Discussion Paper No. 17-056

Contract (Re-)Negotiation with Private and Common Values

Vitali Gretschko and Achim Wambach

ZEW
Zentrum für Europäische
Wirtschaftsforschung GmbH
Centre for European
Economic Research

Discussion Paper No. 17-056

Contract (Re-)Negotiation with Private and Common Values

Vitali Gretschko and Achim Wambach

CONTRACT (RE-)NEGOTIATION WITH PRIVATE AND COMMON VALUES

VITALI GRETSCHKO AND ACHIM WAMBACH
ZEW MANNHEIM AND UNIVERSITY OF MANNHEIM

ABSTRACT. *We analyze the contracting problem of a principal who faces an agent with private information and cannot commit to not renegotiating a chosen contract. We model this by allowing the principal to propose new contracts any number of times after observing the contract choice of the agent. We propose a characterization of renegotiation-proof states of this (re-)negotiation and show that those states are supported by a perfect Bayesian equilibrium of an infinite horizon game. The characterization of renegotiation-proof states provides a tool, which is both powerful and simple to use, for finding such states in specific environments. We proceed by applying the results to adverse selection environments with private and common values. We show that with private values and common values of the 'Spence' type only, fully efficient and separating states can be renegotiation-proof. With common values of the "Rothschild-Stiglitz" type inefficient and (partial) pooling states may be renegotiation-proof.*

JEL classification: C73, C78, D82

Keywords: Principal-Agent models, renegotiation, Coase-conjecture

1. INTRODUCTION

The solution to the screening problem of a principal who is endowed with all the bargaining power and wishes to contract with a privately informed agent is well known. The principal proposes a menu of contracts that is designed such that the agent optimally chooses one of the contracts according to his type. Typically, the chosen contract is inefficient and the choice of the agent reveals information. In this case, both parties are able to benefit from immediate renegotiation after information has been revealed. As such renegotiation will be anticipated by the agent, it may distort his ex-ante incentives to accept any given contract

Previous versions of this article were circulated as "Renegotiation before Contract Execution" (CEPR Discussion Paper No.2189). We thank Geir Asheim, Patrick Bolton, Vitor Farinha-Luz, Georg Nöldecke, Ray Rees, Ariel Rubinstein, Klaus Schmidt, Monika Schnitzer, Roland Strausz, Thomas Tröger and the seminar participants at HU Berlin, Cologne University, Cologne Procurement Workshop (2015), EARIE (Munich 2015), Edinburgh University (2016), EEA conference (Mannheim 2015), JEI (Alicante 2015) and SED conference (Istanbul 2015) for helpful comments and suggestions. Christoph Gross-Boelting provided excellent research assistance. Financial support from the German Science Foundation (DFG) through the research unit "Design and Behavior" and the Fulbright Commission is gratefully acknowledged.

in the first place. Thus, the optimality of the solution to the screening problem crucially depends on the assumption that the chosen contract will not be renegotiated.

We introduce an extension of the one-shot screening problem in which the principal has limited commitment power. We model this by assuming that there are no physical costs of renegotiation and that any signed contract can be renegotiated any number of times. In each round of the (re-)negotation, the agent can decide to retain his current contract or to choose one of the new offers made by the principal.

Our main contribution is to characterize the set of renegotiation-proof states of the negotiation. A state of the negotiation is a tuple consisting of the current signed contract of the agent and the current belief of the principal that was formed by observing the previous choices of the agent. We focus on states rather than contracts as whether the principal would like to renegotiate the currently signed contract will crucially depend on her belief. Renegotiation-proof states are not identified one-by-one but simultaneously as a set. The key insight is that whether a state is renegotiation-proof or not will depend on if it can be improved by other renegotiation-proof states. That is, renegotiation-proof states can't be improved by other renegotiation-proof states. States that are not renegotiation proof can be improved by renegotiation-proof states. This implies that negotiation can stop even if there is room for Pareto improvement. This is always the case if this improvement leads to a state that is not renegotiation-proof as contracts in such states will be renegotiated and thus should not block the original contract choice.

The characterization of the set of renegotiation-proof states is based on two simple properties. First, for every renegotiation-proof state there is no other renegotiation-proof state that would make the principal strictly better off (*internal consistency*). Second, in any state of the negotiation game it is feasible to reach a renegotiation-proof state in a single round of further negotiations (*external consistency*).[1] Both properties reflect the sequential rationality of the principal. Suppose the negotiation game reaches a renegotiation-proof state and the principal proposes new contracts that would make her and the agent better off. External

[1]That renegotiation-proof states can be reached in a single round of further negotiations is without loss of generality. Without frictions, every history of the negotiation can be represented as a single-stage mechanism.

consistency ensures that in a further round of negotiations she will renegotiate these contracts. Internal consistency implies that the resulting outcomes do not make the principal better off than the original outcome would have.[2]

The proposed solution concept based on renegotiation-proof states is not meant to provide an alternative for the standard equilibrium techniques. Rather, it is consistent with them. We demonstrate this by deriving a perfect Bayesian equilibrium of an infinite horizon negotiation game that arises naturally from the principals lack of commitment.[3] In equilibrium, the strategy of the principal prescribes to end the negotiation and implement the current signed contract whenever a renegotiation-proof state was reached. Whenever the game reaches (off path) a state that is not renegotiation-proof the principal proposes a menu of contracts that leads to renegotiation-proof states given the agent chooses the contract optimally. The equilibrium strategy of the agent is then to choose the optimal contract given that choosing optimally induces beliefs of the principal that lead to a renegotiation-proof state.[4]

One of the main advantages of a general characterization of renegotiation-proof states is that it provides a powerful tool for the analysis of specific instances of the general problem.[5] We apply the characterization of renegotiation-proof outcomes to screening problems with private and common values. We show that while only fully separating and efficient states can arise with private values and common values of 'Spence' type, inefficient and pooling states can be renegotiation-proof with common values of the "Rothschild-Stiglitz" type.[6] To this end, we use internal and external consistency to derive a simple but useful necessary condition for a state to be renegotiation-proof. Given the information revealed in a renegotiation-proof state there should not exist a single feasible (pooling) state that would make the principal and the agent better off, irrespective of the type of agent. If such a state would exist, the principal could just offer the corresponding contract and all types of the agent could accept without revealing any additional information. Thus, this pooling state would make both

[2]Similar approaches to mechanism design without commitment were proposed by Asheim and Nilssen (1997) and Vartiainen (2013). We will comment on the similarities and differences in the literature overview below.
[3]Note that the number of potential negotiation rounds is unlimited. We cannot therefore use backward induction as in Bester and Strausz (2001) to apply a (modified) revelation principle.
[4]Even though the general reasoning is straightforward, many technical difficulties arise. For example, in our set-up, neither the one-shot deviation nor the revelation principle do hold. Thus, the construction of an equilibrium is one of the main contributions of this article.
[5]This is somewhat similar to Beaudry and Poitevin (1993) who first characterize the potential outcomes by a set of constraints and construct perfect Bayesian equilibria that achieve those outcomes.
[6]The used nomenclature was first introduced by Beaudry and Poitevin (1993).

parties better off without changing the strategic incentives in the negotiation game. This simple insight enables us to prove that with private values only fully separating and efficient states can be renegotiation-proof. This is a result of the fact that in any inefficient state, the indifference curve of the principal is either steeper or flatter than the indifference curves of the different types of agent. Thus, for every inefficient state there exists a single contract that would make the principal and all of the types of the agent better off without revealing additional information.

For common values we must distinguish two cases: common values of the "Spence" type and common values of the "Rothschild-Stiglitz" type. Common values of the 'Spence' type represent a situation in which the ranking of the marginal trade-offs between types of the agent is the same for the agent and the principal. This situation corresponds, for example, to the education model in Spence (1973) where education is both marginally more productive and less costly for the high type of the agent. In this case, for any inefficient state, the indifference curves of the principal are either both steeper or both flatter than the indifference curves of both types of the agent. Thus, the same logic as with private values applies and for any inefficient state a single contract exists that would result in the principal and all types of the agent being better off. It follows that only fully separating and efficient states can be renegotiation-proof.

The situation changes if common values of the "Rothschild-Stiglitz" type are considered. Common values of the "Rothschild-Stiglitz" type represent the situation that the marginal trade-offs between the types of agent are ranked differently for the agent and the principal. This situation corresponds, for example, to the insurance model developed by Rothschild and Stiglitz (1976) where insurance is marginally less costly but also marginally less valuable for the low risk type of agent. In this case, the logic described above is not applicable as there are inefficient states such that no single pooling contract would see both parties better off. Moreover, it may be the case that all pairs of efficient states would result in the principal being strictly worse off than he would be in the initial situation. The set of all efficient states therefore lacks the external consistency property. We proceed by constructing a set of renegotiation-proof states that results in inefficient (partial) pooling states. Interestingly, inefficient states can be sustained even if there are pairs of efficient states that would be weakly more advantageous to the principal and the agent.

Relation to the Literature. The majority of previous analyzes of renegotiation have typically taken one of two approaches. Either renegotiation-proof states were characterized by ad-hoc assumptions or renegotiation was limited to finite negotiation protocols. The first approach usually yields clear-cut results in complex settings and serves as a powerful tool for the analysis of specific problems. However, the lack of foundation of renegotiation-proof states as equilibrium outcomes of a non-cooperative game may raise doubts.[7] The second approach allows for equilibrium analysis but still leaves the principal with a considerable amount of commitment power.[8] In our frictionless setting for example, 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 subsequently propose the optimal contracts. Our approach combines the clarity and power of an axiomatic approach with the equilibrium analysis of an infinite horizon negotiation protocol. 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.

---

[7]For examples of such an approach see Asheim and Nilssen (1997), Neeman and Pavlov (2013), and Vartiainen (2013).

[8]For examples of such an approach see Evans and Reiche (2015), Fudenberg and Tirole (1990b), Hart and Tirole (1988), Hörner and Samuelson (2011), Skreta (2006), or Skreta (2015).

In Gretschko and Wambach (2016) we apply the characterization of renegotiation-proof states to a model with private values and and infinite type space. We show that with an infinite type space even in the private values case, inefficient states can be sustained as renegotiation-proof. This is a consequence of the revenue equivalence theorem. In both cases, with discrete and continuous types, efficient and separating states need to be in any set of renegotiation-proof states. After any history the principal can use VCG mechanisms where the agent's payment is the cost of provision to implement the efficient solution. By the revenue equivalence theorem, this payment schedule is unique up to an additive constant for the continuous type space. Hence, the principal makes a zero profit along the efficient path. Thus, the principal can commit to some inefficient outcomes, as renegotiating to efficient outcomes would generate no additional profit for her. For discrete types, revenue equivalence fails and the maximal profit is positive. Thus, renegotiating to efficient outcomes makes the principal always better off.

Our set-up is closely related to that considered by Beaudry and Poitevin (1993) who study the effects of immediate and unlimited renegotiation in a general signaling model. In contrast to our work, therefore, they consider the situation in which the informed agent is able to make the contract offers. In this case, separating but inefficient contracts are sustained by the threat that if the agent who signed an inefficient contract proposes a new contract, a switch in beliefs takes place and the uniformed party will assume that the agent is of the undesired type. However, this can only work if it is the informed party who makes the offers. We complement Beaudry and Poitevin (1993) by extending the analysis to a screening model. That is, we assume that the uninformed party makes all the offers.

Deneckere and Liang (2006) consider an uninformed buyer who bargains about the price of an object with a seller who has perfect information about the common value of this object. They consider an infinite horizon bargaining game in which the uninformed buyer makes all the offers and show that when bargaining frictions disappear the solution does not necessarily converge to be efficient. Gerardi et al. (2014) consider the same model but assume that the informed party makes all the offers. They find that the buyer's inability to commit before observing the terms of trade is what precludes efficiency. Our analysis complements these articles by generalizing their environment to different models of private and common values.

Strulovici (2016) analyzes an infinite horizon negotiation protocol for a set-up with private values. In contrast to our work, negotiations may exogenously break down. In other words, renegotiation is not frictionless. Strulovici (2016) shows that if such friction disappears, efficient and fully separating contracts arise in any perfect Bayesian equilibrium of the negotiation game. Our approach complements his analysis by considering more general environments with private and common values and showing that with common values frictionless renegotiation can lead to an inefficient solution.

Using 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 contracts arise in equilibrium.

With respect to the characterization of the renegotiation-proof states, our study is related to Asheim and Nilssen (1997) and Vartiainen (2013). Asheim and Nilssen (1997) consider a monopolistic insurance market and Vartiainen (2013) an auction without commitment. In both cases, the assumptions used for the characterization of renegotiation-proof states resemble those used in our study. That is, both rely on properties similar to internal and external consistency in order to characterize renegotiation-proof states. In both cases, this approach proves to be very useful in deriving clear results for otherwise very complex problems. We extend their analysis by providing a foundation of renegotiation-proof states as a perfect Bayesian equilibrium of a very general negotiation game and applying the results to settings not considered by those authors.

Krasa (1999) applies a slightly different, but also axiomatic approach. Krasa (1999) defines a state as unimprovable (renegotiation-proof) if agents would not want to deviate from it either by changing the allocation or by revealing (additional) information. Application of this characterization of renegotiation-proof states to an insurance monopoly with two types of agent, reveals thats agents either reveal information fully or not at all. This results in either full separation or full pooling of types. This is not the case in our model as in the "Rothschild-Stiglitz" case, partial pooling can be supported.

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 has to be ex-post efficient to survive renegotiation under any renegotiation procedure. This assumption places more restrictions on equilibrium outcomes than our approach in which a specific renegotiation procedure is fixed. Thus, in contrast to Neeman and Pavlov (2013), our approach allows for inefficient outcomes as can be seen in the case of common values of the "Rothschild-Stiglitz" type.

