Renegotiation-Proof Contracts with Persistent Private Information

Bruno Strulovici

October 11, 2020

Abstract

This paper introduces a new approach to model renegotiation in stochastic games and applies it to study how renegotiation shapes long-term contracts in principal-agent relationships with persistent private information. The structure of players’ payoff and state dynamics generates an algebraic structure over contractual equilibria and determines the set of alternatives considered through renegotiation. Using recent advances in functional stochastic differential equations, the paper derives an Observability Theorem and a Revelation Principle to address asymmetric information over persistent variables. Truthful renegotiation-proof contracts are characterized by a single number—their sensitivity to the agent’s report—and are self-correcting off the equilibrium path. The sensitivity of the optimal contract is increasing in information persistence and decreasing in players’ patience.

1 Introduction

Financial bailouts, fiscal policies, and corporate pay cuts during recessions are instances in which regulators, firms, and other economic agents have the possibility to renegotiate explicit or implicit contracts in reaction to shocks in their environment. In these instances, moreover, state variables such as agents’ revenue or productivity are often privately observed and correlated over time.

Despite its relevance for economics, renegotiation has proved challenging to model and analyze. Even for repeated games, there is no universally accepted concept of renegotiation-proof equilibrium. One concept, internal consistency, is often perceived as a minimal requirement, and defined as follows: an equilibrium is internally consistent if it does not have a continuation equilibrium that is Pareto dominated by a second continuation equilibrium. Presumably, players facing the first continuation equilibrium could think of moving to the second, Pareto superior continuation.

*Northwestern University. This paper subsumes the 2011 working paper entitled “Renegotiation-Proof Contracts with Moral Hazard and Persistent Private Information.” I am particularly grateful to Alex Bloedel, Hector Chade, Roozbeh Hosseini, R. Vijay Krishna, Nikolai Krylov, Alessandro Pavan, as well as to Bruno Biais, Mehmet Ekmekci, Daniel Garrett, Brett Green, Christian Hellwig, Johannes Hörner, Francisco Ruiz-Aliseda, Larry Samuelson, Ilya Segal, Itai Sher, Juuso Toikka, and various seminar audiences for questions and comments. Part of this research was conducted while I was visiting Kyoto University, who hospitality is gratefully acknowledged.
equilibrium, which would destroy the initial equilibrium.\footnote{Internal consistency is due to Bernheim and Ray (1989) and closely related to Farrell and Maskin’s (1989) weakly renegotiation-proof equilibria. Although internal consistency is a concept mild, Abreu and Pearce (1991) and Asheim (1991) point out that it does assume time invariance.} Internal consistency is weak because it does not impose any comparison with any other equilibrium. For example, indefinitely playing a bad Nash equilibrium is internally consistent.

At the extreme opposite, strongly renegotiation proof equilibria (Farrell and Maskin (1989)) require comparisons across all internally consistent equilibria and may fail to exist. Other concepts involve a fixed point problem determining the set of equilibria that are renegotiation proof (see Asheim (1991)) and are subject to existence and multiplicity issues.

**State Consistency:** These problems would seem even more severe in environments, pervasive in economics, in which agents are subject to persistent shocks. In these environments, however, a strengthening of internal consistency can prove powerful enough to drastically reduce the set of renegotiation-proof equilibria and yield sharp predictions about their properties. The idea is to compare continuation payoffs of a given equilibrium not only across histories leading to the same state, but also across histories leading to distinct states, thus creating a concept of state consistency.

For example, if a state variable describes the current productivity of a firm, one cannot directly compare continuation equilibria starting from distinct states since the physical environment (production frontier) is different across states. However, one may be able to transform a continuation equilibrium starting from a given state into another continuation equilibrium starting from a different state, for example by using some homotheticity argument. One may then compare the continuation payoffs of equilibrium transformations starting from the same state and use Pareto dominance to rule out some of these continuations. State consistency is weaker than strongly renegotiation proofness and provides a natural way to construct equilibria in stochastic games: it does not require that players envision radically different equilibria, but only that they think by analogy with the equilibrium that they are playing.

State consistency imposes an algebraic structure on renegotiation-proof equilibria. By transforming an equilibrium starting from one state into equilibria starting from other states, one generates an orbit within the set of equilibria, in a group-theoretic sense that the paper makes precise. This orbit has a structure that reflects the nature of the transformation. State consistency imposes a comparison between continuations of an equilibrium and the orbit that it generates. As a result, a state-consistent equilibrium inherits many properties of its orbit.

Together with an additional comparison (to convex combinations of equilibria), state consistency pins down the structure of all renegotiation-proof equilibria and reduces the family of renegotiation-proof equilibria.
proof equilibria to a simple, low-dimensional family, which may be viewed as the quotient of the set of renegotiation-proof equilibria with respect to their orbits.

**Private Information and Renegotiation:** An additional complexity arises if some persistent state variables are observed only by some party. Even if parties can communicate, they may misreport private information, and parties could have to renegotiate under asymmetric information. A simple solution is to restrict attention to equilibria in which parties truthfully reveal their information. In a truthful equilibrium, all parties know the state variables at all times and state consistency can be used to study truthful, renegotiation-proof equilibria.

With renegotiation, however, focusing on truthful equilibria may *a priori* be restrictive, since the absence of commitment precludes the use of the standard Revelation Principle. However, renegotiation also incites parties to try and erase any inefficiency in ongoing agreements, which requires some exchange of information. When parties can communicate frequently, parties can renegotiate away any inefficiency stemming from information asymmetries and cannot commit not to do so. If efficiency requires full information disclosure, focusing on truthful equilibrium may thus be without loss of generality when parties can continually communicate.²

**Observability and Revelation Principle in a Diffusion Model:** This paper takes a different approach to address private information. It studies a continuous-time model, in which communication continually takes place. In the model, due to Williams (2011), an agent privately observes a stochastic cash flow process, which he reports to a principal. The agent can lie: if the true cash flow process is \( \{X_t\}_{t \geq 0} \), the agent reports a cash flow process \( \{Y_t\}_{t \geq 0} \) where \( Y_t \) is a function of \( \{X_s\}_{s \leq t} \) and of the agent’s reporting strategy \( L \).

The agent’s equilibrium strategy, \( L \), is known to the principal. Therefore, given a reporting history \( \{Y_s\}_{s \leq t} \) the principal’s belief about \( X_t \) need not coincide with \( Y_t \). For example, if the agent often underreports his cash flow in equilibrium, the principal’s belief about \( X_t \) will be higher than the report \( Y_t \). The question, then, is to determine how much information the principal can back out about \( X \) from observing the process \( Y \) and knowing the reporting strategy \( L \). Since \( L \) can depend

²The main theorem of Strulovici (2017) formalizes this intuition: it shows in a principal-agent model with asymmetric information that when (i) allocation efficiency requires full information disclosure and (ii) parties can communicate with arbitrarily high frequency before the allocation takes place, information is fully revealed and the transaction is efficient. Maestri (2017) shows a similar result when parties transact in each period: information is revealed arbitrarily quickly relative to the parties’ discount rate as parties become arbitrarily patient. These papers consider an explicit game of renegotiation—with offer and acceptance decisions—whereas renegotiation in the present paper is modeled using a set-theoretic approach.

