.e. there is no ambiguity). Of course, in addition, they should be relevant in the context of the transaction, that is, correlated to the cost and values for the parties. That said, such a contract does not fully guarantee against aggrievement, since aggrievement is determined by the best outcome expected by the parties, independently of the context they face (that is among all the contingencies present in the contract). In concrete terms, it means that the parties must implement a price range that does not allow optimal adaptation, or accept the risk of shading. Taken to its limit, a rigid contract avoids shading. The optimal contract design obviously depends upon the characteristics of the situation. The more uncertainty and the higher impact of shading that exist, the more rigid the contract should be.

**Discretion to adapt when pure allocative challenges**

In case of uncertainty about the quality, the optimal contract is rigid for the variables on which there are only distributive issues/zero-sum game – in the theory the price – but flexible on *ex post* adaptation of quality if there is a significant asymmetry in sensitivity to quality among the parties. Typically, a right to adapt the quality (in the framework of a pre-established set) should be granted to the party to whom it matters more. Such a contract is relevant only if uncertainty is moderate and if the cost of shading is also limited. If the parties are equally sensitive to quality, and if uncertainty is moderate, then a rigid contract is the best option. If uncertainty is high, last-minute negotiations are better but the probability is then high to miss opportunities of trade/wealth creation.

**iv. Property rights channelling co-operation**

Lastly, the CRPT sheds a new light on the role of ownership/property rights on assets. In both the transaction cost literature and the incomplete contract approach, the focus is on the distribution of rights of decision among the parties. The idea is that their primary role is to secure investments when necessary, that is, when investment specific to the relationship (i.e. not re-deployable in alternative trade) has to be made by one of the
parties. In that case, authority should be transferred to the party that could be held up by the other. Both the size of payoffs (in the transaction cost literature) and the sensitivity of payoffs to investments (in the property rights literature) are the drivers of the distribution of property rights understood as the right to decide what the other party should provide (in Transaction Costs Economics) or how to share the surplus (in the Incomplete Contract Theory).

Property of assets to reduce the other party’s propensity to hold-up the surplus...

Hart insists on the idea that the distribution of property rights on assets plays another role: impacting on outside options of players, and therefore on the likelihood they can be held up by the other party. Hold-up happens when parties can unilaterally play on some non-contractible dimensions of their contribution to a given transaction, to benefit from a favourable renegotiation of the condition of exchange. Typically, a party (P) could be held up if the state of the world is very favourable to P ex post much beyond what was expected (for instance P derives much higher value from trade or incurs much lower cost). The other party, then, is incited to extort P’s share of the unexpected rent, which will lead to either underperformance of the held-up party, who is likely to retaliate against the non-cooperative attitude of the opportunistic party (O), or to break-up the relationship, allowing P to exploit its outside options. Incentives to hold-up are therefore bounded by the comparative benefits for the victim of the hold-up to perform the transaction while being held up, and the value from the outside option. In this context, ownership of the assets involved in the transaction (owner will be able to use those assets for alternative transactions) increases the value of her outside option whilst limiting the extraction capability of the other party. Moreover, ownership of assets is a mechanism that indexes the value of the exit option to the level of the rent generated by the state of the world. Hence, when incentives to extort surplus rise as a consequence of an increasing surplus, the level of protection of the potential victim rises as well.

Thus, the distribution of property rights upon assets, whether they are tangible, intangible, or human capital, is a tool to impact upon the parties’ willingness to keep co operating within the contractual relationship and therefore avoid hold-up and shading. Moreover, Hart highlights that the two parties can benefit from attribution of property rights. The logic is to grant the parties with property rights of the assets that are idiosyncratic to them, which means they are relevant for their business, and not for the business of the other part. Thus, the logic is inversed compared to the logic of the manipulation of hostages often found in the literature, and in particular in the transaction cost approach. In the latter, assets ownership should go to the party that could be expropriated by the opportunistic behaviour of the other. It is a dissuasion tool. Here the logic is to enhance the bargaining power of the parties, which might benefit from windfall profit due to a change in the environment, in order to discourage hold-up, as hold-up risks destroying the co-operative spirit between both parties. The manipulation of asset ownership is then understood as a tool aimed at maintaining the incentives of the parties to co-operate, which can lead to distribution of property rights to both parties if they could benefit from windfall profits due to potential changes in the environment.

...that can, however, reduce incentives to invest
In case idiosyncratic investments by one party have to be encouraged, because the collective outcome would be greatly improved by these investments, the will to protect the other party against the potential opportunism of the investor goes against the incentive to stimulate investments. In such cases, it might be relevant to not protect the second party through property right manipulation, even if there is uncertainty on her payoff. In fine, the optimal distribution of property rights is depends upon each party’s payoffs uncertainty given a change in the environment and upon the sensitivity of the collective wealth to each of the parties’ investments and of the respective magnitude of the two effects.

2. Relevance of the CRPT to analyse contracting among levels of government

A quite generic theoretical framework...

We now explore why the approach and its essential takeaways are relevant for understanding the contractual relationship among levels of government. At first sight, indeed, this very abstract vision of contracts applies well to trade among two firms or two individuals. Moreover, the only examples that are referred to by the authors of the various contributions to the theory quote essentially trade between individuals (and most often hiring labour) and the only empirical evidence relies on experiments ran with university students playing games about hiring the service of another party.

... fitting to the analysis of contracting between two mutually dependent parties...

At the same time, the purpose is clearly to shed new light on a principal-agent relationship where a principal (referred in the models as a buyer) contracts to procure a service or a good provided by an agent (referred in the models/experiments as a seller). The theory insists on the fact that both parties’ behaviours impact upon the efficient outcome of the relationship while these behaviours matter for the efficiency of joint action, while there are non-contractible dimensions of these actions and while there are uncertainties about the optimal adaptation of both parties to external conditions that might impact ex post on their respective first best.

...that might be applied to contracts among levels of government

Because it is very generic and abstract, this setting clearly fits with the type of problem of contracting between a higher level of government and a subnational one, especially in the case where the higher level of government is considered as a principal because it contracts with the subnational government for the provision of a given policy (or a component of a given policy) promoted by the higher level. Both sovereign entities can indeed renge ex post on their co-operative spirit, and they can even consider not complying with the letter of the initial contract and prefer to pay damages but stop execution in order to perform alternative policies in case significant changes occurred that lead at least one of the parties to consider dedicating its means and efforts to “outside options” that are more politically/economically relevant given the new circumstances: a major political crisis or a severe economic depression and its consequences in terms of fiscal constraint.