2. The Setup

A principal (she) and an agent (he) negotiate over a contract. A contract is a tuple $w \in \mathbb{R}^2$. Let $\theta \in \{L, H\}$ denote the type of the agent. The type is private knowledge to the agent and the principal has a prior characterized by $\mu_0 = Pr(\theta = H)$. When a contract w is signed by an agent of type θ, the utility of the principal amounts to $v(w, \theta)$. The utility of the agent is then $u(w, \theta)$. If no contract is implemented, both parities receive the outside option contract w_0. Both, $v(w, \theta)$ and $u(w, \theta)$ are assumed to be quasiconcave in w. Let $v_i(w, \theta)$ and $u_i(w, \theta)$ denote the partial derivative with respect to the i-th component of w. The principal prefers smaller values of w_1 and larger values of w_2, whereas the opposite is true for the agent, that is,

$$v_1(w, \theta) < 0, \;\; v_2(w, \theta) > 0, \;\; u_1(w, \theta) > 0, \;\text{ and } u_2(w, \theta) < 0.$$

The functions $u(w, \theta)$ satisfy the standard single-crossing condition, that is,

$$-\frac{u_2(w, L)}{u_1(w, L)} > -\frac{u_2(w, H)}{u_1(w, H)}. \tag{1}$$

A contract is θ-efficient if it is the cheapest contract providing an agent of type θ with a given utility level. That is, the iso-utility curve of the principal is tangent to the iso-utility curve of the agent in any such contract. For each θ, denote by ξ_θ the set of all θ-efficient contracts. Sometimes we will refer to ξ_θ as the efficient contract curve.

Whenever v is independent of θ, that is, $v(w, \theta) \equiv v(w)$, we will refer to private values. In this case, due to the single-crossing property, the efficient contract curve ξ_L lies to the left of ξ_H. Whenever the utility of the principal explicitly depends on the type of the agent, we will refer to common values. For common values we will distinguish two cases: the 'Spence' case and the "Rothschild-Stiglitz" case.

"Spence" case: common values of the 'Spence' type represent the situation that the ranking of the marginal trade-offs between types is the same for the agent and the principal, i.e.,

$$-\frac{v_2(w,L)}{v_1(w,L)} < -\frac{v_2(w,H)}{v_1(w,H)}. \tag{2}$$

This situation corresponds, for example, to the education model in Spence (1973), where education is both marginally more productive and less costly for the H type. It follows from equation (1) and equation (2) that the efficient contract curves ξ_θ do not cross and ξ_L lies to the left of ξ_H.

"Rothschild-Stiglitz" case: common values of "Rothschild-Stiglitz" type represent situations where the marginal trade-offs are ranked differently for the informed and the uninformed player, i.e.,

$$-\frac{v_2(w,L)}{v_1(w,L)} > -\frac{v_2(w,H)}{v_1(w,H)}. \tag{3}$$

This situation corresponds, for example, to the insurance model by Rothschild and Stiglitz (1976), where insurance is marginally less costly but also marginally less valuable for the high type of agent, i.e., the lower risk type of agent. Given this generality, in the "Rothschild-Stiglitz" case it may be the case that the efficient contract curves cross and the efficient contract curve of the L type lies to the right of the efficient contract curve of the H type. In the majority of the applications of models with common values of the "Rothshild-Stiglitz" type, the efficient contract curves do not cross. We therefore assume that in the "Rothschild-Stiglitz" case, the efficient contract curve of the L type lies weakly to the left of the efficient contract curve of the H type.[9] We conjecture that most of the analysis can be extended to the cases in which this assumption does not hold. However, this comes at the cost of severely complicating the exposition of the main results.

We partition the contract space into three regions. We say that contracts that are left of ξ_L are in the 'H-Rent' configuration, contracts that are to the right of ξ_H are in the 'L-Rent' configuration, contracts that are in the inner region between ξ_L and ξ_H are in the 'No-Rent'

[9]This greatly simplifies the the exposition of the the strategies of the agent in the proof of Proposition 1 ensuring that only one of the types of the agent uses a mixed strategy. That the efficient contract curve of the L type lies to the left of the efficient contract curve of the H type can be ensured by assuming, for example, that the utility functions of the principal and the agent are additively separable, i.e., $u(w,\theta) = f(w_2,\theta) + \lambda_\theta w_1$ and $v(w,\theta) = g(w_2,\theta) - w_1$ with $\lambda_\theta \in \mathbb{R}_+$ and that if $-f_1(w_2,L)/\lambda_L = g_1(w_2,L)$ then $-f_1(w_2,H)/\lambda_H < g_1(w_2,\theta_H)$.

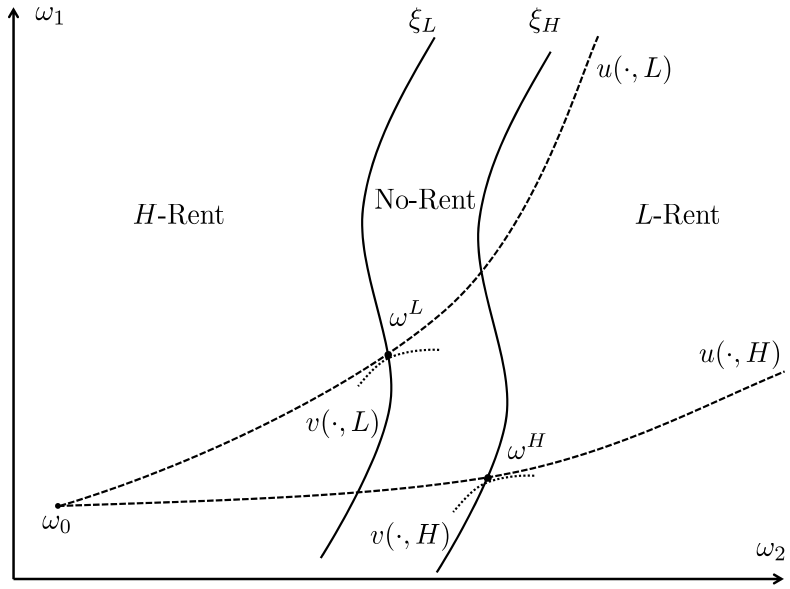

FIGURE 1. L-type and H-type efficient contracts that lie on the same indifference curves as the outside option of the agent.

configuration. Essentially, the partition of the contract space reflects that once a contract has been signed in the 'H-Rent' configuration for example, further negotiation will leave the H type with a positive rent whereas the L type gains nothing from further negotiation.[10] Figure 1 depicts an illustration of this set-up.

Mechanisms. To elicit information from the agent and implement a contract the principal uses a mechanism. A mechanism is a tuple $\mathcal{M} = (\mathcal{Z}, w(\cdot))$ consisting of a set of messages $\mathcal{Z}$ and a function $w : \mathcal{Z} \to \mathbb{R}^2$. A mechanism works as follows. The agent chooses a message $z \in \mathcal{Z}$. When the message is sent, $w(z)$ generates a contract. We will consider direct communication only.[11] Thus, it is without loss of generality that every mechanism can be written as a menu of contracts, $\mathcal{M} = \{w(z)\}_{z \in \mathcal{Z}} \subset \mathbb{R}^2$ from which the agent chooses directly.

The Problem. The problem for the principal is that she cannot commit to not renegotiating w after the the agent choose a contract w. That is, after the agent chooses from $\mathcal{M}$ and the principal observes w, she will update her belief about the type of the agent. After observing w the principal may propose a new menu of contracts $\mathcal{M}'$. The agent can then decide whether

[10]The designation 'H-Rent', 'L-Rent', and 'No-Rent' configuration provides a very vivid definition and is taken from Strulovici (2016).
[11]For an analysis of contracting with renegotiation an mediated communication, see for example Pollrich (2016) or Strausz (2012).

he wants to choose a new contract or whether he wants to hold on to the initial contract. In other words, w is the new outside-option contract of the agent. If the agent decides to choose from $\mathcal{M}'$, the principal again observes the contract, updates her belief, and may again renegotiate this contract by proposing a new menu of contracts. Overall, the principal is not able to commit to not renegotiating any contract. Whenever the agent choses a contract, the principal may propose a new menu of contracts and the agent may decide to either hold on to his current contract or to choose a new one. Thus, we are concerned with the question of what menus of contracts the principal will not renegotiate at the ex-post stage.

To be more precise, consider the following game. In $t = 0$ the agent observes his type θ. In each following round $t \in \mathbb{N}_+$ the principal offers a menu $\mathcal{M}_t$ of contracts, with $\mathcal{M}_t$ a subset of $\mathbb{R}^2_+$. The number of contracts $|\mathcal{M}_t|$ is bounded by an arbitrary constant $K \geq 2$ that is fixed throughout the negotiation.[12] The agent chooses a contract in $\mathcal{M}_t$ or decides keep the contract he chose in round $t-1$. We denote by w_t the contract that the agent chose in round t and by w_0 the initial contract, that is the normalized outside option contract of the agent. The game ends if at time t the principal does not propose new contracts, that is, if $\mathcal{M}_t = \emptyset$. In this case w_{t-1} is executed. Thus, the timing of the negotiation game is as follows.

(i) In $t = 0$ the agent learns his type θ.

(ii) In $t > 0$

  (a) the principal offers a menu of contracts $\mathcal{M}_t \subset \mathbb{R}^2_+$ with $|\mathcal{M}_t| \leq K$

  (b) the agent chooses a contract $w_t \in \mathcal{M}_t$ or decides to keep the current contract w_{t-1}.

(iii) The game ends if $\mathcal{M}_t = \emptyset$, in this case the last chosen contract w_{t-1} is executed.

Denote a potential history realized before the principal moves in round t as

$$h^p(t) = \{(\mathcal{M}_1, w_1), (\mathcal{M}_2, w_2), \ldots, (\mathcal{M}_{t-1}, w_{t-1})\}.$$

A potential history realized before the agent moves is

$$h^a(t) = \{(\mathcal{M}_1, w_0), (\mathcal{M}_2, w_1), \ldots, (\mathcal{M}_t, w_{t-1})\}.$$

[12]As contracts can be infinitely renegotiated there is no guarantee that proposing only two contracts is without loss of generality, that is, the results obtained by Bester and Strausz (2001) do not apply to the setting at hand.

In round 1 there is no relevant history for principal, so $h^p(1) = \emptyset$. Denote by $h_k(t)$ the restriction of $h(t)$ to the first k rounds. Let $\mathcal{H}_a^t$ be the set of all histories in round t and $\mathcal{H}_a$ be the set of all terminal histories for the agent and let $\mathcal{H}_p^t$ be the set of all histories in round t and $\mathcal{H}_p$ be the set of all terminal histories for the principal.[13] That is, $h(t) \in \mathcal{H}$ if $\mathcal{M}_t = \emptyset$. Moreover, all infinite histories are terminal histories. That is, $h(\infty)$ is an infinite history if there exist a series of histories such that $h(t) \to h(\infty)$ and $\mathcal{M}_t = \emptyset$ is not in $h(t)$ for all t. For every $h(t) \in \mathcal{H}$, we define the payoff of the agent and the principal as $u(h(t), \theta) = u(w_{t-1}, \theta)$ and $v(h(t), \theta) = v(w_{t-1}, \theta)$ if $t < \infty$ and as $u(h(\infty), \theta) = v(h(\infty), \theta) = -\infty$.

By modeling renegotiation in this way we intend to capture the effects of unlimited, immediate and frictionless renegotiation where the uninformed party makes all the offers. We therefore complement the work of Beaudry and Poitevin (1993) who model a similar renegotiation game for the case in which the informed party makes all the offers. The negotiation consists of four elements. Firstly, the uninformed party makes both the initial offer and all following propositions of renegotiation. Secondly, after the agent chooses one of the offers there is at least one more round of offers and neither the principal nor the agent can commit not to renegotiate. Thus, the renegotiation process can potentially last for an infinite number of rounds. Thirdly, however, the negative pay-off at infinite histories prevents the principal from stalling the negotiation indefinitely.[14] Finally, only the final signed contract is pay-off relevant and there is no discounting in-between negotiation rounds. Hence, the focus is on the effects of renegotiation rather than long-term relationships.

Strategies and beliefs. Before we discuss the solution of the game it is useful to define strategies and beliefs of the principal and the agent. Denote by $\mathcal{A}$ the set of all subsets of $\mathbb{R}_+^2$ with at most K elements. A behavior strategy σ^p of the principal prescribes in each round t a distribution over contract menus $\mathcal{M}_t \in \mathcal{A}$ conditional on the history $h^p(t)$.[15] That is, σ^p

[13]We will drop the superscript from h^p and h^a whenever we refer to both or whenever it is unambiguous whose history is used.

[14]The definition of the pay-off of infinite histories seems harsh at first glance. However, with an unrestricted contracting space, off equilibrium path, the negotiation could reach contracts that yield an arbitrary low pay-off to the principal. In this case, the principal is better off by stalling the negotiation ad infinitum if the pay-off of such a strategy is bounded from below. Alternatively, we can assume that pay-offs of infinite histories are bounded from below if we assume that the contracting space is bounded. However, this yields technical difficulties close to the bounds of the contracting space that would greatly complicate the notation in the proofs. Thus, for the sake of clarity of exposition we decided to opt for the current model specification.

[15]We endow $\mathcal{A}$ with the Borel sigma-algebra and denote by $\Delta(\mathcal{A})$ the set of all probability measures over $\mathcal{A}$.

is a sequence of maps σ_t^p with

$$\sigma_t^p(h^p(t)) : \ \mathcal{H}_p^t \to \Delta(\mathcal{A}).$$

A behavior strategy σ^θ of an agent of type θ prescribes in each round t a probability distribution over contracts in $\mathcal{M}_t \cup \{w_{t-1}\}$ conditional of the history $h^a(t)$. That is, σ^θ is a sequence of maps σ_t^θ with

$$\sigma_t^\theta(h^a(t)) : \ \mathcal{H}_a^t \to \Delta\left(\mathcal{M}_t \cup \{w_{t-1}\}\right).$$

In this case, $\Delta(\mathcal{M}_t \cup \{w_{t-1}\})$ denotes the set of all probability distributions over $\mathcal{M}_t \cup \{w_{t-1}\}$. A continuation strategy $\sigma_+^{\{\theta,p\}}(t)$ is a truncated strategy. For example, $\sigma_+^p(t) = \left\{\sigma_t^p, \sigma_{t+1}^p, \sigma_{t+2}^p, \ldots\right\}$.