³Mathematically, the lying process \( L \) is adapted to the filtration generated by \( X \) and is assumed to enter the reporting process \( Y \) through its drift: \( dY_t = dX_t + L_t dt \) until Section 7, which considers reporting jumps.
arbitrarily on history, the question would seem a priori hard to answer.

Using recent advances in the analysis of functional stochastic differential equations, this paper provides the following Observability Theorem: (i) Under very mild regularity conditions on $L$, the principal can infer the true process $X$, (ii) Even without such conditions, there always exists a belief equation based on the principal’s information, which, whenever it has a well-defined solution, reveals the true cash flow $X$.

The Observability Theorem implies that the principal knows the agent’s cash flow and continuation utility at all times and that parties have symmetric information at all times. This addresses the difficulty raised by private information and validates the state consistency approach.

Building on this result, the paper provides the following Revelation Principle: any state-consistent equilibrium is outcome equivalent to a truthful state-consistent equilibrium, as long as the transformation group used to define state consistency satisfies a simple monotonicity condition.

**Renegotiation-Proof Contracts with Persistent Private Information:** The application studied in this paper revisits a central question in macroeconomics. The agent who generates the cash flow is risk averse and insured by a risk-neutral principal. The agent reports and transfers his cash flow to the principal and receives in exchange a transfer from the principal (equivalently, the agent is subsidized or taxed depending on his cash flow report). The principal proposes a contract (a transfer process adapted to the agent’s report process) to the agent, and an equilibrium consists of a contract and a reporting strategy for the agent. The principal wishes to give the agent some expected lifetime utility at the smallest possible cost. The basic tension is that, in order to properly insure the agent, the principal must know the agent’s cash flow, but the agent may benefit from underreporting cash flow to get a higher subsidy (or lower tax) or, conversely, overreport his cash flow to get rewarded by a higher continuation utility.

From the principal’s perspective there are two state variables: the current cash flow and the agent’s continuation utility. When the agent has exponential utility, a natural class of equilibrium transformations from one state to another emerges, which creates a rich and well structured class of challengers for the equilibrium and its continuations.

Any truthful renegotiation-proof contract (i.e., a contract with which truth-telling forms an equilib-

---

4A functional SDE is an SDE whose drift and volatility at time $t$ depend on the path of the solution until $t$.

5The theorem provides two independent sufficient conditions: a local Lipschitz condition and an “arbitrarily small delay” condition, which means that the agent’s lying strategy cannot depend on what he observed in the last $\epsilon$ units of time, where $\epsilon$ is arbitrarily small. This condition is particularly mild because the process $X$ is continuous. The theorem also requires a local boundedness condition.

6See Thomas and Worrall (1990), Williams (2011), Bloedel, Krishna, and Leukhina (2020) and references therein.
rium) is characterized by a single “sensitivity” parameter, which determines the agent’s incentive to truthfully report his cash flows. For such a contract, all contractual variables have exact formulas as a function of the sensitivity parameter and the continuation utility of the agent. The sensitivity parameter describes how the agent’s continuation utility varies with his reports, and can take any value between 0 and the coefficient of absolute risk aversion of the agent.

The class of renegotiation-proof contracts contains the contract studied by Williams (2011) and is equivalent to the class of stationary contracts studied by Bloedel, Krishna, and Strulovici (2020), who provide a self-insurance implementation of these contracts. The optimal renegotiation-proof contract is obtained by maximizing a closed-form objective with respect to the sensitivity parameter, which makes it easy to derive comparative statics of the optimal sensitivity parameter with respect to information persistence, discounting, and risk. The agent’s flow utility optimally has a negative drift for all parameters of the model, but the occurrence of immiserisation depends on the strength of cash flow persistence and volatility (Bloedel, Krishna, and Strulovici (2020)).

**Reporting Incentives and Reporting Jumps:** Reporting incentives are linear for any arbitrary contract, which implies that the agent either is indifferent between telling the truth and lying, or wishes to lie at maximal (infinite) rate, either upwards or downwards. To account for this, the agent’s strategy space is enlarged to allow jumps in the agent’s reports. For the contracts characterized in this paper, there is a natural way to specify how such jumps affect the agent’s continuation utility. The agent’s incentives are characterized by a Hamilton-Jacobi-Bellman equation with an impulse response component, which provides a new (in the contracting literature, to the author’s knowledge) and simple way of dealing with the possibility of unbounded drift of the reporting process. This technique is used to derive in closed form the agent’s value function not only on the equilibrium path, but also after any possible deviation.

With truthful renegotiation-proof contracts, the agent wants to report cash flows truthfully not only on the equilibrium path, but also after any possible deviation. If, there was any mistake in the report, it is strictly optimal for the agent to immediately correct this mistake.\(^7\)

**Literature Review:** This paper proposes a way to analyze renegotiation in stochastic games and contracting environments with persistent types,\(^8\) inspired by the concepts introduced by Bernheim  

\(^7\)This feature is interesting, for instance, if the agent is a newly-arrived CEO who discovers, upon taking the job, that the financial situation of his firm is worse than what outsiders think. The contracts characterized here give the agent the incentive to correctly book a nonrecurring loss on the firm’s accounts.

\(^8\)Hart and Tirole (1988), Laffont and Tirole (1990) Dewatripont (1989), Fudenberg and Tirole (1990), and Battaglini (2007) study contract renegotiation with private information in finite period models, which impose de facto constraints on the frequency of communication. Maestri (2017) studies an infinite horizon version of Hart and Tirole (1988) and finds that information is revealed arbitrarily quickly relative to the discount rate as parties become
and Ray (1989) and Farrell and Maskin (1989). Gromb (1994) studies a binary-state model of debt contracts and compares payoffs across the two states, similarly to the approach of this paper. Ray (1994) proposes a reinforcement of internal consistency, which may be interpreted as follows: if parties are aware of all continuations of an equilibrium (as they should under internal consistency), they could build new equilibria recursively by using the set of continuation payoffs to incentivize current-period actions (see also Van Damme (1991)). Maestri (2017) studies explicit, dynamic renegotiation between a principal and an agent with a private type.

