Moral Hazard and Renegotiation: Multi-Period Robustness

Abraham L. Wickelgren
Federal Trade Commission[1]

April 2003

[1] I thank Chris Adams, Jae Nahm, and Jennifer Reinganum for helpful comments. This paper does not represent the views of the Federal Trade Commission or any individual Commissioner.

Abstract

Is the second best outcome of static agency models renegotiation proof? In models with one period of renegotiation, Fudenberg and Tirole (1990) answer no when the principal makes the offer, while Ma (1994) and Matthews (1995) answer yes when the agent makes the offer. This paper analyzes the robustness of these two claims when there are more periods of renegotiation. With a known number of periods, if the principal makes at least one offer, even if the agent makes the offer in every other period, the equilibrium is identical to Fudenberg and Tirole equilibrium. With an uncertain number of periods, the agency problem is even more severe than in the Fudenberg and Tirole model.

1 Introduction

Consider the standard principal-agent model: a risk neutral principal contracts with a risk averse agent to induce the agent to exert unobservable effort to increase the principal's profit. The second best solution to this problem occurs when the parties can commit not to renegotiate the contract (Holmstrom (1979), Shavell (1979), Grossman and Hart (1983)). In many cases, however, the realization of the principal's profit occurs sometime after the agent chooses his effort level. (Throughout the paper, I will use female pronouns for the principal and male pronouns for the agent.) A product's profitability occurs long after the product development effort of the manager; there are often many weeks between a sharecropper's farming effort and the realization of the field's crop yields. In these, and many other similar, situations, there is ample opportunity for the principal and agent to renegotiate the original incentive contract after the agent has already chosen his effort level. The question then arises, is the standard commitment solution to the principal-agent problem renegotiation proof?

In the current literature, there are two conflicting answers to this question. Fudenberg and Tirole (1990) show that, when the principal can make a take it or leave it offer to the agent after the agent has chosen his action, the principal cannot obtain the outcome that she could obtain if she could commit not to renegotiate. In fact, the principal cannot induce the agent to exert high effort with probability one. To see this, imagine the agent did exert high effort. After the action was taken, the principal would have an incentive to completely insure the agent. Knowing this would occur, the agent would have no incentive to exert high effort in the first place. On the other hand, Ma (1994) and Matthews (1995) show that, when the agent makes the take it or leave it offer in the renegotiation period, the second best outcome remains an equilibrium.[1] The reason the agent can be induced to take the high effort action when the agent, rather than the principal, makes the renegotiation offer, is that the principal can make inferences about the agent's action from the agent's renegotiation offer. If the agent offers a contract that provides more insurance than the original contract, the principal can infer that the agent did not exert high effort. This deters a high effort agent from renegotiating, which maintains his incentives to exert high effort in the initial stage.

In this paper, I investigate the robustness of these results when there is more than one period

[1]Hermalin and Katz (1991) also analyze renegotiation of principal-agent contracts, but they consider the case of renegotiation after observation of a non-verifiable signal of output.

of renegotiation. First, I analyze the case where the number of rounds of renegotiation is known with certainty. Contrary to what one might think, the equilibrium does not depend on who makes the final offer, or the first offer. I find that so long as the principal can make at least one offer, no matter what round it is in, the equilibrium is identical to that described in Fudenberg and Tirole (1990). Thus, it is this equilibrium, not the one described in Ma (1994) and Matthews (1995), that is robust to more than one period of renegotiation, unless the agent makes the offer in every round of renegotiation.

The reason for this is that whenever it is the principal's turn to offer, if she believes the agent exerted high effort with probability one, she has an incentive to offer a full insurance contract to the agent. This means that prior agent offers cannot signal that the agent has not deviated by exerting low effort since the agent can expect to get a full insurance offer in a later round. Similarly, subsequent offers by the agent cannot restore the incentives to exert high effort since, to do so, they must expose the agent to more risk, which cannot make both the principal and agent better off ex post.

Second, I consider the case where the number of available rounds for renegotiation is unknown. This will happen if there is some uncertainty as to when the outcome will occur, so that every round could be, but is not necessarily, the final round. In this case, the agency problem can be even more severe than with just one round of renegotiation. That is, uncertainty over when the outcome can occur lowers the maximum probability that the agent exerts high effort and increases the expected cost to the principal of implementing any feasible probability of high effort.

To understand intuition for this result, notice that the equilibrium contract in the Fudenberg and Tirole (1990) model is a menu from which the agent selects a particular contract. If the outcome could occur after this contract is accepted, then the agent must select from the menu that period. If the outcome does not occur, this selection provides information about the action the agent took, which alters the optimal contract that the principal will offer in a subsequent period. That is, the optimal renegotiation proof contract with one period of renegotiation is no longer renegotiation proof after the agent has chosen from the menu in the contract.[2] This increases the principal's cost of implementing any given probability distribution over effort because it gives an agent that exerted low effort a greater incentive to choose the contract meant for the high effort agent. This means that the low effort agent is exposed to some risk with positive probability and, to deter (at least partially) the low effort agent from pooling, the high effort agent is exposed to

[2]In their conclusion, Fudenberg and Tirole (1990) mention this fact and informally discuss this model.

even more risk than in the one period contract.

The plan of the paper is as follows. Section 2 describes the model. Section 3 analyzes the case where the number of renegotiation rounds is known with certainty. Section 4 analyzes the case where the number of rounds is uncertain. Section 5 concludes.

2 Model

Consider a simple two effort, two outcome model (very similar to the one in Fudenberg and Tirole (1990)). The two effort levels are $\overline{e} > \underline{e}$. The two outcomes are $g > b$. The probability of outcome g is given by $p(e)$. As shorthand, I call this $\overline{p}$ when the agent exerts effort $\overline{e}$ and $\underline{p}$ when the agent exerts effort $\underline{e}$. The agent's utility function for income w and effort e is additively separable: $V(w, e) = U(w) - D(e)$ where $U' > 0$ and $U'' < 0$ (the agent is risk averse). Further, let $\Phi(U)$ be the inverse function corresponding to U. The agent's cost of effort is increasing, $D(\overline{e}) > D(\underline{e})$. The agent's utility for not working for the principal is normalized to zero. The principal is risk neutral. She maximizes the difference between her expected revenue, $I(e) = p(e)g + (1 - p(e))b$, and her expected wage bill, $E(w \mid e)$. I also normalize the principal's utility from not employing the agent to zero.