The belief system of the principal is a sequence $\{\mu_0, \mu_1, \ldots\}$ where $\mu_{t-1} \in [0,1]$ are the beliefs held after a history $h^p(t)$ that the agent is of type H.[16]

States and outcome functions. There are two additional concepts which it is useful to define before we turn to the solution of the game. These are the state of the negotiation and the outcome function. A state of the negotiation in round t for a given pair of strategies and belief system is $C_t = (w_{t-1}, \mu_{t-1})$. w_{t-1} denotes the current signed contract and μ_{t-1} the belief of the principal. The set of all states is denoted by Γ and $\pi(C_t) = (1-\mu_{t-1})v(w_{t-1}, L) + \mu_{t-1}v(w_{t-1}, H)$ denotes the expected utility of the principal in a given state if the negotiation where to end. For a given history and pair of strategies the outcome function $f(h^p(t), \sigma_+^p(t), \sigma_+^\theta(t)) \subset \Gamma$ gives the set of states after which the negotiation ends. That is, $C = (w, \mu)$ is in $f(h^p(t), \sigma_+^p(t), \sigma_+^\theta(t))$, if there exists a $t' \geq t$ such that a state $C_{t'} = C$ is reached with positive probability starting from $h^p(t)$ and $\sigma^p(h^p(t')) = \emptyset$.

3. Solution concept: Renegotiation-proof states

In this section we will identify renegotiation-proof states. That is, states $C_t = (w_t, \mu_t)$, such that the principal will not renegotiate the contract w_t. It is crucial to focus on renegotiation-proof *states* (w_t, μ_t) rather then renegotiation-proof *contracts* w_t as wether the principal will

[16]We slightly abuse notation as we suppress that different histories in period t might lead to a different posterior.

want to renegotiate will not only depend on the current contract but also crucially on the current belief.

We start by observing that every subgame after some history $h(t)$ can be represented by a single stage mechanism that the principal will not renegotiate. In any subgame after $h(t)$ the strategy of the principal together with the strategy of the agent defines a distribution over potential states $\bigcup_{t' \geq t} f(h^p(t'), \sigma_+^p(t'), \sigma_+^\theta(t'))$. Instead of playing out the game according to those histories, the principal can just offer a (countable) menu of contracts $\mathcal{M}_{t+1}$ that includes all the contracts that are part of the states in $\bigcup_{t' \geq t} f(h^p(t'), \sigma_+^p(t'), \sigma_+^\theta(t'))$ and the agent chooses between those contracts (possibly mixing) in a way that induces the same probability distribution over states as playing out the game would. Thus, we will focus on strategies of the principal that prescribe to offer a single-stage mechanism after each history. The problem then boils down to identify states after a single-stage mechanism that the principal would not like to renegotiate given that the agent chooses optimally in a single-stage mechanism. We will call such states renegotiation-proof states, denote the set of renegotiation-proof states by Ω and derive its properties in the following.

Agent's incentives. Suppose at history $h(t)$ that the current state of negotiation is $C_t = (w_t, \mu_t)$ and the principal offers a menu of contracts M_{t+1} that all result in renegotiation-proof states given the optimal choice of the agent. That is, the game will be over if the agent chooses a contract. Then the agent should choose optimally in the single-stage mechanism. We define which states can be induced by a menu of contracts in which the agent chooses optimally. That is, we define conditions on a set of states such that there exists a menu of contracts that, if the agent choose optimally, generates this set of states. Those conditions then constitute necessary conditions for renegotiation-proof states.

Definition 1 (*Feasibility*)**.** Let $\mathcal{Z}$ denote a countable index set. We call $\{C^z = (w^z, \mu^z) : z \in \mathcal{Z}\} \subset \Gamma$, feasible starting from $C_t = (w_t, \mu_t)$ if the following conditions are satisfied:

(i) *(Individual rationality of the agent) For all $\theta \in \{L, H\}$ there exists a $z \in \mathcal{Z}$ such that $u(w^z, \theta) \geq u(w, \theta)$*

(ii) (*Incentive compatibility*) If there exists a z and a z' in $\mathcal{Z}$ such that $u(w^z, H) > u(w^{z'}, H)$, then $\mu^{z'} = 0$ or if there exists a z and a z' in $\mathcal{Z}$ such that $u(w^z, L) > u(w^{z'}, L) \Rightarrow \mu^{z'} = 1$

(iii) (*Bayesian consistency*) There exists $\{p^z : z \in \mathcal{Z}\}$ with $p^z \in [0,1]$ such that $\sum_{z\in\mathcal{Z}} p^z \mu^z = \mu$.

We define by $IC(C_t) : \Gamma \to 2^\Gamma$ the mapping from some C_t to all feasible states starting from this C_t. That is, $\{C^z = (w^z, \mu^z) : z \in \mathcal{Z}\}$ is in $IC(C_t)$ if it satisfies conditions (i) to (iii).

For $\{C_z = (w_z, \mu_z) : z \in \mathcal{Z}\}$ to be states that can be generated a mechanism starting from C_t it is necessary that the agent is weakly better off compared to the initial situation in state C_t (requirement (i)). As the resulting states should be renegotiation-proof and the principal will end the negotiation, the agent optimally chooses the contract that is most desirable to him. The principal must therefore take this into account when updating her belief (requirement (ii)). From the ex-ante point of view of the principal, the probability of reaching state C^z is some $p^z \in [0,1]$ that should be consistent with bayesian updating (requirement (iii)) and the strategy of the agent.

Principal's strategy. We will take the view that bygones are bygones and consider only strategies that are history independent other than with respect to the current outside option contract of the agent and the current belief of the principal. That is, we will consider strategies such that if $h(t)$ and $h'(t)$ lead to the same state C_t then $\sigma^p(h(t) = \sigma^p(h'(t))$. From now on we will drop $h(t)$ from the notation and simply denote the strategy of the principal by $\sigma^p(C_t)$. As argued above, in our solution concept, we focus on strategies of the principal that prescribe her to offer a single-stage mechanism that results only in renegotiation-proof states and end the negotiation afterwards. Let Ω denote the set of renegotiation-proof states, that we have yet to define. Thus, we focus on strategies such that if $C_t \notin \Omega$, then $\sigma(C_t) = \mathcal{M}_t = \{w^z : z \in \mathcal{Z}\}$ for some index set $\mathcal{Z}$ with $\{C^z = (w^z, \mu^z) : z \in \mathcal{Z}\} \in IC(C)$ and $\{C^z = (w^z, \mu^z) : z \in \mathcal{Z}\} \subset \Omega$. If $C_t \in \Omega$, then $\sigma(C_t) = \emptyset$. We now turn our attention to the definition of Ω, the set of renegotiation-proof states, such that the principal can indeed commit to not renegotiating those states. To characterize renegotiation-proof states it is convenient to introduce some notation on which feasible states make the the principal weakly better off when proposing an new menu of contracts in state C_t.

Definition 2. Let $\mathcal{Z}$ denote a countable index set. A feasible set of states

$$\{C^z = (w^z, \mu^z) : z \in \mathcal{Z}\} \in IC(C_t)$$

makes the principal weakly better off starting from $C_t = (w_t, \mu_t)$ if

$$\pi(C_t) \leq \sum_{z \in \mathcal{Z}} p^z \pi(C^z). \tag{4}$$

We define by $X(C_t) : \Gamma \to 2^\Gamma$ the mapping from some C_t to all feasible states that make the principal better off starting from C_t. That is, $\{C^z = (w^z, \mu^z) : z \in \mathcal{Z}\}$ is an element of $X(C_t)$ if $\{C^z = (w^z, \mu^z) : z \in \mathcal{Z}\}$ is an element of $IC(C_t)$ and satisfies inequality (4).

We are now in the position to define renegotiation-proof states. Renegotiation-proof states are not identified one-by-one but simultaneously as a set. The key insight is that whether a state is renegotiation-proof or not will depend on if it can be improved by other renegotiation-proof states. That is, renegotiation-proof states can't be improved by other renegotiation-proof states. States that are not renegotiation proof can be improved by renegotiation-proof states. This implies that negotiation can stop even if there is room for Pareto improvement. This is always the case if this improvement leads to a state that is not renegotiation-proof as contracts in such states will be renegotiated and thus should not block the original contract choice. Thus, the set of renegotiation-proof states should have two properties. First, independent of the current state of the negotiation it should be feasible to reach a renegotiation-proof state. This is to ensure that our solution concept is well defined. Second, whenever the negotiation has reached a renegotiation-proof state, that is a state after which negotiation should end, the principal should not be better off by renegotiating to another renegotiation-proof state after which the negotiation should end. This is to ensure that the principal will end the negotiation after reaching a renegotiation proof state. The following definition formalizes these conditions.

Definition 3 (*Renegotiation-proofness*)**.** Let $\mathcal{Z}$ denote a countable index set. $\Omega \subset \Gamma$ is a set of *renegotiation-proof states* if the following holds true.

(i) *(External consistency)* For all $C \in \Gamma$ there exist

$$\{C^z = (w^z, \mu^z) : z \in \mathcal{Z}\} \in X(C)$$

with $C^z \in \Omega$. That is, there is $\{C^z = (w^z, \mu^z) : z \in \mathcal{Z}\} \in IC(C)$ that makes the principal better off and leads only to renegotiation proof states.

(ii) *(Internal consistency)* From $C \in \Omega$ follows $\pi(C) \geq \sum_{z\in\mathcal{Z}} p^z \pi(C^z)$ for all

$$\{C^z = (w^z, \mu^z) : z \in \mathcal{Z}\} \in IC(C_t)$$

with $C^z \in \Omega$. That is, there is no $\{C^z = (w^z, \mu^z) : z \in \mathcal{Z}\}$ in Ω that makes the principal better off.

The restrictions we place on the set of renegotiation-proof states reflect sequential rationality of the principal. Thus, our solution concept is not meant to provide an alternative for the standard equilibrium techniques. Rather, it is consistent with them. Indeed, in Section 4 below, we will show that for each set of renegotiation-proof states Ω there exists a perfect Bayesian equilibrium of the negotiation game, such that the game ends if and only if the game has reached a renegotiation-proof state in Ω. To see how external and internal consistency imply sequential rationality suppose that the negotiation has reached a renegotiation-proof state and the principal deviates by proposing new feasible contract that make her better off. External consistency ensures that if the principal would follow her equilibrium strategy after the deviation, there are contracts that make her even better off. Internal consistency then implies that the resulting contracts make her not better off than the original state would have.[17]

The key idea is that states that are not renegotiation proof but can be improved by renegotiation-proof states shall not obstruct the principal's contract choice. That is, the principal will not renegotiate a menu of contracts even if the choices of the agent lead to states that could be improved by further renegotiation. This is the case if those states can be further improved by offering a menu of contracts that will lead to renegotiation-proof states. In particular, our solution concept allows for inefficient states to be renegotiation proof. In fact, as we will show below, the optimal mechanism from the point of view of the principal will be inefficient.

[17]The definition of the set of renegotiation-proof states 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. With a similar logic, internal consistency ensures that payoffs of different feasible states that are in Ω cannot make the principal strictly better off without making the agent strictly worse off.

4. Renegotiation-proof states in a perfect Bayesian equilibrium

In this section we will show that renegotiation-proof states arise naturally in a perfect Bayesian equilibrium of the negotiation game as defined in Section 2. For each set of renegotiation-proof states we will construct a perfect Bayesian equilibrium of the negotiation game in which the principal stops the negotiation whenever a renegotiation-proof state was reached and the agent chooses optimally between the offered contracts. In order to construct an equilibrium of the negotiation game we work our way backwards. For this we need to define the strategy of the principal in a way such that the principal behaves sequentially optimal. Thus, suppose an Ω that satisfies the conditions of Definition 3 exists and define for each state $C = (w, \mu)$

$$\begin{aligned} s(C) = &\arg \max_{(C^L, C^H)} p\pi \left(C^L\right) + (1-p)\, \pi \left(C^H\right) \\ \text{s.t. } &\left(C^L, C^H\right) \in X\left(C\right) \\ &C^\theta = \left(w^\theta, \mu^\theta\right) \in \Omega \end{aligned}$$

as the optimal feasible states in Ω starting from state C for each type of the agent.[18] In what follows we will slightly abuse notation and for $s(C) = ((w^H, \mu^H), (w^L, \mu^L))$ define $s^\theta(C) := w^\theta$.

To ensure that the problem is well-behaved we make three assumptions. Imposing these assumptions at this point is convenient as it greatly simplifies notation and allows us to make the statement of the main theorem as general as possible. In Section 5 we show that all of the assumptions are satisfied for all of the considered applications.

Assumption 1. *For all C in Γ, s(C) exists.*

Assumption 2. *For every $C = (w, \mu)$ one of the following holds true*

(i) If w is in the 'H-Rent' configuration, $u(s^L(C), L) = u(w, L)$ and $u(s^H(C), H) \geq u(w, H)$.

(ii) If w is in the 'L-Rent' configuration, $u(s^H(C), H) = u(w, H)$ and $u(s^L(C), L) \geq u(w, L)$.

[18] If $C \in \Omega$, $s(C) = (C, C)$.

(iii) If w is in the 'No-Rent' configuration, $u(s^L(C), L) = u(w, L)$ and $u(s^H(C), H) = u(w, H)$.

Assumption 3. *For $\theta \neq \theta' \in \{L, H\}$ and any two contracts $(w^\theta, w^{\theta'})$ and belief $\mu \in (0, 1)$, if*

$$u(s^\theta(w^\theta, 1), \theta) < u(s^\theta(w^{\theta'}, 0), \theta)$$

and

$$u(s^\theta(w^\theta, 1), \theta) > u(s^\theta(w^{\theta'}, \mu), \theta)$$

then there exists $\rho \in (0, 1)$ such that

$$u(s^\theta(w^\theta, 1), \theta) = u(s^\theta(w^{\theta'}, \rho\mu), \theta).$$

Assumption 1 ensures that $s(C)$ exists. As a result of external consistency, Assumption 1 is always satisfied whenever Ω is a closed set. Assumption 2 states that depending on the current state, the optimal renegotiation-proof state leaves at least one of the types without additional rent. Assumption 3 requires some more explanation. Suppose the principal offers two (out of equilibrium) contracts w^θ and $w^{\theta'}$ and the agent chooses one of the contracts. After observing the choice of the agent the principal updates her belief, offers the optimal contracts that lead to renegotiation-proof states and the negotiation subsequently ends. Suppose furthermore that in this case it is optimal for a forward-looking agent to choose $w^{\theta'}$ with probability one if he is of type θ'. Assumption 3 then ensures that if it is not strictly profitable for an agent of type θ to take either w^θ or $w^{\theta'}$ with probability one there must exist a mixing probability such that if the principal believes that the agent mixes with such a probability, the agent is indifferent between $s^\theta((w^\theta, 1))$ and $s^\theta(w^{\theta'}, \rho\mu)$.