The paper also builds on the literature on dynamic contracting with persistent private information initiated by Fernandes and Phelan (2000). Most modeling features of this paper are based on Williams (2011) who focuses on full commitment. The model is related to optimal insurance and taxation models studied by Green (1987), Thomas and Worrall (1990), Golosov, Kocherlakota, and Tsyvinski (2003), Farhi and Werning (2013). Golosov and Iovino (2020) study optimal insurance without commitment. Pavan, Segal, and Toikka (2014) study persistence in discrete time through “impulse response functions.” Contracting with persistent private information also arises in delegated experimentation models (Bergemann and Hege (2005), Garfagnini (2011), Hörner and Samuelson (2013)), as an agent’s private effort to learn about some technology may result in persistent superior information.

The paper is organized as follows. Section 2 presents the contractual setting. Section 3 introduces state consistency, a concept of renegotiation-proofness for stochastic games, and establishes an Observability Theorem and a Revelation Principle to address private information. Section 5 characterizes truthful, renegotiation-proof contracts. Section 6 studies how the sensitivity of the optimal contract to the agent’s reports varies with persistence, discounting, and risk. Section 7 provides a necessary and sufficient condition for truthfulness on and off the equilibrium path. Section 8 discusses several extensions.

---

12 See also Guo (2016) and Halac, Kartik, and Liu (2016). Similarly in DeMarzo and Sannikov (2016), principal and agent both learn the agent’s skill but the agent privately observes his effort. Sannikov (2014) considers a moral hazard problem in which the agent’s action have long-term consequences. Garrett and Pavan (2012, 2015) study managerial compensation contracts when the type of the manager is persistent. These papers do not consider renegotiation.
2 Setting

An agent generates cash flow $X_t \in \mathbb{R}$ at time $t \geq 0$, which evolves according to the dynamic equation\(^{13}\)

$$dX_t = [\xi - \lambda X_t] \, dt + \sigma dB_t$$  \hspace{1cm} (1)

where $B$ is the standard Brownian motion. The cash flow has a mean-reversion component with speed $\lambda$ and long run average $\xi / \lambda$. A low (high) mean-reversion speed $\lambda$ results in high (low) persistence of the cash flows and, hence, of the agent’s private information. $\lambda$ is the rate at which a shock in the current cash flow decays over time. Uncertainty is modeled with a probability space $(\Omega, \mathcal{F}, P)$ satisfying the usual conditions and whose outcomes $\omega$ are identified with the paths of $B$.

The agent reports and transfers to the principal a cash flow $Y_t$ that obeys the dynamic equation

$$dY_t = dX_t + L_t \, dt = [\xi - \lambda X_t + L_t] \, dt + \sigma dB_t$$  \hspace{1cm} (2)

where $L_t$ is the rate at which the agent lies about the increment $dX_t$ of the true cash flow.\(^{14}\)

The gap $G_t = Y_t - X_t$ between reported and actual cash flows satisfies

$$G_t = \int_0^t L_s ds.$$

One could impose additional constraints on the agent’s lies without affecting the main results.\(^{15}\)

The agent’s initial cash flow is assumed to be known by the principal.

Assumption 1 (i) The principal observes the report process $\{Y_t\}_{t>0}$ but not the actual cash flow process $\{X_t\}_{t>0}$. (ii) $Y_0 = X_0$.

Contract: A contract is a stochastic process $C = \{C_t\}_{t \geq 0}$ that is adapted to the filtration $\mathbb{F}^Y = \{\mathcal{F}_t^Y\}_{t \geq 0}$ generated by the report process $Y$. The process $C$, discounted at rate $r$, is assumed to be integrable when the agent tells the truth: $E[\int_0^\infty e^{-rt} |C_t| \, dt \mid L = 0] < \infty$.

\(^{13}\)The model is due to Williams (2011). An earlier version of this paper (Strulovici (2011)) includes moral hazard: the agent privately chooses some effort level $A_t \in \mathbb{R}$ that affects the cash flow dynamics according to $dX_t = (A_t + \xi - \lambda X_t) \, dt + \sigma dB_t$. This addition does not affect the main results, as explained in Section 8.2.

\(^{14}\)The rate at which the agent lies may be unbounded. To address this, Section 7 allows jumps in the agent’s report process and proposes a natural extension of the contract to this case.

\(^{15}\)One could require that $Y_t \leq X_t$ (the agent cannot transfer more than he earns) or that $G_t \leq \bar{g}$ for some arbitrary $\bar{g} > 0$ (the agent cannot overreport more than a fixed amount). The contracts derived in this paper remain truthful in the presence of such constraints and for any $\bar{g} > 0$, no matter how small, the necessary conditions for truth-telling are the same as if $G$ is unconstrained.
A contract $C$ requires the principal to make a transfer $C_t \in \mathbb{R}$ to the agent at time $t$: the agent gives $Y_t$ to the principal and the principal gives $C_t$ to the agent. An alternative interpretation is that the agent reports and keeps $Y_t$ and the principal gives a (possibly negative) subsidy $C_t - Y_t$ to the agent. These interpretations are formally equivalent and the former is used throughout the paper for consistency.

We rule out private savings and assume that the agent immediately consumes the transfer that he receives from the principal, as well as any difference between his real and reported cash flows:

**Assumption 2** The agent’s consumption at time $t$ is equal to $C_t + (X_t - Y_t) = C_t - G_t$.

The agent’s strategy consists of a lying process $L$ adapted to the agent’s information, which corresponds to the filtration $\mathbb{F}^X = \{\mathcal{F}_t^X\}_{t \geq 0}$ generated by $X$.\(^{17}\)

Given a strategy $L$, the agent’s expected discounted utility is\(^{18}\)

$$V_0(L) = E\left[ \int_0^{\infty} e^{-rt} \left(u(C_t + X_t - Y_t)\right) dt \right] \tag{3}$$

where $u$ is strictly concave. In computations to follow, the agent will be assumed to have exponential utility $u(c) = -\exp(-\theta c)$ for some risk-aversion coefficient $\theta > 0$.

A strategy $L$ is admissible if expression (3) is well defined. Letting $V_t(L)$ denote the agent’s continuation utility at time $t$ given strategy $L$, admissibility requires that the following transversality condition be satisfied:\(^{19}\)

$$\lim_{T \to +\infty} E \left[ e^{-rT} | V_T(L) | \right] = 0 \ a.s. \tag{4}$$

**Definition 1** A contractual equilibrium $(C,L)$ consists of a contract $C$ and a strategy $L$ for the agent that solves

$$V_0 = \sup_L \left\{ E \left[ \int_0^{\infty} e^{-rt} \left(u(C_t + X_t - Y_t)\right) dt \right] \right\}$$

over all admissible strategies.

**Objective of the Principal:** Given a contractual equilibrium $(C,L)$, the principal’s expected payoff is

$$\Pi = E \left[ \int_0^{\infty} e^{-rt} (Y_t - C_t) dt \right]$$

\(^{16}\)See Bloedel, Krishna, and Strulovici (2020) for the case of private savings.