The four specificities of contracts among levels of government
Of course, then, the relationship between two levels of government differs in many ways of the situation dealt with in the canonical models and experiments. First, the two parties are in a bilateral monopoly relationship. Second, it is not an inter-individual relationship, but a relationship between two (complex) organisations. The behavioural assumptions on which the theory is based might not apply. Third, these organisations are driven by political goals rather than by the search for economic efficiency. And so forth. The following sections explain to what extent and under what conditions the CRPT brings useful insight to the analysis of contracts across levels of government:

- Contracts among (levels of) government(s) should be self-enforceable.

- Policy makers care about the spirit of contractual relationships and henceforth are led to retaliate in case they feel the other party is not co-operative.

- The institutions and procedures in the frame of which the contract is established impact on its legitimacy and therefore on its credibility and finally on its ability to guarantee a successful co-operation.

- While flexibility might be needed, it has a cost, and especially in the long run when the search for immediate adjustment destroys the co-operative capability in the long run. Avoiding the negative consequences of contractual flexibility and ambiguity is thus essential for steering the relationship between governments.

a. The logic of self-enforcing agreements

One of the specificities of a contract among levels of government is that it is between two sovereigns, which limits the ability of a third party to enforce them. Of course, administrative and constitutional courts exist, but their actual enforcement capabilities are severely limited; including because their capability to operate is in last resort depending upon the higher level of government. This is, therefore, a typical situation where the credibility of bilateral commitments is based on self-enforcement: the parties should write contracts that they will have interest to comply with ex post without being forced by a third party to do so, even if a third party can be called upon to help the parties to manage their agreements.

**Considering the bounded enforcement capability of third parties**

Self-enforcing agreements are key in several contractual approaches. The point was initially made by Oliver Williamson, who insisted on the idea that, even for commercial transactions, courts are often poor enforcers since they take time to document cases, might make mistakes (that justify appeal systems which increase the duration of a decision) and cannot always have their own decisions enforced because enforcement is costly and hence bounded, and because players have many strategies to bypass the enforcement capabilities of the judicial system (e.g. convicted people/organisations can move outside of the jurisdiction). Thus, even in countries with reasonable public and judicial institutions, economic agents have incentives to design agreements that are (at least partly) self-enforcing. Hart’s approach to contracting clearly fits with such approaches in which the parties contract while anticipating the limited enforcement capabilities of
Courts. Courts, however, play a role since a contract makes sense only when external enforcement is at play.

The limitation of enforcement capabilities is traditionally explained by two main factors: first, the cost of exercising punishments; second, the limited information and cognitive capabilities of third parties, like judges. These two dimensions are discussed below.

**The costs of complying vs. renegotiating vs. breaching**

In the case of relations across levels of governments it is obvious that punishment is limited, not only by cost, but also by constitutional/political boundaries. Sovereigns can be submitted to checks and balances, but the very principle of sovereignty limits *de facto* the range of applicable punishments. In any case, as private contractors consider the costs of enforcement, parties to a contract between governments always consider the benefit/cost to exploit any outside option to the benefit/cost to execute the arrangement as it was originally established, and to the benefit/cost of renegotiating in some way this arrangement. When they design an arrangement, they must anticipate the respective cost of the three options, and eventually manipulate them *ex ante* to make sure that in the future the parties would have interest in sticking to the initial agreement. At least they have to consider the benefits and costs of all alternatives: parties complying with the agreement, parties agreeing on a new agreement (i.e. continuing to co-operate, while renegotiating the conditions), or parties stopping to co-operate (either because one party does not accept the terms of a potentially renegotiated contract or because there is no longer interest on either side in co-operating on the considered issue).

**Cognitive limitations of contract guarantors as the source of aggrievement and shading**

While their enforcement capabilities are bounded, judges and courts are useful because a contract makes sense only if there is a third party to assess the existence of a mutual commitment and to assess whether the parties fulfill their mutual obligations. If such an independent third party does not exist, then contracts are useless since *ex post* this would be always the party with the highest bargaining power – either because it has more valuable outside options or because is more capable of exercising violence as a last resort – that would unilaterally decide what should be done by both parties, making the initial agreement void, and therefore useless *ex ante*. Thus, courts and judges matter because they contribute to the credibility of the mutual commitment. However, it might be the case that either they cannot get all the relevant information to assess whether parties comply with their commitments, or they do not have the right expertise to interpret this information. Of course, judges can request information from the parties, manage investigations and benefit from interpretation by experts, but all these strategies result in costs that limit the actual ability of a third party to oversee the behaviours of contracting parties. Thus, these parties can shade on their contractual obligations, since the third party responsible for assessing compliance will be unable to check some dimensions of their actions. When they write a contract, parties should anticipate this by avoiding contracting on variables that are relevant for their transactions but costly to observe by a third party – qualified in the theory as unverifiable variables. The CRPT relies in particular on the idea that the second limitation of contractual enforcement matters a lot for the understanding of contracts. It explains that the theory focuses on the distinction between the spirit of the contract (that requests consummate/ non verifiable performance) and
the letter of the contract (which implements perfunctory/verifiable performance). Such a distinction allows focusing on situations where (consummate) performance that can never be judicially enforced matters for the performance of a relationship.

Sources of inverifiability in co-operative relationships between levels of government

There are many possible interpretations of the difference between consummate and perfunctory performance, and some of them apply particularly well to contracts among levels of government. First is the idea of the spirit of contracting. Contracts are a way to establish co-operation among parties. This is especially needed when they join their actions in complex projects which are difficult to fully describe ex ante and which several relevant dimensions might be revealed ex post. Mutually beneficial adjustments/development will therefore be requested ex post. The contract design should care about not implementing provisions that could lead to an extinction of the co-operative spirit among the parties. Second, and more straightforward, is the idea that there are relevant dimensions of a complex transaction – and policy implementations are complex issues – that are hardly describable and measurable, then hardly verifiable and contractible. Third is the idea that there might exist sophisticated forms of opportunism by which one of the parties tries to renegotiate ex post the splitting of the burden (or of the benefits) of a transaction for fear of providing verifiable rather than requested level of performance; i.e. work to rule. Typically, this might occur in relations among governments and the theory deals with this type of sophisticated strategy by also considering its effect through the idea of soured relationship: i.e. the non-cooperative compliance to an agreement that results in poor performance. Clearly relations among governments can encompass this type of cynical non-cooperative co-operation.