Proposition 1. *Suppose Assumptions 1 to 3 hold true and let $(C^L, C^H) = s(C_0)$ with $C^\theta = (s^\theta(w_0, \mu_0), \mu^\theta)$. For each Ω that satisfies Definition 3 there exists a Perfect Bayesian Equilibrium with equilibrium strategies σ^p and σ^θ such that $\Omega = \bigcup_{t\geq 0} f(h^p(t), \sigma^p_+(t), \sigma^\theta_+(t))$ the equilibrium path is characterized by:*

(i) The principal offers in the first round $\sigma^p(C_0) = \{s^L(w_0, \mu_0), s^H(w_0, \mu_0)\}$.

(ii) The agent of type L chooses contract $s^L(w_0, \mu_0)$ with probability

$$p^L = \frac{(1 - \mu^L)(\mu^H - \mu_0)}{(1 - \mu_0)(\mu^H - \mu^L)},$$

that is, $\sigma^L(\{s^L(w_0,\mu_0), s^H(w_0,\mu_0), w_0\}, C_0) = (p^L, 1-p^L, 0)$.[19]

(iii) The agent of type H chooses contract $s^H(w_0,\mu_0)$ with probability

$$p^H = \frac{\mu^H(\mu_0 - \mu^L)}{\mu_0(\mu^H - \mu^L)},$$

that is, $\sigma^H(\{s^L(w_0,\mu_0), s^H(w_0,\mu_0), w_0\}, C_0) = (1-p^H, p^H, 0)$.

(iv) The negotiation ends in the following round as the principal does not propose a new contract.

Proof. The proof can be found in Appendix A. □

The proof of Proposition 1 is a direct consequence of Definition 3. Firstly, the strategy of the principal prescribes that in any current state of the negotiation that is not in Ω she offers the optimal contracts that lead to feasible states in Ω. Whenever the negotiation reaches a state in Ω the principal ends the negotiation. If the agent observes an offer that on the equilibrium path would lead to a state in Ω, he chooses the contract that is optimal given his type (possibly mixing when indifferent). If the agent observes an offer that would lead to a state which is not in Ω, he chooses the contract that will lead to the optimal offer for him in the next period. The key at this point is to establish that for any deviating proposal of the principal there is one type of agent who is best off choosing one of the contracts with probability one. This is a direct consequence of single crossing and Assumption 2. For the other type of agent it is then optimal to mix between this contract and at most one other contract from the proposal (Lemma 5 in the appendix). Secondly, we show that to prove that the proposed strategies form a perfect Bayesian equilibrium it is sufficient to consider only one-stage deviations of the principal. This is the case despite the fact that the game is not continuous at infinity (Lemma 6 in the appendix). Thirdly, from external consistency of Ω it follows that once the principal has deviated and offered a contract that leads to a state that is not in Ω, there exists a feasible state in Ω that would make her better off. From internal consistency of Ω it follows that once a state in Ω has been reached, there is no profitable deviation that would lead to another state in Ω. Overall, any deviation can be improved by negotiating to a feasible state in Ω. Thus, rather than reaching a state via a deviation, the principal could have negotiated directly to this state. Hence, offering a menu of contracts

[19]Recall that σ^θ is a mapping from $(\mathcal{M}_t, w_{t-1})$ to $\Delta(\mathcal{M}_t \cup \{w_{t-1}\})$.

other than those prescribed by the equilibrium strategy does not make the principal better off.

5. Applications

One of the main advantages of our approach is that in order to apply the results of Theorem 1 to specific principal-agent problems we merely need to construct the set Ω of renegotiation-proof states. That is, we need to construct a set Ω that is internally and externally consistent. In the following we provide three helpful results which will facilitate the construction of Ω.

Definition 4. A state C of the negotiation is called efficient if $X(C) = \{C\}$.

That is, in an efficient state C there does not exist a set of states which both the principal and the agent would weakly prefer. Thus, such a state must be part of any set of renegotiation-proof states:

Lemma 1. *For every set Ω that satisfies the conditions of Definition 3 and every efficient state C it holds that C is an element of Ω.*

Proof. Follows directly from external consistency. If C is the only element of $X(C)$, then C must be in Ω. □

Note that Lemma 1 implies that if $w^H \in \xi_H$, $(w^H, 1)$ is a renegotiation-proof state. By the same token, if $w^L \in \xi_L$, $(w^L, 0)$ is a renegotiation-proof state.

Lemma 2. *A state of the negotiation $C = (w, \mu)$ with w in the 'No-Rent' configuration is not renegotiation-proof. That is, for every Ω that satisfies the conditions of Definition 3, $C \notin \Omega$.*

Proof. Suppose the state of the negotiation is $C = (w, \mu)$ with w in the 'No-Rent' configuration. Thus, $w^L \in \xi_L$ and $w^H \in \xi_H$ exist such that $\big((w^L, 0)\,, (w^H, 1)\big)$ is feasible starting from C. Moreover, $(1-\mu)\pi\big((w^L, 0)\big) + \mu\pi\big((w^H, 1)\big) > \pi(C)$. As by Lemma 1 $\big((w^L, 0)\,, (w^H, 1)\big)$ are renegotiation-proof states, C cannot constitute a renegotiation-proof state. This would violate internal consistency of Definition 3. □

If the current state of the negotiation is in the 'No-Rent' configuration, the optimal states that leave both types of agent with the same payoff as the current state of the negotiation are feasible. Moreover, those states are efficient and thus renegotiation-proof. It follows from internal consistency of sets of renegotiation-proof states that the current state of the negotiation cannot be renegotiation-proof.

Lemma 3. *Let Ω satisfy the conditions of Definition 3. If a state of the negotiation C is in Ω, there does not exist a single feasible state that yields larger profit for the principal, that is,*

$$\forall C \in \Gamma, \text{ if } C' \in X(C) \text{ exists such that } \pi(C') > \pi(C), \text{ then } C \notin \Omega.$$

Proof. Suppose a $C \in \Omega$ and a $C' \in X(C)$ with $\pi(C') > \pi(C)$ exist. It follows from internal consistency (Definition 3 (i)) that $C' \notin \Omega$. Following from external consistency (Definition 3 (ii)) therefore $(C^L, C^H) \in X(C')$ with $C^\theta \in \Omega$, $\theta \in \{L, H\}$ exists. As, (C^L, C^H) are feasible starting from C' and (C', C') is feasible starting from C, (C^L, C^H) are also feasible starting from C. Feasibility together with $p\pi(C^L) + (1-p)\pi(C^H) \geq \pi(C') > \pi(C)$ implies that internal consistency must be violated and thus $C \notin \Omega$. □

Lemma 1 implies that if the type of agent is known ($\mu \in \{0, 1\}$) the negotiation will lead to an efficient state. Lemma 3 incorporates this result and generalizes it for application to a case where the type of the agent is unknown ($\mu \in (0, 1)$). In this case, for every state of the negotiation a pooling state which would make the principal and both types of the agent better off cannot exist. If such a state were to exist, the principal would be able to simply negotiate towards this state and both types of agent could accept the offer without revealing any additional information. The pooling state would therefore make both parties better off without changing the strategic incentives.

5.1. **Private values.** In the private values case, the utility of the principal is independent of the type of agent involved, that is, $v(w, \theta) = v(w)$. The situation corresponds, for example, to a monopolist selling different quantities (or qualities) of a good to a buyer with heterogeneous valuations for the good.

Examples.

(i) *Financial contracts.* The principal is a lender who provides a loan of size L to a borrower, the agent. The lender faces a capital cost of RL with R being the risk-free interest rate. The lender receives a transfer t from the borrower. The objective function of the lender can be written as $t - RL$. The profit of the borrower depends on a productivity shock $\theta \in \{\theta^L, \theta^H\}$ that is private knowledge to him. The objective function of the borrower can be written as $\theta f(L) - t$ with a function f such that $f' > 0$ and $f'' < 0$. Contracts are then tuples (t,L).

(ii) *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 t. The costs $C(q)$ of the manufacturer are increasing and convex, his objective function is $t - C(q)$. The retailer faces a demand $D(p, \theta)$ with p being the resale price and $\theta \in \{\theta^L, \theta^H\}$ being a demand shock that is private knowledge to the retailer. The objective function of the agent is $p \cdot \min(D(p, \theta), q) - t$. Assuming $D(p, \theta^H) > D(p, \theta^L)$ ensures that condition (1) is satisfied. Contracts are then tuples (q, t).

Analysis. In the standard single-period model without the potential for renegotiation, the principal offers two contracts. The contract for the H type of agent is efficient and he is indifferent between his contract and the contract of the L type. The contract of the L type is inefficient and provides him with the same utility as in his outside option contract. The exact position of the described contracts depends on the prior μ_0 of the principal. However, such contracts will lead to renegotiation in the sense defined above. With private values, only efficient states are renegotiation-proof:

Proposition 2. *With private values, the unique set of renegotiation-proof states of the negotiation is*

$$\Omega = \left\{(w^L, 0); w^L \in \xi_L\right\} \cup \left\{(w^H, 1); w^H \in \xi_H\right\}.$$

Proof. It follows from Lemma 1 that the proposed Ω must be a subset of any set of renegotiation-proof states, as every efficient state of the negotiation is renegotiation-proof. It must now be shown that no other state can be renegotiation-proof.

Suppose the state of the negotiation is $C^1 = (w^1, \mu^1)$ with w^1 in the 'H-Rent' configuration. As w^1 is then to the left of both efficient contract curves and the principal's utility is independent of the type of the agent, the indifference curve of the principal is steeper than

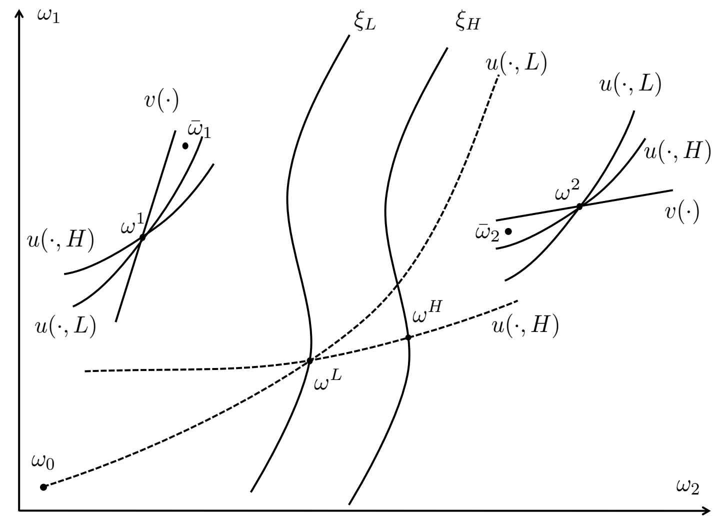

FIGURE 2. Efficient contracts arise in the equilibrium of the negotiation game with private values.

the indifference curves of both of the agents. Thus, independent of μ^1, there exists a single contract $\bar{w}^1$ that makes the principal and both types of agent better off. It follows from Lemma 3 that C^1 cannot constitute a renegotiation-proof state.

Suppose the state of the negotiation is $C^2 = (w^2, \mu^2)$ with w^2 in the 'L-Rent' configuration. As w^2 is then to the right of both efficient contract curves and the principal's utility is independent of the type of the agent, the indifference curve of the principal is flatter than the indifference curves of both of the agents. Thus, independent of μ^2, there exists a single contract $\bar{w}^2$ that makes the principal and both type of agent better off. It follows from Lemma 3 that C^2 cannot constitute a renegotiation-proof state. The case that in the state of the negotiation the current signed contract is in the 'No-Rent' configuration is covered by Lemma 2. □

Proposition 2 is the direct consequence of the fact that if the utility of the principal is independent of the type of the agent and the contract in the current state is in the 'H-Rent' ('L-Rent') configuration, there exists a pooling state that would make the principal and the agent strictly better off. Lemma 3 then implies that no such state can be renegotiation-proof. Given the set of renegotiation-proof states it is easy to verify that Assumption 1 to 3 hold and Proposition 1 applies.

Corollary 1. *With private values, there exists a perfect Bayesian equilibrium of the negotiation game such that in the first round the principal offers the contracts $w^L \in \xi_L$ and $w^H \in \xi^H$ with $u(w^L, L) = u(w_0, L)$ and $u(w^H, H) = u(w^L, H)$. The agent of type θ chooses w^θ with probability 1 and the principal ends the negotiation in the following round.*

The principal offers efficient contracts such that the L-type of agent receives the same utility as with his outside option contract and the H-type of agent is indifferent between his efficient contract and the contract of the L type. It is remarkable that in contrast to the single-period model, the result is independent of the prior of the principal. The results are illustrated in Figure 2.

This result can be considered a generalization of the Coase conjecture. Coase argued that a monopolist cannot keep prices high after the high valuation types of agent have bought the good. That is to say, if the monopolist knows that the non-buyers are the low valuation types of agent, she is not able to refrain from reducing the price to attract these buyers. From a conceptual perspective, this situation is identical to that in which the monopolist offers two contracts: One with a high price at which she agrees to sell the good, and another with a low price and no good exchanged.[20] The second contract, however, is inefficient and would be renegotiated in equilibrium.

5.2. **Common Values: The "Spence" case.** We turn our attention to the case that the utility of the principal depends on the type of agent who signs the contract. We start by considering the "Spence" case. The distinctive characteristic of the "Spence" case it that the principal and the agent agree on the marginal trade-off's between types. Thus, the indifference curve of the principal for the H type of agent is steeper than that for the L type of agent. For example, in the education model in Spence (1973), the L type of agent has larger costs of providing more effort. At the same time, the additional productivity gained through additional education of the L type of agent is smaller than that of the H type of agent.