\(^{17}\)The agent also observes $Y$. However, since $X$ determines $Y$, given the agent’s strategy, $X$ is a sufficient statistic for the agent’s information.

\(^{18}\)The strategy $L$ affects the probability measure over the paths of $Y$, which affects the expectation.

\(^{19}\)Admissibility rules out, for instance, strategies in which the agent continually underreports his cash flow, which permits him to get a higher immediate transfer by the principal but leads to an ever-decreasing continuation utility, with a decrease rate that exceeds $r$, akin to a Ponzi scheme.
where $Y$ is given by (2). The objective of the principal is to maximize $\Pi$ subject to giving the agent some minimal expected lifetime utility $w$ (i.e., $V_0 \geq w$) and to renegotiation-proofness constraints that are the subject of the next section.

3 Concept of Renegotiation

Any model of renegotiation entails a comparison between some current agreement and alternative agreements. The key is to determine the set of challengers that parties may consider as valid alternatives to the current agreement. This section proposes an approach to do this when there are persistent state variables and in the presence of private information.

3.1 Cash Flow and Continuation Utility as State Variables

In our setting, one of the state variables is the current cash flow $X_t$ generated by the agent. Since this cash flow is persistent, it affects the set of achievable payoffs for the players at any given time. In addition, it is useful convenient to treat the agent’s continuation utility $V_t$ as a state variable. We take the perspective of the principal and consider contractual equilibria that give the agent the same utility $V_t$ as the current equilibrium but increase the principal’s payoff $\Pi_t$ given the current cash flow $X_t$.

Reconstructing the Agent’s True Cash Flow: An Observability Theorem

The agent may lie about his cash flow. However, given an equilibrium strategy $L$ that is $\mathbb{F}^X$ adapted, the principal can, under mild regularity conditions on the agent’s strategy $L$, reconstruct the true cash flow $X_t$ and the agent’s continuation utility $V_t$ from the agent’s report process $Y$. To see this, notice that the agent’s lying process is $\mathbb{F}^X$-adapted and may thus be expressed as a functional

$$L_t = \mathcal{L}(t, X_s : s \leq t),$$

which depends, at each time $t$ on the path of $X$ until time $t$. Given the functionals $\{\mathcal{L}(t, \cdot)\}_{t \geq 0}$—which the principal knows in equilibrium—and the report process $Y$, we can construct a process $\tilde{X}_t$ defined by the equation

$$d\tilde{X}_t = dY_t - \mathcal{L}(t, \tilde{X}_s : s \leq t)dt$$

and the initial condition $\tilde{X}_0 = Y_0 = X_0$. Equation (6) is a functional stochastic differential equation whose unknown is the process $\tilde{X}$ and source of uncertainty is the process $Y$, whose quadratic variation is constant, equal to $\sigma$. From (2), the true cash flow process $X$ is a weak solution of (6). Moreover, any strong solution of (6) is, by definition, adapted to $Y$. If (6) has a unique solution,
the principal can back out the process $X$ by observing $Y$ and using his knowledge of the agent’s equilibrium strategy. This intuition is formalized by Theorem 1, proved in Appendix A.

**Theorem 1 (Observability Theorem)** Consider a contractual equilibrium $(C, L)$ such that $L$ is locally bounded: for each $T > 0$, there exists $M(T) > 0$ such that $\sup_{t \leq T} |L_t| \leq M(T)$ a.s. Then, the following holds:

1. If Equation (6) has a strong solution, then this solution is unique, and $X$ is $\mathbb{F}^Y$-adapted.

2. If $L$ is locally Lipschitz continuous,\(^{20}\) then Equation (6) has a unique strong solution and $X$ is $\mathbb{F}^Y$-adapted.

3. If there exists some $\varepsilon > 0$ such that for each $t > 0$, the functional $\mathcal{L}(t, \cdot)$ is independent of $\{X_s\}_{s \in [t-\varepsilon, t]}$, then Equation (6) has a unique strong solution and $X$ is $\mathbb{F}^Y$-adapted.

It is important here to distinguish a contract $C$, whose transfers depend on the agent’s report, from the principal’s belief, which may be quite different from the reports. As long as a contract $C$ is in place, the agent can affect his transfers by overreporting or underreporting his cash flows: these reports mechanically feed into the transfer specified by $C$, regardless of their effect on the principal’s belief. Theorem 1 says that the principal’s belief process matches the true cash flow as long as the principal’s belief equation (6) has a strong solution (Part 1), which is guaranteed if the agent’s strategy exhibits some arbitrarily small lag (Part 3) or if it is Lipschitz for some arbitrary Lipschitz constant (Part 2).

This distinction is conceptually important in the context of renegotiation: it means that in equilibrium the principal knows the agent’s continuation utility under the current contract when renegotiating with the agent, even if the agent has fed lies (i.e., inaccurate reports) into the contract $C$. Thus, it is legitimate to treat the actual cash flow, $X_t$, and agent continuation value $V_t$ as state variables for the purpose of analyzing renegotiation, as long as the agent’s strategy satisfies the assumptions of Theorem 1, which we impose from now on as part of the admissibility requirement.

### 3.2 Internal and State consistency

**Internal Consistency:** In the context repeated games, two well-known and essentially identical concepts of renegotiation are *internal consistency* (Bernheim and Ray (1989)) and *weakly renegotiation-proofness* (Farrell and Maskin (1989)). According to these concepts, an equilibrium

\(^{20}\)This means that for each $T > 0$, there exists $\bar{L}(T) > 0$ such that all functionals $\{\mathcal{L}(t, \cdot)\}_{t \leq T}$ are $\bar{L}(T)$-Lipschitz continuous over their respective domain.
is renegotiation-proof if it has no continuation equilibrium that is Pareto dominated by another continuation equilibrium. The challengers of a given equilibrium thus consist of all continuations of this equilibrium.

Internal consistency can be embedded in stochastic games: say that an equilibrium is internally consistent if there do not exist two histories leading to the same state such that the players’ continuation payoffs following the first history Pareto dominate those following the second history.

**Beyond Internal Consistency:** Internal consistency does not exploit any structure of the problem, such as the shape of player’s utility function or the dynamics of the state process, and therefore has limited power. To see how it may be strengthened to analyze stochastic games, it is useful to think about its rationale.

Internal consistency presumes that, after observing some history, the principal is able to recognize that he could use the continuation equilibrium following another history and achieve a higher payoff. This cognitive ability should extend to other natural comparisons.