A *feasible contract* is a pair of compensation schemes $\{c(\underline{e}), c(\overline{e})\}$ from which an agent chooses before the outcome is realized.[3] A *feasible compensation scheme* $c(e)$ specifies two utility levels $\{U_g(e), U_b(e)\}$, that is, the principal pays the agent $w_g(e) = \Phi(U_g(e))$ if outcome g occurs and $w_b(e) = \Phi(U_b(e))$ if outcome b occurs. Let C denote the set of feasible contracts.

The extensive form of the game is as follows. In period 0, the principal offers the agent an initial contract $c^0 \in C$ and the agent accepts or rejects. If the agent accepts, in period 1/2 he chooses to exert effort $\overline{e}$ with probability $x = x(c^0)$. The principal does not observe x or e. All subsequent periods $1, 2, ...$ are renegotiation periods. In each renegotiation period, either the principal makes a new offer or the agent makes a new offer. Who will make the offer in any period $t \geq 1$ is common knowledge in period 0. In each period $t \geq 1$, an offer is made and either accepted or rejected. If it is rejected, then the contract from the preceding period is in effect. If it is accepted, then the agent chooses a compensation scheme from that contract. Then the outcome occurs with probability $q_t > 0$. If the outcome is determined (g or b is realized), the game ends and the agent receives his

[3] In section 4, when there is uncertainty as to when the outcome is reached, this definition is generalized to be a pair of compensation schemes for every history of the renegotiation process.

wage according to the compensation scheme in place. If the outcome does not occur, then there is another period with the same structure. Notice that discounting will not affect either party's incentive to agree given that the actual time of payment is only a function of when the outcome occurs, not when any agreement is reached. Thus, for simplicity, I assume no discounting.

3 Known Number of Periods

In this section, I assume that $q_t = 0$ for all $t < T$ and $q_T = 1$. That is, both the principal and the agent know that the contract in effect in period T will be the contract that determines the agent's compensation. In this model, one might expect that the final contract would only depend on who makes the final offer. In other words, one might think that if the principal makes the final offer, then this model will generate results identical to that of Fudenberg and Tirole (1990), while if the agent makes the final offer the model's results will be identical to that of Ma (1994). As the first proposition shows, however, this conjecture is incorrect. In fact, so long as the principal makes an offer in at least one period $\hat{t} \leq T$, both the final contract and the maximum probability that the agent exerts high effort are identical to what they are in Fudenberg and Tirole (1990).[4]

Before proceeding to show that the Ma (1994) and Matthews (1995) result about the feasibility of the second best with renegotiation is not robust to any offers by the principal, it is useful to analyze why this result holds. If the agent proposes any contract other than the second best efficient contract under commitment, then the principal believes he exerted low effort. Thus, the high effort agent cannot get more insurance through renegotiation, making it optimal for him to exert high effort. Even though more insurance is ex post Pareto improving, asymmetric information prevents the high effort agent from offering this. When the principal makes at least one offer, however, this is not the case.

First, I define what I call a F-T contract.

Definition *A* **F-T contract** *for any given distribution* $(x^*, 1 - x^*)$ *is a feasible contract* $c = \{U_g(e), U_b(e)\}_{e=\overline{e},\underline{e}}$ *with the following properties:*

(a) $U_g(\underline{e}) = U_b(\underline{e}) = \underline{U}$

(b) $U_g(\overline{e}) \geq U_b(\overline{e})$

[4] If the identity of the proposer were randomly determined each period, the equilibrium would approach the Fudenberg and Tirole equilibrium so long as the probability of the principal making at least one offer approached one.

(c) $\underline{p}U_g(\underline{e}) + (1-\underline{p})U_b(\underline{e}) = \underline{p}U_g(\overline{e}) + (1-\underline{p})U_b(\overline{e})$

(d) $\Phi'(U_g(\overline{e})) - \Phi'(U_b(\overline{e})) \leq \frac{(\overline{p}-\underline{p})(1-x)\Phi'(\underline{U})}{(1-\overline{p})\overline{p}x}$

That is, a F-T contract is a contract that satisfies the properties of the optimal renegotiation proof contract in the Fudenberg and Tirole model when the principal want to induce the agent to work hard with positive probability. Condition (a) says the optimal contract when the agent exerts low effort is a full insurance contract. Condition (b) says that to induce high effort, the high effort contract must give at least as much utility when the good outcome occurs as when the bad outcome occurs. Condition (c) requires that the expected utility of the low effort agent from choosing compensation scheme $\underline{e}$ must equal his expected utility if he chooses compensation scheme $\overline{e}$. That is, the low effort agent's interim incentive compatibility constraint is binding. The last condition, (d), says that the difference between the payoff for the good and bad outcome cannot be so large, in scheme $\overline{e}$, that the principal can lower her compensation costs by providing more insurance for high effort agents even though this requires paying low effort agents more to maintain incentive compatibility. Fudenberg and Tirole (1990) show that any renegotiation proof contract (that induces high effort with positive probability) in the one period game where the principal makes the renegotiation offer must satisfy these conditions.

Second, I show that one can restrict attention to renegotiation proof contracts. The following lemma shows that one only need to consider initial contracts that do not get renegotiated until after the principal's final offer.

Lemma 1 *Let t be the period where the principal makes her final renegotiation offer. If there is a Nash equilibrium where the distribution over effort levels is $(x^*, 1-x^*)$ with $x^* > 0$, the initial contract is $\widehat{c}$ and the contract in effect after period t* is *c^*, then there is an equilibrium with the same distribution over efforts where c^* is the initial contract as well as the period t contract.*

Proof. Say that the principal offers $c^0 = c^*$ in period zero and that that this contract gets renegotiated so that the contract in effect after period t is $\widetilde{c}$. At this time, the principal must continue to believe that the agent exerted high effort with probability x^*. To see this, imagine that this were not the case. Then actions of each type of agent up to this point must not be identical. For different actions to be optimal, this requires that there exists some x^H and x^Lsuch that the expected utility under the final contract of a high type agent is larger when the principal believes at t that the agent chose $\overline{e}$ with probability x^H, but the expected utility of a low type agent is larger when the principal believes the agent chose $\overline{e}$ with probability x^L. If there are two different sequences of offers and responses by the agent that induce the principal to have beliefs x^H and x^L, then if the agent chooses $\overline{e}$ he will choose the sequence that produces x^H, and if the agent chooses $\underline{e}$ then he will choose the sequence that produces x^L. Thus, one must have $x^H = 1$ and $x^L = 0$. In this case, there must be a flat wage contract that the principal can offer in period t that the agent would accept and that would not be renegotiated. Knowing that the final contract would be a flat wage contract, the agent would not choose high effort, which contradicts $x^* > 0$.