[20]The original Coase conjecture can also be mapped in the framework at hand. If we interpret w_1 as the probability of trade, then the efficient contract curves will overlap. Thus, in equilibrium the principal will offer the good to both types of agent at the willingness to pay of the low type ($w_2 = \theta^L$). In related work (Gretschko and Wambach, 2016) we discuss the Coase conjecture with a continuous type space and are able to rederive the "gap - no gap" result, in which there is a sharp difference between the case where the costs of the principal lie in or outside the range of valuations.

Examples.

(i) *Equity.* The principal is an investor who provides an investment I to a firm, the agent. The investor faces a capital cost of RI where R represents the risk-free interest rate. The firm offers the investor a share of the firm as compensation for the investment. The profit of the firm $\pi(I,\theta)$ depends on a privately known parameter $\theta \in \{\theta^L, \theta^H\}$ and fulfills the following conditions $\pi_I(I,\theta) > 0, \pi_{II}(I,\theta) < 0, \pi(I,\theta^L) > \pi(I,\theta^H)$. The objective function of the investor is $E\pi(I,\theta) - RI$. The objective function of the firm is $(1-E)\pi(I,\theta)$. Suppose that the objective functions fulfill conditions (1) and (2).

(ii) *Regulating a monopolist.* The agent is a (natural) monopolist that produces a quantity q of a good and faces costs $C(q,\theta)$ that depend on an efficiency parameter $\theta \in \{\theta^L, \theta^H\}$ and are increasing and convex in q and increasing in θ. The principal is a regulator that decides on a subsidy t. The objective function of the monopolist is $t - C(q,\theta)$, the objective function of the regulator is $(S(q) - t) + \alpha(t - C(q,\theta))$ in which $S(q) - t$ denotes the consumer surplus. Suppose that the objective functions fulfill conditions (1) and (2). Contracts are tuples (q,t).

Analysis. As in the private values case, without the potential for renegotiation, the principal offers two contracts. The contract for the H type of agent is efficient and he is indifferent between his contract and the contract of the L type of agent. The contract of the L type is inefficient and provides him with the same utility as in his outside option. The exact position of the described contracts depends on the prior μ_0 of the principal. However, as in the private values case, only efficient contracts are renegotiation-proof.

Proposition 3. *For common values of the 'Spence' type, the unique set of renegotiation-proof states of the negotiation is*

$$\Omega = \left\{(w^L, 0); w^L \in \xi_L\right\} \cup \left\{(w^H, 1); w^H \in \xi_H\right\}.$$

Proof. It follows from Lemma 1 that the proposed Ω has to be a subset of any set of renegotiation-proof states, as every efficient state of the negotiation is renegotiation-proof. It must now be shown that no other state can be renegotiation-proof.

Suppose the state of the negotiation is $C^1 = (w^1, \mu^1)$ with w^1 in the 'H-Rent' configuration. As w^1 is then to the left of both efficient contract curves, the indifference curve of the principal for the L type is steeper than the indifference curve of the L type. Moreover, the indifference curve of the L type is steeper than the indifference curve of the H type. In the 'Spence' case the indifference curve of the principal for the H type is steeper than the indifference curve for the L type. It follows that in w^1 both indifference curves of the principal are steeper than both indifference curves of the agent. Thus, independent of μ^1, there exists a single contract $\bar{w}^1$ that makes the principal and both types of agent better off. It follows from Lemma 3 that C^1 cannot constitute a renegotiation-proof state.

Suppose the state of the negotiation is $C^2 = (w^2, \mu^2)$ with w^2 in the 'L-Rent' configuration. As w^2 is then to the right of both efficient contract curves, the indifference curve of the H type is steeper than the indifference curve of the principal for the H type. Moreover, the indifference curve of the L type is steeper than the indifference curve of the H type. As above, the indifference curve of the principal for the H type is steeper than the indifference curve for the L type. It follows that both indifference curves of the agents are steeper than both indifference curves of the principal. Thus, independent of μ^2, there exists a single contract $\bar{w}^2$ that makes everyone better off. It follows from Lemma 3 that C^2 cannot constitute a renegotiation-proof state. The case that the state of the negotiation is in the 'No-Rent' configuration is covered by Lemma 2. □

Interestingly, changing from private values to common values 'Spence' type does not change the fact that only efficient contracts are renegotiation-proof. As with private values, whenever the contract of the current state is in the 'H-Rent' ('L-Rent') configuration, there exists a pooling contract that would make the principal and the agent strictly better off. This is a direct consequence of the fact that both indifference curves of the principal are steeper (less steep) than both indifference curves of the agent. Lemma 3 thus implies that no such state can be renegotiation-proof. This is illustrated in Figure 3. Given the set of renegotiation-proof states it is easy to verify that Assumption 1 to 3 hold and Proposition 1 applies.

Corollary 2. *With common values of the "Spence" type, there exists a perfect Bayesian equilibrium of the negotiation game such that in the first round the principal offers the contracts* $w^L \in \xi_L$ *and* $w^H \in \xi^H$ *with* $u(w^L, L) = u(w_0, L)$ *and* $u(w^H, H) = u(w^L, H)$*. The agent of*

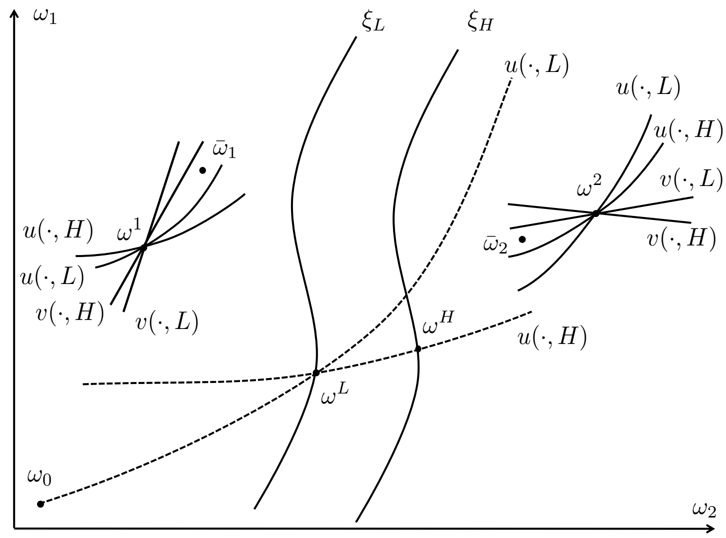

FIGURE 3. The equilibrium contracts of the negotiation game with common values 'Spence' type.

type θ chooses w^{θ} with probability 1 and the principal ends the negotiation in the following round.

It could be argued that this result does not come as a surprise as efficient contracting is to be expected given that frictionless renegotiation should lead to an exploitation of all gains from trade. However, as we will show in the following section, this is not true if the agent and the principal rank the marginal trade-offs between types differently.

5.3. **Common Values: The "Rothschild-Stiglitz" case.** The distinctive characteristic of the "Rothschild-Stiglitz" case is that the agent and the principal do not agree on the marginal trade-offs between types of agent. For example, in the insurance model in Rothschild and Stiglitz (1976), the L type of agent has a lower risk probability. On one hand, for a marginal increase in premium the agent must be compensated by a larger increase in indemnity than the H type. Thus, the L type indifference curve is steeper than the indifference curve of the H type. On the other hand, a marginal increase in indemnity is more costly for the principal if the agent is of the H type than if the agent is of the L type.

Examples.

(i) *Financing contracts.* The agent is a firm that needs financing I to undertake a project with return $R(\theta, I)$ with a risk-spread of $\theta \in \{\theta^L, \theta^H\}$. That is, let $R(\theta, I)$ be normally distributed with $\mathrm{E}(R(\theta^H, I)) > \mathrm{E}(R(\theta^L, I))$ and $\mathrm{Var}(R(\theta^H, I)) > \mathrm{Var}(R(\theta^L, I))$ such that $\mathrm{E}(R(\theta^H, I))$ is increasing in I. The risk spread θ is private information to the firm. The principal is an investor who provides an investment I and receives $\tilde{t} = \min\{t, R(\theta^H, I)\}$. The objective function of the firm is $\max\left\{\mathrm{E}(R(\theta^H, I)) - t, 0\right\}$ and the objective function of the investor is $\mathrm{E}\left(\min\{t, R(\theta^H, I)\}\right) - I$. Suppose that the objective functions fulfill conditions (1) and (3). Contracts are tuples (I, t).

(ii) *Green procurement.* The principal is a public authority that wants to procure a quantity q of a good for the exchange of a transfer t. The production of a quantity q of the good generates a utility of $S(q)$ and creates harmful pollution of P(q,θ) with $\theta \in \left\{\theta^L, \theta^H\right\}$ being the degree of pollution. That is, $P(q, \theta^H) > P(q, \theta^L)$. The agent is a firm that produces the good with an exogenously given technology $\theta \in \left\{\theta^L, \theta^H\right\}$ at costs $C(q, \theta)$ such that technology θ^L creates less pollution but is associated with higher production costs than technology θ^H. That is, $C(q, \theta^L) > C(q, \theta^H)$. The used technology is private knowledge to the agent. The objective function of the principal is $S(q) - P(q, \theta) - t$. The objective function of the agent is $t - C(q, \theta)$. Suppose that the objective functions fulfill conditions (1) and (3). Contracts are tuples (t, q).

Analysis. Under the assumption that the efficient contract curve of the L type of agent lies weakly to the left of the efficient contract curve of the high type, the solution of the one-shot negotiation is similar to the solution of the one-shot negotiation in the private values and the 'Spence' case: the principal offers two contracts. The contract for the H type of agent is efficient and the agent is indifferent between his contract and the contract of the L type of agent. The contract of the L type of agent is inefficient and provides the agent with the same utility as his outside option contract. The exact position of the described contracts depends on the prior μ_0 of the principal.

If renegotiation is taken into account, the renegotiation-proof states in the "Rothschild-Stiglitz" case can be remarkably different from those in the "Spence" and private values case as a set containing only efficient states may not be renegotiation-proof.

Lemma 4. *Let $w^L \in \xi_L$ and $w^H \in \xi^H$. In the "Rothschild-Stiglitz" case, if $w \in \mathbb{R}^2_+$ exists such that $u(w^L, L) = u(w, L)$ and $u(w^H, H) = u(w^L, H)$ and $v(w, H) > v(w^H, H)$, then $\left\{(w^L, 0); w^L \in \xi_L\right\} \cup \left\{(w^H, 1); w^H \in \xi_H\right\}$ is not a set of renegotiation-proof states.*

Proof. As $v(w,H) > v(w^H,H)$ there exist an $\mu < 1$ such that $(1-\mu)v(w,L) + \mu v(w,H) > (1-\mu)v(w^L,L) + \mu v(w^H,H)$. Thus, starting from state (w,μ), $(w^L,0)$ and $(w^H,1)$ are not feasible. Moreover, as $(w^L,0)$ and $(w^H,1)$ are the optimal efficient states there do not exist other efficient states that are feasible. Thus, $\left\{(w^L,0); w^L \in \xi_L\right\} \cup \left\{(w^H,1); w^H \in \xi_H\right\}$ violates external consistency of Definition 3. □

The main difference between the "Rothschild-Stiglitz" case and other cases is that in the latter cases it is always feasible to reach efficient allocations from any state of the negotiation game. This is not true for the "Rothschild-Stiglitz" case. To clarify this we can consider some state with a contract in the 'H-Rent' configuration. As, for example, in the Spence case the H type indifference curve of the principal lies above the L type indifference curve, the optimal efficient contracts for the principal lie both on higher indifference curves than the contract of the original state. This not need be the case in the "Rothschild-Stiglitz" case as the H type indifference curve of the principal is below the L type indifference curve. In this case it could be that the optimal efficient contracts are such that the principal receives a strictly lower utility from contracting with the H type of agent than in the contract of the original state. Thus, if the probability of facing the H type of agent in the original state is high, the principal is strictly worse off with the efficient contracts. It follows that a set containing only efficient states may not be renegotiation-proof.

Lemma 4 could be seen as an impossibility result. However, to show that inefficient equilibria of the negotiation game exist, we must show that a set of renegotiation-proof states Ω does indeed exist. To do so we use internal and external consistency to construct such a set. For the construction we impose some more structure on our very general set-up by making the following assumption:

Assumption 4. *The utility functions of the principal and the agent are additively separable. That is,*

(i) $u(w,\theta) = f(w_2,\theta) + \lambda_\theta w_1$ *and* $v(w,\theta) = g(w_2,\theta) - w_1$ *with* $\lambda_\theta \in \mathbb{R}_+$ *and*

(ii) if $-f_1(w_2,L)/\lambda_L = g_1(w_2,L)$ *then* $-f_1(w_2,H)/\lambda_H < g_1(w_2,\theta_H)$.

Assumption 4 guarantees that the utility of the principal and the agent are additively separable and thus indifference curves are parallel. If Assumption 4 holds true, we can use internal and external consistency to construct a set of renegotiation-proof states:

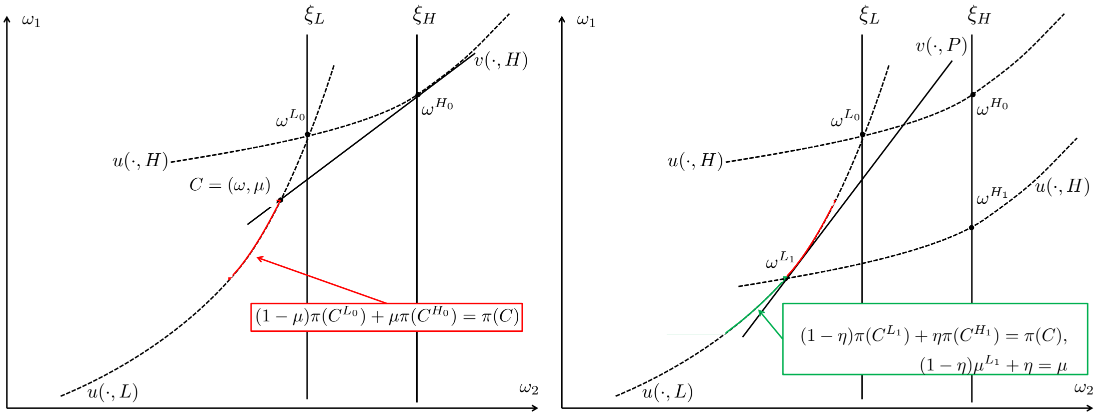

FIGURE 4. Construction of renegotiation-proof Ω in the Rothschild-Stiglitz case.