One such comparison, often used in economics, is convexification: if two distinct continuation equilibria give the same expected utility to the agent, and the agent has a concave utility, it is natural for the principal to consider convex combinations of these continuation equilibria, as they may achieve the same expected utility for the agent at a lower cost for the principal. This convexification comparison is exploited later (Section 3.4). For now, we focus on comparisons across states, which constitute the main conceptual innovation of the paper.

**Consistency Across States:** Consider a contractual equilibrium \((C, L)\) that gives expected utility \(v_1\) to the agent starting from initial cash flow \(x\) and a payoff \(\pi_1\) to the principal. Also suppose, for the sake of illustration, that the agent has linear utility: \(u(c) = c\). Consider a new contract, \(\tilde{C}\), obtained by translating \(C\) by a fixed amount, to \(\tilde{C}_t = C_t + r(v_2 - v_1)\) for some constant \(v_2\). In many settings, translating consumption uniformly does not affect the agent’s incentives, so suppose—again for the sake of illustration—that \((\tilde{C}, L)\) also forms a contractual equilibrium starting from \(x\). This equilibrium provides expected utility \(v_2\) to the agent since it translates the agent’s original utility stream by a constant amount \(r(v_2 - v_1)\) and, hence, increases the agent’s discounted lifetime utility by \(v_2 - v_1\). Let \(\pi_2\) denote the principal’s payoff in the new contractual equilibrium.

Suppose that, starting from contractual equilibrium \((C, L)\), parties reach a history at which the agent’s continuation utility is \(V_t = v_2\), the principal’s continuation payoff is \(\Pi_t < \pi_2\), and the current cash flow is \(X_t = x\). In this scenario, the principal can reason that, by replacing the continuation contract by \(\tilde{C}\), reset from time 0, he can implement the contractual equilibrium \((\tilde{C}, L)\) from scratch and achieve the strictly higher payoff \(\pi_2\) while providing the expected utility \(v_2\) to the agent. This
kind of comparison enlarges the class of challengers to a given continuation contract by adding transformations of the initial contract that make it compatible with another state (here, a different level of continuation utility).

Such comparisons seem realistic as long as the principal can notice them (for example, translating a contract seems easy enough). Moreover, the challengers obtained through such transformations form a \textit{subset} of the class of challengers obtained by stronger concepts of renegotiation, such as a suitably adapted version of Farrell and Maskin’s concept of strongly renegotiation-proof (SRP) equilibrium. In the present setting, a weakly renegotiation proof (WRP) equilibrium is SRP if it has no continuation that is Pareto dominated by another WRP equilibrium starting at the same cash flow. An SRP equilibrium must withstand the comparisons described in this paper.

\textbf{A Transformation Group:} To formalize this idea, let

- \( S \) denote the state space. In our application, \( S \) consists of all pairs \((v, x)\) of continuation utility and cash flow over the relevant domain.

- \( \mathfrak{G} \) denote a group—in the algebraic sense—that exerts a \textit{left action} on \( S \). This means that for any \( g \in \mathfrak{G} \) and any \((v, x)\in S\), one can associate an element \((\tilde{v}, \tilde{x}) = g \circ (v, x) \in S\) that respects the group’s structure, as follows:
  
  - \( e \circ (v, x) = (v, x) \) for the identity element \( e \) of \( \mathfrak{G} \)
  
  - \((gg') \circ (v, x) = g \circ (g' \circ (v, x))\) for any \( g, g' \in \mathfrak{G} \).

Now consider the set \( \mathcal{E} \) of all contractual equilibria \((C, L)\). Each contractual equilibrium is associated with an initial state \((v, x)\).

To define a left action group on \( \mathcal{E} \), we associate for each \( g \in \mathfrak{G} \) and \((C, L) \in \mathcal{E}\), a new contractual equilibrium \( \mathcal{G}_g(C, L) \) that respects the structure of the group, as follows:

\textbf{Definition 2} \textit{The mapping from} \( \mathfrak{G} \times \mathcal{E} \rightarrow \mathcal{E} \textit{defined by} (g, (C, L)) \rightarrow \mathcal{G}_g(C, L) \textit{is a transformation group if it satisfies the following axioms:}

\textbf{Axiom 1.} \textit{For any} \((C, L) \in \mathcal{E}, \mathcal{G}_e(C, L) = (C, L).}

\textbf{Axiom 2.} \textit{For all} \(g, g' \in \mathfrak{G} \textit{and} (C, L) \in \mathcal{E}, \mathcal{G}_{g'}(\mathcal{G}_g(C, L)) = \mathcal{G}_{g'g}(C, L).}

\textbf{Axiom 3.} \textit{For any} \(g \in \mathfrak{G} \textit{and} (C, L) \in \mathcal{E} \textit{starting in state} (v, x), \textit{the contractual equilibrium} \mathcal{G}_g(C, L) \textit{starts in state} (\tilde{v}, \tilde{x}) = g \circ (v, x).}
A transformation group describes analogies that a principal can make to compare equilibria across states. The definition applies equally well to equilibria in discrete-time and continuous-time settings. Axioms 1 and 2 are the defining properties of a left group action: Axiom 1 says that if the state is unchanged (i.e., the group identity \( e \) is applied), then the transformation is the identity mapping over \( \mathcal{E} \), and Axiom 2 is an associativity axiom that will be used shortly, applied to group elements that are inverse of each other. Axiom 3 requires that the transformed contractual equilibria be consistent with the effect that the group \( \mathcal{G} \) has on the state: in the translation example above, for instance, it means that if \( g \) transforms the state \( v \) into \( \tilde{v} \) when \( g \) operates on the state space \( \mathcal{S} \), then \( g \) should also transform any contractual equilibrium that gives expected utility \( v \) to the agent into one that gives him expected utility \( \tilde{v} \).

We now define a concept of renegotiation-proofness, according to which challengers are obtained through the transformation group \( \mathcal{G} \). Given any contractual equilibrium \( (C, L) \), let \( \Pi(C, L) \) denote the principal’s expected payoff.

**Definition 3** A contractual equilibrium \( (C, L) \) starting in state \( (v, x) \) is state consistent (relative to \( \mathcal{G} \)) if, after any history leading to some state \( (\tilde{v}, \tilde{x}) = g \circ (v, x) \) and continuation contractual equilibrium \( (\tilde{C}, \tilde{L}) \), we have

\[
\Pi(\tilde{C}, \tilde{L}) \geq \Pi(\mathcal{G}_g(C, L)),
\]

and, reciprocally,

\[
\Pi(C, L) \geq \Pi(\mathcal{G}_{g^{-1}}(\tilde{C}, \tilde{L})).
\]

State consistency entails two requirements: first, each continuation of \( (C, L) \) must sustain the comparison with the transformation of \( (C, L) \) to the state corresponding to this continuation.21

Second, the initial contractual equilibrium must sustain the comparison with each continuation equilibrium transformed back to the initial state, i.e., corresponding to the group element \( g^{-1} \), since \( g^{-1} \circ (g \circ (v, x)) = e \circ (v, x) = (v, x) \).

