

Gretschko, Vitali; Wambach, Achim

**Conference Paper**

## CONTRACT DESIGN WITH LIMITED COMMITMENT

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2017: Alternative Geld- und Finanzarchitekturen - Session: Contract Theory, No. G09-V2

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Gretschko, Vitali; Wambach, Achim (2017) : CONTRACT DESIGN WITH LIMITED COMMITMENT, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2017: Alternative Geld- und Finanzarchitekturen - Session: Contract Theory, No. G09-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/168269>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# CONTRACT DESIGN WITH LIMITED COMMITMENT

VITALI GRETSCHKO AND ACHIM WAMBACH

ZEW MANNHEIM AND UNIVERSITY OF MANNHEIM

*ABSTRACT.* We consider the problem of a principal who wishes to contract with a privately informed agent and is not able to commit to not renegotiating any outcome of any mechanism. That is, we allow the principal, after observing the outcome of a mechanism to renegotiate the outcome without cost by proposing a new mechanism any number of times. This implies that for any outcome of a mechanism to be implementable it must be robust against renegotiation to outcomes that are themselves robust against renegotiation to outcomes that are themselves robust against... We provide a general characterization of renegotiation-proof outcomes. The proposed solution dispenses with the details of the particular mechanism selection strategy and provides an effective and easy-to-use tool to analyze contracting problems with limited commitment. We apply the solution to a setting with a continuous type space, private values and non-linear contracts. We find that the optimal renegotiation-proof outcomes for the principal are pooling outcomes and satisfy a “no-distortion-at-the-bottom” property.

*JEL classification:* C72, C73, C78, D82

*Keywords:* Principal-Agent models, renegotiation, commitment, consistency, one-deviation property, Coase-conjecture

## 1. INTRODUCTION

**Motivation and results.** Consider the problem of a principal who is endowed with all the bargaining power and wishes to contract with a privately informed agent. As a consequence of the revelation principle, we can usually dispense with the details of the particular procedure that the principal may use to close the contract and focus on direct revelation mechanisms (Myerson, 1979). This approach is valid only if the principal honors the rules of the proposed mechanism and the agent trusts that this is the case. By playing the mechanism, however, the agent reveals information and the outcomes of optimal mechanisms are typically inefficient. Both parties therefore can benefit if the outcome is renegotiated, that is, if the principal

---

We would like to thank Geir Asheim, Roman Inderst, Georg Nöldeke, Ray Rees, Klaus Schmidt, Ina Taneva, Thomas Tröger, and the seminar participants at the Edinburgh University for their helpful comments and suggestions. We would also like to gratefully acknowledge the financial support from the German Science Foundation (DFG) through the research unit “Design and Behavior”.

proposes a new mechanism after observing the outcome of the original mechanism. In this case, the agent may decide whether to participate in the new mechanism or whether to stick to the original outcome. If he decides to participate in the new mechanism, the outcome of this new mechanism can still also be subject to renegotiation. If the principal cannot commit to not renegotiating any outcome and there is no deadline that puts an end to the renegotiation, backward induction cannot be applied and the revelation principle is hard to restore.

We adapt and extend the approach Vartiainen (2013) used to analyze auctions without commitment and provide a solution to the contract design problem of a principal who cannot commit to not renegotiate the outcomes of the mechanism. To this end, we capture the mechanism selection behavior of the principal through a selection strategy which identifies for every potential outcome the mechanism which the principal will choose to renegotiate this outcome. An outcome in this context is the contract chosen by the agent and the (updated) belief of the principal about the type of the agent after observing his choice. We show that it is sufficient to focus directly on the set of outcomes that such mechanism selection strategies produce, rather than the strategies themselves. Our solution concept therefore allows us to dispense with the details of the particular mechanism selection strategy of the principal and directly characterize the outcomes that the principal can commit to. Much like the revelation principle, this provides a tool to analyze contracting problems with limited commitment which is both effective and easy to use.

We apply the solution concept to a setting with a continuous type space, private values and non-linear contracts. This setting encompasses many applications in which it is natural for the principal to suffer from the inability to rule out renegotiation like selling when price and quality matter, procurement, or franchising. Without commitment the set of optimal outcomes for the principal has the following features. Firstly, the principal offers a *countably* infinite number of contracts. Secondly, each contract is signed by a *pool* of types of the agent that is of positive measure. Thirdly, the lowest type in each pool receives an efficient contract, every other type in each pool receives an *inefficient* contract. Thus, the result

differs markedly from the full-commitment benchmark. With full commitment, it is optimal for the principal to offer a continuum of contracts, the types fully separate with only the highest type receiving an efficient contract.

The key to finding the solution is to identify which outcomes the principal could observe without being tempted to renegotiate and propose a new mechanism. We define an outcome of a mechanism as the chosen contract of the agent and the resulting belief of the principal. Hence, in order to be able to derive the belief of the principal and talk about outcomes in a meaningful way, we need to define the behavior of the agent first. To this end, we assume that the agent will play the mechanism truthfully, that is, if a mechanism is offered, the agent chooses, possibly mixing if he is indifferent, the message which results in the optimal contract from his point of view. Observe that if the principal can commit to not proposing a new mechanism after observing any outcome, given that the agent plays truthfully, it is indeed optimal for the agent to play the mechanism truthfully. Thus, the problem boils down to what outcomes the principal may observe without being tempted to renegotiate and propose a new mechanism. Whether the principal wants to renegotiate depends above all on the mechanisms that are available to her ex-post, given her ex-post belief and the contract signed in the original mechanism. As at the ex-post stage the principal faces the same commitment problem as at the ex-ante stage, the outcomes that the principal can commit to not renegotiate need to be identified simultaneously.

Similar to Vartiainen (2013), we specify the conditions that are imposed on the mechanism selection strategy of the principal by sequential rationality. We thereby identify the selection strategies that the principal can commit to.<sup>1</sup> In particular, we require that a mechanism selection strategy is *consistent* and satisfies a *one-deviation property*. The former condition implies that employing the mechanism selection strategy ex-ante should not contradict employing it ex-post. The later condition implies optimality: the principal should choose a

---

<sup>1</sup>As Vartiainen (2013) considers a situation with a total lack of commitment, his approach needs to be adapted to our setting in which we focus on renegotiation. With a total lack of commitment the agent does not have the option to retain the contract which was generated by a previous mechanism instead of playing the new mechanism.

mechanism that maximizes her pay-off among the mechanisms which are consistent with the selection strategy.

The resulting selection strategies may be rather complex. To facilitate application, we therefore extend the approach of Vartiainen (2013) and shift our focus from selection strategies to outcomes of mechanisms which are selected by a selection strategy which the principal can commit to. For any given starting point, we identify each mechanism with the set of outcomes it produces assuming that the mechanism is played truthfully by the agent given his type. That is, we identify the mechanisms with the generated contracts and the induced ex-post beliefs of the principal. Every mechanism selection strategy can then be represented by the set of the outcomes produced by the selected mechanisms from any potential starting point. A set of renegotiation-proof outcomes is characterized by two simple conditions – internal and external consistency. Firstly, for every renegotiation-proof outcome there is no set of feasible renegotiation-proof outcomes that would make the principal strictly better off (*internal consistency*). Secondly, it should be feasible to reach a set of renegotiation-proof outcomes from any starting point (*external consistency*). We show that every mechanism selection strategy that the principal can commit to generates a set of outcomes which satisfies internal and external consistency. Moreover, for every set of outcomes that satisfies internal and external consistency there exists a mechanism selection strategy which the principal can commit to and that generates this set.

Both properties reflect sequential rationality of the principal’s mechanism selection strategy: suppose the mechanism proposed by the principal reaches a renegotiation-proof outcome and the principal deviates from her mechanism selection strategy and proposes a new mechanism that would potentially make her and the agent better off. External consistency ensures that if the principal follows her mechanism selection strategy after the deviation, there exists a mechanism that will make her even better off. Internal consistency implies that the resulting outcomes cannot make her better off than the original outcome.

One of the main advantages of a general characterization of renegotiation-proof outcomes is that it provides an effective and easy-to-use tool to analyzing specific instances of the

general problem. We apply the solution concept to a setting with a continuous type space, private values and non-linear contracts. As the principal example of the setting we use that of a seller selling a good to a privately informed buyer. The contracts consist of two dimensions: price and quality.

In contrast to our findings, the optimal mechanism with full commitment yields a full separation of the types of the buyer. That is, each type of buyer receives a different contract. Moreover, apart from the highest type, each buyer will receive a contract with inefficient quality. This is the well-known “no-distortion-at-the-top” result. Clearly this is not sustainable if the principal is not able to commit. If there is full separation of types with inefficient contracts, the principal must know the type of the agent and could propose a new, strictly better, mechanism offering efficient quality to each type.

We start by showing that with renegotiation, renegotiation-proof outcomes must be either efficient and separating or pooling so that one of the types in the pool receives his efficient quality.<sup>2</sup>

Having defined the general structure of renegotiation-proof outcomes, we show that, optimally, the principal offers a mechanism that leads to pooling outcomes only. Moreover, in each pool, the lowest type will receive his efficient quality. That is, there is “no distortion at the bottom”. This is due to the fact that efficient and separating contracts would imply a high information rent to the agent which can be reduced by offering pooling contracts instead. Thus, efficient and separating outcomes cannot be optimal. Moreover, if the principal needs to provide efficient quality to one type in each pool, it is optimal to do so for the lowest type. This allows the principal to reduce the information rent of the higher types in the other pools.

As well as solving the problem of the principal with non-linear contracts, we also demonstrate the cutting power of the solution concept by analyzing the famous Coase (1972) conjecture and rederiving the “gap - no gap” result in a simple way. That is, we analyze

---

<sup>2</sup>Pooling outcomes where one of the agent types receives his efficient quality can be renegotiation-proof if the only other renegotiation-proof outcomes that can be reached, starting from the pooling outcome, are efficient and separating. This is due to the fact that such outcomes do not make the principal strictly better off in comparison to the pooling outcome.

the problem of a seller selling a durable good to a buyer who has private information about his valuation where price is the only relevant dimension. We show that if there is a gap between the cost of the seller and the lowest value of the buyer – the “gap” case – the seller can at most charge a price equal to the lowest valuation of the buyer. If the cost of the seller, however, is above the lowest valuation – the “no-gap” case – the seller can charge the monopoly price (Ausubel and Deneckere, 1989).

Vartiainen (2013) also analyzes the Coase conjecture and writes that in principle it is possible to construct mechanism selection strategies that are consistent and satisfy the one-deviation property for the no-gap case. However, those selection strategies would be complex. Thus, he refrains from construction and makes the additional assumption that mechanism selection strategies need to be stationary. In this case, even in the no-gap case, the seller sells at a price equal to his costs. As we are able to rederive the “gap – no gap” result in a simple way, we demonstrate that shifting the focus from mechanism selection strategies to outcomes simplifies the application of the solution concept.

**Related literature.** Gretschko and Wambach (2016) provide a particular game form for the mechanism design game and show how renegotiation-proof outcomes arise as a perfect Bayesian equilibrium of this game. Furthermore, they apply the solution concept used in this article to different models with a discrete type space. They find that with private values and common values “Spence” type only efficient outcomes are renegotiation-proof. With common values of the “Rothschild-Stiglitz” type, however, inefficient outcomes can be renegotiation-proof. This is different to the results provided in this manuscript, given that with a continuous type space, inefficient outcomes can be renegotiation-proof even in the case of private values.

Asheim and Nilssen (1997) consider a monopolistic insurance market with a finite type-space. They use assumptions regarding the characterization of renegotiation-proof outcomes which resemble our characterization. That is, they rely on properties similar to internal and external consistency to characterize renegotiation-proof outcomes. As in our case, this

approach proves to be very useful in deriving clear results for an otherwise very complex problem.

Neeman and Pavlov (2013) argue that for outcomes of a mechanism to be renegotiation-proof under any renegotiation procedure there must be no Pareto improvements to the outcomes of the mechanism. That is, they take the view that if the mechanism designer is agnostic about the specific renegotiation game that is played after the mechanism, the outcome of the mechanism must be ex-post efficient to survive renegotiation under any renegotiation procedure. The conceptual problem with this approach is that it permits all Pareto-improving outcomes to be obstructive, even if those outcomes are themselves subject to renegotiation. In our approach, we require outcomes to be renegotiation-proof only with respect to outcomes that are themselves renegotiation-proof. In contrast to Neeman and Pavlov (2013), our approach allows for inefficient outcomes.