So, there are two cases to consider: either $\widetilde{c}$ is offered by the agent or by the principal. If the principal offers $\widetilde{c}$, then doing so must reduce the principal's expected compensation costs below what they would be if c^* were in effect at t and must give each type of agent at least as much expected utility as he would receive if c^* were in effect at t. In this case, the principal would have offered $\widetilde{c}$ in period t when the initial contract is $\widehat{c}$. If $\widetilde{c}$ is offered by the agent, then doing so must give the agent more expected utility than he would receive if c^*were in effect at t and must not increase the principal's expected compensation costs above what they would be if c^*were in effect at t. In this case, the agent would have offered $\widetilde{c}$ prior to t when the initial contract was $\widehat{c}$ and the principal would have accepted. Moreover, the agent would never accept a later offer by the principal of c^*. So, c^*could not be in effect at t when the initial contract is $\widehat{c}$, a contradiction.

Since the agent's effort depends only on the final contract, which will be identical so long as the contract in effect at t and the principal's beliefs at t are unchanged, $(x^*, 1 - x^*)$ remains an equilibrium effort distribution. Q.E.D.

Next, I show that the principal will not offer a contract that the agent will renegotiate after the principal's final renegotiation offer.

Lemma 2 *Let t be the period where the principal makes her final renegotiation offer. If there is a Nash equilibrium where the distribution over effort levels is $(x^*, 1 - x^*)$ with $x^* > 0$, the initial contract is $\widehat{c}$, the contract in effect after period t is $\widehat{c}$, and the final contract is c^* then there is an equilibrium with the same distribution over efforts where c^* is the initial contract and the contract in effect after period t as well as the final contract.*

Proof. Say the principal offers an initial contract of c^* which remains in effect after period t. First, assume the principal's beliefs about the probability the agent chose $\overline{e}$ do not change after period t. If there exists a contract $\widetilde{c}$ that the agent will offer after period t that the principal will accept, then this contract must give each type of agent more utility than c^* and not increase the principal's expected compensation costs. If so, consider the contract $\widetilde{c}'$ that is identical to contract $\widetilde{c}$ except that it reduces the payments under each outcome for each compensation scheme by $\varepsilon > 0$. For a small enough ε, this contract must still increase the expected utility of the agent over contract c^* and must reduce the principal's expected compensation costs relative to c^*. In this case, the principal could have offered $\widetilde{c}'$ in period t when the initial contract was $\widehat{c}$ and the agent would have accepted it, contradicting the assumption that the period t contract is $\widehat{c}$ when the initial contract is $\widehat{c}$.

Now, consider what happens if the principal's beliefs about the probability the agent chose $\overline{e}$ do change. The argument in the proof of Lemma 1 establishes that, in this case, the principal either believes the agent took $\overline{e}$ with probability one or zero. Call the contract offer by the high effort agent $c^H = \{c^H(\underline{e}), c^H(\overline{e})\}$ and the one by the low effort agent $c^L = \{c^L(\underline{e}), c^L(\overline{e})\}$. Since each agent type could mimic the other, the principal's expected compensation costs under these contracts are equivalent to a contract that both types offer of $\widetilde{c} = \{c^L(\underline{e}), c^H(\overline{e})\}$. This contract must make each agent type better off than under c^* and not increase the principal's expected compensation costs. Again, consider the contract $\widetilde{c}'$ that is identical to contract $\widetilde{c}$ except that it reduces the payments under each outcome for each compensation scheme by $\varepsilon > 0$. For a small enough ε, this contract must still increase the expected utility of both types of agent over contract c^* and must reduce the principal's expected compensation costs relative to c^*. In this case, the principal could have offered $\widetilde{c}'$ in period t when the initial contract was $\widehat{c}$ and the agent would have accepted it, contradicting the assumption that the period t contract is $\widehat{c}$ when the initial contract is $\widehat{c}$.

Since the agent's effort depends only on the final contract, $(x^*, 1 - x^*)$ remains an equilibrium effort distribution. Q.E.D.

Together, Lemmas 1 and 2 imply that one can restrict attention to renegotiation proof contracts when analyzing equilibria that induce the agent to exert high effort with a positive probability. This considerably simplifies that task of showing that the agent cannot exert high effort with probability one in the T period game if the principal gets to make one renegotiation offer. The intuition for this is that, if the agent did exert high effort with probability one, the principal, when it is her turn to make the offer, will offer a full insurance contract that the agent would accept. Because full insurance contracts are ex post pareto optimal, once in place they are never renegotiated. In fact, because the principal's incentives to renegotiate are stronger than are the agent's (because the principal need not worry about adverse inferences about her type based on her offer), the impact of the renegotiation proof constraint is to ensure that the original contract must be a F-T contract, despite the fact that the agent can make up to $T-1$ of the renegotiation offers.

Lemma 3 S*ay there is at least one period $\hat{t} \leq T$ when the principal makes an offer. For any given distribution $(x^*, 1-x^*)$ with $x^* > 0$, if a contract c is renegotiation proof, then c is a F-T contract.*

Proof. Say c is not a F-T contract because either (a), (b), or (c) does not hold. Then, by lemma 2.1 in Fudenberg and Tirole (1990), either it the interim utility of an agent who chooses low-effort is greater than that of the agent who chooses high effort, in which case the agent would not choose to work hard ($x^* = 0$), or the principal can offer different contract when it is her turn to make a renegotiation offer that gives each type of agent as much utility as the original contract and lower the principal's expected payment. The agent will only reject this new offer if, by doing so, it can expect the principal to accept an offer (call this contract $\widetilde{c}$) that he will make in a later bargaining round. If this is the case, then c is not renegotiation-proof. Now say c violates (d). Then, by lemma 2.2 in Fudenberg and Tirole (1990), either (a), (b), or (c) does not hold, or c does not minimize the principal's expected payments subject to each type of agent receiving at least as much expected utility as he receives from c. Thus, there is another contract, c^*, that lowers the principal's expected payments while leaving each type of agent's expected utility unchanged. The principal with offer c^* when she can make a renegotiation offer. The agent will only reject this new offer if, by doing so, it can expect the principal to accept an offer (call this contract $\widetilde{c}$) that the agent will make in a later bargaining round. If this is the case, then c is not renegotiation-proof. Q.E.D.