b. Behavioural assumptions: Aggrievement and shading

Expectations of fairness

This leads to revisit the behavioural assumptions at the root of the theory. The fundamental idea is that parties take their mutual commitment seriously. If they do not get what they expect from the contract, they feel aggrieved and shade/underperform, which results in lower collective outcome/inefficiencies. The important point is that aggrievement can occur even if what the parties get remains within the boundaries of the formal contract. Dissatisfaction is not anchored in any form of irrationality. It is linked to two logics. The first one is that reserving/dedicating capabilities to perform the transaction, and a fortiori investing to manage it, is costly and therefore requests protection against the risk of ex post exit of the relationship by one of the parties due to changing conditions that makes alternative options of trade more profitable. The second one is linked to social norms of fairness. In most societies, even with variations, norms of fairness are deeply anchored in the education and in the culture, in particular to promote co-operation among individuals. Co-operative behaviours are supported and rewarded, while non-cooperation has a bad name and can be punished in many ways, starting by ostracisation (e.g. Binmore [2010]). In any case, individuals tend to punish uncooperative behaviour and this happens in contractual relationships as well as when one party thinks that the other tends to be non-cooperative. This is clearly highlighted by the experimental results obtained by Fehr, Hart and Zehnder [2009, 2011, 2015]. However, to what extent this applies to relationships among governments which could be seen, at first sight, as rational and cold players, who
should therefore act as the homo-oeconomicus of the standard economic/game theory approach and consider that once sunk costs have been lost, it is rational to restart the relationship from scratch and negotiate to optimally adapt to the situation (since past behaviours are not predictors and should not be the driver of future behaviours). There are two levels of responses to that feeling of inadaptation of behavioural assumptions to relationships among public authorities.

Why policy makers might care about contractual terms?

First, it can be argued that aggrievement and relations against uncooperative/unfair behaviours are particularly relevant to characterise the behaviours of governments/political decision makers. The latter are usually involved in long-term repeated interactions among each other and used to the dynamic of coalitions and alliances. It is well known, including by practitioners, that tit-for-tat strategies – i.e. being co-operative and punishing non cooperative behaviour when it arises incidentally, and becoming uncooperative with parties that are repeatedly non-cooperative – are in this context dominant strategies (see Axelrod [1984]). Thus, the propensity to retaliate against non-cooperative behaviours that is observed by behavioural economists among individuals is probably also quite relevant to describe relationships among governments and their leaders. In addition, as pointed out by Spiller and various co-authors (e.g. Spiller [2008], Spiller & Moszoro [1984]), policy makers’ decisions are scrutinised by the public opinion and political opponents. This leads them to highlight the notion of third party opportunism. It describes the idea that there are players in social systems that might be interested in the failure of decision makers, which impacts on contracting behaviours. Political opponents, in particular, check if decision makers deviate from the rules, including the contracts they signed (with private firms, with other governments, etc.) to criticise them if these deviations could seem suboptimal, unfair or simply illegitimate. Anticipating this, political decision makers have strong incentives to stick to the content of the agreement and avoid (even efficient) adaptions/renegotiations in order to limit the risk of losing political support. Thus, both higher levels of government as well as subnational ones could take the contractual expectations seriously and react negatively if they are not reached; which is precisely the behaviour considered by the CRP literature.

Shading as a general approach for the cost of flexibility

Second, as argued by Hart, shading can have different interpretations. In reality, it is a way to take into consideration the cost of flexibility in the various models. Of course, as already argued, flexibility can induce aggrievement and retaliation in the form of costly underperformance. However, flexibility can also result in efforts of persuasion to convince the other that the proposed interpretation of the contract is relevant, or that the intended adaptation is fair, etc. Moreover, flexibility can result in haggling. Also, contractual flexibility can open the door to complex strategies of influence, since decisions have to be taken ex post, necessitating access to information by the decision maker(s). Another perspective is open when one considers rent-seeking behaviours. Especially in the political and administrative spheres, all kinds of interest groups and private interests can attempt to exploit contractual flexibility to benefit from rents.

Thus, the general insight carried by the theoretical approach proposed by Hart that triggers a flexibility-rigidity trade-off seems of particular relevance to analyse contracting among (levels of) government.
c. Institutions and procedures

Relationships among (levels of) government, while among sovereigns, are highly “regulated” by all kinds of procedural rules that de facto bind freedom of contracting and limit contract enforceability. A theory relying on the assumptions that, ex ante, parties are totally free to negotiate a contract (in a competitive context) outside from any type of hierarchical relations could therefore be considered as irrelevant to understand what is going on when governments are negotiating contracts. In addition, a theory that builds models assuming that parties’ feelings about their entitlement are controlled entirely by the contracts they have written could also be considered as a bad candidate to explain relationships among entities whose entitlements are influenced by many other considerations, such as the constitutional order, the balance among political coalitions, the various pressures of the public opinion, etc. Again, there are two “levels” of replies to this type of considerations.

Fairness of the initial contracting conditions...

First, the contracts as reference points approach precisely considers the role of the environment in which the parties interact. In particular, the importance of the conditions in which the initial contract is agreed upon are at the core of its legitimacy and its role as reference points (see above). Moreover, developments consider the role of external reference point (see annex) and highlight how external reference points, as the contracts governing the same type of transactions, might impact on parties’ expectations about the fairness of the contract, the need for adaptation, the possibly more efficient solutions to their transactional or relational challenge. Again, the models are generic enough to apply to the specific case of arrangements among entities contracting in an institutional context. It is simply a question of “calibrating” the models to take into account the specificities of each situation.