Bester and Strausz (2001), Hörner and Samuelson (2011), Evans and Reiche (2015), Skreta (2006), and Skreta (2015) limit renegotiation to finite procedures. This approach allows for interesting equilibrium analysis but still leaves the principal with a considerable amount of commitment power. In our frictionless setting, limiting the renegotiation to  $n$  opportunities would allow the principal to implement the full commitment outcome. She could simply pass on  $n - 1$  opportunities and then propose the optimal contracts. Evans and Reiche (2015), for example, assume that after an initial mechanism is played, the principal can offer a new mechanism and the agent may choose whether to retain the outcome of the original mechanism or to participate in the new mechanism. They assume that there is no friction in-between the mechanism proposals, as do we. After the new mechanism is played, the renegotiation is over and there is no scope for further offers from the principal. In this setting, the optimal mechanism from the point of view of the principal is easy to implement if she proposes the null mechanism in the first round and the optimal mechanism in the second round. What makes the analysis of Evans and Reiche (2015) interesting is the fact that they allow a third party whose goals are not aligned with the principal to propose the initial mechanism. This third party must then take into account that the outcome of



the mechanism may be subject to renegotiation. Moreover, their analysis also encompasses situations in which the designer is the principal and she might, as in the hold-up problem, want to propose a mechanism *ex ante* to improve investment incentives. This will not, in general, be the same as the mechanism which is optimal for her once investment is undertaken and the state of the world is realized. At this point the initial purpose of the mechanism is served and the parties will have an incentive to renegotiate the existing contract.

Strulovici (2016) shows that if in a specific infinite-horizon bargaining protocol friction disappears, efficient and fully separating contracts arise in any Perfect Bayesian Equilibrium if values are private and the type space is discrete. This differs from the results provided in this article, given that with a continuous type space inefficient outcomes can be renegotiation-proof even in the case of private values.

In a similar set-up, Maestri (2012) uses a refinement that in any subgame the principal induces the continuation equilibrium that maximizes her payoffs. As in Strulovici (2016), when frictions disappear, only efficient outcomes arise in equilibrium.<sup>3</sup>

The article proceeds as follows. In Section 2, we introduce the general model and the commitment problem of the principal. In Section 3, we derive the solution concept and the optimization problem of the principal. In Section 4, we apply the solution concept to a setting with a continuous type space, private values and non-linear contracts. In Section 5, we demonstrate how our solution concept can be used to retrieve the Coase conjecture and the “gap – no gap” result. Section 6 concludes.

## 2. THE SETUP

**Preferences.** A principal (she) and an agent (he) want to implement a contract  $\omega$  from a metric space of contracts  $\mathcal{W}$ . If a contract  $\omega$  is implemented, the utility of the principal amounts to  $v(\omega)$  where  $v : \mathcal{W} \rightarrow \mathbb{R}$  is a von Neumann-Morgenstern utility function. The utility of the agent is given by  $u(\omega, \theta)$  where  $u : \mathcal{W} \times \Theta \rightarrow \mathbb{R}$  is a von Neumann-Morgenstern utility function and depends on the agent’s type  $\theta$ . The agent’s type is private information to the agent and is drawn from a metric space  $\Theta$ . Endow  $\Theta$  with the Borel  $\sigma$ -algebra and

---

<sup>3</sup>See also Beaudry and Poitevin (1993) for the case in which the agent has all the bargaining power.

denote by  $\Delta(\Theta)$  the set of all probability measures on the Borel  $\sigma$ -algebra. Let  $d(\cdot, \cdot)$  be a metric on  $\Theta$ . We will denote by  $\theta_\epsilon$  the  $\epsilon$ -neighborhood of  $\theta$ . That is, the set of all  $\theta' \in \Theta$  such that  $d(\theta, \theta') < \epsilon$ . The principal's prior about the type of the agent is characterized by  $\mu_0(\cdot) \in \Delta(\Theta)$ . The prior is common knowledge between the agent and the principal. For  $\mu \in \Delta(\Theta)$ , we will denote by  $\text{supp}(\mu)$  the support of  $\mu$ . That is,  $\theta$  is in  $\text{supp}(\mu)$  whenever  $\mu(\theta_\epsilon) > 0$  for all  $\epsilon > 0$ . If no contract is implemented, both parties receive the outside option contract denoted by  $\omega_0 \in \mathcal{W}$ .

**Mechanisms.** To elicit information from the agent and implement a contract the principal uses a mechanism. A mechanism is a tuple  $M = (\mathcal{Z}, \omega(\cdot))$  consisting of a metric space of messages  $\mathcal{Z}$  and a function  $\omega : \mathcal{Z} \rightarrow \mathcal{W}$ . A mechanism works as follows. The agent chooses a message  $z \in \mathcal{Z}$ . When the message  $z$  is sent,  $\omega(z)$  generates a contract.<sup>4</sup> Denote by  $\mathcal{M}$  the set of all mechanisms.

**The problem.** The problem for the principal is that she cannot commit to not renegotiating  $\omega(z)$  after the mechanism has been played. That is, after  $M$  is played and the principal observes  $z$ , she will update her belief about the type of the agent to  $\mu(\cdot : z)$ . We call  $C = (\omega(z), \mu(\cdot : z))$  an outcome of the mechanism and  $\Gamma = \mathcal{W} \times \Delta(\Theta)$  the set of all possible outcomes. After observing the outcome the principal may propose a new mechanism  $M' = (\mathcal{Z}', \omega'(\cdot))$ . The agent can then decide whether he wants to play the new mechanism or whether he wants to hold on to the initially generated contract  $\omega(z)$ . In other words,  $\omega(z)$  is the new outside-option contract of the agent. If the agent decides to play  $M'$ , the principal again observes the message and the contract, updates her belief, and may again renegotiate this outcome by proposing a new mechanism. Overall, the principal is not able to commit to not renegotiating any outcome produced by any mechanism. Whenever a mechanism is played, the principal may propose a new mechanism and the agent may decide to either hold on to his current contract or to participate in the new mechanism. Thus, we are concerned

---

<sup>4</sup>Here it is assumed that communication is direct. For an analysis of contracting with renegotiation and mediated communication, see for example Pollrich (2016) or Strausz (2012).

with the question of what mechanisms the principal can commit to given that any outcome of a mechanism can potentially be renegotiated at the ex-post stage.

### 3. THE SOLUTION CONCEPT

We will start by defining truthful play of the agent given that the principal can commit to not renegotiating the outcome of the proposed mechanism. We then identify conditions under which the principal can indeed commit not to renegotiate the outcome given that the agent plays truthfully.

**Agent's incentives.** In order to study mechanisms that induce truthful play of the agent, suppose that the current outside option of the agent is  $\omega$  and that the principal offers a mechanism  $M = (\mathcal{Z}, \omega(z))$  whose outcomes she will not renegotiate. To simplify notation we will include the outside option  $\omega$  into  $\omega[\mathcal{Z}]$  and a message  $z'$  into  $\mathcal{Z}$  such that  $\omega(z') = \omega$ . Endow  $\mathcal{Z}$  with the Borel  $\sigma$ -algebra and denote by  $\Delta(\mathcal{Z})$  the set of all probability measures on the Borel  $\sigma$ -algebra. Denote by  $\sigma^a(\theta) : \Theta \rightarrow \Delta(\mathcal{Z})$  a (mixed) strategy of the agent.

- (i) *Individual rationality.* If the principal can commit to not renegotiate the mechanism, the agent with type  $\theta$  will participate in the mechanism if there exists a  $z \in \mathcal{Z}$  such that  $u(\omega(z), \theta) \geq u(\omega, \theta)$ , otherwise he will choose the outside option. In particular, by including  $\omega$  into  $\omega[\mathcal{Z}]$  we can assume without loss of generality that for every mechanism of interest there exist for all  $\theta$  a  $z \in \mathcal{Z}$  such that  $u(\omega(z), \theta) \geq u(\omega, \theta)$ .
- (ii) *Incentive compatibility.* An agent with type  $\theta$  will choose  $z \in \mathcal{Z}$  (possibly mixing between messages) such that  $u(\omega(z), \theta) \geq u(\omega(z'), \theta)$  for all  $z' \in \mathcal{Z}$ . In particular,  $u(\omega(z), \theta) > u(\omega(z'), \theta)$  for all  $z' \in \mathcal{Z} / \{z\}$  implies that  $\sigma^a(\theta)[A] = 1$  for all  $\sigma^a(\theta)$ -measurable  $A \subset \mathcal{Z}$  such that  $z \in A$ .

Truthful play by the agent will only be induced if the principal will not renegotiate any outcome of the mechanism. Thus, we turn our attention to mechanisms for which the principal is able to commit to not renegotiating its outcomes.

**Principal's incentives.** We start by stating the payoff of the principal from using mechanism  $M = (\mathcal{Z}, \omega(z))$  given that the agent plays truthfully. Denote by  $\nu^M \in \Delta(\mathcal{Z})$  the

measure on  $\mathcal{Z}$  that is induced by the strategy  $\sigma^a(\theta)$  of the agent and the belief of the principal  $\mu$ . That is, for all  $A$  in the Borel  $\sigma$ -algebra on  $\mathcal{Z}$  it must hold  $\nu^M(A) = \int_{\Theta} \sigma^a(\theta) [A] d\mu$ . The payoff of the principal can then be written as

$$V(M) = \int_{\mathcal{Z}} v(\omega(z)) d\nu^M.$$

After the mechanism is played, the principal updates her belief to  $\mu(\cdot : z)$ . In particular, incentive compatibility implies that for every  $\theta$  in  $\text{supp}(\mu)$ , whenever there exists a  $z$  and  $z'$  such that  $u(\omega(z), \theta) > u(\omega(z'), \theta)$  it must hold that  $\theta \notin \text{supp}(\mu(\cdot : z'))$ . If the mechanism is played truthfully by the agent, the principal should not believe that the agent chose a message  $z'$  if by choosing message  $z$  he would be strictly better off. Moreover, the belief of the principal should be consistent with Bayes' rule. That is, for all  $\theta$  and all  $\epsilon > 0$  it must hold

$$\int_{\mathcal{Z}} \mu(\theta_{\epsilon} : z) d\nu^M = \mu(\theta_{\epsilon}).$$

**Mechanism selection strategies.** In order to identify which outcomes will not be renegotiated by the principal, we define the seller's mechanism selection strategy as

$$\sigma^p : \mathcal{W} \times \Delta(\Theta) \rightarrow \mathcal{M}.$$

The selection strategy specifies for each combination of principal's belief and outside option  $(\omega, \mu)$  the mechanism  $\sigma^p(\omega, \mu) \in \mathcal{M}$  that the principal will choose. Thus, for every possible outcome a selection strategy specifies whether the principal will propose a new mechanism or choose not to renegotiate. Not renegotiating is equivalent to proposing  $M = (\mathcal{Z}, \omega(z) \equiv \omega)$ .<sup>5</sup> For a given mechanism  $M = (\mathcal{Z}, \omega(\cdot))$  and a given outcome  $C$  we will denote by  $f(M, C, \sigma^a) \subset \Gamma$  the set of outcomes generated by this mechanism starting from  $C$ . That is, an outcome  $C' = (\omega', \mu')$  is in  $f(M, C, \sigma^a)$  if – given the strategy  $\sigma^a$  of the agent – there exists a message  $z \in \mathcal{Z}$  that is chosen by some type of the agent with positive probability such that  $\omega(z) = \omega'$  and  $\mu(\cdot : z) = \mu'$ .

---

<sup>5</sup>To simplify the exposition we will call the initial outside option and belief  $(\omega_0, \mu_0)$  also an outcome.

We will now adapt the approach by Vartiainen (2013) and identify the criteria a selection strategy should satisfy. He argues that sequential rationality of the principal, and the agent's knowledge of this, requires that  $\sigma$  reflects consistency and optimization. To be able to define these conditions, we first need to develop some concepts.

Suppose the principal can commit to following his mechanism selection strategy. We say a principal can commit to a mechanism  $M$  today given that  $\sigma^p$  is followed in the future if for all  $C' = (\omega', \mu') \in f(M, C, \sigma^a)$  it holds

$$v(\omega) \geq V(\sigma^p(C')).$$

That is, after every possible ex-post outcome of mechanism  $M$ , the selection strategy of the principal would prescribe to implementing a mechanism with a weakly lower expected surplus. Thus, if the principal follows his selection strategy, no outcome of  $M$  would be renegotiated ex post. Such a mechanism would be truthfully played by the agent given the mechanism selection strategy of the principal.