The intuition for this lemma is as follows. If condition (a) is not satisfied, then the principal can renegotiate the contract by offering the low effort agent a flat wage contract that makes him just as well off and is cheaper for the principal (since the agent is risk averse). If condition (b) is not met, then the agent will choose $x^* = 0$. Condition (c) is necessary since the compensation scheme intended for the low effort agent cannot provide this agent with less utility than he would get if he chose the compensation scheme intended for the high effort agent because that would increase the cost of compensating low effort agents (because of risk aversion). It cannot provide him with strictly more utility or else the principal could renegotiate the contract to reduce the risk facing a high effort agent (reducing the cost of compensating this agent) while still inducing the low effort agent to choose the full insurance scheme. Similarly, if condition (d) is not satisfied, then the probability of facing a high effort agent is large enough that principal can reduce her expected compensation costs when it is her turn to offer by decreasing the riskiness of the high effort contract even though this requires paying the low effort agent more.

If the original contract must be a F-T contract, then the constraints facing the principal in her period zero contracting problem are identical as in the game where the principal makes the only renegotiation offer. Thus, as the next proposition establishes, the unique equilibrium is identical also.

Proposition 1 *Say there is at least one period $\hat{t} \leq T$ when the principal makes an offer. The optimal final contract that induces $x^* > 0$ and the agent's effort distribution $(x^*, 1 - x^*)$ is identical to the optimal final contract and the agent's effort distribution $(x^*, 1-x^*)$ when there is only one period of renegotiation and the principal makes the take it or leave it renegotiation offer.*

Proof. By Lemmas 1 and 2, the optimal final contract can be implemented as an initial contract that is renegotiation-proof. By Lemma 3, a renegotiation-proof initial contract must be a F-T contract. Fudenberg and Tirole (1990) show that the optimal contract that induces $x^* > 0$ with one period of renegotiation where the principal makes a take it or leave it renegotiation offer must also be an F-T contract. Thus, the principal in this game faces the same problem as the principal in the game of Fudenberg and Tirole (1990), hence the final contract and effort distribution is identical. Q.E.D.

Proposition 1 demonstrates that a more general renegotiation process does not alter the results in Fudenberg and Tirole (1990) so long as the principal gets to make at least one offer and there is no uncertainty about when the outcome will occur (and, thus, about the number of periods of renegotiation). On the other hand, it shows that the Ma (1994) and Matthews (1995) results that renegotiation need not undermine the optimal commitment contract is much less robust. The following corollary gives the most important implication.

Corollary 1 *If there is any period in which the principal makes a renegotiation offer, then the principal cannot induce the agent to choose high effort with probability one.*

Proof. Follows from Proposition 1 and lemma 3.1 of Fudenberg and Tirole (1990). Q.E.D.

If the agent works hard with probability one, then, in the period where the principal makes a renegotiation offer, she will optimally offer a full insurance contract to the high effort agent. Of course, because this is a full insurance contract, it offers the same expected utility to the agent regardless of his effort. But, if the agent works hard with probability one, then the fact that low effort agents receive high effort compensation does not affect the principal's expected compensation, making a full insurance contract for high effort agents optimal. Because the agent will expect such a renegotiation offer, however, he has no incentive to work hard. Thus, there can be no equilibrium where the agent works hard with probability one. Agent offers cannot undermine this outcome since full insurance is ex post optimal.

What this corollary makes clear is that the reason renegotiation does not undermine the commitment outcome in the models of Ma (1994) and Matthews (1995) is not because the agent can make a renegotiation offer in these models; rather, it is because the principal cannot make a renegotiation offer. When there is only one period of renegotiation, there is no distinction between the two. Allowing for multiple rounds of bargaining, however, makes clear that the situation where renegotiation does not undermine the commitment solution to the principal-agent problem is much more limited than the situation where it does.

4 Unknown Number of Periods

The last section generalized the renegotiation process by adding more periods of bargaining. Since the number of periods remained fixed, however, the agent never had to choose among the menu of contracts until after the last period. In addition, only the contract in effect after the last period affected the payoffs for the principal or the agent. In this section, I consider the case where neither the principal nor the agent knows exactly when the outcome will occur. As a result, the contract that is in effect at the end of any given period could be the final contract if the outcome occurs before the next bargaining period. Whenever a contract could be the final contract, the agent must choose from that menu of contracts immediately after that contract is agreed upon. If the outcome does not occur, this choice may provide information to the principal about the agent's type. As will become clear, this greatly affects the renegotiation bargaining.

In this section, I consider the case where only the principal makes offers. If the agent does make offers, then there is always an equilibrium where the agent always re-offers the existing contract because the principal will believe the agent is a low type otherwise. Thus, the case where only the principal makes offers is identical to the case where the agent also makes offers when this equilibrium obtains. The only adjustment is that the probability of the game ending after any given period when the principal makes offers is the sum of the probability of the game ending in any of the periods in between the last offer by the principal and her next offer.

First, consider the case with two periods of renegotiation. That is, the principal offers a contract $c^0 = \{U_g^0(e), U_b^0(e)\}_{e=\overline{e},\underline{e}}$ in period 0. If the agent accepts, then the agent chooses an action in period 1/2. The principal makes a renegotiation offer, $c^1 = \{U_g^1(e), U_b^1(e)\}_{e=\overline{e},\underline{e}}$, in period 1, which the agent accepts or rejects. After this period, the outcome occurs with probability $q_1 = q$. Since the game could end after this period, the agent must also choose a compensation scheme from the menu of the contract in effect (either c^0 or c^1). With probability $1 - q$, the outcome does not occur, and there is a period 2 where the principal makes another renegotiation offer, $c^2 = \{U_g^2(e), U_b^2(e)\}_{e=\overline{e},\underline{e}}$. After the agent accepts or rejects this offer and chooses a compensation scheme, the outcome occurs with probability $q_2 = 1$.