...and its translation in terms of contracting between governments

However, and this is the second line of reply, the fact that the theory is quite precise about the conditions in which a (self-enforcing) contract can optimally respond to the co ordination challenges between two parties, allows these models to be interpreted. As pointed out above, when the conditions in which a contract is established and potentially revised are not those of a perfectly competitive market, then the institutional setting should be designed to allow the establishment of legitimate contracts (among levels of government) and to revise them if needed. This triggers recommendations in terms of the optimal setting to establish efficient and efficiently managed contractual agreements. It also allows understanding why in some institutional settings, a contractual agreement would not result into the expected outcome. Take the example of the public debt crisis. Agreement among governments were made in the absence of transparency, of public release of the relevant information, and were characterised by strong asymmetries among the parties around the table, etc. Thus, and independently of the substance of these agreements, most of the conditions for the settlement of a self-enforcing agreement were not met and many parties (as well as the public opinion who might influence them or be instrumentalised) do not trust the agreement and therefore are locked into a stable non cooperative equilibrium.

d. Centrality of the stability/flexibility trade-off
Thus, despite the very abstract nature of this approach, we believe it helps to understand contacts among levels of government and to propose innovative ways to address the stability/flexibility trade-offs specific to this context. Indeed, the objective of the theory is fundamentally to understand the issue of \textit{ex post} adaptation of contractual commitments. The theory highlights that contracts inevitably establish reference points that impact upon parties’ expectations and their behaviours. To a large extent this is exactly the challenge raised nowadays in the EU and in many countries throughout the world because, on the one hand, policies of contracting among levels of governments have been developing everywhere as they were vectors of objectivisation/rationalisation/ modernisation of governance in decentralised contexts and, on the other hand, most countries’ multi-level policies have been strongly impacted by the necessity to deal with major shocks with the financial crisis, then the public debt crisis, not speaking of the consequences of the international disorder of other global issues. The CRP approach focuses to a large extent on the question of optimal contract design and revision given various factors, which include the magnitude of shocks, the volatility of the environment, the degree of divergence of interest among the parties, and points out that in any event a contract, while allowing \textit{ex post} adaptations, including breaking off relations, influences the room for maneuver of the parties.

Also, the theory allows dealing with the various types of contractual agreements that are usually contracted among levels of government. Take the example of the recognised differences between transactional and relational contracts between central and local governments. They correspond to the relationships analysed by the theory as dealing essentially with uncertainty on monetary values and on quality respectively.

\section*{3. Lessons for contractual design}

On the basis of the analysis developed above on the main takeaways of the literature on the contract as reference points, and of the conditions to apply them to the analysis of contracting among level of governments, the following principles can be highlighted:

- Independent agencies might be needed to guarantee the legitimacy and the adaptations of contractual agreements.

- Contracts could be shorter than the expected duration of a project/co-operation. A series of successive contracts can indeed allow mutual commitments to be reset.

- Renewable contracts are a way to reduce the costs of renegotiation when a series of contracts is needed.

- When necessary, renegotiations should always be yielded, and steered by a mechanism independent of the parties.

- Decision rights should be granted to one of the parties when the considered decisions have low distributive effects only. The central government should consider the complementarity/substitutability of the contributions of local governments to a policy to decide whether it is appropriate to be granted with these decision rights.
• Granting property rights on the assets resulting from the co-operative process between levels of governments can be a way to control the propensity to co-operate, and henceforth to avoid shading.

• The cumulative perspective should always be considered when attempting to adapt to new circumstances. Explicitly complying with the spirit of the agreement will facilitate renegotiation and co-operation in the long run.

a. Recognising and establishing the legitimacy of contractual agreements: Toward independent agencies overseeing the contractual process

Contractual agreements can result in harsh constraints for (one of) the parties and quite unequal distribution of the surplus that is generated, as long as they are considered legitimate by the parties. The conditions in which a contract is initially settled are key. The negotiation of the initial contract requests equal footing in bargaining capabilities and as much informational transparency as possible. These elements could result in the establishment of independent third parties – independent agencies or regulators – responsible:

• First, for the collection and publication of information which is relevant for the ex post assessment of contractual performance and for grounding potential indexation (e.g. information about market price of relevant technologies, service provision, etc.). This should come with systematic and mandatory reporting both on contractual arrangement and contractual performance by public authorities, both higher and subnational levels.

• Second, for the supervision of the negotiation process. Negotiation of contracts should be made under the scrutiny of a de facto regulator of contractual practices among levels of government that should have the ability, and even obligation, to publish an opinion on the process of negotiation. This opinion could trigger the provision of assistance to a party that would be in a weak position, especially in terms of information/competence. Of course, such a third party would not alone be able to rebalance the bargaining power between the two entities, but the publication of its opinion would be a way to limit the ability of a dominant party to exploit its dominant position, especially because of the potential impact in terms of reputation of unfair behaviour in the political game and in the public opinion.

In the context of contracts among level of governments, and given its role, it is obvious that such an agency should be independent from the central government. Typically it should be built on the model of independent regulators relied upon to regulate competition and industries, and also the relationship between the government and private operators (as in the case of the provision and exploitation of public infrastructure).

While it might at first appear as a reduction of the sovereignty of the contracting parties, it is a necessary condition to ensure the credibility and the legitimacy of the initial contract, which is itself a necessary condition to guarantee the performativity of that contract, and its “revisability” ex post, if it needs to be adapted to unanticipated shocks or evolutions.
b. Contract duration: Regular renewal rather than long-term commitments

The theory clearly calls for adapting contractual arrangements to uncertainty by shortening their duration. Short-term commitments mean, in fact, frequent renegotiation, and while renegotiations can be harsh and costly, they are preferable to adaptation within the contract that bound the capabilities of adaption – since the parties rely on the contract as a reference point – and this results in shading in addition to poor adaptation to the new state of the play. In case of a high level of uncertainty/volatile environment, long-term commitments are justified only when incentives to invest in specific assets are necessary, while they are difficult to manipulate in the context of shading.

As pointed out below, the drawback of short-term arrangements (i.e. the cost of renegotiation) can be partly mitigated by the conditions in which a contract is renewed.

c. Renewable contracts: Mitigating the cost of renegotiations

A way to mitigate both the rigidity of long-term contracts and the cost of frequent negotiations of short-term contracts is to renew relatively short-term arrangements. The conditions in which the contract is renewed matter, because they impact on the actual flexibility of the initial arrangement that remains the central reference point. Halonen-Akatwijuka and Hart (2015) highlight a continuum in terms of increasing flexibility:

- Starting by long-term agreement.
- Followed by short-term, renewable-for-cause contracts: that is, contracts that are automatically renewed unless one of the parties calls for a non-renewal (in the initial conditions) and justifies his/her claim. The claim should be addressed to both the third party overseeing the performance of the agreement and to the other party; which can then go the court if the motivations for the claim of rupture seem weak/irrelevant.
- Then by renewable contracts: i.e. contracts that continue if both parties agree, while the terms under which the contract is renewed are specified in advance; which means they are either the same or revised automatically according to various possible principles (e.g. indexation, benchmarking, etc.).
- Ending with continuing contracts; that is, contracts that are renegotiated between the parties only on the dimension that the party considered necessary to change. Consequently, the initial contract remains a strong reference point.