We denote by  $\mathcal{M}^{\sigma^p}(C)$  the set of all mechanisms that are truthfully playable given that the principal follows her mechanism selection strategy  $\sigma^p$ . That is,

$$\mathcal{M}^{\sigma^p}(C) = \{M \in \mathcal{M} : v(\omega) \geq V(\sigma^p(C')) \text{ for all } C' = (\omega', \mu') \in f(M, C, \sigma^a)\}.$$

We are now in the position to formally specify conditions that sequential rationality imposes on the mechanism selection strategy. The first condition requires consistency. That is, employing  $\sigma^p$  ex ante should not contradict employing  $\sigma^p$  ex post. The second condition implies optimality. Given  $\sigma^p$  and  $C$ , the principal should choose a mechanism that maximizes her payoff in the set  $\mathcal{M}^{\sigma^p}(C)$ . In particular, under the hypothesis that  $\sigma$  can be committed to in the future, the principal does not want to change  $\sigma$  for any  $C$ .

**Definition 1.** A mechanism selection strategy  $\sigma^p$

- (i) is *consistent* if  $\sigma^p(C) \in \mathcal{M}^{\sigma^p}(C)$  for all  $C \in \Gamma$ .

- (ii) satisfies the *one-deviation property* if  $V(\sigma^p(C)) \geq V(M)$  for all  $M \in \mathcal{M}^{\sigma^p}(C)$  for all  $C \in \Gamma$ .

A thorough discussion of this approach can be found in Vartiainen (2013). He summarizes why Definition 1 implies that the principal can commit to the selection strategies that are consistent with it as follows. Note that the one-deviation property together with consistency reflects optimization. This is due to the fact that a selection strategy  $\sigma^p$  with these properties maximizes the principal's payoff in  $\mathcal{M}^{\sigma^p}$ . That all outcomes of  $\sigma^p(C)$  are in  $\mathcal{M}^{\sigma^p}(C)$  for every  $C$  guarantees that this act of optimization is consistent with foresight. That is, since  $\sigma^p$  is obeyed in the future,  $\sigma^p \in \mathcal{M}^{\sigma^p}$  guarantees that  $\sigma^p$  will not be redesigned and thus can be committed to.

Mechanism selection strategies can still be quite complex. Thus, to facilitate application, we turn our attention to the outcomes that result from the selection strategies.

**Identifying mechanism selection strategies with outcomes.** Each mechanism selection strategy of the principal together with the strategy of the agent induces a set of possible outcomes  $\Omega = \bigcup_{C \in \Gamma} f(\sigma^p(C), C, \sigma^a) \subset \Gamma$ . Instead of verifying directly whether a particular selection strategy  $\sigma^p$  is consistent and satisfies the one-deviation property (Definition 1), we will identify all selection strategies with the set of outcomes that they generate. We then define conditions on the set of outcomes which characterize selection strategies that the seller can commit to, that is, selection strategies that are consistent and satisfy the one-deviation property.

We start the analysis by defining which outcomes can be induced by truthfully played mechanisms which make the principal better off, starting from some outcome  $C$ . That is, we define conditions on a given set of outcomes  $\{C(z) = (\omega(z), \mu(\cdot, z)) : z \in \mathcal{Z}\}$  such that there exists a mechanism  $M = (\mathcal{Z}, \omega(\cdot))$  that, if played truthfully, generates this set of outcomes and makes the principal weakly better off as compared to  $C$ . These conditions are directly derived from the agent's and principal's incentives discussed above.

**Definition 2** (*Feasibility*). Let  $\mathcal{Z}$  be a metric space. We call a set of outcomes

$$\{C(z) = (\omega(z), \mu(\cdot, z)) : z \in \mathcal{Z}\}$$

feasible starting from  $C = (\omega, \mu)$  if the following conditions are satisfied.

- (i) (*Individual rationality of the agent*) For all  $\theta \in \text{supp}(\mu)$  there exists a  $z \in \mathcal{Z}$  such that  $u(\omega(z), \theta) \geq u(\omega, \theta)$ .
- (ii) (*Incentive compatibility*) For any  $\theta \in \text{supp}(\mu)$ , if there exists a  $z$  and a  $z'$  in  $\mathcal{Z}$  such that  $u(\omega(z), \theta) > u(\omega(z'), \theta)$  then  $\theta \notin \text{supp}(\mu(\cdot : z'))$ .
- (iii) (*Bayesian consistency*) There exists a probability measure  $\nu \in \Delta(\mathcal{Z})$  such that

$$\int_{\mathcal{Z}} \mu(\theta_\epsilon : z) d\nu = \mu(\theta_\epsilon).$$

- (iv) (*Individual rationality of the principal*) For the same probability measure  $\nu$  it holds

$$v(\omega) \leq \int_{\mathcal{Z}} v(\omega(z)) d\nu.$$

We define by  $X(C)$  the mapping from some  $C$  to all feasible outcomes starting from  $C$ .  $\{C(z) : z \in \mathcal{Z}\}$  is an element of  $X(C)$  if  $\{C(z) : z \in \mathcal{Z}\}$  satisfies the conditions of Definition 2. That is,  $X(C)$  is the set of sets of outcomes that is generated by all mechanisms that – if played truthfully – do not make the principal and the agent worse off than  $C$ . For any  $C \in \Gamma$ ,  $X(C)$  is not empty as  $\{C\} \in X(C)$ .

We turn our attention to the set of outcomes  $\Omega = \bigcup_{C \in \Gamma} f(\sigma^p(C), C, \sigma^a) \subset \Gamma$  which is generated by a mechanism selection strategy  $\sigma^p(C)$  of the principal that satisfies the conditions of Definition 1, that is, a mechanism selection strategy that satisfies consistency and the one-deviation properties. If  $\Omega$  is generated by such a selection strategy, we will call outcomes in  $\Omega$  renegotiation-proof outcomes.

To reflect the consistency and one-deviation property of  $\sigma^p$ , a set of renegotiation-proof outcomes should have two properties. Firstly, for every state  $C$  it should be feasible to reach a renegotiation-proof outcome. This is just to ensure that a given set of outcomes  $\Omega$  is generated by a well-defined mechanism selection strategy  $\sigma^p$ . Secondly, if an outcome  $C$  is

in  $\Omega$  there should be no feasible set of outcomes that are all in  $\Omega$  and make the principal better off. This ensures that if outcome  $C$  in  $\Omega$  is reached, proposing  $\sigma^p(C)$  would not make the principal better off. We summarize this in the following definition.

**Definition 3** (Renegotiation proofness).  $\Omega \subset \Gamma$  is a set of *renegotiation-proof outcomes* if the following holds true.

- (i) (*External consistency*) If  $C$  is not in  $\Omega$ , there exists  $\{C(z) : z \in \mathcal{Z}\} \in X(C)$  such that  $\{C(z) : z \in \mathcal{Z}\} \subset \Omega$ .
- (ii) (*Internal consistency*) For all  $C = (\omega, \mu)$  in  $\Omega$  and for all  $\{C(z) : z \in \mathcal{Z}\} \in X(C)$  such that  $\{C(z) : z \in \mathcal{Z}\} \subset \Omega$  it holds that  $v(\omega) = \int_{\mathcal{Z}} v(\omega(z))d\nu$ .

We are now in a position to identify mechanism selection strategies that the principal can commit to with their outcomes.

**Proposition 1.** *The following holds true*

- (i) *Let  $\sigma^p$  be a mechanism selection strategy that is consistent and satisfies the one-deviation property. In this case,  $\bigcup_{C \in \Gamma} f(\sigma^p(C), C, \sigma^a)$  is internally and externally consistent.*
- (ii) *For every  $\Omega$  that is internally and externally consistent such that for all  $C \in \Gamma$  the solution to*

$$\begin{aligned} \max_{\{C(z)\}} \quad & \int_{\mathcal{Z}} v(\omega(z))d\nu \\ \text{s.t.} \quad & \{C(z)\} = (\omega(z), \mu(\cdot : z) : z \in \mathcal{Z}) \in X(C) \\ & \{C(z)\} = (\omega(z), \mu(\cdot : z) : z \in \mathcal{Z}) \subset \Omega. \end{aligned}$$

*exists, there is a mechanism selection strategy of the principal  $\sigma^p$  that is consistent and satisfies the one deviation property such that  $\bigcup_{C \in \Gamma} f(\sigma^p(C), C, \sigma^a) \subset \Omega$ .*<sup>6</sup>

*Proof.* The proof is relegated to the appendix. □

---

<sup>6</sup>Whether for a given  $\Omega$  the solution to the maximization problem exists, can be verified directly for each setting at hand. In our applications in Sections 4 and 5 this will be straightforward.



It is crucial to understand that the restrictions we place on the set of renegotiation-proof outcomes reflect sequential rationality. Thus, the solution concept is not meant to provide an alternative for the standard equilibrium techniques. Rather, it is consistent with them. Indeed, in a slightly simpler setting, Gretschko and Wambach (2016) explicitly construct an extensive form game in which the seller, at each round, designs a mechanism and then chooses whether to implement the outcome of the mechanism after it was played by the agent, or whether to design a new mechanism. They show that for every set of renegotiation-proof outcomes there exists a perfect Bayesian equilibrium of the proposed game that implements the optimal outcome in the set.

To see how renegotiation-proofness implies sequential rationality, suppose the principal – instead of implementing a renegotiation-proof outcome – deviates and proposes a new mechanism  $M$ . By external consistency, the outcomes of the new mechanism can be further improved by renegotiation-proof outcomes. Thus, the principal benefits from proposing a new mechanism  $M'$  that leads to renegotiation-proof outcomes. It follows from internal consistency that proposing  $M$  could not have been profitable in the first place.

The same argument applies to any finite deviation by the principal. This is due to the fact that starting from some outcome  $C$ , any finite sequence of proposed mechanisms can be interpreted as just one feasible mechanism. Following the same argument as above, deviation to this mechanism cannot be profitable. In terms of infinite long deviations, Gretschko and Wambach (2016) show that, with the appropriate assumptions regarding the payoffs of infinite terminal histories, every infinite deviation can be improved by a finite one.<sup>7</sup>

We are now in the position to derive the principal's optimization problem.

---

<sup>7</sup>The definition of the set of renegotiation-proof outcomes is related to the concept of weakly renegotiation proof equilibrium as proposed by Farrell and Maskin (1989). An equilibrium of an infinitely repeated game is called weakly renegotiation proof if equilibrium payoffs of different subgames cannot be strictly Pareto ranked. Following a similar logic, internal consistency ensures that payoffs of different feasible outcomes that are in  $\Omega$  cannot make the principal strictly better off without leaving the agent strictly worse off.

**Corollary 1** (Principal's problem). *Denote by  $C_0 = (\omega_0, \mu_0)$  the initial state and by  $\Omega$  some set of renegotiation-proof outcomes. The principal's optimization problem can be written as*

$$\begin{aligned} & \max_{\Omega, \{C(z)\}} \int_{\mathcal{Z}} v(\omega(z)) d\nu \\ & \text{s.t.} \quad \{C(z) = (\omega(z), \mu_0(\cdot : z)) : z \in \mathcal{Z}\} \in X(C_0) \\ & \quad \{C(z) = (\omega(z), \mu_0(\cdot : z)) : z \in \mathcal{Z}\} \subset \Omega. \end{aligned}$$

**Some useful results.** One of the main advantages of identifying mechanism selection strategies with the set of outcomes is that in order to apply our solution we can abstract from the details of the selection strategy. We simply need to construct sets  $\Omega$  of renegotiation-proof outcomes. That is, we need to construct sets  $\Omega$  that are internally and externally consistent. Once we have constructed such a set, Proposition 1 allows us to retrieve a mechanism selection strategy which implements the set of outcomes. Before we turn our attention to specific applications of the solution concept, we will state the following two results that facilitate the construction of renegotiation-proof  $\Omega$ .

**Lemma 1.** *If  $X(C) = \{C\}$ , then  $C$  is in any renegotiation-proof  $\Omega$  which satisfies the conditions of Definition 3.*

*Proof.* Follows directly from external consistency. If  $\{C\}$  is the only element of  $X(C)$ , then  $C$  must be in any  $\Omega$ .  $\square$

In particular, Lemma 1 implies that any Pareto-efficient allocation must be in  $\Omega$ .