A first observation is that a state-consistent contractual equilibrium is internally consistent as long as \( \mathcal{G} \) satisfies a natural monotonicity assumption.

**Definition 4** The transformation group \( \mathcal{G} \) is monotone if for any two contractual equilibria \( (C, L) \) and \( (C', L') \) starting in some common state \( (v, x) \) and giving payoffs \( \Pi(C, L) \leq (<) \Pi(C', L') \) to the principal and any \( g \in \mathcal{G} \), the principal’s payoffs with the transformed contractual equilibria \( (\tilde{C}, \tilde{L}) = \mathcal{G}_g(C, L) \) and \( (\tilde{C}', \tilde{L}') = \mathcal{G}_g(C', L') \) satisfy \( \Pi(\tilde{C}, \tilde{L}) \leq (<) \Pi(\tilde{C}', \tilde{L}') \).

21 This definition does not distinguish between on-path and off-path continuations. This distinction is irrelevant for the setting of Section 2 since the agent’s lies are absolutely continuous and the report process is always on path.
Monotonicity means that the ranking of the principal’s payoffs is preserved under transformations to a different state.

**Proposition 1** Suppose that \( G \) is monotone and that the contractual equilibrium \((C, L)\) starting in state \((v, x)\) is state consistent with respect to \( G \). Then, after any finite history leading to state \((\tilde{v}, \tilde{x}) = g \circ (v, x)\), the continuation payoff for the principal is equal to his initial payoff in the contractual equilibrium \( G_\theta(C, L) \).

This proposition, proved in Appendix A, implies the following result.

**Corollary 1** If \( G \) is monotone, any state-consistent equilibrium is internally consistent.

The proof is immediate: if \((C, L)\) is state consistent, Proposition 1 implies that any two histories leading to the same state \((\tilde{v}, \tilde{x}) = g \circ (v, x)\) give the same continuation payoff \(\tilde{\pi}\) to the principal and, hence that the players’ continuation payoffs \((\tilde{v}, \tilde{\pi})\) of both parties are not Pareto ranked.

**Strength, Stability, and Uniqueness of the State-Consistency Concept:** In principle, there may be several transformations groups to consider. The group \( \mathcal{G} \) operating on the state space is transitive (in a group-theoretic sense) if for any states \((v, x)\) and \((v', x')\) there exists some \( g \in \mathcal{G} \) such that \((v', x') = g \circ (v, x)\). This is the case that we consider in our application paper (Section 4). Intuitively if \( \mathcal{G} \) is defined with respect to a transitive group \( \mathcal{G} \), it means that we can compare contractual equilibria across any two states, and is thus “maximal.”

The theory of state consistency does not rely on transitivity, however: the theory is well defined for any group \( \mathcal{G} \). This observation is useful if, for instance, in settings for which it is hard to establish comparisons across all pairs. The next proposition establishes two results that concern non-transitive groups. Given a group \( \mathcal{G} \) operating on \( S \) and a transformation group \( \mathcal{G}' \) defined with respect to \( \mathcal{G} \), we can consider, for any subgroup \( \mathcal{G}' \) of \( \mathcal{G} \) the transformation group \( \mathcal{G}' \) that is the restriction of \( \mathcal{G} \) with respect to \( \mathcal{G}' \). A particular subgroup is the trivial group \( \mathcal{G} \) which has only the identity element \( e \). With respect to this group, \( \mathcal{G} \) reduces to the identity transformation over contractual equilibria. We have the following result.

**Proposition 2** (i) Suppose that \( \mathcal{G} \) is a transformation group with respect to \( \mathcal{G} \) and that \( \mathcal{G}' \) is a subgroup of \( \mathcal{G} \). A contractual equilibrium is state consistent with respect to \( \mathcal{G} \) only if it is state consistent with respect to \( \mathcal{G}' \). (ii) If \( \mathcal{G}' \) is the trivial group, a contractual equilibrium is state consistent with respect to \( \mathcal{G}' \) if and only if it is internally consistent.

Part (i), whose proof is straightforward, establishes a nestedness condition for concepts of state-consistency. Regarding Part (ii), we note that the identity transformation group is clearly mono-
tonic, so internal consistency follows from Corollary 1. The reverse direction follows immediately from the definition of state consistency.

From Proposition 2, the stronger concepts of state consistency are obtained for transitive groups. Fixing a group \( \mathfrak{G} \), there may a priori exist multiple transformation groups with respect to \( \mathfrak{G} \). However, the next result shows that, for a contractual equilibrium to be state consistent with respect to several monotone transformation groups, these groups must, taken individually, yield the same concept of state consistency. Consider two monotone transformation groups \( \mathcal{G} \) and \( \tilde{\mathcal{G}} \). Given a contractual equilibrium \((C, L)\) starting at \((v, x)\), suppose that there is a state \((v', x') = g \circ (v, x)\) for which the groups yield different payoffs: \(\Pi(\mathcal{G}_g(C, L)) > \Pi(\tilde{\mathcal{G}}_g(C, L))\). Applying the transformations again from \((v', x')\) to \((v, x) = g^{-1}(v', x')\) and using monotonicity, we get

\[
\Pi(\tilde{\mathcal{G}}_g^{-1}(\mathcal{G}_g(C, L))) > \Pi(C, L).
\]

Therefore, \((C, L)\) is dominated (from the principal’s perspective) by a simple composition of elements in the two groups. In such a case, we will say that \((C, L)\) is unstable with respect to \((\mathcal{G}, \tilde{\mathcal{G}})\), otherwise, \((C, L)\) is called stable. Intuitively, instability implies that the principal’s ability to consider transformations that yield different payoffs after a given history prevents the existence of a state-consistent contractual equilibrium.

**Proposition 3 (Concept Equivalence for Stable Contractual Equilibria)** Let \( \mathcal{G} \) and \( \tilde{\mathcal{G}} \) be two monotone transformation groups with respect \( \mathfrak{G} \) and \((C, L)\) be a contract equilibrium starting in state \((v, x)\) that is stable with respect to \( \mathcal{G} \) and \( \tilde{\mathcal{G}} \). Then, \((C, L)\) is state consistent with respect to \( \mathcal{G} \) if and only if it is state consistent with respect to \( \tilde{\mathcal{G}} \). Moreover, the principal’s payoffs for the \( \mathcal{G} \) and \( \tilde{\mathcal{G}} \) transformations of \((C, L)\) to any state \((v', x')\) are identical.

### 3.3 Revelation Principle

**Definition 5** A contract \( C \) is truthful if \((C, L \equiv 0)\) is a contractual equilibrium. A contractual equilibrium \((C, L)\) is truthful if \( L \equiv 0 \).