These alternative mechanisms provide the contracting parties with tools to subtly/optimally manage the flexibility/rigidity trade-off, by allowing adaptation while controlling the cost of renegotiation and the level of aggrievement. They can therefore adapt to the actual degree of uncertainty of each relationship.

d. Revision: Preventing the manipulability of renegotiation

The complementary strategy to impact upon the flexibility/rigidity trade-off is to allow renegotiation of a rigid contract. Indeed, the theory highlights that in case parties anticipate
the need for future adaptation because there is volatility, neither the multiplication of contingent obligations to adapt to the maximum number of possible circumstances calling for adaptation nor the implementation of unilateral decision rights allowing a party to redefine contractual obligations (of course in a range of contractually agreed upon possibilities) are optimal strategies in the context of aggrievement and shading. Contingent contracts trigger conflicts of interpretation. Unilateral decision by one of the parties raises suspicion by the other party. In such a context, a rigid contract that is revisable can achieve the benefits of flexibility without incurring its costs.

As pointed out above, a first strategy to design a revisable contract is to play on its duration. *De facto*, short-term contracts allow renegotiation while avoiding the counterproductive consequences of contractual flexibility (i.e. shading, low credibility, etc.).

**Automated renegotiation**

Second, the contract can implement a mechanism automating its renegotiation when needed because the contractual terms lead the parties to underperform. The decision to renegotiate should not be in the hands of the contracting parties, since any unilateral decision could be considered as unfair or uncooperative, and a collective decision could lead to costly haggling (especially if parties have a self-serving bias). Indexation is most often a poor solution since it is often difficult to identify a relevant basis for indexation that would not be manipulable by the parties.\footnote{Indexation on observable variables related to the transaction triggers incentive/revelation issues.} However the definition of thresholds and ceilings of some key variables for the performance of the relationship is an efficient way to automatically trigger renegotiation and is a good way to compromise between contract rigidity and flexibility.

**Oversight by independent third parties...**

Third, the contract can rely on a third, independent party either to trigger renegotiation or to take unilateral decisions on contracting obligations when the obligations deviate too significantly from what would be optimal. Here again, the existence of an independent “regulator” could be justified.

...*which might be useless in trust-based societies*

Lastly, the parties can negotiate and implement a formal renegotiation mechanism at their discretion in the contract. The one and only condition for this mechanism to be workable is that the parties should trust each other, either because they have a long experience of co-operation together, or because they are embedded in social networks guaranteeing/favouring co-operative relations. These conditions are very rare in relationships among governments. However, there are clearly differences in terms of the dynamics of political life between trust-based societies and those that are characterised by mistrust.

**e. Decision rights**

The bottom line is that when shading is possible a contract should leave as little discretion as possible to the parties, because a co-operative spirit is necessary *ex post* since many sensitive dimensions of the relationship are not contractible (typically because mutual adjustments will be needed, which are too complex to describe and depend on too many
factors), and when aggrievement is potentially high since the parties are highly sensitive to compliance with provisions agreed ex ante; either because this is a sensitive issue due to its impact either on costs or (economic or political) benefits, or because it is submitted to unfriendly scrutiny by third parties. However, when there are asymmetries on shading and aggrievement, it might be relevant to grant one of the parties with decision rights to optimise adaptation. This might occur when parties do not face a distributive issue (since the zero-sum game structure of the problem is likely to generate aggrievement), but rather allocative issues; that is, when ex post adaptation can greatly impact upon the welfare produced, because improving the quality will help to better meet needs, without a great impact on costs.

*Authority to one of the parties only if adaptations are expected to yield a very significant impact on the outcome*

The driver of the implementation of an authority provision in favour of one of the two parties should clearly be the asymmetry in the matter of aggrievement: when one party cares much more than the other about the quality of the service/good provided, then it must be able to decide the quality *ex post*. It is relevant if and only if the decision on the quality of what is provided does not negatively impact on the overall performance of the transaction. More generally, it should not be linked to a principle of revision of the other dimensions of the deal; typically, the commitments made by the other party.

For instance, if the central government is very sensitive about issues it considers to be relevant from a sovereignty (or more broadly from a political) perspective, but contracts with a local government the implementation of a policy that relies partly on the means of the local government, it should keep control of the way the policy is provided if it requests future adaptation. Let’s take the example of public security; the central government might think it should be exclusively provided by civil servants, while the precise way they will operate will request future adaptation because security crises are difficult to foresee. It should therefore keep control of the way the local security forces will be mobilised in the future. In contrast, if the cost of provision is more important than the way it is produced – e.g. by civil servants vs. private contractors – then the central government should leave a lot of discretion to the local government and only establish objectives in terms of level and cost of provision of a given public policy/goods.

*Considering the complementarity/substitutability among contributions to the public good provision*

To make the right choice, the central government should take into consideration the nature of the policy/public good, as well as the different capabilities of the local governments, since determining their costs and the ability to take the right decision about implementation request both knowledge and administrative capabilities. It might also be important to consider the so-called aggregation technology for the considered public good/policy. The notion of “aggregation technology” was initially introduced by Hirschleifer (1983) and Cornes and Sandler (1984) to highlight that individual (in our case, local) contribution to the provision of a public good (in our case a national policy) does not necessarily impact in the same way for all types of public goods. With summation goods, each unit contributed to the public good adds identically and cumulatively to the overall level of the good available. This is, for instance, the case for contributions to reducing
CO2 emissions. Since individual contributions are substitutes for each other it is essential to get the contribution of the largest possible contributors; i.e. the large polluters in our example. Neither unanimity nor majority are required to ensure minimal provision of the good. This is obviously not the case when dealing with weakest-link public goods, where the smallest contribution fixes the quantity of the public good available for the entire group, as in security control. In such cases, unanimity is required while at the same time each decision maker has a strong incentive to contribute since free-riding could be costly if it is observable by the other providers involved. In case of best-shot public goods – for which the overall level of the good is determined by the largest individual provision, as it is the case for innovation policy where the essential challenge is to make sure that the most likely to innovate are facilitated and incited to do so. Of course there is a wide set of intermediary cases, such as weighted-sum public goods (where different contributions can have different impacts, as in the case of investments in transportation infrastructures since geography makes some regions and cities – typically hubs – more central for the overall impact of the investments); or threshold public goods (where there are local benefits to the provision of a public good, while much lower than the aggregate benefits at the national level, which results in a suboptimal level of provision). Aggregation technology reflects then the sensitivity of a national policy to the quality of the local provision. The central government should certainly be cautious about granting freedom of adaptation to local governments in case of weakest link or threshold public good, while choices are clearly more open in case of summation or best-shot goods.