**Lemma 2.** *Let  $\Omega$  satisfy the conditions of Definition 3. For any  $C = (\omega, \mu)$  if there exists  $\{C' = (\omega', \mu)\}$  in  $X(C)$  such that  $v(\omega') > v(\omega)$ , then  $C$  is not in  $\Omega$ .*

*Proof.* Suppose to the contrary there exist  $C \in \Omega$  and a  $C' = (\omega', \mu)$  such that  $\{C'\} \in X(C)$  and  $v(\omega') > v(\omega)$ . In this case, internal consistency implies that  $C'$  is not in  $\Omega$ . External consistency implies that there exists  $\{C(z) : z \in \mathcal{Z}\} \in X(C')$  with  $\{C(z) : z \in \mathcal{Z}\} \subset \Omega$ . This implies that  $\int_{\mathcal{Z}} v(\omega(z)) d\nu \geq v(\omega')$ . As  $\{C(z) : z \in \mathcal{Z}\}$  is feasible starting from  $C'$ ,  $\{C(z) : z \in \mathcal{Z}\}$  is also feasible starting from  $C$ . Together with  $v(\omega') > v(\omega)$  this violates

internal consistency and it follows that neither  $C \notin \Omega$  nor  $\Omega$  is a set of renegotiation-proof outcomes.  $\square$

Lemma 2 has an intuitive interpretation. For any potential outcome, if there exists a single contract that would be accepted by the agent independent of his type and makes the principal strictly better off, then the initial outcome cannot be renegotiation-proof. If this were the case, the principal could simply offer a mechanism in which the same contract could be implemented for any message from the agent. Such a mechanism would be played truthfully by the agent since it does not reveal any additional information and would make the principal strictly better off. Thus, the original outcome would be renegotiated.

#### 4. DESIGN OF NON-LINEAR CONTRACTS WITH LIMITED COMMITMENT

We now turn to the main application of our solution concept. We will proceed as follows. Firstly, we will set up the model. Secondly, we will introduce applications of the model. Thirdly, we will state the main result and briefly describe the main intuition. Fourthly, we will introduce a simple example to illustrate the main result. Finally, we will provide a proof of the main result.

**Set up.** Consider a principal who wants to implement a two-dimensional contract  $\omega = (p, q)$  with  $q \in \mathbb{R}_+$  and  $p \in \mathbb{R}$ . If a contract  $(p, q)$  is implemented, the utility of the principal is given by

$$v(\omega) = p - c(q).$$

Denote by  $c_q(\cdot)$  the derivative of  $c(\cdot)$  with respect to  $q$  and by  $c_{qq}(\cdot)$  the second derivative of  $c(\cdot)$  with respect to  $q$ . Assume that  $c_q(\cdot) > 0$  and  $c_{qq}(\cdot) > 0$ .

The utility of the agent is given by

$$\bar{u}(\omega, \theta) = u(q, \theta) - p.$$

The type  $\theta$  of the agent is taken from  $\Theta = [\underline{\theta}, \bar{\theta}]$ . Denote by  $u_q$  the derivative of  $u$  with respect to  $q$  and by  $u_{qq}$  the second derivative of  $u$  with respect to  $q$ . Similarly, denote by

$u_\theta$  the derivative of  $u$  with respect to  $\theta$  and by  $u_{q\theta}$  the cross-derivative of  $u$  with respect to  $q$  and  $\theta$ . Assume that  $u_q > 0$  and  $u_{qq} \leq 0$  and that  $u$  satisfies single crossing. That is,  $u_\theta > 0$  and  $u_{q\theta} > 0$ , a larger type receives larger utility and larger marginal utility from a given  $q$ . The principal's prior about the agent's type is given by  $\mu_0 \in \Delta(\Theta)$ . Assume that  $\text{supp}(\mu_0) = \Theta$ . The initial contract  $\omega_0$  is  $(0, 0)$ .

**Applications.** The initial model can be interpreted to fit, but is not limited to, the following applications.<sup>8</sup>

- (i) *Selling when price and quality matter.* The principal is a seller that sells a good to the agent, the buyer. The contract  $(q, p)$  specifies the quality  $q$  and the price  $p$  of the good. The seller incurs a cost of  $c(q)$  when producing a good of quality  $q$ . A buyer of type  $\theta$  enjoys a utility of  $u(q, \theta)$  when consuming a good of quality  $q$ . Higher types of buyer enjoy a higher utility and a higher marginal utility from consuming the good.
- (ii) *Procurement.* The principal is a buyer who procures a good from the agent, the seller. The contract  $(p, q)$  specifies the quantity  $q$  and the price  $-p$  of the good. The buyer derives a utility of  $-c(q)$  when procuring a quantity  $q$  of the good. A seller of type  $\theta$  incurs a cost of  $-u(q, \theta)$  when producing a quantity  $q$  of the good. Higher types of seller enjoy a lower cost of production and a lower marginal cost of production.
- (iii) *Franchising.* The principal is a manufacturer who produces a quantity  $q$  of a good at cost  $c(q)$  and sells the good to a retailer, the agent, at price  $p$ . The retailer faces a demand of  $D(t, \theta)$  for the good with  $t$  denoting the resale price and  $\theta$  the demand shock that is private knowledge to the retailer. Higher  $\theta$  imply a higher demand and a higher marginal demand. The profit for the retailer from selling the good is  $u(q, \theta) = tD(t, \theta)$ .
- (iv) *Labor contracts.* The principal is a potential employer and the agent is a worker. The contract  $(p, q)$  specifies the effort  $q$  of the agent and his wage  $-p$ . The principal derives a utility of  $-c(q)$  from the effort of the agent. An agent with productivity  $\theta$

---

<sup>8</sup>Applications partially taken from Gretschko and Wambach (2016) and Strulovici (2016).

incurs a cost of effort of  $-u(q, \theta)$ . Higher types of agent enjoy a lower cost of effort and a lower marginal cost of effort.

In what follows we will adopt the “*selling when price and quality matter*” interpretation of the model and call  $q$  the quality of the good and  $p$  the price.

**Useful properties of the model.** Before we turn our attention to the main result, we first state some useful definitions for and properties of the considered model. Denote by  $q^*(\theta)$  the efficient quality for a given type  $\theta$ . The efficient quality is implicitly given by

$$(1) \quad -v_q(q^*(\theta)) = u_q(q^*(\theta)).$$

Given the assumptions we made about  $v$  and  $u$ ,  $q^*(\theta)$  is unique and satisfies

$$q_\theta^*(\theta) = \frac{u_{q\theta}(q^*(\theta), \theta)}{v_{qq}(q^*(\theta)) - u_{qq}(q^*(\theta), \theta)} > 0.$$

**Definition 4.** We will call an outcome

- (i)  $C = ((p, q), \mu_\theta)$  *separating* if for all measurable sets  $A$ ,  $\mu_\theta(A) = 1$  whenever  $\theta \in A$  and  $\mu_\theta(A) = 0$  otherwise. Efficient and separating outcomes are denoted by  $C = ((p, q^*(\theta), \mu_\theta)$ ,
- (ii)  $C = ((p, q), \mu)$  *pooling* if it holds  $\mu_0(\text{supp}(\mu)) > 0$ .

**Lemma 3.** *If the principal can commit to any mechanism, there exists a direct and incentive compatible mechanism that implements the efficient quality. That is, there exists a price function  $p(\theta)$  such that for all types  $\theta$ ,*

$$u(q^*(\theta), \theta) - p(\theta) \geq u(q^*(\hat{\theta}), \theta) - p(\hat{\theta})$$

*for all  $\hat{\theta} \in \Theta$ . This holds true for all  $p(\theta)$  such that*

$$(2) \quad p_\theta(\theta) = u_q(q^*(\theta), \theta)q^*(\theta) > 0.$$

*Proof.* See Fudenberg and Tirole (1990) Theorem 7.3 for a proof. □

**Lemma 4.** *If every type obtains his efficient quality and the prices satisfy (2), then the principal is indifferent between all contracts; she obtains the same profit from all types. That is,  $v(q^*(\theta)) - p(\theta) = k$  for some constant  $k$ .*

*Proof.* This is a consequence of  $p_\theta(\theta) - v_q(q^*(\theta))q_\theta^*(\theta) = (u_q q^*(\theta), \theta) - v_q(q^*(\theta))q_\theta^*(\theta) = 0$ . Due to equation (1) and equation (2).  $\square$

**The main result.** Our main result is that the optimal set of outcomes  $\{C(z)\}_{z \in \mathbb{N}}$  that will not be renegotiated by the seller has the following properties. Firstly, all outcomes that are achieved with positive probability are pooling outcomes. Secondly, for each of these pooling outcomes the lowest type choosing this contract obtains his efficient quality; all other types receive an inefficient quality. Thus, the principal's maximization problem is the following.

**Theorem 1.** *The profit maximizing set of renegotiation-proof outcomes  $\{C(z)\}_{z \in \mathbb{N}}$  is countably infinite and has the following properties*

- (i) (Pooling) All  $C(z) = ((p, q), \mu)$  are pooling outcomes.
- (ii) (No distortion at the bottom)  $C(z) = ((p, q^*(\theta_z), \mu)$  with  $\theta_z = \min(\text{supp}(\mu))$ .

*The maximization problem of the principal becomes:*

$$\begin{aligned}
 (3) \quad & \max_{\{\theta_z\}_{z \in \mathbb{N}}} \sum_{z \in \mathbb{N}} (p_z - v(q^*(\theta_z))\mu_0([\theta_z, \theta_{z+1}])) \\
 & \text{s.t. } \theta_{z+1} > \theta_z, \theta_0 = \underline{\theta}, \text{ and } \theta_z < \bar{\theta} \\
 & u(q^*(\theta_{z+1}), \theta_{z+1}) - p_{z+1} = u(q^*(\theta_z), \theta_{z+1}) - p_z.
 \end{aligned}$$

To gain some intuition for our main result, observe that by Lemma 1 efficient and separating outcomes must be in any set of renegotiation-proof outcomes. Thus, only outcomes that cannot be strictly improved by a set of efficient and separating outcomes can be in any set of renegotiation-proof outcomes. The only outcomes that cannot be improved by a set of efficient and separating outcomes are pooling outcomes in which one of the types in the pool receives his efficient quality. This is due to the fact that the type that received his efficient quality in the pooling outcome, also receives his efficient quality in a feasible set of

efficient and separating outcomes. Thus, by Lemma 4, if the principal were to propose a new mechanism that led to efficient and separating outcomes she would make the same profit as before. The profit maximizing outcome among such pooling outcomes is then a countably infinite set of pooling outcomes with the lowest type in the pool receiving his efficient quality. This stems from the fact that separation of types is only possible with efficient outcomes. However, efficient and separating outcomes imply a higher information rent than, say, two pooling outcomes would. Moreover, if the principal offers the efficient quality to the lowest type in the pool, she reduces the information rent of the higher types that are not in the pool.

Before we turn our attention to the proof of Theorem 1, we shall illustrate our results by means of an example.

**Illustration of the results.** Let  $v(p, q) = p - 1/2q^2$  and  $u(p, q, \theta) = \theta q - p$ . Let  $\Theta = [1, 2]$  and  $\mu_0$  the uniform measure on  $[1, 2]$ . In this case the efficient quality for each type is given by  $q^*(\theta) = \theta$ . We compare four different scenarios.