Proposition 4. *Suppose Assumption 4 holds true. A set Ω of renegotiation-proof states exists in the "Rothschild-Stiglitz" case.*

Proof. The proof is relegated to Appendix A. ☐

The construction proceeds inductively. We work our way along all states with contracts on a particular indifference curve of the agent. We start by including the states that make the principal indifferent between ending the negotiation or negotiating towards fully efficient states. In particular, the belief that makes the principal indifferent is decreasing as we move down the indifference curve. See the left-hand side of Figure 4 for an illustration. At some point, the construction may hit the pooling indifference curve of the principal. At this point, we start including those states in Ω which mean that the principal is indifferent to either negotiating in a way that leads to this state or opting for the corresponding efficient contract for the H type of agent. Again, moving down the indifference curve, the belief of the principal will be decreasing for states in Ω. See the right-hand side of Figure 4 for an illustration. We then use the additive separability of the utility functions to extend the construction to all indifference curves.

As suggested by Lemma 4, the constructed Ω will contain inefficient states. However, whether the optimal state will be inefficient depends on the model parameters:[21]

Corollary 3. *Let $w^L \in \xi_L$ and $w^H \in \xi^H$ such that $u(w^L, L) = u(w_0, L)$ and $u(w^H, H) = u(w^L, H)$. With common values "Rothschild-Stiglitz" type:*

[21]Given the set of renegotiation-proof states constructed in the proof of Proposition 4 it is straightforward to verify that Assumption 1 to 3 hold and Proposition 1 applies.

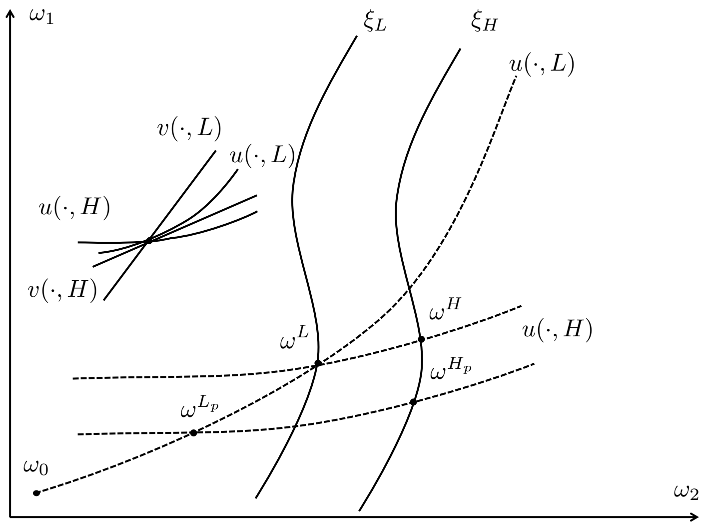

FIGURE 5. Equilibrium contracts of the negotiation game with common values "Rothschild-Stiglitz" type.

(i) If $v(w_0, H) < v(w^H, H)$, there exists a perfect Bayesian equilibrium of the negotiation game such that in the first round the principal offers the contracts w^L and w^H. The agent of type θ chooses w^θ with probability 1 and the principal ends the negotiation in the following round.

(ii) Otherwise, there exist a perfect Bayesian equilibrium of the negotiation game in which the principal offers contracts $w^{L_p} \notin \xi_L$ and $w^{H_p} \in \xi_H$ with $u(w^{L_p}, L) = u(w_0, L)$ and $u(w^{H_p}, H) = u(w^{L_p}, H)$. The L-type agent chooses w^{L_p} with probability 1 and the H-type agent chooses w^{H_p} with a probability strictly below 1. The principal ends the game in the following round.

Proof. Follows immediately from the construction of Ω in the proof of Proposition 4. $\square$

The results are illustrated in Figure 5. The exact position of (w^{L_p}, w^{H_p}) will depend on the initial parameters of the problem. For example, all states $C = (\text{w},\mu)$ with contracts on the indifference curve of the L-type agent that goes through w_0 are renegotiation-proof if the principal is indifferent between to either maintaining w or negotiating toward the efficient contracts (w^L, w^H), that is, if $\pi(C) = (1-\mu)\pi(C^L) + \mu\pi(C^H)$. Thus, the principal will pick the optimal renegotiation-proof state that is feasible starting from (w_0, μ_0). Feasibility in this cases amounts to $\mu \leq \mu_0$. More generally, if the initial contract w_0 is in the 'H-Rent' configuration, the principal solves the following problem:

$$\max_{C^{L_p}} \nu\pi(C^{L_p}) + (1-\nu)v(w^{H_p}, H)$$

$$s.t. \quad \nu\mu^{L_p} + (1-\nu) = \mu_0$$

$$C^{L_p} \in \Omega$$

$$u(w^{H_p}, H) = u(w^{L_p}, H)$$

with the Ω constructed in Proposition 4.

$v(w_0, H) > v(w^H, H)$ does not imply that efficient states are not feasible ex-ante. Hence, the optimal renegotiation-proof states can be inefficient, even if efficient contracts were feasible ex-ante. Even more interesting perhaps is the fact that there may exist states which would make the principal strictly better off then in the renegotiation-proof states, if she were to negotiate towards these states. These states, however, are not renegotiation-proof.

Note that in contrast to the one-shot negotiation which also yields inefficient contracts, a full separation of types of agent does not occur if renegotiation is taken into account. On the equilibrium path, the H type of agent always chooses the contract of the L type of agent with a positive probability.

6. Conclusion

One of the main contributions of this article has been the characterization of a set of renegotiation-proof states. By using internal and external consistency, effective 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 the state has to be efficient and thus allow for inefficient states as in the "Rothschild-Stiglitz" case.

In the case at hand we are able to prove that the renegotiation-proof states can indeed be supported by perfect Bayesian equilibrium of a general negotiation game. Provided that the type space remains finite the presented analysis can be extended. One must however, be careful with the definition of the (mixed) strategies of the agent.[22] However, if type spaces become more complicated, the explicit derivation of strategies in the negotiation game becomes

[22]One way to overcome the problems associated with mixed strategies of the agent is to introduce a mediator that translates (pure-strategy) messages of the agent into a noisy signal to the principal. Examples for this approach include (but are not limited to) Pollrich (2016), Strausz (2012), and Vartiainen (2013).

intractable. Nevertheless, we believe that the presented analysis based on the properties of renegotiation-proof sets remains valid as it captures the sequential rationality of the principal in a dynamic game. Thus, extending the 'axiomatic' analysis of renegotiation-proof states to more complicated settings provides significant insights and should be implemented as part of future research. Examples of this approach include Asheim and Nilssen (1997), Gretschko and Wambach (2016), Vartiainen (2013), and Vartiainen (2014).

Appendix A. Proofs

Proof of Proposition 1.

Proof. The proof proceeds in six steps:

Step 1: We specify the equilibrium strategies of the principal.

Step 2: We prove a Lemma that facilitates the exposition of the strategies of the agent.

Step 3: We specify the equilibrium strategies of the agent.

Step 4: We specify how the beliefs of the principal and the agent are updated

Step 5: We show that to prove that the proposed strategies form a PBE it is sufficient to consider only one-shot deviations.

Step 6: We prove that the proposed strategies form a PBE.

Step 1: Definition of the equilibrium strategy of the principal

The proposed equilibrium strategy of the principal in stage t depends only on the current state $C_{t-1} = (w_{t-1}, \mu_{t-1})$ of the negotiation. If C_{t-1} is in Ω, the principal offers $\mathcal{M}_t = \emptyset$. If C_{t-1} is not in Ω, the principal offers $\mathcal{M}_t = \{s^L(C_{t-1}), s^H(C_{t-1})\}$. That is, the principal ends the game if the current state C_{t-1} is in Ω. Otherwise, she proposes the optimal contracts in Ω starting from state C_{t-1}:

$$\sigma_t^p(h^p(t)) = \sigma^p(C_{t-1}) = \begin{cases} \{s^L(C_{t-1}), s^H(C_{t-1})\} & \text{if } C_{t-1} \notin \Omega \\ \emptyset & \text{if } C_{t-1} \in \Omega \end{cases}.$$

Step 2: A useful result

We establish that for any deviating proposal of the principal there is one type of agent who is best off choosing one of the contracts with probability one given that the principal will play her equilibrium strategy in the continuation game. For the other type of agent it is then optimal to mix between this contract and at most one other contract from the proposal.

Lemma 5. *For any proposal $\mathcal{M}_t$ of the principal at least one of the following holds true:*

(i) there exists a contract $w^L \in \mathcal{M}_t \cup \{w_{t-1}\}$ such that $u(s^L(w^L, \nu), L) \geq u(s^L(w, \mu), L)$, for all $w \in \mathcal{M}_t \setminus \{w^L\} \cup \{w_{t-1}\}$ and all $\nu, \mu \in [0, 1]$.

(ii) there exists a contract $w^H \in \mathcal{M}_t \cup \{w_{t-1}\}$ such that $u(s^H(w^H, \nu), H) \geq u(s^H(w, \mu), H)$, for all $w \in \mathcal{M}_t \setminus \{w^H\} \cup \{w_{t-1}\}$ and all $\nu, \mu \in [0, 1]$.

Proof. Suppose first that all $w \in \mathcal{M}_t \cup \{w_{t-1}\}$ are in the 'H-Rent' configuration. In this case, Assumption 2 (i), ensures that the contract w^L that maximizes $u(\cdot, L)$ over $\mathcal{M}_t \cup \{w_{t-1}\}$ has property (i) as $u(s^L(w, \mu), L) = u(w, L)$ and $u(s^L(w^L, \nu) = u(w^L, L)$. If all $w \in \mathcal{M}_t \cup \{w_{t-1}\}$ are in the 'L-Rent' configuration, Assumption 2 (ii) ensures that the contract w^H that maximizes $u(\cdot, H)$ over $\mathcal{M}_t \cup \{w_{t-1}\}$ has property (ii). If all $w \in \mathcal{M}_t \cup \{w_{t-1}\}$ are in the 'No-Rent' configuration, Assumption 2 (iii) ensures that the contract w^L that maximizes $u(\cdot, L)$ over $\mathcal{M}_t \cup \{w_{t-1}\}$ has property (i) and that the contract w^H that maximizes $u(\cdot, H)$ over $\mathcal{M}_t \cup \{w_{t-1}\}$ has property (ii).

Finally, let $\mathcal{M}_t \cup \{w_{t-1}\}$ be arbitrary. For a contradiction, suppose that (i) and (ii) do not hold true. As (i) does not hold, there must exist a contract $w' \in \mathcal{M}_t \cup \{w_{t-1}\}$ and $\mu' > 0$ such that $u(s^L(w', \mu'), L) > u(s^L(w, 0), L)$ for all w $\in \mathcal{M}_t \cup \{w_{t-1}\}$.[23] In view of Assumption 2, w' must lie in the 'L-Rent' configuration. This same argument also means that there exists $w'' \in \mathcal{M}_t \cup \{w_{t-1}\}$ in the 'H-Rent' configuration and $\mu'' > 0$ such that $u(s^H(w'', \mu''), H) > u(s^H(w, 1), H)$ for all w $\in \mathcal{M}_t \cup \{w_{t-1}\}$. Observe that $u(s^L(w', \mu'), L)$ is bounded above by $u(\bar{w}^H, L)$ where $\bar{w}^H$ is the H-efficient contract that gives the H-type the same utility as w', i.e. $u(w', H) = u(\bar{w}^H, H)$ and $\bar{w}^H \in \xi_H$. This is a result of Assumption 2 and the fact that efficient allocations are in Ω (Lemma 1). As (i) does not hold true it must follow that $u(w'', L) < u(\bar{w}^H, L)$. By the same token, $u(s^H(w'', \mu''), H)$ is bounded above by $u(\bar{w}^L, H)$ where $\bar{w}^L$ is the L-efficient contract that gives the L-type the same utility as w'', i.e. $u(w'', L) = u(\bar{w}^L, L)$ and $\bar{w}^L \in \xi_L$. However, single crossing together with the fact that ξ_H lies to the right of ξ_L implies that $u(\bar{w}^L, H) < u(w', H)$. As $u(\bar{w}^L, H) \geq u(s^H(w'', \mu''), H)$ this constitutes a contradiction to the assumption that (ii) does not hold true. The proof is illustrated in Figure 6.

□

[23] $u(s^L(w, \nu), L)$ is weakly increasing in ν. Thus, if $u(s^L(w', \mu'), L) > u(s^L(w, \nu), L)$ for some $\nu \in [0, 1]$, it follows that $u(s^L(w', \mu'), L) > u(s^L(w, 0), L)$.

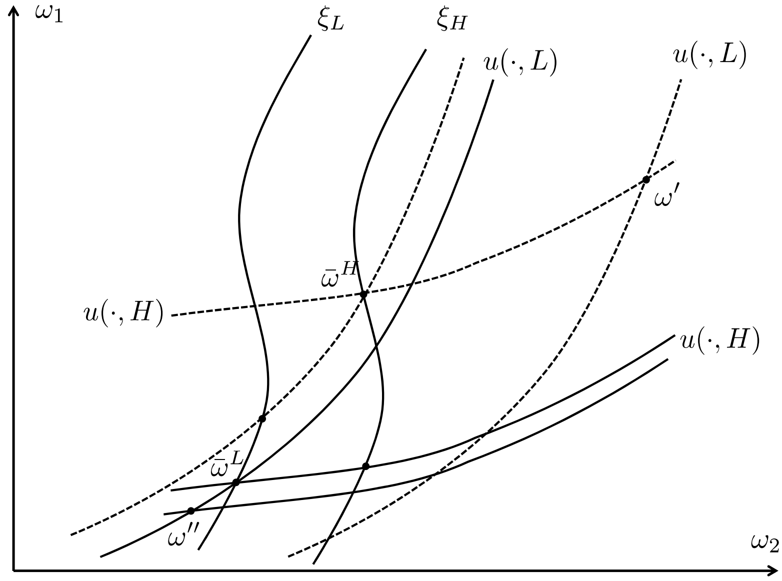

FIGURE 6. Sketch of proof of Lemma 5.