f. Ownership, knowledge and the propensity to remain co-operative

The CRP approach highlights the role of the “co-operative spirit” in a contractual relationship and the role of the distribution of property rights in maintaining it/enlarging the span of co-operation. In short, parties should be less likely to initiate opportunistic renegotiation – that is, a renegotiation aiming at modifying the distribution of the surplus between the two parties, while the realised distribution remains what could have been anticipated ex ante and is not linked to any substantial difference of provision of inputs by both parties – if the potential victim(s) of the blackmail benefit(s) from favourable exit options. In the case of contracts among level of governments, the exit is certainly not contracting with another partner, but should be understood as the political and economic outcome of discontinuing to co-operate on a given project/policy.

Each government should consider this issue ex ante, since it impacts ex post on both the incentives of each party to engage in opportunistic renegotiation and (in consequence) on the spirit of the relationship. It might be the case that the incentives of both parties are well-aligned, in which case opportunistic actions are very unlikely to occur. However, there are several situations in which the central and the local governments could have diverging interests, especially when the political benefits of their co-operation are unevenly distributed. In such a case, the government that gets more political benefits from the joint policy could be blackmailed by the other ex post, resulting into counterproductive shading.

Granting property rights to the level of government benefiting from the greater political benefits

At the opposite of the common wisdom, the party likely to draw more (political) benefits from the co-operation should be granted with some form of protection ex ante. As the
theory points out, this is where property rights distribution can be useful. In the case of transportation policy, for instance, it might be relevant to consider granting the local government a large share of the property rights on infrastructure for local transportation, beyond their share in the investments, for that purpose. It must be pointed out that the necessity to protect parties against future blackmail is linked to uncertainty about the future benefits one of the parties could unexpectedly draw from the relationship. Even when the economic costs and benefits can be well anticipated, political games often result in unexpected political outcomes; precisely because policy makers play on citizens’ feelings and emotions. Thus, the protection of parties thanks to the granting of property rights over the assets and the outcome derived of their co-ordination seems to be highly relevant in the case of relationships among governments. Since both parties are likely to engage in opportunistic renegotiation \textit{ex post} given the randomness of the attribution of political credit by voters to local or central governments, both should benefit from protection. As pointed out in the theoretical developments, in such a setting, each party should own the assets that are the more idiosyncratic/relevant to its business/policy credit. To follow up on the case of infrastructure, those of “national interest” would be best in the hands of the central government, while those dedicated to local services should be owned by the local government, even if all these infrastructures have a hybrid nature and are jointly funded.

\textit{Preferring investments in redeployable assets}

That said, what the CRPT highlights is that what matters is not the ownership of assets per se, but the level of outside options of the party that could be blackmailed. Other levers can be manipulated to impact upon the cost of breaching the co-operation (redeployability of the means engaged, value of outside options). An important issue here is to consider the nature of the resources that are mobilised/invested by each party. Most non-specialised investments – e.g. office buildings, general purpose administrative labour – are easy to redeploy in alternative policies by contrast with investments in highly specialised (physical and human) capital. Co-operating on policies that do not request investments in highly specialised assets, or choosing to ground a relationship on non-specialised investments while developing the general capability of the (local) government, is a relevant way to initiate co-operation without fearing the development of non-cooperative behaviours.

\textit{The case of non-contractible efforts/investments}

That said, it is not always possible to develop policies on non-specialised/flexible investments and the analytical framework of the CRPT insists on the fact that distributed ownership is not always optimal since enhancing the outside option of one party increases its ability to blackmail the other and might have a negative impact on the incentives to invest in assets specific to a transaction/policy by the other party. Thus, if significant non contractible investments should be made by one of the parties, it is better not to share property rights upon the resulting assets. For instance, despite joint investments in physical infrastructure, if the performance of a local infrastructure is strongly dependent on implementation and maintenance operations that are sensitive to the efforts of the local government (because they depend on local conditions/capabilities), then it is better to incentivise it by granting the local authorities with the ownership of the built infrastructure, even if it is infrastructure that is relevant for the central government’s
political objectives.

g. Considering the long-term and cumulative perspective

The last set of lessons to be drawn from the CRP perspective draws specifically from the behavioural insight of the experiments developed to explore the theory more in-depth. Indeed, Fehr, Hart and Zehnder have shown that parties of a contract are particularly sensitive to the signals of non-cooperation sent by the other party, while they are less sensitive to signals of co-operation. Indeed, in a context of renegotiation, they retaliate strongly against what they consider as opportunistic renegotiations, while they do not respond positively to the non-binding signals of co-operation sent in the context of an informal agreement. This is obviously in line with the perspective of the rigidity/flexibility dilemma resulting from the reference point perspective; i.e. parties’ expectations resulting from a contractual relationship. It is therefore essential to situate each contract in a more dynamic perspective, when relevant. In the case of a contract among levels of government, there are two dynamics to consider.

Repeated contracting to build mutual trust

First is the dynamic of the co-operation among the two governments itself. It is clear from the theory that while a contract may be a reference point, other elements of a relationship or of its environment may serve as reference points too. In particular, past experience may be relied upon by each party to assess each others’ propensity to co-operate. Because contracting among levels of governments is structurally a repeated game, building a mutual experience of co-operation through a co-operative attitude in the management of each contract is potentially of high value since in the long-term it should result in signing less detailed contracts, allowing for more flexibility thanks to a climate of mutual trust and confidence. That perspective should be essential for each party in the daily management of each contract, even if it might lead to non-optimal adaptation in the short run.