(i) *First best.* If the principal is able to observe the agent's type, the principal can extract all the surplus from the agent. She will offer contract  $(p, q) = (\theta^2, \theta)$  to an agent of type  $\theta$  and thereby achieve the first-best allocation from her point of view. The overall profit of the principal is then given by

$$\int_1^2 \frac{1}{2} \theta^2 d\theta = \frac{7}{6}.$$

(ii) *Efficient contracting.* If the principal cannot observe the agent's type, she is still able to implement the efficient quality levels. Efficient quality levels are achieved by offering a menu of contracts  $(p, q) = (0.5(1 + q^2), q)$  for  $q \in [1, 2]$ . In this case, for each type  $\theta$  of the agent it is optimal to choose  $(0.5(1 + \theta^2), \theta)$  so that every type of agent obtains his efficient quality. Moreover, the agent with the type  $\theta = 1$  obtains a rent of 0, every other type of agent receives a positive rent. The principal obtains the same profit from every type of agent which amounts to

$$\frac{1}{2} (1 + \theta^2) - \frac{1}{2} \theta^2 = \frac{1}{2}.$$

(iii) *Second best – with commitment.* If the principal cannot observe the agent's type and can fully commit, she can implement distorting contracts to maximize her own profit. The optimal contract for each type  $\theta$  is then given by  $(p, q) = (\theta^2 - 1, 2(\theta - 1))$ . In this case, the agent of type  $\theta = 2$  obtains a contract with his efficient quality. Every other type of agent  $\theta \in [1, 2)$  obtains a contract with a quality that is lower than his efficient one. The result is often called “*no distortion at the top*”. The second-best outcome can be achieved by offering a menu of contracts  $(p(q), q) = (q^2/4 + q, q)$ . The profit of the principal is given by

$$\int_1^2 \theta^2 - 1 - 2(\theta - 1)^2 d\theta = \frac{2}{3}.$$

(iv) *Third best – without commitment.* Using the incentive compatibility constraints in equation (3), that is,  $\theta_{z+1}^2 - p_{z+1} = \theta_{z+1}\theta_z - p_z$ , and setting  $p_0 = 1$ , the program in Theorem 1 can be rewritten to give

$$\begin{aligned} \max \quad & 1 + \sum_{z \in \mathbb{N}} \left( (2 - \theta_{z+1}) \theta_{z+1} - \frac{1}{2} \theta_z^2 \right) (\theta_{z+1} - \theta_z) \\ \text{s.t.} \quad & \theta_{z+1} > \theta_z, \theta_0 = 1, \text{ and } \theta_z < 2 \end{aligned}$$

We approximated this program for  $z \in \{0, \dots, 6\}$  numerically.<sup>9</sup> With six contracts the optimal pooling of types is given by

$$\begin{aligned} [\theta_0, \theta_1] &= [1, 1.63], & (p_0, q_0) &= (1, 1), & u(p_0, q_0, \theta_0) &= 0 \\ [\theta_1, \theta_2] &= [1.63, 1.87], & (p_1, q_1) &= (2.03, 1.63), & u(p_1, q_1, \theta_1) &= 0.63 \\ [\theta_2, \theta_3] &= [1.87, 1.95], & (p_2, q_2) &= (2.48, 1.87), & u(p_2, q_2, \theta_2) &= 1.012 \\ [\theta_3, \theta_4] &= [1.95, 1.98], & (p_3, q_3) &= (2.63, 1.98), & u(p_3, q_3, \theta_3) &= 1.17 \\ [\theta_4, \theta_5] &= [1.98, 1.99], & (p_4, q_4) &= (2.69, 1.98), & u(p_4, q_4, \theta_4) &= 1.23 \\ [\theta_5, \theta_6] &= [1.99, 2], & (p_5, q_5) &= (2.71, 1.99), & u(p_5, q_5, \theta_6) &= 1.25. \end{aligned}$$

The profit of the principal is given by

$$0.577.$$

The overview of the results is depicted in Figure 1.

<sup>9</sup>Adding  $z=7$  and thereby a seventh contract increased the profit of the principal only in the order of  $10^{-5}$ .



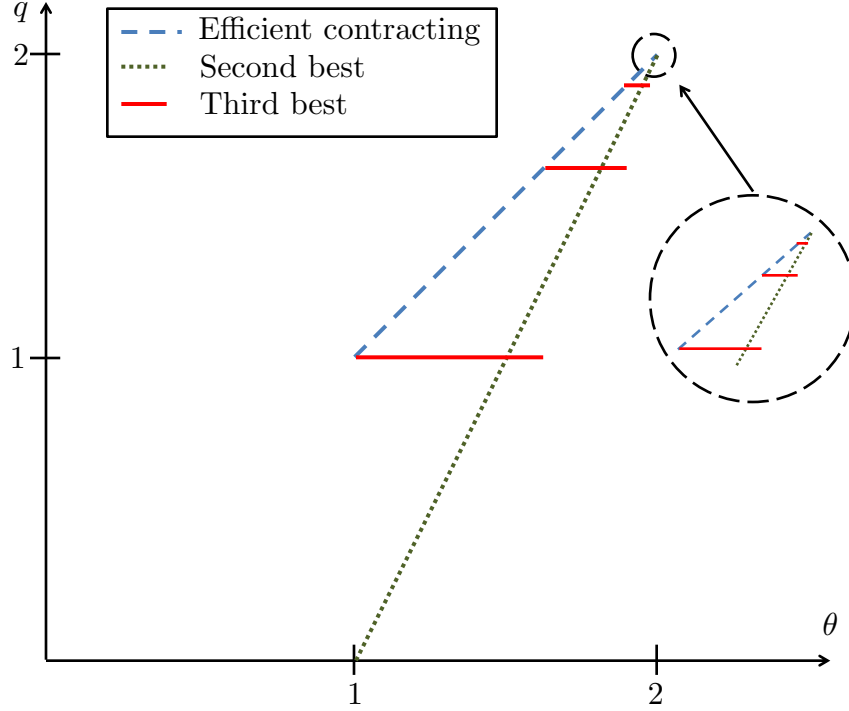


FIGURE 1. The comparison of the efficient quality levels (blue line) with the quality levels with the second-best solution with commitment (green line) and the third-best solution without commitment (red line).

**Outline of the proof of the main result.** We now turn our attention to the proof of Theorem 1. We proceed by a series of lemmata. Firstly, we show that for any outcome  $C$  there exists a set of efficient and separating outcomes that are feasible (Lemma 5). Secondly, we demonstrate that the following outcomes *cannot* be part of any renegotiation-proof  $\Omega$ :

- (i) Inefficient and separating outcomes, that is,  $C = ((p, q), \mu_\theta)$  such that  $q \neq q^*(\theta)$  (Lemma 6).
- (ii) Outcomes with a belief of the principal whose support has a gap, that is,  $C = ((p, q), \mu)$  such that  $\text{supp}(\mu)$  is not connected (Lemma 7).
- (iii) Outcomes with a quality that is efficient for a type that is not in the support of the belief of the principal, that is,  $C = ((p, q), \mu)$  such that  $q = q^*(\theta)$  but  $\theta \notin \text{supp}(\mu)$  (Lemma 8).

Thirdly, we prove that the profit maximizing set of outcomes  $\{C(z)\}$  among the outcomes with a connected support of  $\mu(z)$  and an efficient quality for one type in  $\text{supp}(\mu(z))$  takes the

form described in Theorem 1. That is, the set is countably infinite,  $\mu(z)$  is not degenerate for every  $C(z)$ , and the lowest type in support of  $\mu(z)$  receives his efficient quality (Lemma 9). Fourthly, we demonstrate that the resulting set is internally and externally consistent and thus renegotiation-proof (Lemma 10).

### Proof of the main result.

**Lemma 5.** *For any outcome  $C = ((p, q), \mu)$  there exists a set of efficient and separating outcomes  $\{C(z)\}$  which is feasible. That is,  $\{C(z)\} \in X(C)$ .*

*Proof.* The proof is relegated to the appendix. □

**Lemma 6.** *If  $\theta \neq \theta'$ , a separating  $C = ((p, q^*(\theta), \mu_{\theta'})$  cannot be an element of any set of renegotiation-proof outcomes  $\Omega$ .*

*Proof.* Starting from  $C = ((p, q^*(\theta), \mu_{\theta'})$ , there exists a feasible outcome  $\bar{C} = ((\bar{p}, q^*(\theta')), \mu_{\theta'})$  that makes the principal strictly better off. Lemma 2 implies the result. □

In what follows, we will call outcomes such that  $\text{supp}(\mu)$  is non-degenerated, pooling outcomes.

**Lemma 7.** *If  $\text{supp}(\mu)$  is not connected,  $C = ((p, q), \mu)$  cannot be an element of any set of renegotiation-proof outcomes  $\Omega$ .*

*Proof.* That  $\text{supp}(\mu)$  is not connected implies that there exist  $\theta'$  and  $\theta''$  in  $\text{supp}(\mu)$  with  $\theta'' > \theta'$  such that  $\mu((\theta', \theta'')) = 0$ . Suppose that  $q = q^*(\hat{\theta})$  for some  $\hat{\theta} \leq \theta'$ .<sup>10</sup> We will show that there exists a feasible set of efficient and separating outcomes that make the principal strictly better off. Consider the following menu of contracts  $\{(p(\theta), q^*(\theta)) : \theta \in \text{supp}(\mu)\}$  with  $p_{\theta} = u_q q_{\theta}^*$ . For  $\theta \leq \theta'$  set  $p(\hat{\theta}) = p$ . For  $\theta \geq \theta''$  set  $p(\theta'') = p(\theta') + u(q^*(\theta''), \theta'') - u(q^*(\theta'), \theta'')$ . From Lemma 3 it follows that each type of agent is better off with the contract with his efficient quality. That is, an agent of type  $\theta$  is better off with  $(p(\theta), q^*(\theta))$ . The seller is

---

<sup>10</sup>The case  $\hat{\theta} > \theta'$  works analogously.

strictly better off, since she makes the same profit from all types  $\theta \leq \theta'$  and strictly more profit from types  $\theta \geq \theta''$ . This is a consequence of the efficiency of  $q^*(\theta'')$  and

$$\begin{aligned} [p(\theta'') - v(q^*(\theta''))] - [p(\theta') - v(q^*(\theta'))] = \\ [u(q^*(\theta''), \theta'') - v(q^*(\theta''))] - [u(q^*(\theta'), \theta'') - v(q^*(\theta'))] > 0. \end{aligned}$$

Thus, we constructed a feasible set of efficient and separating outcomes that make the principal better off. As, due to Lemma 5, all efficient and separating outcomes are in every set of renegotiation-proof outcomes, the initial outcome  $C$  could not have been part of any set of renegotiation-proof outcomes.  $\square$

Lemma 5, Lemma 6 and Lemma 7 taken together illustrate why similar models with discrete type spaces lead to efficient outcomes: with discrete type spaces the support of  $\mu$  cannot be connected. Thus, renegotiation-proof outcomes need to be separating and this is only possible with efficient outcomes. In the following we will call outcomes such that  $\text{supp}(\mu)$  is connected, connected outcomes.

**Lemma 8.** *If  $\tilde{\theta} \notin \text{supp}(\mu)$ ,  $C = ((p, q^*(\tilde{\theta})), \mu)$  cannot be an element of any set of renegotiation-proof outcomes.*

*Proof.* From Lemma 7 it follows that we need only to consider connected outcomes  $C = ((p, q^*(\theta)), \mu)$ . Suppose  $\tilde{\theta} < \min \text{supp}(\mu) = \theta'$ .<sup>11</sup> We show that there exists a set of feasible and efficient outcomes that make the principal strictly better off. Consider the following set of contracts:  $(p(\theta), q^*(\theta))$  with  $p_\theta = u_q q_\theta^*$  and  $p(\theta') = p + [u(q^*(\theta'), \theta') - u(q^*(\tilde{\theta}), \theta')] > p$ . It follows from Lemma 3 that each type of agent is better off with the contract with his efficient quality. That is, an agent of type  $\theta$  is better off with  $(p(\theta), q^*(\theta))$ . The seller is strictly better off because she can offer to every agent type the efficient quality at a higher price. Thus, we constructed a set of efficient and separating outcomes that are feasible and leave the principal better off. As, due to Lemma 5, all efficient and separating outcomes are

---

<sup>11</sup>The case  $\tilde{\theta} > \max \text{supp}(\mu)$  works analogously.

in every set of renegotiation-proof outcomes, the initial outcome  $C$  could not have been part of any set of renegotiation-proof outcomes.  $\square$

At this point we have shown that renegotiation-proof outcomes must be either efficient and separating or pooling and connected such that one of the agent types in the support obtains his efficient quality. We now show that the profit-maximizing set of outcomes among those outcomes takes the form as described in Theorem 1.

**Lemma 9.** *The profit-maximizing set of outcomes  $\{C(z)\}_{z \in \mathbb{N}}$  among efficient and separating or pooling and connected outcomes such that one of the agent types in the support obtains his efficient quality is countably infinite and has the following properties:*

- (i) (Pooling) For all  $C(z) = ((p, q), \mu)$  it holds  $\mu_0(\text{supp}(\mu)) > 0$
- (ii) (No distortion at the bottom)  $C(z) = ((p, q^*(\theta_z), \mu)$  with  $\theta_z = \min(\text{supp}(\mu))$ .

*Proof.* The proof is relegated to the appendix.  $\square$

The proof of Lemma 9 proceeds along five steps. Firstly, we show that in the set of optimal outcomes higher types receive higher quality.

Secondly, we demonstrate that the principal optimally makes a higher profit from the higher agent types.