Not surprisingly, one can restrict attention to renegotiation proof contracts.[5] Of course, since the principal will want to offer a different menu of contracts in period 2 than in period 1, a renegotiation proof contract is more complicated than the simple menu of compensation schemes offered in the one period game. That is, a renegotiation proof contract must specify the menu of compensation schemes the agent can choose from in period 1 and a menu of compensation schemes available to the agent in period 2 for each choice of compensation schemes in period 1. That is, $c^{RP} = \{U_g^1(e^1), U_b^1(e^1), U_g^{2e^1}(e^2), U_b^{2e^1}(e^2)\}_{e^1=\overline{e},\underline{e},e^2=\overline{e},\underline{e}}$ where e^1 specifies the compensation scheme chosen in period 1, and e^2 specifies the compensation scheme chosen in period 2.

I analyze this game by comparing it with the one period game, which is governed by the F-T contract $c^* = \{U_g^*(e), U_b^*(e)\}_{e=\overline{e},\underline{e}}$. Of course, the low type agent is offered a fixed wage, so I refer to his utility as $\underline{U}^*$. By condition (c) of the definition of a F-T contract, low types get the same utility from taking the compensation scheme $\underline{e}$ as they would get from pretending to be a high type and taking the compensation scheme $\overline{e}$. Because it is a one period game, one can assume that a low type honestly reveals himself to be a low type by taking the $\underline{e}$ scheme. In a two period game, however, if a low type chooses scheme $\underline{e}$, then the principal will not make the agent any better offer should the game continue to period two. So, a low type can often do better in period two if he chooses scheme $\overline{e}$ in period one. As a result, one has to consider equilibria where low types separate from high types with probability r (in period 1) by taking scheme $\underline{e}$ and pool with probability $1 - r$ by taking scheme $\overline{e}$. The next proposition shows that maximum probability of

[5]The proof of this claim is nearly identical to the proof of Lemma 2.1 (the renegotiation proof lemma) in Fudenberg and Tirole (1990).

high effort that the principal can implement, x^{**}, in the two period game is strictly less than it is in the one period game.

Proposition 2 *Let x^* be the maximum probability of high effort that the principal can implement when the agent receives rent R in the one period game and x^{**} be the maximum probability of high effort that the principal can implement when the agent receives rent R in the two period game. If $q \in (0,1)$, then $x^{**} < x^*$.*

Proof. From Lemma 2.2 of Fudenberg and Tirole (1990), x^* for any given contract is given by:

$$\frac{x^*}{1-x^*} = \frac{(\overline{p}-\underline{p})\Phi'(\underline{U}^*)}{(1-\overline{p})\overline{p}\Phi'(U_g^*(\overline{e})) - \Phi'(U_b^*(\overline{e}))} \tag{1}$$

The agent's rent in this contract is uniquely determined by $\underline{U}^*$. Now consider a renegotiation proof contract in the two period model that gives the agent the same rent. To be renegotiation proof in each period, such a contract must minimize the principal's expected payments among all contracts in that period that satisfy incentive compatibility and individual rationality constraints. In period 1 the binding constraints are the low type's incentive compatibility constraint (he must get the same utility from choosing compensation scheme $\underline{e}$ as he would get from choosing compensation scheme $\overline{e}$) and the high type's individual rationality constraint (the contract must give high type agent at least as much expected utility as he would get from the original contract). These constraints are:

$$q[\underline{p}U_g^1(\overline{e}) + (1-\underline{p})U_b^1(\overline{e})] + (1-q)[\underline{p}U_g^{2\overline{e}}(\overline{e}) + (1-\underline{p})U_b^{2\overline{e}}(\overline{e})] = \underline{U}^* \tag{2}$$

and

$$\begin{aligned} & q[\overline{p}U_g^1(\overline{e}) + (1-\overline{p})U_b^1(\overline{e})] + (1-q)[\overline{p}U_g^{2\overline{e}}(\overline{e}) + (1-\overline{p})U_b^{2\overline{e}}(\overline{e})] \\ = \; & q[\overline{p}\tilde{U}_g^1(\overline{e}) + (1-\overline{p})\tilde{U}_b^1(\overline{e})] + (1-q)[\overline{p}\tilde{U}_g^{2\overline{e}}(\overline{e}) + (1-\overline{p})\tilde{U}_b^{2\overline{e}}(\overline{e})] \end{aligned} \tag{3}$$

This must hold for any feasible contracts $\tilde{c} = \{\tilde{U}_g^1(e), \tilde{U}_b^1(e), \tilde{U}_g^{2\underline{e}}(e), \tilde{U}_b^{2\underline{e}}(e), \tilde{U}_g^{2\overline{e}}(e), \tilde{U}_b^{2\overline{e}}(e)\}_{e=\overline{e},\underline{e}}$. By differentiating (2) and (**??**) with respect to $\underline{U}^*$, one can determine that $\frac{dU_g^{1*}(\overline{e})}{d\underline{U}^*} = -\frac{1-\overline{p}}{q(\overline{p}-\underline{p})}$ and $\frac{dU_b^{1*}(\overline{e})}{d\underline{U}^*} = \frac{\overline{p}}{q(\overline{p}-\underline{p})}$.[6]

The principal's expected compensation costs are:

[6] $U_g^{2\overline{e}*}(\overline{e})$ and $U_b^{2\overline{e}*}(\overline{e})$ do not vary with $\underline{U}^*$. This follows from the fact the period 2 interim incentive compatibility and interim individual rationality constraints do not depend on $\underline{U}^*$.

$$x\{q[\overline{p}\Phi(U_g^1(\overline{e}))+(1-\overline{p})\Phi(U_b^1(\overline{e}))]+(1-q)[\overline{p}\Phi(U_g^{2\overline{e}}(\overline{e}))+(1-\overline{p})\Phi(U_b^{2\overline{e}}(\overline{e}))\}$$
$$+(1-x)\{r\Phi(\underline{U}^*)+ \tag{4}$$
$$(1-r)[q(\underline{p}\Phi(U_g^1(\overline{e}))+(1-\underline{p})\Phi(U_b^1(\overline{e})))+(1-q)\Phi(\underline{p}U_g^{2\overline{e}}(\overline{e})+(1-\underline{p})U_b^{2\overline{e}}(\overline{e}))]\}$$

Differentiating this with respect to $\underline{U}^*$ gives:

$$-\frac{x\overline{p}(1-\overline{p})(\Phi'(U_g^1(\overline{e}))-\Phi'(U_b^1(\overline{e})))}{\overline{p}-\underline{p}} \tag{5}$$
$$+(1-x)\{\frac{(1-r)[\overline{p}(1-\underline{p})\Phi'(U_b^1(\overline{e}))-\underline{p}(1-\overline{p})\Phi'(U_g^1(\overline{e}))]}{\overline{p}-\underline{p}}+r\Phi'(\underline{U}^*)\}$$