Building a win-win atmosphere

Second, the most generic way of envisaging how the perspective of a relationship might influence the propensity to remain co-operative is to take into account the long-term perspective of empowerment by both sides. A local government that understands that a continuing relationship will empower it will lower its propensity to shade. A central government that understands that a continuing relationship will allow it to reinforce its implementation capability in the long run will also reduce its propensity to shade. The issue then is clearly to maintain these perspectives in the long run, and the understanding by both parties of the interests of the other party. Otherwise, in situation of tension (lower value than expected, higher costs incurred, external shocks to absorb), either party could interpret any attempt to renegotiate an agreement as a (tentative) opportunistic renegotiation.

Assessing performance and reporting

In both cases, these elements call for a systematic reporting and publication of the performance of the contractual relationship among levels of governments, including the way adaptations have been managed to allow the parties, but also third parties and the
public opinion, to better understand the general perspective in which these contractual relationships are embedded.

4. Conclusion

*Expectations matter for contractual performance*

The contract as reference points approach highlights that the parties to a contract have expectations about the outcome of a contract, and that these expectations matter more than their objective incentives to provide performance as previous theory did. It impacts on the expected behaviours of parties *ex post* when the contract is operated, and even more so when it is revised. In a sense, it focuses on the very consequence of uncertainty on the outcome of contractual relationships, while previous analyses did assume that contractual provisions can fix *ex ante* most issues raised by *ex post* uncertainty. Hence the focus on the *ex post* dynamics triggered by the *ex ante* arrangement.

*Playing on the spirit of the contract rather than on its letter to retaliate*

The spirit of the analysis is that parties may be disappointed by what they get compared to what they expected (aggrievement) and that they can retaliate by providing less that what they promised (shading), by playing on the spirit and not on the letter of the contract. Consequently, one party (she) might hold-up the other party (he) which, in turn will not get what he expected. The reason why the first party (she) shades is that she interprets the fact that she did not get what she expected because the other party took uncooperative actions. In such context, a rigid contract that would deliver a low outcome for a given party given some unpleasant external conditions would not lead to shading, while a contract that would open the door to adaptation to these unpleasant conditions but based on a decision taken by one of the parties would potentially yield shading.

*Avoiding unilateral decisions to adapt the mutual commitments*

Ex post interventions by one party in the redefinition of the other’s provision should therefore be considered with attention. It can happen both when the contract is performed as it was initially settled or when it is renegotiated because external conditions make it too inadequate to the environment.

*The very specific conditions in which contractual flexibility is to be preferred*

At the stage of the performance of the contract, it is always better to avoid unilateral decisions by one of the two parties and to write contingent contracts based on subjective contingencies. However, if the room for maneuver for shading is low — because it is easy to describe and verify the actual commitments of each party, and because the potential shading does not impact much on the performance of the contract (i.e. has a low cost) — and if one of the parties cares a lot more about a dimension of the performance of the other that has a low distributive impact (i.e. it has a low impact on the other party’s costs), then a contract might grant discretion to the party who cares about, for instance, quality. Thus, flexibility can be implemented in the contract if very specific conditions are met.

*The benefits of rigidity*
aggrievement *ex post* that is difficult to manage because loose contractual obligations bound renegotiation capabilities. Moreover, when a given contract comes to an end or should be renegotiated because it no longer fits with the environment, the negotiation to renew it might not be so costly if there were not any major deviations from the initial contract. There was therefore no conflict, resulting in mutual trust about the co-operative spirit of the other party, which yields a climate of fair repeated negotiations.

*Channeling renegotiations*

At the stage of the revision of a contract, the issue is to avoid the potential the feeling of one party that it is being held up by the other. Thus, triggering renegotiation should as much as possible be based on automatic/external mechanisms to avoid the suspicion that one party is manipulating the contract mechanisms in its favour. Second, once the renegotiation occurs, procedures and the potential intervention of an independent, competent and trusted third party should guarantee that the parties are renegotiating on an equal footing and that no one is taking advantage of a dominant position.

*Short-termism in transactional contracting*

The theory applies both to the case of transactional and relational contracting. When it is easier to describe each party’s provisions and when the horizon of the relations is shorter, then the theory clearly calls for shorter term arrangements, especially in a period of high uncertainty. The frequent renegotiation of contracts at the renewal stage will yield costs that should be inferior to the cost of uncooperative behaviours in longer term arrangements (of course if shading is possible).

*Renegotiations in relational contracting*

When the horizon is longer and the parties’ provision are difficult to forecast *ex ante* because parties are co-operating on a complex evolutionary project, then the focus should be on the conditions in which the objective of the co-operation could be revised. Fundamentally the theory insists on the fact that it cannot be revised unilaterally by one of the parties – e.g. the higher level of government – for fear of destroying the co-operative spirit. This calls for the signature of less flexible relational contracts, implementing, for instance, some rigidities and renegotiation provisions to allow to “objectively/neutrally” end up with a contract and renegotiate a new arrangement from scratch. In the same spirit, shorter term relational contracts (e.g. medium term rather than long term) allow reconsidering the relationship from time to time, permitting adaptions without the fear of raising aggrievement.

*The key role of independent third parties*

Of course, as pointed out by the standard approach to relational contracting, mechanisms of exchange of information among the two parties could also be useful to align their respective visions on the dynamic of the environment. However, the results of experiments managed by behavioural economists highlight that the exchange of information among the parties has a lower impact on mutual trust (and therefore potential aggrievement) than the provision of "external" reference points by independent third parties or the environment; simply because the parties feel that the information in the latter case is less
However, at the stage of contracting, parties must always keep in mind that it could be preferable to sign a more rigid/less adaptable contract to be renegotiated in case it becomes no longer relevant. Too flexible/open a contract could indeed lead to significant aggrievement *ex post* that is difficult to manage because loose contractual obligations bound renegotiation capabilities. Moreover, when a given contract comes to an end or should be renegotiated because it no longer fits with the environment, the negotiation to renew it might not be so costly if there were not any major deviations from the initial contract. There was therefore no conflict, resulting in mutual trust about the co-operative spirit of the other party, which yields a climate of fair repeated negotiations.

*Channeling renegotiations*

At the stage of the revision of a contract, the issue is to avoid the potential the feeling of one party that it is being held up by the other. Thus, triggering renegotiation should as much as possible be based on automatic/external mechanisms to avoid the suspicion that one party is manipulating the contract mechanisms in its favour. Second, once the renegotiation occurs, procedures and the potential intervention of an independent, competent and trusted third party should guarantee that the parties are renegotiating on an equal footing and that no one is taking advantage of a dominant position.