Thirdly, we argue that efficient and separating outcomes are not optimal for the principal. This is due to the fact that if the principal offers the efficient quality to a connected subset of types while satisfying the incentive compatibility constraints, she will not obtain additional rent from higher types. Those types receive a high information rent. In this case, the principal can reduce the information rent by offering pooling contracts.

Fourthly, we show that the lowest type in a pooling and connected outcome will receive his optimal quality in the profit-maximizing set of outcomes. If in the pooling outcome the quality is optimal for some intermediate type, the quality for the lower types is distorted in the wrong direction. That is, they receive too much quality. In this case, the seller prefers to give some types a lower quality which reduces the information rent to higher types. The

best way to achieve this is to give the lowest type in the pool the efficient quality, such that the quality for all types is distorted in the right direction.

Finally, we demonstrate that the set of profit-maximizing outcomes must be countably infinite. To see this, let us suppose to the contrary that the number of outcomes is finite. In this case, there exists an outcome that contains all types above some  $\theta$ . It follows that the principal is better off if she splits this pool into two pools: one where she obtains the same profit as before and that contains the lower types of agent and a second one where she obtains a higher profit and that is chosen by the higher types of agent.

It remains to be shown that a set of outcomes as described in Lemma 9 is indeed renegotiation-proof.

**Lemma 10.** *Let  $\{C(\theta_z) = (p_i, q^*(\theta_z), \mu_z)\}_{z \in \mathbb{N}}$  be a set of feasible outcomes starting from  $C_0$  such that  $\mu_z([\theta_z, \theta_{z+1})) = 1$ . There exists a set of renegotiation-proof outcomes  $\Omega$  such that  $\{C(\theta_z)\}_{z \in \mathbb{N}} \subset \Omega$ .*

*Proof.* Let  $\Omega = \{C : C \text{ is efficient and separating}\} \cup \{C(\theta_z)\}_{z \in \mathbb{N}}$ . Firstly, we consider external consistency. By Lemma 5, for any outcome  $C$  there exists a set of efficient and separating outcomes which is feasible. Thus,  $\Omega$  is externally consistent.

Secondly, we consider internal consistency. Take any  $C \in \Omega$ . If  $C$  is efficient and separating,  $X(C) = \{C\}$ . If  $C = C(\theta_z)$  for some  $i$  then the only feasible set of outcomes in  $\Omega$  is a set of efficient and separating outcomes such that  $\theta_z$  obtains the same contract as before. From Lemma 4, it follows that the principal makes the same profit as before. Thus, internal consistency is not violated.  $\square$

## 5. THE COASE CONJECTURE

In this section we demonstrate how our solution is consistent with other results from non-cooperative bargaining theory. In particular, we will consider the Coase conjecture, which is a special instance of our set-up. The Coase conjecture argues that if a seller is not able to commit to not selling a durable good, she can at most charge a price equal to the lowest valuation of the buyer as long as the cost of the seller is strictly below the buyer's lowest

valuation (gap case). Whenever the cost of the seller is equal or above the lowest valuation of the buyer (no-gap case), the seller is able to charge the monopoly price even without commitment (Ausubel and Deneckere (1989)).

For a discrete type space, Vartiainen (2013) demonstrated how the solution concept can recreate the Coase conjecture in the gap case. For the no-gap case he writes that, in principle, it is possible to construct mechanism selection strategies that are consistent and satisfy the one-deviation property. However, those selection strategies would be complex. Thus, he refrains from constructing any such strategies and makes the additional assumption that mechanism selection strategies need to be stationary. In this case, even in the no-gap case, the seller sells at a price equal to her cost. In this section we will demonstrate the simplicity and cutting power of our approach: to focus on outcomes rather than selection strategies. We will recreate the conceptual difference between the gap and no-gap case without resorting to complex selection strategies.

**Set up.** Consider a monopolistic seller who is selling one object to a buyer. A contract is a tuple  $\omega = (p, q)$  with  $p \in \mathbb{R}$  specifying the price and  $q \in \{0, 1\}$  specifying whether the good is exchanged ( $q = 1$ ) or not ( $q = 0$ ). The seller incurs a cost of  $c$  of producing the good and this cost is common knowledge between the seller and the buyer. Thus, the utility function of the seller is given by

$$v(\omega) = p - cq.$$

The buyer has a valuation of  $\theta \in \Theta = [\underline{\theta}, \bar{\theta}]$  for consuming the good, which is private information to the buyer. Thus, the utility of the buyer is given by

$$u(\omega, \theta) = \theta q - p.$$

The seller's prior about the valuation of the buyer is given by  $\mu_0 \in \Delta(\Theta)$ . To simplify notation, assume that  $\text{supp}(\mu_0) = \Theta$ . The initial contract  $\omega_0$  is  $(0, 0)$ . In the following we will study two cases; the gap case with  $c < \underline{\theta}$  and the no-gap case with  $c \geq \underline{\theta}$ . In the first case, there is a gap between the cost of the seller and the lowest valuation of the buyer so

it is common knowledge that there are gains from trade. In the second case, the buyer can have a valuation that makes trading the good inefficient.

**The gap case.** In what follows we will show that if  $c < \underline{\theta}$ , that is, if there is a gap between the cost of the seller and the lowest valuation of the buyer, the unique set of renegotiation-proof outcomes contains only outcomes in which the good is traded. This implies that the seller will charge a price equal to the lowest valuation and not propose a new mechanism after the buyer has accepted.

**Proposition 2.** *Let  $c < \underline{\theta}$ , the unique set of renegotiation-proof outcomes  $\Omega$  that satisfies the conditions of Definition 3 is  $\Omega = \{C = ((p, q), \mu) : q = 1\}$ .*

*Proof.* We show that for a given  $\Omega$  that is internally and externally consistent,  $((p, q), \mu)$  is in  $\Omega$  if and only if  $q = 1$ .

“*if*”: For any outcome  $C = ((p, 1), \mu)$  there exists no set of outcomes that would make both parties weakly better off. Thus,  $X(C) = \{C\}$ . It follows from Lemma 1 that  $C$  must be in  $\Omega$ .

“*only if*”: Let  $C = ((p, 0), \mu)$ . Consider the contract  $(p + c + 1/2(\underline{\theta} - c), 1)$ . The buyer strictly prefers this contract to  $(p, 0)$  independent of his type. Moreover the seller is strictly better off than with contract  $(p, 0)$ . Thus,  $\{((p + c + 1/2(\underline{\theta} - c), 1), \mu)\}$  is feasible and in  $X(C)$ . Lemma 2 gives the desired result.  $\square$

Proposition 2 implies that starting from a status quo in which the good is not traded, any final outcome cannot imply a price charged for the good that is above the lowest valuation that is in the support of the current belief of the principal. If this is not the case, some of the types of buyer would optimally choose not to buy the good, which would lead to  $q = 0$  in some of the outcomes. Such outcomes, however, are not in the unique set of renegotiation-proof outcomes.

That the set of renegotiation proof outcomes  $\Omega = \{C = ((p, q), \mu) : q = 1\}$  is unique implies the following optimization problem of the seller.

$$\begin{aligned} \max_p \quad & p - c \\ \text{s.t.} \quad & p \leq \underline{\theta} \end{aligned}$$

The solution of this problem is the outcome  $((\underline{\theta}, 1), \mu_0)$ .

**Corollary 2.** *If  $c < \underline{\theta}$  and the seller cannot commit to a mechanism, she can only charge a price of  $p = \underline{\theta}$ .*

This result is in agreement with the literature on the Coase conjecture: the monopolist competes herself down to the lowest valuation.<sup>12</sup>

**The no-gap case.** We now consider  $c \geq \underline{\theta}$ , the no-gap case. In this case, we construct an internally and externally consistent  $\Omega$  such that in the optimal set of outcomes within this  $\Omega$ , the seller can charge the monopoly price. As the problem of the seller without commitment is a more constrained version of the problem of a seller with full commitment, charging the monopoly price must also be the optimal solution without commitment.

**Proposition 3.** *Let  $p^M$  denote the monopoly price. That is,  $p^M = \arg \max_p (1 - \mu_0(p))(p - c) > c$ . An  $\Omega$  as defined below is internally and externally consistent.*

$C = ((p, q), \mu)$  is in  $\Omega$  if one of the following holds true

- (i)  $q = 1$  and for all  $\theta$  in  $\text{supp}(\mu)$  it holds  $\theta \geq c$ .
- (ii)  $q = 0$  and for all  $\theta$  in  $\text{supp}(\mu)$  it holds  $\theta \leq c$ .
- (iii)  $q = 0$ ,  $\max(\text{supp}(\mu)) = p^M$ , and there exists an  $\epsilon > 0$  such that for all  $\delta \leq \epsilon$  it holds  $\mu[(c - \delta, c + \delta) \cap \text{supp}(\mu)] > 0$ .

*Proof.* The proof is relegated to the appendix □

The set of renegotiation-proof outcomes in Proposition 3 is constructed in such a way that the seller proposes a mechanism that leads to two outcomes: either the object is traded at a price  $p^M$  and the seller believes that all buyers have a valuation above  $p^M$ , or the object is not traded and the seller believes that all buyers have a valuation below  $p^M$ . Whenever the

<sup>12</sup>See for example Fudenberg et al. (1985), or more recently Strulovici (2016).



object is not traded, the only other feasible renegotiation-proof outcomes are that the good is traded at a price equal to  $c$  or not traded at all. Thus, the seller cannot profit from proposing a new mechanism whenever the object is not traded. As charging the monopoly price is the optimal mechanism under full commitment it must also be the optimal mechanism without commitment.

**Corollary 3.** *If  $c \geq \theta$ , the principal can charge the monopoly price  $p^M$  even without commitment.*

This result is in agreement with the non-cooperative bargaining literature. Ausubel and Deneckere (1989) show that in the no-gap case, in an infinite-horizon bargaining game, the seller can sustain the monopoly price in equilibrium if the frictions go to zero.

## 6. CONCLUSION

The main contribution of this manuscript is to construct the optimal mechanism for the screening problem of a principal who is not able to commit to not renegotiating the outcome of any mechanism she proposes. Interestingly, the optimal mechanism is inefficient given that it leads to a pooling of types. This is different from previous work on contract design without commitment that find that with private values the principal can only implement efficient outcomes.<sup>13</sup> The main difference between our approach and previous approaches is that we consider a continuous type space. Only with a continuous type space is the principal able to pool types and thus extract more rent from higher types.

To arrive at the results we characterize the set of renegotiation-proof outcomes. By using internal and external consistency, powerful yet simple results are achieved. The main advantage of this approach is that, in contrast to other definitions of renegotiation-proofness, we do not assume that renegotiation-proof outcomes must be efficient. This is due to the fact that we do not consider outcomes as blocking which themselves are subject to renegotiation.

We demonstrate the cutting power of our concept by considering the Coase conjecture. We show that, if there is a gap between the costs of the seller and the lowest valuation

---

<sup>13</sup>For example see Gretschko and Wambach (2016), Maestri (2012), Strulovici (2016), and Vartiainen (2013).

of the buyer, the seller can only charge a price equal to the lowest valuation of the buyer. However, if there is no such gap, the seller is able to charge the monopoly price even without commitment.

## APPENDIX

### Proof of Proposition 1.

*Proof.* Ad (i): Let  $\sigma^p$  be a mechanism selection strategy that is consistent and satisfies the one-deviation property. Let  $\Omega = \bigcup_{C \in \Gamma} f(\sigma^p(C), C, \sigma^a)$ . External consistency follows, by definition, from the fact that for every  $C$  in  $\Gamma$ ,  $f(\sigma^p(C), C, \sigma^a)$  is in  $X(C)$  and a subset of  $\Omega$ . It remains to check that  $\Omega$  is internally consistent. Let  $C \in \Omega$  and  $\{C(z) = (\omega(z), \mu(z))\}_{z \in \mathcal{Z}} \in X(C)$  with  $C(z) \in \Omega$  for all  $z \in \mathcal{Z}$  and some metric space  $\mathcal{Z}$ . Denote by  $M = (\mathcal{Z}, \omega(z))$  the mechanism that is induced by  $\{C(z)\}_{z \in \mathcal{Z}}$ . As  $\{C(z)\}$  is in  $X(C)$ ,  $M$  exists and is well defined. Due to consistency of  $\sigma^p$  and the fact that  $C(z)$  is in  $\Omega$  it follows that  $\sigma^p(C(z)) \leq v(\omega(z))$ . Thus,  $M$  is in  $\mathcal{M}^{\sigma^p}$ . Due to the one-deviation property of  $\sigma$  it follows that  $V(\sigma^p(C)) \geq V(M)$ . Thus, again by consistency,  $V(\sigma^p(C)) \leq v(\omega)$  and thereby  $V(M) \leq v(\omega)$ .