For the contract to be renegotiation proof, this must be non-negative (otherwise, the principal could offer to increase $\underline{U}^*$ in period 1 and lower her expected payments). (5)≥ 0 is equivalent to:

$$\frac{x^{**}}{1-x^{**}}\leq\frac{(\overline{p}-\underline{p})\Phi'(\underline{U}^*)-(1-r)\{\overline{p}(1-\underline{p})[\Phi'(\underline{U}^*)-\Phi'(U_b^1(\overline{e}))]+\underline{p}(1-\overline{p})[\Phi'(U_g^1(\overline{e}))-\Phi'(\underline{U}^*)]\}}{\overline{p}(1-\overline{p})(\Phi'(U_g^1(\overline{e}))-\Phi'(U_b^1(\overline{e}))} \tag{6}$$

Say $x^{**}\geq x^*$. The term in curly braces is clearly positive, making the right hand side numerator weakly (strictly for $r<1$) smaller than the numerator on the right hand side of (1). Thus, $x^{**}\geq x^*$ requires that the right hand side denominator be weakly smaller (strictly for $r<1$) than the denominator on the right hand side of (1). This holds if and only if $\Phi'(U_g^1(\overline{e}))-\Phi'(U_b^1(\overline{e}))\leq \Phi'(U_g^*(\overline{e}))-\Phi'(U_b^*(\overline{e}))$. This is true if and only if $U_g^1(\overline{e})-U_b^1(\overline{e})\leq U_g^*(\overline{e})-U_b^*(\overline{e})$. For any $r\geq 0$, it must be that $U_g^{2\overline{e}}(\overline{e})-U_b^{2\overline{e}}(\overline{e})\leq U_g^*(\overline{e})-U_b^*(\overline{e})$ (and strictly for $r>0$) since condition (1) applies to the second period of the two period game where x^* is replaced by $\frac{x^{**}}{1-r(1-x^{**})}\geq x^*$. But $U_g^{2\overline{e}}(\overline{e})-U_b^{2\overline{e}}(\overline{e})\leq U_g^*(\overline{e})-U_b^*(\overline{e})$ (and strictly for $r>0$) together with (2) implies that $U_g^1(\overline{e})-U_b^1(\overline{e})\geq U_g^*(\overline{e})-U_b^*(\overline{e})$ (and strictly for $r>0$) so long as $q\in(0,1)$. This contradicts $x^{**}\geq x^*$. Q.E.D.

When there is uncertainty about the number of periods, adding more periods of renegotiation can continue to undermine the principal's ability to induce high effort. The problem is that the principal's incentive to offer the agent a less risky contract in period one is now even greater than before. In the one period game, doing so reduced payments to a high effort agent but increased them to a low effort agent. Now, renegotiating the contract to make it less risky also reduces payments (this period) when one is facing a low effort agent who is pretending to be a high effort agent in order to get a more favorable contract offer should the game continue to the next period.

Adding uncertainty to amount of possible renegotiation has an even more fundamental effect. Recall that Fudenberg and Tirole (1990) show that with one period of renegotiation, renegotiation will often only affect the set of feasible distributions the principal can implement. If the agent's degree of risk aversion does not decrease too fast with his wealth, then renegotiation does not affect the optimal contract that implements a given effort distribution. The next proposition shows that this result does not carry over when the number of renegotiation periods is uncertain.

Before stating the proposition, I further examine the principal's contracting problem. Since the agent is risk averse, a flat wage scheme is the cheapest way for the principal to give the agent any given level of utility. So, in any contract, scheme $\underline{e}$ will be a flat wage scheme. If the agent takes this scheme in period 1, then there is no scope for renegotiation in period 2. The principal will simply re-offer the same scheme. Thus, the contracts the principal will offer can be narrowed down to $\{U_g^1(e), U_b^1(e), U_g^{2\overline{e}}(e), U_b^{2\overline{e}}(e)\}_{e=\overline{e},\underline{e}}$. These contracts are subject to some constraints. First, notice that if the agent takes scheme $\overline{e}$ in period 1, then in period 2 the principal faces the same situation as she faces in period 1 of the one period game. At this point, she has no incentive to offer the high type agent any more rent than she would receive from the period 1 contract. That is, scheme $\overline{e}$ in period 2 must satisfy:

$$\overline{p}U_g^1(\overline{e}) + (1-\overline{p})U_b^1(\overline{e}) = \overline{p}U_g^{2\overline{e}}(\overline{e}) + (1-\overline{p})U_b^{2\overline{e}}(\overline{e}) \tag{7}$$

In addition, there is an analog to Fudenberg and Tirole's (1990) renegotiation proof constraint for period 2. The only difference is that the fact that the agent chose scheme $\overline{e}$ in period 1 changes the probability that he is the high type from x (the original probability the agent exerted high effort) to $\frac{x}{1-r(1-x)}$ (the probability the agent is the high type in period two given that low types choose scheme $\overline{e}$ in period 1 with probability $(1-r)$). This constraint holds at equality since the principal wants to induce the highest probability of high effort that she can.[7]

$$\Phi'(U_g^{2\overline{e}}(\overline{e})) - \Phi'(U_b^{2\overline{e}}(\overline{e})) = \frac{(\overline{p}-\underline{p})(1-x)(1-r)\Phi'(\underline{U^{2\overline{e}}})}{(1-\overline{p})\overline{p}x} \tag{8}$$

The last two constraints come from Proposition 2. It shows that for any given x that the

[7] Fudenberg and Tirole (1990) show that principal can make this largest x the unique equilibrium.

principal wants to implement, she must give the agent at least as much rent as he would receive in the one-period game. The following constraints, then, represent the best case for the principal: she gives each agent type exactly the same expected utility as they would receive in the one period game.

$$q[\overline{p}U_g^1(\overline{e}) + (1-\overline{p})U_b^1(\overline{e})] + (1-q)[\overline{p}U_g^{2\overline{e}}(\overline{e}) + (1-\overline{p})U_b^{2\overline{e}}(\overline{e})] = \overline{p}U_g^*(\overline{e}) + (1-\overline{p})U_b^*(\overline{e}) \tag{9}$$

$$q[\underline{p}U_g^1(\overline{e}) + (1-\underline{p})U_b^1(\overline{e})] + (1-q)[\underline{p}U_g^{2\overline{e}}(\overline{e}) + (1-\underline{p})U_b^{2\overline{e}}(\overline{e})] = \underline{p}U_g^*(\overline{e}) + (1-\underline{p})U_b^*(\overline{e}) = \underline{U} \tag{10}$$

If the principal cannot implement a given x at the same or lower costs as in the one period game subject to these constraints, then her costs of implementing x are necessarily higher than in the one period game. The following proposition shows that this is the case.