If \( C \) is truthful, the agent finds it optimal to report his cash flow truthfully. As with standard applications of the Revelation Principle, truthful contracts simplify the analysis, so it is natural to ask whether focusing on them is without loss of generality. In the absence of commitment, it is well known that the Revelation Principle need not hold.\(^2\) However, a version of the principle holds in our setting, as described next.

\(^2\)See, e.g., Bester and Strausz (2001).
Say that two contractual equilibria starting from the same state are *outcome equivalent* if they induce the same net transfer after all histories.

**Theorem 2** Suppose that $\mathcal{G}$ is monotone. Then, any state-consistent contractual equilibrium is outcome-equivalent to a truthful state-consistent contractual equilibrium.

**Proof.** Consider any contractual equilibrium $(C, L)$. As with the standard Revelation Principle, we can replace the contract $C$ by a contract $\hat{C}$ in which: (i) the agent truthfully reports his cash flow, (ii) the principal simulates the lies that the agent would have made in the initial equilibrium and makes the same net transfer to the agent that he would have made in the initial equilibrium. By construction, the truthful contractual equilibrium $(\hat{C}, \hat{L} \equiv 0)$ is outcome equivalent to the initial contractual equilibrium.

We must show that if $(C, L)$ is state consistent with respect to $\mathcal{G}$, then so is $(\hat{C}, 0)$. Let $(v, x)$ denote the initial state (which is the same for both contractual equilibria) and consider any history leading to some state $(v', x') = g \circ (v, x)$. Let $(\hat{C}', 0)$ denote the continuation equilibrium of $(\hat{C}, 0)$ following this history. To this history corresponds a history under the equilibrium $(C, L)$, which is the history used by the contract $\hat{C}$ to calculate the agent’s net transfer as function $C$. Let $(C', L')$ denote the continuation equilibrium of $(C, L)$ following this corresponding history.

By outcome equivalence, we have

$$\Pi(\hat{C}', 0) = \Pi(C', L') \quad (7)$$

and both histories lead to the same state, $(v', x') = g \circ (v, x)$. Again by outcome equivalence, we have

$$\Pi(C, L) = \Pi(\hat{C}, 0)$$

for the contracts evaluated at time 0. Since $\mathcal{G}$ is monotone, the previous equality implies that

$$\Pi(\mathcal{G}_g(C, L)) = \Pi(\mathcal{G}_g(\hat{C}, 0)). \quad (8)$$

Comparing (7) and (8) and using the fact that $(C, L)$ is state consistent then shows that $(\hat{C}', 0)$ also satisfies the definition of state consistency (Definition 3).

**3.4 Convexification and Renegotiation-Proof Contracts**

As noted earlier, when the agent’s utility is concave and a contractual equilibrium has multiple continuations giving the agent the same expected utility, it is natural for the principal to consider convexifications of these continuations, as they may be cheaper.
For expositional simplicity, convexification is defined for truthful contracts. Consider two truthful contracts $C_1, C_2$ starting from the same cash flow $x$ and giving the same utility $v$ to the agent. For any $\lambda \in [0,1]$, consider the contract $C^\lambda$ defined as follows: for any $t$ and report history $\{Y_s : s \leq t\}$, 

$$u(C^\lambda_t) = \lambda u(C^1_t) + (1 - \lambda) u(C^2_t).$$

(9)

The consumption level $C^\lambda_t$ is uniquely defined if $u$ is strictly increasing and continuous, which is the case for exponential utility functions. In words, $C^\lambda$ provides a flow utility to the agent that is a fixed convex combination of the flow utilities that he gets under $C^1$ and $C^2$. If $C^\lambda$ is truthful, it is straightforward to see that it provides expected utility $v$ to the agent, starting from $x$, because so do $C^1$ and $C^2$ and all three contracts generate the same distribution for the report process $Y$.

The next result shows that $C^\lambda$ is indeed truthful and formalizes this result.

**Proposition 4** Suppose that $u(c) = -\exp(-\theta c)$ for some parameter $\theta > 0$. For any $\lambda \in [0,1]$ and initial cash flow $x$, the contract $C^\lambda$ is truthful and gives the agent expected utility $v$.

Given a truthful contract $C$ and state $(v,x)$, let $\mathcal{K}(C)$ denote the set of all continuation contracts of $C$ and

$$\mathcal{K}(v,x) = \left\{ G_g(\hat{C}) : (g, \hat{C}) \in \mathfrak{G} \times \mathcal{K}(C) \quad \text{s.t.} \quad g \circ (v_{\hat{C}}, x_{\hat{C}}) = (v, x) \right\}$$

where $(v_{\hat{C}}, x_{\hat{C}})$ denotes the starting state of contract $\hat{C}$. In words, $\mathcal{K}(v,y)$ consists of all contracts starting in state $(v,x)$ that are transformations of continuation contracts of $C$.

Finally, let

$$\text{Conv}(\mathcal{K}(v,x)) = \{ C^\lambda : \lambda \in [0,1], C^1, C^2 \in \mathcal{K}(v,x) \}$$

denote the set of convex combinations of contracts in $\mathcal{K}(v,x)$.

The set $\text{Conv}(\mathcal{K}(v,x))$ defines the class of all challengers that the principal can consider, after any history leading up to state $(v,x)$, to replace the current continuation contract.

**Definition 6** A truthful contract $C$ is renegotiation-proof if (i) the pair $(C, L \equiv 0)$ forms a state-consistent contractual equilibrium and (ii) for any history leading to any state $(v,x)$, the continuation contract $\hat{C}$ satisfies

$$\Pi(\hat{C}) \geq \Pi(C')$$

for all $C' \in \text{Conv}(\mathcal{K}(v,x))$.\footnote{Convexification can be defined for non-truthful contracts by considering the truthful contracts that are outcome equivalent to these contracts and applying the definition of this section.}
This definition captures the following intuition: if the principal can come up with some challengers to a contract, possibly by comparing contracts across states, then he can also consider the convex combination of these challengers. Convexification is used in Lemma 1 to pin down contractual variables as a function of the state (Proposition 9).

4 Transformation Group: Explicit Construction

This section constructs a transformation group $G$ explicitly for the setting of Section 2 that compare equilibria across all pairs of states.

4.1 Transforming Contracts Across Cash Flow Levels

Given a contractual equilibrium $(C, L)$ starting from state $(v, x)$, let us construct a new contractual equilibrium starting from $(v, \hat{x})$, for any $\hat{x} \neq x$. Since $C$ is $FY$-adapted, it can be expressed as $C_t = C(Y_s : s \leq t)$ for some functional $C$ (for notational simplicity, $C$ is defined over paths of $Y$ of varying time lengths, which implicitly allows $C$ to depend $t$).