Ad (ii): Let  $\Omega$  be externally and internally consistent and let the solution to

$$\begin{aligned}
 (4) \quad & \max_{\{C(z)\}} \int_{\mathcal{Z}} v(\omega(z)) d\nu \\
 & \text{s.t.} \quad \{C(z) = (\omega(z), \mu(\cdot : z)) : z \in \mathcal{Z}\} \in X(C) \\
 & \quad \{C(z) = (\omega(z), \mu(\cdot : z)) : z \in \mathcal{Z}\} \subset \Omega.
 \end{aligned}$$

exist. Define  $\sigma^p$  as follows. Choose for every  $C = (\omega, \mu) \in \Omega$ ,  $\sigma^p(C) = (\mathcal{Z}, \omega(z) \equiv \omega)$  and for  $C \notin \Omega$ ,  $\sigma^p(C) = (\mathcal{Z}, \omega(z))$  such that  $\{C(z) : z \in \mathcal{Z}\} \in X(C)$ ,  $\{C(z) : z \in \mathcal{Z}\} \subset \Omega$  and  $\{C(z)\}$  is a solution to Problem 4. This is always possible due to external consistency and the fact that a solution exists. For any  $C \in \Gamma$ ,  $f(\sigma^p(C), C, \sigma^a)$  is a subset of  $\Omega$ . Thus, by definition, for every  $C' = (\omega', \mu') \in f(\sigma^p(C), C, \sigma^a)$ ,  $v(\omega') \geq V(\sigma^p(C')) = v(\omega')$ . It follows that  $\sigma^p(C)$  is in  $\mathcal{M}^{\sigma^p}(C)$  and thereby  $\sigma^p$  is consistent. For the one-deviation property take any  $M = (\mathcal{Z}, \omega(z))$  in  $\mathcal{M}^{\sigma^p}(C)$ . Observe that  $\{C(z)\}_{z \in \mathcal{Z}} = f(M, C, \sigma^a)$  is feasible. First, suppose that  $\{C(z)\}_{z \in \mathcal{Z}} \subset \Omega$  it follows from the definition of  $\sigma^p(C)$  as the solution

to Problem (4) that  $V(\sigma^p(C)) \geq V(M)$ . Thus, suppose that there exists  $\mathcal{Z}' \subseteq \mathcal{Z}$  such that for all  $z' \in \mathcal{Z}'$ ,  $C(z') \notin \Omega$ . Consider the following set of outcomes  $\mathcal{C} = \{C(z)\}_{z \in \mathcal{Z} \setminus \mathcal{Z}'} \cup \bigcup_{z' \in \mathcal{Z}'} f(\sigma(C(z')), C(z'), \sigma^a)$ . By definition of  $\sigma^p$ ,  $\mathcal{C}$  is feasible starting from  $C$ ,  $\mathcal{C} \subset \Omega$  and  $V(\mathcal{C}) \geq V(M)$ . It follows from the definition of  $\sigma^p$  as the solution of Problem (4) that  $V(\sigma^p(C)) \geq V(\mathcal{C}) \geq V(M)$ .  $\square$

### Proof of Lemma 5.

*Proof.* Three cases are relevant. Either there exists a type  $\tilde{\theta}$  in  $\text{supp}(\mu)$  such that  $q = q^*(\tilde{\theta})$  or for  $\theta' = \min(\text{supp}(\mu))$  it holds  $q < q^*(\theta')$  or for  $\theta'' = \max(\text{supp}(\mu))$  it holds  $q > q^*(\theta'')$ . We focus on the first case only, the other cases can be proven analogously.

Let  $q = q^*(\tilde{\theta})$  for some  $\tilde{\theta} \in \text{supp}(\mu)$ . Consider the following set of efficient and separating outcomes  $\{C(\theta) = ((q^*(\theta), p(\theta)), \mu_\theta)\}$  such that  $p(\tilde{\theta}) = p$  and  $p_\theta(\theta) = u_q(q^*(\theta), \theta)q^*(\theta)$ . An agent of type  $\tilde{\theta}$  will receive the initial contract and an agent of type  $\theta \neq \tilde{\theta}$  will receive a contract that makes him strictly better off. Thus, this set of outcomes is individually rational for each type of the agent (condition (i) of Definition 2). Moreover,  $p(\theta)$  satisfies equation 2. Thus, this set of outcomes is incentive compatible (condition (ii)) and satisfies Bayesian consistency (condition (iii)). As a consequence of

$$p_\theta(\theta) - v_q(q^*(\theta))q_\theta^*(\theta) = (u_q(q^*(\theta), \theta) - v_q(q^*(\theta)))q_\theta^*(\theta) = 0,$$

the principal is indifferent between all outcomes in  $\{C(\theta) = ((q^*(\theta), p(\theta)), \mu_\theta)\}$ . In particular, she is then indifferent between the original outcome  $C$  and any outcome in

$$\{C(\theta) = ((q^*(\theta), p(\theta)), \mu_\theta)\}.$$

Hence, the proposed set of outcomes satisfies individual rationality of the principal (condition (iv)). Overall, this implies that

$$\{C(\theta) = ((q^*(\theta), p(\theta)), \mu_\theta)\} \in X(C).$$

$\square$

**Proof of Lemma 9.** The proof is divided into five steps. We show that

Step 1 *In the set of optimal outcomes, higher types obtain higher quality*

Step 2 *In the set of optimal outcomes, the seller achieves a weakly higher profit from higher types*

Step 3 *Efficient and separating outcomes are not in the set of optimal outcomes*

Step 4 *For every outcome in the set of optimal outcomes, the lowest type in the support of the belief of the principal receives his efficient quality.*

Step 5 *The set of optimal outcomes is countably infinite.*

Denote by  $\{C(z)\}$  the set of feasible, profit-maximizing outcomes for the seller starting from  $C^0$  and by  $\{\omega(z)\}$  the set of corresponding contracts.

*Step 1: In the set of optimal outcomes, higher types obtain higher quality.*

*Proof.* The principal maximizes among efficient and separating or pooling and connected outcomes such that one of the types receives his efficient quality. Thus, for  $\theta_2 > \theta_1$  with  $q_2$  and  $q_1$  denoting the quality that type  $\theta_2$  respectively  $\theta_1$  receives, two cases are relevant. Firstly, both types obtain their efficient quality  $q_2 = q^*(\theta_2)$  and  $q_1 = q^*(\theta_1)$ . In this case, as  $q^*$  is an increasing function,  $q_2 > q_1$ . Secondly, both types are in different pooling outcomes.<sup>14</sup> In this case, as one of the types in each pool receives his efficient quality, the pools are connected and due to the fact that  $q^*$  is an increasing function,  $q_2 > q_1$ . Thirdly, both types are in the same pooling outcome. In this case,  $q_2 = q_1$ . Summing up, it follows that if  $\theta_2 > \theta_1$ ,  $q_2 \geq q_1$ , that is, higher types receive a higher quality.  $\square$

*Step 2: In the set of optimal outcomes, the seller achieves a weakly higher profit from higher types.*

*Proof.* We will show that for any type  $\theta$  there exists an  $\epsilon > 0$  such that for all types  $\theta' \in (\theta, \theta + \epsilon)$  the seller realizes an equal or higher profit than with type  $\theta$ . Let  $C = ((p, q), \mu)$  such that  $\theta \in \text{supp}(\mu)$ . If  $\theta < \max\{\text{supp}(\mu)\}$ , that is, if  $C$  is a pooling and connected outcome and  $\theta$  is not the largest type in the pool, there exists an  $\epsilon > 0$  such that all types

<sup>14</sup>Or one of the types receives his efficient quality and the other is in a pooling outcome.

$\theta' \in (\theta, \theta + \epsilon)$  receive the same contract. Thus, the seller makes the same profit with all these types. Assume that  $\theta = \max \{\text{supp}(\mu)\}$ , that is, assume that  $\theta$  is the largest type in the pool or that  $C$  is an efficient and separating outcome. All types  $\theta' > \theta$  obtain a different contract than type  $\theta$ . If there exists an  $\epsilon > 0$  such that almost all types in  $\theta' \in (\theta, \theta + \epsilon)$  receive their efficient contract, then the seller makes the same profit from all of these types (Lemma 4). If type  $\theta$  also obtains his efficient quality, the seller makes the same profit from  $\theta$  and any  $\theta'$ . If type  $\theta$  does not obtain his efficient quality, then the seller makes a strictly larger profit from any type  $\theta'$ . Thus, assume that for all  $\epsilon > 0$  the amount of types in  $(\theta, \theta + \epsilon)$  that do not receive their efficient contract is of positive measure. In this case there exists an  $\epsilon > 0$  such that all types in  $(\theta, \theta + \epsilon)$  are in the same pooling outcome with a contract  $\hat{\omega} = (\hat{p}, q^*(\hat{\theta}))$  for some  $\hat{\theta}$ . Call this outcome  $\hat{C} = (\hat{\omega}, \hat{\mu})$ . If the seller obtains the same profit or higher with  $(\hat{p}, q^*(\hat{\theta}))$  than with  $(p, q)$ , we are done. Thus, assume that the seller makes less profit with  $(\hat{p}, q^*(\hat{\theta}))$ . We will show that there exists a set of outcomes that makes the seller strictly better off. In this case there exists a type  $\theta'$  with  $\theta < \theta' < \hat{\theta}$  and a contract  $\omega' = (p', q^*(\theta'))$  such that  $\bar{u}((p, q), \theta) = \bar{u}((p', q^*(\theta')), \theta)$ ,  $\bar{u}((\hat{p}, q^*(\hat{\theta})), \hat{\theta}) > \bar{u}((p', q^*(\theta')), \hat{\theta})$ , and  $v((p', q^*(\theta'))) > (\hat{p}, q^*(\hat{\theta}))$ . That is, we can find a contract with an efficient quality for type  $\theta' \in (\theta, \hat{\theta})$ . Moreover, type  $\theta$  is indifferent between his original contract and the new contract, type  $\hat{\theta}$  strictly prefers his original contract, and the principal makes a higher profit from the new contract. Now consider the following set of outcomes  $\{C(z)\} \setminus \hat{C} \cup \{C', C''\}$  with  $C' = (\omega', \mu')$  and  $C'' = (\hat{\omega}, \mu'')$ . By construction, there exists a type  $\theta'' \in (\theta, \hat{\theta})$  such that all types between  $\theta$  and  $\theta''$  prefer  $\omega'$  to  $\omega$  and all types between  $\theta''$  and  $\max \{\text{supp}(\hat{\mu})\} = \tilde{\theta}$  prefer  $\hat{\omega}$ . Thus, set  $\mu'$  such that it is Bayesian consistent and satisfies  $\text{supp}(\mu') = (\theta, \theta'']$  and  $\mu''$  such that it is Bayesian consistent and satisfies  $\text{supp}(\mu'') = (\theta'', \tilde{\theta})$ . It follows that the constructed set of outcomes,  $\{C(z)\} \setminus \hat{C} \cup \{C', C''\}$ , is feasible and leaves the principal better off. Thus, the original set of outcomes,  $\{C(z)\}$ , could not have been optimal.  $\square$

*Step 3: Efficient and separating outcomes are not in the set of optimal outcomes.*

*Proof.* We will show that there is no interval  $[\theta', \theta''] \subset \Theta$  with  $\mu_0([\theta', \theta'']) > 0$  such that all types  $\theta \in [\theta', \theta'']$  obtain their efficient quality  $q^*(\theta)$ . Denote the set of outcomes that

contain those types by  $\{C(\theta)\} \subseteq \{C(z)\}$ . We will find a different set of outcomes that is feasible starting from  $C_0$  and makes the principal strictly better off. Denote by  $p'$  the price of the contract that  $\theta'$  obtains. Consider the following two contracts  $\omega_1 = (p', q^*(\theta'))$  and  $\omega_2 = (\hat{p}, q^*(\hat{\theta}))$  with  $\hat{\theta} = (1/2)(\theta'' + \theta')$  and  $\hat{p} = \bar{u}(q^*(\theta'), \hat{\theta}) - p'$ . Now consider the following set of outcomes  $\{C(z)\} \setminus \{C(\theta)\} \cup \{C_1, C_2\}$  with  $C_1 = (\omega_1, \mu_1)$  and  $C_2 = (\omega_2, \mu_2)$ . By construction, all types not in  $[\theta', \theta'']$  still prefer their old contract  $\omega(z)$  to both  $\omega_1$  and  $\omega_2$ . Types in  $[\theta', \hat{\theta}]$  prefer  $\omega_1$  to any other contract in  $\{\omega(z)\} \setminus \{\omega(\theta)\}$ . Types in  $[\hat{\theta}, \theta'']$  prefer either  $\omega_2$  or some contract in  $\{\omega(z)\} \setminus \{\omega(\theta)\}$  that is also preferred by a higher type. Thus, there exist Bayesian consistent  $\mu_1$  and  $\mu_2$  such that the constructed set of outcomes  $\{C(z)\} \setminus \{C(\theta)\} \cup \{C_1, C_2\}$  is feasible starting from  $C_0$  and makes the principal strictly better off.<sup>15</sup>  $\square$