*Short-termism in transactional contracting*

The theory applies both to the case of transactional and relational contracting. When it is easier to describe each party’s provisions and when the horizon of the relations is shorter, then the theory clearly calls for shorter term arrangements, especially in a period of high uncertainty. The frequent renegotiation of contracts at the renewal stage will yield costs that should be inferior to the cost of uncooperative behaviours in longer term arrangements (of course if shading is possible).

*Renegotiations in relational contracting*

When the horizon is longer and the parties’ provision are difficult to forecast *ex ante* because parties are co-operating on a complex evolutionary project, then the focus should be on the conditions in which the objective of the co-operation could be revised.\(^3\)

Fundamentally the theory insists on the fact that it cannot be revised unilaterally by one of the parties – e.g. the higher level of government – for fear of destroying the co-operative spirit. This calls for the signature of less flexible relational contracts, implementing, for instance, some rigidities and renegotiation provisions to allow to “objectively/neutrally” end up with a contract and renegotiate a new arrangement from scratch. In the same spirit, shorter term relational contracts (e.g. medium term rather than long term) allow reconsidering the relationship from time to time, permitting adaptions without the fear of raising aggrievement.

---

3 The standard approach to relational contracting focuses on the mechanism allowing trust and mutual capacities to be built to allow adaptation to the evolution of the project that cannot be fully described/decided *ex ante*. It, however, assumes that the main objective – e.g. building an efficient infrastructure and the capabilities to manage it – is clear and would not change *ex post*. The CRPT insists on the idea that, if in the framework of a relationship between two levels of government the objective is unilaterally changed – e.g. shifting from investments in public equipment project to more immediate social infrastructures – this might result in significant aggrievement which will not be manageable by the contractual mechanisms aimed at allowing the mutual adjustment of means, not of ends.
The key role of independent third parties

Of course, as pointed out by the standard approach to relational contracting, mechanisms of exchange of information among the two parties could also be useful to align their respective visions on the dynamic of the environment. However, the results of experiments managed by behavioural economists highlight that the exchange of information among the parties has a lower impact on mutual trust (and therefore potential aggrievement) than the provision of “external” reference points by independent third parties or the environment; simply because the parties feel that the information in the latter case is less manipulable than in the former case. Thus, the more reliance on independent regulators, on peers to practice benchmarking, on international organisations, on non-governmental organisations, the better the guarantee of efficient _ex post_ adaptations between parties, either in the framework of a contract or at the renegotiation stage.

Deepening knowledge on public contractors’ behaviours

As pointed out in the introduction, however, the theory is still developing and needs to be adapted/calibrated to result in clearer recommendations for policy makers. Its conclusions crucially depend upon the behavioural assumptions in matters of aggrievement (i.e. its magnitude), shading (i.e. the propensity to retaliate and be harmful to the relationship), and also regarding how parties envisage the legitimacy of contractual arrangements and therefore built expectations about their outcome. While this paper highlights that there are good reasons to think that all of these elements matter in the case of relationships among levels of government, the size of the effects and the precise identification of the conditions triggering them still have to be measured. This certainly calls for future research and impact analysis. In the meantime, decision makers should keep in mind this issue of credibility of mutual commitment and how inefficient _ex post_ adaptation mechanisms could destroy it, to the cost of the effectiveness and efficiency of arrangements.
Annex A.

Glossary

**Aggrievement**: Disappointment experienced by a party to a contract when it does not get the expected outcome from the performance of the contract. Aggrievement refers to the loss in expected gains/adaptation to the needs. See also shading.

**Asset specificity**: see specific investment.

**Asymmetric information**: see observability/information symmetry.

**Consummate performance**: see perfunctory vs. consummate performance.

**Contractible/uncontractible variable**: see verifiability/unverifiability.

**Contingent contract**: a contract that specifies what the obligations of both parties are given (different events/state of the world) a change in circumstances/environment.

**Flexible/rigid contract**: a flexible contract allows parties to adapt to future contingencies, whereas a rigid contract specifies ex ante what future provisions should be.

**Hold-up**: characterises in a contractual relationship any action that allows a party to capture part of the surplus in its favour, playing on the fact that the other party might experience higher costs if he does not accept that capture.

**Idiosyncrasy/idiosyncratic asset**: an asset is idiosyncratic to a party if it increases the productivity/economic value of the other assets owned by this party. Idiosyncrasy refers therefore to a party, not to a relationship (see specific investment).

**Observability/information symmetry**: a variable is observable when the two parties to a contract can, in a cost-free way, get the relevant information about it. When information is unobservable to a party, the party can attempt to implement a revelation mechanism to try to fix the issue. A revelation mechanism relies on incentives/conditional payments to reveal the truth. Such mechanisms propose to the informed part a larger share of the surplus if it tells the truth – e.g. about its efforts, its costs, etc. – than if it does not reveal information or lies.

**Perfunctory vs. consummate performance (of contractual obligations)**: this distinction reflects the difference between the verifiable and the non-verifiable (while useful) contractual obligations. The letter of a contract implements perfunctory/verifiable performance, while the spirit of the contract call for consummate/non-verifiable performance that is essential to the successful performance of the relationship (i.e. co-operative behaviour).

**Principal-agent relationship**: describes a relationship where a principal (which can be also referred as a buyer) contracts to get a service or a good provided by an agent (also referred to a seller). [Not to be confused with the principal-agent theory, which refers to a situation in which the principal-agent relationship is characterised by an asymmetry of information where the agent knows variables that are relevant for but unknown by the principal. This is therefore equivalent to the more generic designation of incentive theory]
**Shading**: retaliation exercised by a party that does not get what it expected from the other in the framework of a contract. These retaliations do not, however, breach the letter of the contract. By extension, shading refers to the cost incurred by the parties to prevent retaliation, such as the cost of convincing the other party of the mutual benefits of adjusting to new circumstances. See also aggrievement.

**Specific investment**: an investment is specific to a transaction when it is costly to redeploy the related assets to manage a transaction with another party. Asset specificity refers therefore to a bilateral relationship between two agents.

**Verifiability/unverifiability**: information or a variable relevant for a bilateral relationship is verifiable when it can be observed without any bias or doubt and in a cost-free way by a third party responsible for supervising/reporting on the relationship between the parties. When a variable is unverifiable, parties cannot contract on it, even if both of them can observe it (there is no information asymmetry).
References


Heider, F (1958). *The psychology of interpersonal relations*. New York: Wiley