Proposition 3 *If $q \in (0,1)$, then the cost of implementing any given probability of high effort x is strictly greater when there are (potentially) two periods of renegotiation than it is in the one period game.*

Proof. The difference in the cost of implementing any given x is given by:

$$\begin{aligned}
&x\{q[\overline{p}\Phi(U_g^1(\overline{e})) + (1-\overline{p})\Phi(U_b^1(\overline{e}))] + (1-q)[\overline{p}\Phi(U_g^{2\overline{e}}(\overline{e})) + (1-\overline{p})\Phi(U_b^{2\overline{e}}(\overline{e}))] \\
&-[\overline{p}\Phi(U_g^*(\overline{e})) + (1-\overline{p})\Phi(U_b^*(\overline{e}))]\} \\
&+(1-x)(1-r)\{[q(\underline{p}\Phi(U_g^1(\overline{e})) + (1-\underline{p})\Phi(U_b^1(\overline{e}))) + (1-q)\Phi(\underline{p}U_g^{2\overline{e}}(\overline{e}) + (1-\underline{p})U_b^{2\overline{e}}(\overline{e}))] \\
&-\Phi(\underline{p}U_g^*(\overline{e}) + (1-\underline{p})U_b^*(\overline{e}))\}
\end{aligned} \tag{11}$$

I show that this expression is always positive. First, I show that the term in curly braces that is multiplied by $(1-x)(1-r)$ is strictly positive. By (10):

$$\Phi(\underline{p}U_g^*(\overline{e}) + (1-\underline{p})U_b^*(\overline{e})) =$$

$$\Phi(q[\underline{p}U_g^1(\overline{e}) + (1-\underline{p})U_b^1(\overline{e})] + (1-q)[\underline{p}U_g^{2\overline{e}}(\overline{e}) + (1-\underline{p})U_b^{2\overline{e}}(\overline{e})]$$

Since Φ is convex, this implies that:

$$\Phi(\underline{p}U_g^*(\overline{e})+(1-\underline{p})U_b^*(\overline{e}))<$$
$$q(\underline{p}\Phi(U_g^1(\overline{e}))+(1-\underline{p})\Phi(U_b^1(\overline{e})))+(1-q)\Phi(\underline{p}U_g^{2\overline{e}}(\overline{e})+(1-\underline{p})U_b^{2\overline{e}}(\overline{e}))$$

This proves that this term is positive.

Now, consider the expected payments to high types, the term in curly braces that is multiplied by x. Constraints (7) and (9) imply that the high type's expected utility in each period is identical to his expected utility in the one period game. By (10), the low type's expected utility from the period 1 compensation scheme for high types must be less than $\underline{U}$ if and only if $\underline{U^{2\overline{e}}}>\underline{U}$ (he gets more utility in period 2). Together, these imply that $\underline{U^{2\overline{e}}}>\underline{U}$ if and only if:

$$U_g^1(\overline{e})-U_b^1(\overline{e})>U_g^*(\overline{e})-U_b^*(\overline{e})>U_g^{2\overline{e}}(\overline{e})-U_b^{2\overline{e}}(\overline{e}) \tag{14}$$

If $r=0$, however, then (8) implies that $U_g^*(\overline{e})-U_b^*(\overline{e})>U_g^{2\overline{e}}(\overline{e})-U_b^{2\overline{e}}(\overline{e})$ if and only if $\underline{U^{2\overline{e}}}<\underline{U}$. So, when $r=0$ one must have $\underline{U^{2\overline{e}}}=\underline{U}$ and $U_g^*(\overline{e})-U_b^*(\overline{e})=U_g^{2\overline{e}}(\overline{e})-U_b^{2\overline{e}}(\overline{e})$. The fact that the high type's expected utility in each period is identical to his expected utility in the one period game means that, if $r=0$, the utility from the good (and bad) outcome is identical in all three cases. This, of course, means that if $r=0$, the difference in the principal's expected payments to the high type (between the two period game and the one period game) is zero. So, if I can show that the difference in expected payment to the high type is strictly increasing in r, then I have proven that the difference in expected payments to the high types is always non-negative (which proves the result).

By differentiating the four constraints, (7), (8, (9), and (10) with respect to r, one can determine how $U_g^1(\overline{e}),U_b^1(\overline{e}),U_g^{2\overline{e}}(\overline{e}),U_b^{2\overline{e}}(\overline{e})$ vary with r:

$$\frac{dU_g^1(\overline{e})}{dr}=\frac{(\overline{p}-\underline{p})(1-x)(1-q)\Phi'(\underline{U})}{\overline{p}qx[\overline{p}\Phi''(U_b^{2\overline{e}}(\overline{e}))+(1-\overline{p})\Phi''(U_g^{2\overline{e}}(\overline{e}))]} \tag{15}$$

$$\frac{dU_b^1(\overline{e})}{dr}=-\frac{(\overline{p}-\underline{p})(1-x)(1-q)\Phi'(\underline{U})}{(1-\overline{p})qx[\overline{p}\Phi''(U_b^{2\overline{e}}(\overline{e}))+(1-\overline{p})\Phi''(U_g^{2\overline{e}}(\overline{e}))]} \tag{16}$$

$$\frac{dU_g^{2\overline{e}}(\overline{e})}{dr} = -\frac{(\overline{p} - \underline{p})(1-x)\Phi'(\underline{U})}{\overline{p}x[\overline{p}\Phi''(U_b^{2\overline{e}}(\overline{e})) + (1-\overline{p})\Phi''(U_g^{2\overline{e}}(\overline{e}))]} \tag{17}$$

$$\frac{dU_b^{2\overline{e}}(\overline{e})}{dr} = \frac{(\overline{p} - \underline{p})(1-x)\Phi'(\underline{U})}{(1-\overline{p})x[\overline{p}\Phi''(U_b^{2\overline{e}}(\overline{e})) + (1-\overline{p})\Phi''(U_g^{2\overline{e}}(\overline{e}))]} \tag{18}$$