Step 3: Definition of the equilibrium strategy of the agent

The strategy of the agent at stage t is more elaborate and depends on the current contract w_{t-1}, the set of proposed contracts of the principal $\mathcal{M}_t$ and the belief of the principal μ_t. We start by defining the strategy of the agent if μ_{t-1} is the current belief of the principal, the current signed contract is w_{t-1} and the principal proposes $\mathcal{M}_t = \left\{s^L((w_{t-1},\mu_{t-1})), s^H((w_{t-1},\mu_{t-1}))\right\}$. In this case, let $(C^L,C^H) = s(C_0)$ with $C_\theta = (s^\theta(w_o,\mu_0),\mu^\theta)$. The agent of type L chooses contract $s^L(w_{t-1},\mu_{t-1})$ with probability

$$p^L = \frac{(1-\mu^L)(\mu^H-\mu_{t-1})}{(1-\mu_{t-1})(\mu^H-\mu^L)},$$

that is, $\sigma_t^L(h^a(t)) = \sigma^L(\left\{s^L(w_{t-1},\mu_{t-1}), s^H(w_{t-1},\mu_{t-1}), w_{t-1}\right\},\mu_{t-1}) = (p^L, 1-p^L, 0)$.[24] The agent of type H chooses contract $s^H(w_{t-1},\mu_{t-1})$ with probability

$$p^H = \frac{\mu^H(\mu_{t-1}-\mu^L)}{\mu(\mu^H-\mu^L)},$$

that is, $\sigma_t^H(h^a(t)) = \sigma^H(\left\{s^L(w_{t-1},\mu_{t-1}), s^H(w_{t-1},\mu_{t-1}), w_{t-1}\right\},\mu_{t-1}) = (1-p^H, p^H, 0)$.

Now suppose again that the current signed contract is w_{t-1}, the current belief of the principal is μ_{t-1} but $\mathcal{M}_t \neq \left\{s^L(C_{t-1}), s^H(C_{t-1})\right\}$. In this case, Lemma 5 ensures that

[24]Recall that σ^θ is a mapping from $(\mathcal{M}_t, w_{t-1})$ to $\Delta(\mathcal{M}_t \cup \{w_{t-1}\})$.

there exists a type θ and a contract w^θ such that choosing w^θ with probability 1 maximizes continuation pay-off given the equilibrium strategy of the principal. To save on notation we will describe only the case $\theta = L$.[25] In this case, the equilibrium strategy prescribes that the L type chooses w^L with probability 1. To define the equilibrium strategy of the H type we distinguish 3 cases:

(i) If there exists w^H such that $u(s^H(w^H,1),H) > u(s^H(w^L,0),H)$, choose $w^H \in \mathcal{M}_t \cup \{w_{t-1}\}$ with probability 1 that maximizes $u(s^H(\cdot,1),H)$

(ii) If $(s^H(w,1),H) \leq u(s^H(w^L,\mu_{t-1}),H)$ for all $w \in \mathcal{M}_t \cup \{w_{t-1}\}$, choose w^L with probability 1.

(iii) If neither of the two hold, choose with probability $(1-\rho)/(1-\rho\mu_{t-1})$ $w^H \in \mathcal{M}_t \cup \{w_{t-1}\}$ that maximizes $u(s^H(\cdot,1),H)$ where ρ is such that $u(s^H(w^H,1),H) = u(s^H(w^L,\rho\mu_{t-1}),H)$, choose w^L with the complementary probability.[26]

Step 4: Belief updating

Suppose the principal offered $\mathcal{M}_t = \left\{s^L(C_{t-1}), s^H(C_{t-1})\right\}$. If the agent chooses, $s^\theta(C_{t-1})$, the principal updates her belief to $\mu_t = \mu^\theta$.

As above, we define belief updating for the case (i) of Lemma 5 only, that is, the case in which the L-type agent chooses w^L with probability one. The other case is defined in an analogous manner. If the principal observes that the agent chooses w^H, she updates her belief to $\mu_t = 1$, i.e., she believes that the agent is definitely of the H type . If the principal observes that contract w^L is chosen, she updates her belief to $\mu_t = \tilde{\rho}\mu_{t-1}$ with $\tilde{\rho} = 0$ in case (i), $\tilde{\rho} = 1$ in case (ii), and $\tilde{\rho} = \rho$ in case (iii). If the principal observes a choice $w \notin \left\{w^L, w^H\right\}$, she updates her belief to $\mu_t = 1$, i.e., she believes that the agent is certainly of the H-type.

This construction defines the strategies and beliefs for every possible history of the game.

Step 5: One-shot deviations

As the negotiation game is played without discounting, it is not continuous at infinity and the standard one-shot deviation principle does not apply directly.[27] However, we will show that irrespective of the strategy of the agent for every profitable deviation of the principal that may result in an infinite history there is another profitable deviation strategy that only results in finite histories. In this case, the standard finite-deviations argument applies and

[25] The case $\theta = H$ proceeds in exactly the same manner.

[26] Assumption 3 guarantees the existence of such a ρ.

[27] For a statement of the one-shot deviation principle see e.g. Fudenberg and Tirole (1990a) p. 108ff.

to establish equilibrium it is sufficient to consider only one-shot deviations of the principal. For the agent, we will show that given the equilibrium strategy of the principal there is no deviation that results in an infinite history and makes the agent better off.

Lemma 6. *For every deviation strategy of the principal or the agent that results in an infinite history, there exist another profitable deviation strategy that only results in finite histories.*

Proof. **Principal:** Let σ^p denote the strategy of the principal. Suppose there exists another strategy $\bar{\sigma}^p$ that may result in an infinite history and improves on the expected pay-off of σ^p by at least some $\epsilon > 0$. Two cases are relevant.

Case 1: There exists a history $h'(t')$ of the game such that given $\bar{\sigma}^p$ an infinite history $h(\infty)$ with $h_{t'}(\infty) = h'(t')$ will be reached with probability higher than some fixed $\delta > 0$. In this case, with probability higher than δ, the final payoff of the principal will be $v(h(\infty)) = -\infty$. The expected continuation pay-off starting from $h'(t')$ is therefore $-\infty$. It follows that replacing $\bar{\sigma}^p(h'(t'))$ by $\mathcal{M}_{t+1} = \emptyset$ yields a strictly higher payoff. There exists therefore a strategy $\tilde{\sigma}^p$ that results only in finite histories and improves the payoff of the principal as compared to σ^p by at least ϵ.

Case 2: There exists no history $h'(t')$ of the game such that given σ an infinite history $h(\infty)$ with $h_{t'}(\infty) = h'(t')$ will be reached with probability higher then some fixed $\delta > 0$.[28] In this case, for a given history $h'(t')$ denote by $\mathcal{H}_{h'(\bar{t})} = \{h(t) \in \mathcal{H} \,|h_{t'}(t) = h'(t'),\ t > t'\}$ the set of all terminal histories that contain $h'(t')$ and denote by $P_{\bar{\sigma}^p}$ the probability distribution on $\mathcal{H}_{h'(t')}$ induced by $\bar{\sigma}^p$. It follows that for every $\delta > 0$ there exist a $t_\delta > t'$ such that for every history $\bar{h}(t_\delta)$ that coincides with $h'(t')$ up to stage t', $P_{\bar{\sigma}^p}\left(\left\{h(t) \in \mathcal{H}_{h'(t')} \text{ and } h_{t_\delta}(t) = \bar{h}(t_\delta)\right\}\right) < \delta$. That is, the ex-ante probability at $h'(t')$ that the game has not ended by stage t_δ is smaller than δ. It is therefore possible to choose a sufficiently small δ such that the change in expected payoff starting from $h'(t')$ of the principal from replacing $\bar{\sigma}^p(\bar{h}(t_\delta))$ by $\mathcal{M}_{t_\delta+1} = \emptyset$ is at most $\epsilon/2$. Hence, there exists a strategy $\tilde{\sigma}^p$ that results only in finite histories and improves the payoff of the principal as compared to σ^p by at least $\epsilon/2$. Using the one-stage deviation principle for finite games yields the result.

[28]In particular, this case covers all histories of arbitrary length which are finite with probability one.

Agent: Given the equilibrium strategy of the principal, the only deviation of the agent that would result in an infinite history is to choose $w_t = w_0$ in every stage t. This is clearly not a profitable deviation for the agent. □

Step 6: Verification of equilibrium

As above, we will verify the equilibrium for case (i) of Lemma 5.

Agent: Suppose the negotiation reached round t, the current signed contract is w_{t-1}, the belief of the principal is μ_{t-1}, and the offer of the principal is $\mathcal{M}_t = \{s^L(C_{t-1}), s^H(C_{t-1})\}$ with $C_{t-1} = (w_{t-1}, \mu_{t-1})$. Given the strategy of the principal, if the agent chooses either $s^L(C_{t-1})$ or $s^H(C_{t-1})$, the negotiation will be over in round $t+1$. Thus, it is optimal for the agent to choose the best contract available. Whenever the agent is indifferent between $s^L(C_{t-1})$ and $s^H(C_{t-1})$ any mixing between the contracts is optimal. Thus, choosing contract $s^\theta(C_{t-1})$ with probability p^θ is optimal as by Definition 1, $p^\theta < 1$ if and only if $u(s^\theta(C_{t-1}), \theta) = u(s^{\theta'}(C_{t-1}), \theta)$. If the agent chooses w_{t-1}, the principal updates his belief to $\mu_t = 1$ and proposes $\mathcal{M}_{t+1} = \{s^L(w_{t-1}, 1), s^H(w_{t-1}, 1)\}$. Neither type of agent is strictly better off with choosing one of those contracts as compared to $s^L(C_{t-1})$ or $s^H(C_{t-1})$. Thus, choosing w_{t-1} in round t cannot constitute a profitable deviation.

Now suppose the offer of the principal is $\mathcal{M}_t \neq \{s^L(C_{t-1}), s^H(C_{t-1})\}$. In this case, for whatever contract the agent chooses the principal will update her belief, the new state will be C_t and the principal will propose $\mathcal{M}_{t+1} = \{s^L(C_t), s^H(C_t)\}$ and end the negotiation afterward. Thus, it is optimal of the agent of type θ to choose a contract in $\mathcal{M}_t \cup \{w_{t-1}\}$ such that $s^\theta(C_t)$ is maximized. Precisely this logic is reflected in the definition of the strategy of the agent. Thus, the agent behaves optimally given the strategy of the principal.

Principal: As established above in Lemma 6, it is sufficient to consider only one-shot deviations of the principal. We will consider two cases: the current state of the negotiation is not in Ω and the current state is in Ω. The main idea of the proof is that any deviation of the principal can be improved further by negotiating to a feasible state that is in Ω. Thus, rather than reaching a final state via a deviation, the principal could have negotiated directly to this state. Hence, offering a menu of contracts other than those prescribed by the equilibrium strategy does not make the principal better off due to internal consistency.

Case 1: The negotiation reached round t and the state is $C_{t-1} \notin \Omega$. Ending the negotiation by proposing $\mathcal{M}_t = \emptyset$ is dominated by the equilibrium strategy of proposing

$\mathcal{M}_t = \{s^L(C_{t-1}), s^H(C_{t-1})\}$. This follows directly from the definition of $s(C_{t-1})$. Thus, suppose the principal offers a nonempty $\mathcal{M}_t \neq \{s^L(C_{t-1}), s^H(C_{t-1})\}$ in round t and subsequently follows her equilibrium strategy. In this case, the agent will choose either contract $w^L \in \mathcal{M}_t \cup \{w_{t-1}\}$ or contract $w^H \in \mathcal{M}_t \cup \{w_{t-1}\}$ as defined in the strategy of the agent above. In round $t+1$ the principal will update her belief and the state of the negotiation is either $(w^H, 1)$ or $(w^L, \tilde{\rho}\mu_{t-1})$. Observe in that case replacing w^H with $\bar{w}^H \in \xi^H$ such that $u(\bar{w}^H, H) = u(w^H, H)$ makes the principal strictly better off. There is therefore no costs associated with assuming that $(w^H, 1) \in \Omega$. Two cases are relevant. Firstly, $(w^L, \tilde{\rho}\mu_{t-1}) \in \Omega$, in this case following her equilibrium strategy, the principal ends the negotiation. From the definition of $s(C)$ it follows that this cannot make the principal strictly better of than following the proposed equilibrium strategy in round t. Secondly, $(w^L, \tilde{\rho}\mu_{t-1}) \neq \Omega$. In this case, after w^L was chosen by the agent, the principal, according to her equilibrium strategy, proposes $\mathcal{M}_{t+1} = \{s^L(C_t), s^H(C_t)\}$, the agent chooses one of the contracts according to his equilibrium strategy and the principal updates her belief such that $C_{t+1} \in \Omega$ and the negotiation ends. If $s(w^L, \tilde{\rho}\mu_{t-1}) = (C^L, C^H)$ is feasible starting from C_{t-1}, by definition of $s(\cdot)$ the principal cannot be better off than by having followed the proposed equilibrium strategy and proposed $\mathcal{M}_t = \{s^L(C_{t-1}), s^H(C_{t-1})\}$. Thus, it remains to establish that $s(w^L, \tilde{\rho}\mu_{t-1}) = (C^L, C^H)$ is feasible starting from C_{t-1}. To do so suppose that w_t is in the 'H-Rent' configuration.[29] Let $C^\theta = (\tilde{w}^\theta, \tilde{\mu}^\theta)$. Due to single crossing and Assumption 2 it follows that $u(s^H(w^L, \tilde{\rho}\mu_{t-1}), H) > u(s^L(w^L, \tilde{\rho}\mu_{t-1}), H)$. Thus, it follows from the definition of feasibility that $\tilde{\mu}^H = 1$. To establish feasibility of (C^L, C^H) starting from C_{t-1} we check condition (iii) of Definition 1.[30] That is, we must show a p exists such that $p\tilde{\mu}^L + (1-p)\tilde{\mu}^H = \mu_{t-1}$. As $\tilde{\mu}^H = 1$ it suffices to show that $\tilde{\mu}_L \leq \mu_{t-1}$. As $\mu_t = \tilde{\rho}\mu_{t-1}$, it follows that $\mu_{t-t} \geq \mu_t$. However, as (C^L, C^H) is feasible starting from C_t it follows that $\mu_{t-1} \geq \mu_t \geq \mu^L$.