*Step 4: For every outcome in the set of optimal outcomes, the lowest type in the support of the belief of the principal receives his efficient quality.*

*Proof.* Suppose to the contrary that there exists a  $C(\hat{z}) = (\omega(\hat{z}), \mu(\hat{z})) = ((p, q), \mu(\hat{z})) \in \{C(z)\}$  such that  $\min \text{supp}(\mu) = \theta$  but  $q > q^*(\theta)$ . Consider the following contract  $(p', q^*(\theta))$  with  $p' = \bar{u}(q^*(\theta), \theta) - \bar{u}(q, \theta) + p$  and the following set of outcomes  $\{C(z)\} \setminus \{C(z)\}_{z \geq \hat{z}} \cup C' \cup \{\hat{C}(z)\}_{z > \hat{z}}$  with  $C' = ((p', q^*(\theta)), \mu')$  and  $\hat{C}(z) = ((\omega(z), \hat{\mu}(z)))$ . That is, construct a set of outcomes such that all contracts which are in the new set of outcomes are the same as in  $\{C(z)\}$  with the exception that  $(p, q)$  is swapped for  $(p', q^*(\theta))$ . By construction, all types below  $\theta$  still prefer their old contract. Moreover, there exists a  $\theta' > \theta$  such that all types in  $[\theta, \theta'] \subseteq \text{supp}(\mu)$  prefer  $(p', q^*(\theta))$  and all types in  $[\theta', \max \text{supp}(\mu)]$  prefer one of the contracts  $\omega(z)$  with  $z > \hat{z}$ . By construction, the principal makes a strictly higher profit from types in  $[\theta, \theta']$ . Step 1 implies that the principal makes at least as much profit from types in  $[\theta', \max \text{supp}(\mu)]$  as if the set of outcomes were  $\{C(z)\}$ . Thus, there exist Bayesian

<sup>15</sup>The idea of this proof is straightforward. If the principal gives each type of agent in some interval his efficient quality and satisfies the incentive compatibility constraints, she does not earn any additional rent from higher types by increasing the information rent they earn. If the principal offers two pooling contracts instead, the information rent to the high types is reduced.

consistent  $\mu'$  and  $\{\hat{\mu}(z)\}$  such that the constructed set of outcomes is feasible starting from  $C_0$  and makes the principal strictly better off.<sup>16</sup>  $\square$

*Step 5: The set of optimal outcomes is countably infinite.*

*Proof.* So far we have shown that the optimal set of outcomes partitions  $[\underline{\theta}, \bar{\theta}]$  in connected intervals of strictly positive measure. This implies that if the set of outcomes is infinite it must be countable. Thus, we merely need to show that the set of optimal outcomes is not finite. Suppose to the contrary that there exists a  $\theta \in \Theta = [\underline{\theta}, \bar{\theta}]$  and  $C = ((p, q^*(\theta)), \mu) \in \{C(z)\}$  such that  $\text{supp}(\mu) = [\theta, \bar{\theta}]$ . Now take any  $\theta' \in (\theta, \bar{\theta})$ , consider the contract  $\omega' = ((p', q^*(\theta'))$  with  $p' = \bar{u}(q^*(\theta'), \theta') - \bar{u}(q^*(\theta), \theta') + p$ , and construct a new set of outcomes  $\{C(z)\} \setminus C \cup \{C_1, C_2\}$  with  $C_1 = ((p, q^*(\theta)), \mu_1)$  and  $C_2 = ((p', q^*(\theta')), \mu_2)$ . Types in  $[\theta, \theta')$  prefer  $(p, q^*(\theta))$  and types in  $[\theta', \bar{\theta}]$  prefer  $(p', q^*(\theta'))$ . The principal obtains a strictly higher profit from  $(p', q^*(\theta'))$ . Thus, there exist Bayesian consistent  $\mu_1$  and  $\mu_2$  such that the constructed set of outcomes is feasible starting from  $C_0$  and makes the principal strictly better off.  $\square$

### Proof of Proposition 3.

*Proof.* We start by showing that  $\Omega$  is externally consistent. Let  $C = ((p, q), \mu)$  be not  $\Omega$ . Consider the following two outcomes  $C^1 = ((p + (1 - q)c, 1), \mu_1)$  and  $C^2 = ((p - qc, 0), \mu_2)$  such that  $\mu_1(\cdot) = \mu(\cdot : \theta \geq c)$  and  $\mu_2(\cdot) = \mu(\cdot : \theta < c)$ . That is, an outcome in which the good is traded at a price of  $p + (1 - q)c$  and an outcome in which the object is not traded and the price is  $p - qc$ . We check now whether the conditions of Definition 2 are satisfied and the proposed outcomes are feasible.

(i) Individual rationality. Outcome  $C^1$  is individually rational for buyers with a valuation above  $c$  and outcome  $C^2$  is individually rational for buyers with a valuation below  $c$ .

(ii) Incentive compatibility. Outcome  $C^1$  is strictly preferred by buyers with a valuation above  $c$  to outcome  $C^2$ . Outcome  $C^2$  is strictly preferred by buyers with a valuation below

---

<sup>16</sup>The idea of the proof is the following. If a pooling outcome contains a contract where  $q$  is optimal for some intermediate type, the lower types are distorted in the wrong direction. They obtain too much quality. The seller thus prefers to give some of those types a contract with a lower quality.

$c$ . The seller does not believe that she is facing buyers with a valuation below  $c$  in outcome  $C^1$  and buyers with a valuation above  $c$  in outcome  $C^2$ .

(iii) Bayesian consistency. The probability of reaching outcome  $C^1$  is  $\mu((\underline{\theta}, c])$  and of reaching  $C^2$  is  $\mu([c, \bar{\theta}])$ . Thus,  $\mu((\underline{\theta}, c])\mu_1 + \mu([c, \bar{\theta}])\mu_2 = \mu$ .

(iv) Individual rationality of the seller. The seller is indifferent between  $C$ ,  $C^1$ , and  $C^2$ .

Thus,  $\{C^1, C^2\} \in X(C)$  is feasible and by definition of  $\Omega$ ,  $C^1$  and  $C^2$  are in  $\Omega$ .

Now that we have shown that  $\Omega$  is externally consistent we turn our attention to internal consistency. We show that for all outcomes  $C \in \Omega$  there is no set of feasible outcomes that is also in  $\Omega$  and makes the seller strictly better off.

For every outcome with  $q = 1$  and  $\theta \geq c$  or  $q = 0$  and  $\theta \leq c$  for all  $\theta$  in  $\text{supp}(\mu)$  there exists no set of outcomes which would make the seller strictly better off and the buyer not worse off. Thus, any such outcome cannot be improved.

We turn our attention to outcomes  $C = ((0, p), \mu)$  such that  $\max(\text{supp}(\mu)) = p^M$  and there exists an  $\epsilon > 0$  such that for all  $\delta \leq \epsilon$  it holds  $\mu[(c - \delta, c + \delta) \cap \text{supp}(\mu)] > 0$ . Due to incentive compatibility of the buyer any feasible set of outcomes starting from  $C$  consists of at most two outcomes. One outcome where the object is exchanged ( $q = 1$ ) and one where the object is not exchanged ( $q = 0$ ). The set of outcomes that merely consists of one outcome in  $\Omega$  with  $q = 0$  either does not make the seller better off or is not individually rational to the buyer starting from  $C$ . The set of outcomes that merely consist of one outcome in  $\Omega$  with  $q = 1$  and a belief such that all  $\theta \geq c$ , either does not make the seller better off, due to  $\mu[(c - \delta, c + \delta) \cap \text{supp}(\mu)] > 0$  or does not satisfy individual rationality of the buyer or does not satisfy Bayesian consistency. Thus, again due to  $\mu[(c - \delta, c + \delta) \cap \text{supp}(\mu)] > 0$  the only set of feasible outcomes that are in  $\Omega$  is  $\{C^1 = ((p + c, 1), \mu_1), C^2 = ((p, 0), \mu_2)\}$  with  $\mu_1(\cdot) = \mu(\cdot : \theta \geq c)$  and  $\mu_2(\cdot) = \mu(\cdot : \theta < c)$ . However, the seller is indifferent as to whether he does not sell at a price of  $p$  or sells at a price of  $p + c$ . Thus, starting from  $C = ((0, p), \mu)$  as defined above, there exists no set of feasible outcomes that is also in  $\Omega$  and makes the seller better off. This implies that  $\Omega$  is internally consistent.  $\square$



## REFERENCES

- ASHEIM, G. AND T. NILSSEN (1997): “Insurance Monopoly and Renegotiation,” *Economic Theory*, 354, 341–354.
- AUSUBEL, L. M. AND R. J. DENECKERE (1989): “Reputation in Bargaining and Durable Goods Monopoly,” *Econometrica*, 57, 511–531.
- BEAUDRY, P. AND M. POITEVIN (1993): “Signalling and Renegotiation in Contractual Relationships,” *Econometrica*, 61, 745–782.
- BESTER, H. AND R. STRAUZ (2001): “Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case,” *Econometrica*, 69, 1077–1098.
- COASE, R. (1972): “Durability and Monopoly,” *Journal of Law and Economics*, 15, 143–149.
- EVANS, R. AND S. REICHE (2015): “Contract Design and Non-Cooperative Renegotiation,” *Journal of Economic Theory*, 157, 1159–1187.
- FARRELL, J. AND E. MASKIN (1989): “Renegotiation in Repeated Games,” *Games and Economic Behavior*, 1, 327–360.
- FUDENBERG, D., D. LEVINE, AND J. TIROLE (1985): “Infinite-Horizon Models of Bargaining with One-Sided Incomplete Information,” in *Game Theoretical Models of Bargaining*, ed. by A. Roth, Cambridge University Press, 73–98.
- FUDENBERG, D. AND J. TIROLE (1990): *Game Theory*, MIT Press.
- GRETSCHKO, V. AND A. WAMBACH (2016): “Contract (Re-)Negotiation with Private and Common Values,” *mimeo*.
- HÖRNER, J. AND L. SAMUELSON (2011): “Managing Strategic Buyers,” *Journal of Political Economy*, 119, 379–425.
- MAESTRI, L. (2012): “Dynamic Contracting under Adverse Selection and Renegotiation,” *mimeo*.
- MYERSON, R. B. (1979): “Incentive Compatibility and the Bargaining Problem,” *Econometrica*, 47, 61–74.
- NEEMAN, Z. AND G. PAVLOV (2013): “Ex-Post Renegotiation-Proof Mechanism Design,” *Journal of Economic Theory*, 148, 473–501.

- POLLRIK, M. (2016): “Mediated Audits,” *RAND Journal of Economics*, *forthcoming*.
- SKRETA, V. (2006): “Sequentially Optimal Mechanisms,” *The Review of Economic Studies*, 73, 1085–1111.
- (2015): “Optimal Auction Design under Non-Commitment,” *Journal of Economic Theory*, 159, 854–890.
- STRAUSZ, R. (2012): “Mediated contracts and mechanism design,” *Journal of Economic Theory*, 147, 1280–1290.
- STRULOVICI, B. (2016): “Contract Negotiation and the Coase Conjecture,” *mimeo*.
- VARTIAINEN, H. (2013): “Auction Design Without Commitment,” *Journal of the European Economic Association*, 11, 316–342.

ZENTRUM FÜR EUROPÄISCHE WIRTSCHAFTSFORSCHUNG (ZEW), L7, 1, 68161, MANNHEIM  
*E-mail address*: [vgretschko@gmail.com](mailto:vgretschko@gmail.com)