Using these, differentiating the difference in expected payment to the high types with respect to r gives:

$$\frac{(\overline{p} - \underline{p})(1-x)(1-q)\Phi'(\underline{U})\{[\Phi'(U_g^1(\overline{e})) - \Phi'(U_b^1(\overline{e}))] - [\Phi'(U_g^{2\overline{e}}(\overline{e})) - \Phi'(U_b^{2\overline{e}}(\overline{e}))]\}}{\overline{p}qx[\overline{p}\Phi''(U_b^{2\overline{e}}(\overline{e})) + (1-\overline{p})\Phi''(U_g^{2\overline{e}}(\overline{e}))]} \tag{19}$$

The sign of this is the sign of:

$$[\Phi'(U_g^1(\overline{e})) - \Phi'(U_b^1(\overline{e}))] - [\Phi'(U_g^{2\overline{e}}(\overline{e})) - \Phi'(U_b^{2\overline{e}}(\overline{e}))] \tag{20}$$

By (14), this is negative only if $\underline{U^{2\overline{e}}} < \underline{U}$. But $\underline{U^{2\overline{e}}} < \underline{U}$ and $r > 0$ and (8) imply that (20) is positive, a contradiction. This proves the result. Q.E.D.

Unlike the effect of adding one period of certain renegotiation, adding the possibility of a second period not only reduces the maximum probability of high effort, it also makes any given probability of high effort more expensive. Because the optimal renegotiation proof contract with one period of renegotiation is not renegotiation proof when another period is added, the principal has to use the first period contract to screen. This means that the first period contract is riskier than in the one period game and that a low effort type does not choose the full insurance contract in the first period with probability one. So, both types of agents face more first period risk, increasing the cost to the principal. The second period contract for high effort types is safer, but this just creates a lottery among two different contracts for the high effort agent. Since he is risk averse, this lottery necessarily increases the cost of compensating him. So, even though the uncertainty as to the timing of the outcome does not create any increase in uncertainty as to total output (there is no discounting here), because the principal cannot commit to a given contract, it has the same effect as an increase in the riskiness of the output.

Not surprisingly, this effect applies when there is the potential to be even more than two periods of renegotiation. In a more general model, where there can be anywhere from two to T periods of renegotiation, Proposition 3 continues to hold.

Proposition 4 *Let $T \geq 2$. If $q_t \in (0,1)$ for $t < T$ and $q_T = 1$, then the cost of implementing any given probability of high effort x is strictly greater than it is in the one period game.*

Proof. The proof is by induction. Proposition 3 showed that the result holds for $T = 2$. Assume it is true for $T-1$ periods. Any T period setting is like a $T-1$ period setting with the last period split into two. So, if costs could be lower with T periods, then consider the $T-1$ setting where $q_{T-1} = 1$ but everything else is identical to the T period setting. Now consider the same contract sequence for the first $T-2$ periods as the T period contract sequence that gives lower costs than the one period game. For period $T-1$, offer the optimal F-T contract that gives each type the same expected utility he would get from the last two periods in the T period case. Then, collapsing these periods together does not change incentives to select among the menus of the prior contracts. By the argument above for the two period case, collapsing the last two periods into one is strictly cheaper for the principal. So, if the T period setting reduced compensation costs, then the $T-1$ period setting must also, which contradicts the inductive hypothesis. Q.E.D.

Whenever there is any uncertainty about when the outcome will occur and multiple periods of renegotiation are possible, then it is more costly to induce any distribution of effort than when there is only one (or any fixed number) of periods of renegotiation. In many settings, this is likely to be the case. In sharecropping contracts, the exact day that the crop will yields are realized is unlikely to be known with certainty. When the principal has hired an agent to develop or run a new project, often the profits from this project will come at an uncertain date. This is especially true when the project involves selling a good, since sales are usually realized at least somewhat stochastically. In fact, if contracts can be renegotiation fairly quickly, there need not be much uncertainty as to when the outcome will occur for the results in this section to apply.

5 Conclusion

With only one round of renegotiation, the effect of renegotiation on agency contracts has been found to be very sensitive to which party, the principal or the agent, makes the offer. In this paper, I analyzed the robustness of these two different models by allowing for multiple rounds of renegotiation. With a fixed number of rounds, I showed that as long as the principal can make at least one renegotiation offer, offers by the agent do not affect the optimal contract. The equilibrium is identical to the one where the only offer is made by the principal, in which renegotiation does undermine the optimal contract with commitment. This suggests that the key difference in the one period models is not which party makes the offer, but whether the principal gets to make an offer. Unless the principal will not ever get the chance to make a renegotiation offer, renegotiation does prevent the principal from inducing the agent to work hard with probability one.

I then analyze the case where there is uncertainty as to when the verifiable outcome will occur, which generates uncertainty as to the number of available rounds for renegotiation. I found that in this setting, the renegotiation problem is even more severe than in the one period game. Allowing for an uncertain number of rounds both further limits the distributions of actions the principal can induce and raises her cost of inducing any given distribution. That is, not only is the pessimistic result in Fudenberg and Tirole (1990) about the effect of renegotiation more robust than the more optimistic one in Ma (1994), but the negative effects of renegotiation may often be even stronger than in the Fudenberg and Tirole model.

References

[1] Drew Fudenberg and Jean Tirole. Moral hazard and renegotiation in agency contracts. *Econometrica*, 58:1279–1319, 1990.

[2] Sanford J. Grossman and Oliver D. Hart. An analysis of the principal-agent problem. *Econometrica*, 51:7–46, 1983.

[3] Benjamin E. Hermalin and Michael L. Katz. Moral hazard and verifiability: The effects of renegotiation in agency. *Econometrica*, 59:1735–1754, 1991.

[4] Bengt Holmstrom. Moral hazard and observability. *Bell Journal of Economics*, 10:74–91, 1979.

[5] Ching-To Albert Ma. Renegotiation and optimality in agency contracts. *The Review of Economic Studies*, 61:109–129, 1994.

[6] Steven A. Matthews. Renegotiation of sales contracts. *Econometrica*, 63:567–589, 1995.

[7] Steven Shavell. Risk sharing and incentives in the principal and agent relationship. *Bell Journal of Economics*, 10:55–73, 1979.