Case 2: The negotiation reached round t and the state is $C_{t-1} \in \Omega$. Suppose the principal deviates from the proposed equilibrium strategy and instead of ending the negotiation proposes a nonempty set of contracts $\mathcal{M}_t$ and follows his equilibrium strategy afterward. The agent will choose either contract $w^L \in \mathcal{M}_t \cup \{w_{t-1}\}$ or contract $w^H \in \mathcal{M}_t \cup \{w_{t-1}\}$ as defined in the strategy of the agent above. In round $t+1$ the principal will update his belief and the

[29]This is consistent with case (i) of Lemma 5. The case that w_t is in the 'No-Rent' or 'L-Rent' configurations would proceed in an anologous manner.

[30]Conditions (i), (ii), and (iv) in Definition 1 are satisfied trivially.

state of the negotiation is either $(w^H, 1)$ or $(w^L, \tilde{\rho}\mu_{t-1})$. Again there are no costs associated with assuming that $(w^H, 1) \in \Omega$. Two cases are relevant. First, $(w^L, \tilde{\rho}\mu_{t-1}) \in \Omega$, in this case, according to her equilibrium strategy, the principal ends the negotiation. As $(w^L, \tilde{\rho}\mu_{t-1}) \in \Omega$ and $(w^L, \tilde{\rho}\mu_{t-1})$ is feasible starting from C_{t-1}, it follows from internal consistency (Definition 3 (i)) that executing $(w^L, \tilde{\rho}\mu_{t-1})$ cannot make the principal strictly better off compared to following his equilibrium strategy and executing C_{t-1}. Second, $(w^L, \tilde{\rho}\mu_{t-1}) \notin \Omega$. In this case, following her equilibrium strategy, the principal proposes $\mathcal{M}_{t+1} = \{s^L(C_t), s^H(C_t)\}$, the agent chooses one of the contracts according to his equilibrium strategy, the principal updates her belief such that $C_{t+1} \in \Omega$ and the negotiation ends. As in Case 1 above, $s((w^L, \tilde{\rho}\mu_{t-1})) = (C^L, C^H)$ is feasible starting from C_{t-1}. It therefore follows from internal consistency (Definition 3 (i)) that proposing (C^L, C^H) cannot make the principal better off compared to following her equilibrium strategy and executing C_{t-1}. □

Proof of Proposition 4.

Proof. The construction of the the set of renegotiation-proof states draws on ideas introduced in Asheim and Nilssen (1997). The following proof indicates, however, that is differs in two respects. Firstly, Asheim and Nilssen (1997) consider an insurance market. Thus, the efficient contract curves for the L type and the H type are the same. We extend the analysis to problems with potentially distinct efficient contract curves. Secondly, they assume that no overinsurance takes place. Thus, their construction does not extend to the 'L-Rent' configuration which is the case for our construction.

To show that a set of renegotiation-proof states exists, we will construct such a set. The main idea of the construction is to make use of internal consistency of renegotiation-proof states in Definition 3. We start by working our way along all states with contracts on a particular indifference curve. Inductively, we will include those states in Ω that cannot be strictly improved upon by feasible states that are already included in Ω.

We start the construction for states with contracts that are in the 'H-Rent' configuration. Take some $w^{L_0} \in \xi_L$. Denote by $w^{H_0} \in \xi_H$ the H-efficient contract that gives the H type the same utility as w^{L_0}. That is, $u(w^{H_0}, H) = u(w^{L_0}, H)$. From Lemma 1 it follows that $C^{L_0} = (w^{L_0}, 0)$ and $C^{H_0} = (w^{H_0}, 1)$ are in Ω. Observe that (w, μ) such that $u(w, L) = u(w^{L_0}, L)$, $u(w, H) \leq u(w^{H_0}, H)$ and $v(w, H) < v(w^{H_0}, H)$ cannot constitute a renegotiation-proof state

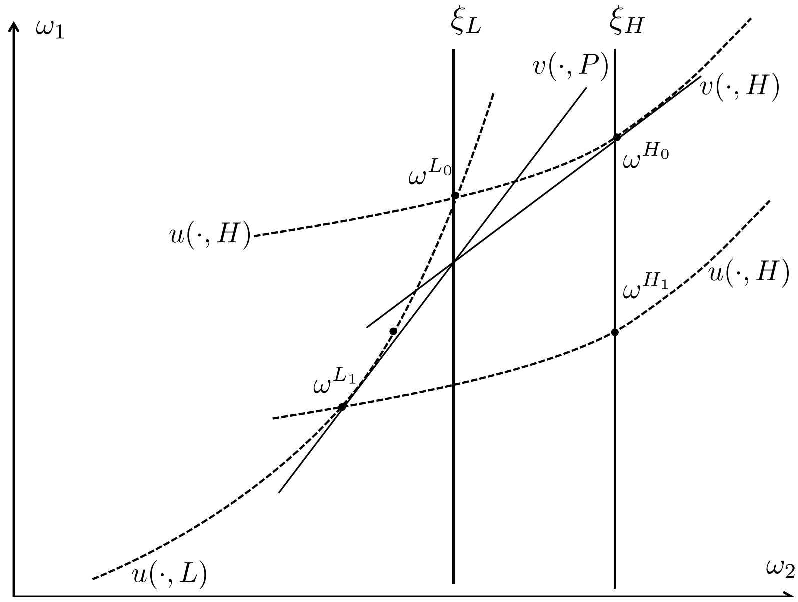

FIGURE 7. Construction of Ω in the case of common values "Rothschild-Stiglitz" type.

as (C^{L_0}, C^{H_0}) is feasible and makes the principal and both types of agent strictly better off. This would violate internal consistency. Thus, for the construction of Ω, we will consider only states such that given her belief, the principal is indifferent as whether she retains w or negotiates towards (C^{L_0}, C^{H_0}). That is, we consider:

$$\Gamma^{-}(C^{L_0}) := \left\{ C = (w, \mu); u(w, L) = u(w^{L_0}, L),\ \mu < 1,\ (1-\mu)\pi(C^{L_0}) + \mu\pi(C^{H_0}) = \pi(C) \right\}.$$

For all $C = (w, \mu)$ and $\bar{C} = (\bar{w}, \bar{\mu})$ in $\Gamma^{-}(C^{L_0})$ it holds that whenever $u(w, H) < u(\bar{w}, H)$ it follows $\mu < \bar{\mu}$. That is, moving along the indifference curve of the L type, μ decreases for all C that are in $\Gamma^{-}(C^{L_0})$. It follows from Lemma 3 that only states for which no pooling state exists which would make all parties better off can be renegotiation-proof. Hence, whenever there exist a state $(w^{L_1}, \mu^{L_1}) \in \Gamma^{-}(C^{L_0})$ such that the pooling indifference curve of the principal for μ^{L_1} is tangent to the L type indifference curve in w^{L_1}, all states with $C = (w, \mu) \in \Gamma^{-}(C^{L_0})$ and $u(w, H) \leq u(w^{L_1}, H)$ cannot constitute a renegotiation-proof state. Thus, we include all states $C = (w, \mu) \in \Gamma^{-}(C^{L_o})$ in Ω with $u(w, H) > u(w^{L_1}, H)$. Define $C^{L_1} = (w^{L_1}, \mu^{L_1})$ and $C^{L_1} = (w^{H_1}, 1)$ with $w^{H_1} \in \xi_H$ and $u(w^{H_1}, H) = u(w^{L_1}, H)$. Continue the construction by considering

$$\Gamma^-(C^{L_1}) := \big\{C = (w,\mu);\ u(w,L) = u(w^{L_1},L),$$
$$\mu < 1,\ (1-\eta)\pi(C^{L_1}) + \eta\pi(C^{H_1}) = \pi(C),\ (1-\eta)\mu^{L_1} + \eta = \mu\big\}\,.$$

As above, there may exist a state $C^{L_2} = (w^{L_2}, \mu^{L_2}) \in \Gamma^-(C^{L_1})$ such that the pooling indifference curve of the principal for μ^{L_2} is tangent to the L type indifference curve. In this case, include all states $C = (w, \mu) \in \Gamma^-(C^{L_o})$ in Ω with $u(w, H) > u(w^{L_1}, H)$. Proceed by constructing $\Gamma^-(C^{L_2})$. The construction stops at $\Gamma^-(C^{L_n})$ if in $\Gamma^-(C^{L_n})$ there is no state such that the pooling indifference curve of the principal is tangent to indifference curve. In this case, include all states in $\Gamma^-(C^{L_n})$ in Ω. The construction is sketched in Figure 7.

So far we have used internal consistency to construct a set of renegotiation-proof states along one particular indifference curve of the L type agent. In the next step, we extend the construction to all potential states in the 'H-Rent' configuration. Again, the main idea is on one hand to include those states in Ω that cannot be strictly improved upon by feasible states that are already in Ω, and on the other hand, to only include states that would not strictly improve on states that are already in Ω. Recall that we assumed that the utility functions of the agent and the principal are additively separable of the form $u(w, \theta) = f(w_2, \theta) + \lambda_\theta w_1$ and $v(w, \theta) = g(w_2, \theta) - w_1$. Thus, for all $((w_1, w_2), \mu)$ that are in Ω so far we also include $((\lambda w_1, w_2), \mu)$ in Ω for all $\lambda > 0$. From additive separability of the utility functions which implies that indifference curves shift in parallel and the construction above, for all w internal consistency of Definition 3 is not violated for all states $C = (w, \mu)$ in Ω along the indifference curve of the L-type agent that contains w. Moreover, in the 'H-Rent' configuration, for each state $C = (w, \mu)$ whenever there is no feasible state $C' = (w', \mu')$ with w' on the same L type indifference curve as w that strictly improves on C, there is also no feasible state which gives the L type a larger utility. Thus, the constructed Ω does not violate internal consistency of Definition 3.

Before we proceed with the construction in the 'L-Rent' configuration, we verify that the constructed Ω does not violate external consistency of Definition 3 for states with contracts in the 'H-Rent' configuration. That Ω does not violate internal consistency is a direct consequence of the construction. Take some state $C = (w, \mu)$. Two cases are relevant. First, there exist a μ' such that $(w, \mu') \in \Gamma^{-}(C^{L_k}) \cap \Omega$ for some k. In this case, whenever $\mu < \mu'$, C^{L_k} and C^{H_k} are feasible. Whenever, $\mu > \mu'$, (w, μ') and $(\bar{w}, 1)$ with $u(\bar{w}, H) = u(\mathrm{w}, H)$ and $\bar{w} \in \xi_H$ are feasible. Second, no such μ' exist. In this case C^{L_0} and C^{H_0} are feasible.

The construction of Ω for states with contracts in the 'L-Rent' configuration mirrors the construction above for the 'H-Rent' configuration. Take some $w^{H_0} \in \xi_H$. Denote by $w^{L_0} \in \xi_L$ the L-efficient contract that gives the L type the same utility as w^{H_0}. That is, $u(w^{L_0}, H) = u(w^{\mathrm{H}_0}, H)$. As above inductively define, C^{H_k} and C^{L_k} as well as

$$\Gamma^{-}(C^{H_k}) := \big\{C = (w, \mu);\ u(w, H) = u(w^{H_k}, H),$$
$$\mu > 0,\ (1-\eta)\pi(C^{L_k}) + \eta\pi(C^{H_k}) = \pi(C),\ (1-\eta)\mu^{H_k} = \mu\big\}\,.$$

The rest of the construction proceeds in exactly the same manner as the construction of renegotiation-proof states with contracts in the 'H-Rent' configuration and is therefore omitted. From Lemma 2 it follows that states with contracts in the 'No-Rent' configurations cannot be part of a set of renegotiation-proof states. This completes the construction. □

References

Asheim, G. and T. Nilssen (1997): "Insurance Monopoly and Renegotiation," *Economic Theory*, 354, 341–354.

Beaudry, P. and M. Poitevin (1993): "Signalling and Renegotiation in Contractual Relationships," *Econometrica*, 61, 745–782.

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

Deneckere, R. and M. Y. Liang (2006): "Bargaining with Interdependent Values," *Econometrica*, 74, 1309–1364.

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. and J. Tirole (1990a): *Game Theory*, MIT Press.

——— (1990b): "Moral Hazard and Renegotiation in Agency Contracts," *Econometrica*, 58, 1279–1319.

Gerardi, D., J. Hörner, and L. Maestri (2014): "The Role of Commitment in Bilateral Trade," *Journal of Economic Theory*, 154, 578–603.

Gretschko, V. and A. Wambach (2016): "Contract Design with Limited Commitment," *mimeo.*

Hart, O. and J. Tirole (1988): "Contract Renegotiation and Coasian Dynamics," *The Review of Economic Studies*, 55, 509–540.

Hörner, J. and L. Samuelson (2011): "Managing Strategic Buyers," *Journal of Political Economy*, 119, 379–425.

Krasa, S. (1999): "Unimprovable Allocations in Economies with Incomplete Information," *Journal of Economic Theory*, 87, 144–168.

Maestri, L. (2012): "Dynamic Contracting under Adverse Selection and Renegotiation," *mimeo.*

Neeman, Z. and G. Pavlov (2013): "Ex-Post Renegotiation-Proof Mechanism Design," *Journal of Economic Theory*, 148, 473–501.

Pollrich, M. (2016): "Mediated Audits," *RAND Journal of Economics, forthcoming.*

Rothschild, M. and J. Stiglitz (1976): "Equilibrium in Competitive Insurance Markets : An Essay on the Economics of Imperfect Information," *The Quarterly Journal of Economics*, 90, 629–649.

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.

Spence, M. (1973): "Job Market Signaling," *The Quarterly Journal of Economics*, 87, 355–374.

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.

——— (2014): “Mechanism Design without Commitment with an Application to Bilateral Bargaining,” *mimeo.*

ZEW Mannheim, L7, 1, 68161 Mannheim, Germany
E-mail address: vgretschko@gmail.com