Starting from any cash flow level $\hat{x} \neq x$, and given a report process $\hat{Y}_t$, suppose that the principal delivers the consumption process $\hat{C}_t = C(\hat{Y}_s : s \leq t)$, where $\hat{Y}$ is constructed as follows: $\hat{Y}_0 = \hat{x}$ and

$$d\hat{Y}_t = d\hat{Y}_t - (\xi - \lambda \hat{Y}_t)dt + (\xi - \lambda \hat{Y}_t)dt.$$ 

The process $\hat{Y}$, on which contract $\hat{C}$ is based, may be described as a virtual report process: it is the report process that the agent would have produced if the initial cash flow had been $x$ instead of $\hat{x}$ and the agent had followed the same strategy $L$ that generates $\hat{Y}$ when starting from $\hat{x}$.

Intuitively, by proposing contract $\hat{C}$, it is as if the principal ignored the initial cash flow level $\hat{x}$. By insulating the agent’s compensation process from the initial cash flow, the principal separates the agent’s lying incentives from the initial cash flow. This separation works when the dynamic equation of the report process $Y$ is linear, as shown the next result.

**Proposition 5** Suppose that the contract $C$ defined by $C_t = C(Y_s : s \leq t)$ together with strategy $L$ forms a contractual equilibrium starting at $(v, x)$. Then, the contract $\hat{C}$ defined by $\hat{C}_t = C(\hat{Y}_s : s \leq t)$ together with strategy $L$ forms a contractual equilibrium starting at $(v, \hat{x})$.

Proposition 5 allows us define the first part of the transformation $G$: we set for any $\hat{x} \in \mathbb{R}$

$$G_{(1,\hat{x}-x)}(C, L) = (\hat{C}, L).$$
The reason for the “1” subscript has to do with how the group $G$ operating on the state space will be defined and will become clear in the next subsection.

4.2 Transforming Contracts Across Promised-Utility Levels

From now on, we assume that the agent’s utility function is given by $u(c) = -\exp(-\theta c)$ for some risk aversion parameter $\theta > 0$. In particular, the agent’s flow utility $u(C_t - G_t)$ is always negative and, hence, so are his promised and continuation utilities at all times.

Consider a contractual equilibrium $(C, L)$ starting from state $(v_0, x)$, with $v_0 < 0$, and consider any alternative utility level $v_1 = \beta v_0$ for some $\beta \in (0, \infty)$. We define a new contract $\hat{C}$ as follows:

$$\hat{C}_t = C_t - \frac{\log(\beta)}{\theta}$$

for all $t$. We have the following result:

**Proposition 6** $(C, L)$ is a contractual equilibrium starting at $(v_0, x)$ if and only if $(\hat{C}, L)$ is a contractual equilibrium starting at $(v_1, x)$.

**Proof.** Let $V(L|C)$ denote the agent’s expected utility when he follows strategy $L$ given contract $C$. Since $L$ is optimal for the agent, we have

$$V(L|C) = v_0 \geq v(L'|C)$$

for all $(L')$. For any $\hat{L}'$, let $L' = \hat{L}'$. Then, it is straightforward to check from (3) that

$$V(\hat{L}'|\hat{C}) = \beta V(L'|C) \leq \beta V(L|C) = \beta v_0 = v_1$$

and the inequality is tight if $\hat{L}' \equiv L$, which shows that $L$ is optimal given $\hat{C}$ and provides utility $\beta v_0$ to the agent.

For any contractual equilibrium $(C, L)$ and $\beta > 0$, we set

$$G(\beta, 0)(C, L) = (\hat{C}, L).$$

4.3 Action Group

Since the agent’s flow utility is negative, the relevant state space for the agent is $S = (-\infty, 0) \times \mathbb{R}$: it consists of all possible pairs $(v, x)$ of expected utility and cash flows.

Let $\mathfrak{G} = (0, +\infty) \times \mathbb{R}$ denote the algebraic group defined by the binary operation $(\beta, \delta)(\beta', \delta') = (\beta \beta', \delta + \delta')$ for all $(\beta, \delta), (\beta', \delta') \in \mathfrak{G}$. It is straightforward to check that $\mathfrak{G}$ is an Abelian group with identity element $(1, 0)$. 

19
Moreover, the operation \((β, δ) \circ (v, x) = (βv, x + δ)\) defines a left action of \(G\) on the state space \(S\): the associativity and identity axioms are straightforward to check.

Finally, the group \(G\) defines a left action on the set of contractual equilibria, as follows:

1. Sections 4.1 and 4.2 define \(G_{(β,0)}(C, L)\) and \(G_{(1,δ)}(C, L)\) for all \(β > 0, δ \in \mathbb{R}\) and contractual equilibrium \((C, L)\).

2. For any \((β, δ) ∈ G\), we define \(G_{(β,δ)}(C, L)\) by
\[
G_{(β,δ)}(C, L) = G_{(β,0)}(G_{(1,δ)}(C, L))
\]

3. It is straightforward to check that this definition is consistent with 1. and that \(G\) satisfies Axioms 1,2,3 of Definition 2, which shows that \(G\) is a transformation group.\(^{24}\)

4. Finally, it is also straightforward to check that \(G\) is monotone: if \(Π(C, L) \geq Π(C', L')\) for any two contractual equilibria starting from the same state, then the constructions of Sections 4.1 and 4.2 preserve this payoff relation.

These observations are summarized by the following proposition.

**Proposition 7** The mapping \(G : G \times E → E\) is a monotone transformation group.

## 5 Characterization of Renegotiation-Proof Contracts

Proposition 7 and Theorem 2 guarantee that we can focus without loss of generality on truthful contracts to analyze the payoffs and outcomes that can be achieved by renegotiation-proof contracts. We do so for the remaining of the paper.

**Truthful Contracts and Promised Utility:** Given a truthful contract \(C\) the agent’s continuation utility at time \(t\) is equal to his promised utility
\[
W_t = E \left[ \int_t^∞ e^{-r(τ-t)} (u(C_τ)) dτ \right] \bigg| \mathcal{F}_t^Y.
\]
Since the promised utility process \(W = \{W_t\}_{t≥0}\) is adapted to the filtration \(\mathcal{F}_t^Y\), the Martingale Representation Theorem implies that \(W\) satisfies the dynamic equation\(^{25}\)
\[
dW_t = (rW_t - u(C_t)) dt + σS_t d\tilde{B}_t \tag{11}
\]
\(^{24}\)To prove associativity, one must show that consumption translation commutes with the virtual cash flow construction of Section 4.1, which is immediate from the construction.

where the process $S_t$ is $\mathbb{F}^Y$-adapted and $\tilde{B}_t$ is an $\mathbb{F}^Y$-adapted Brownian motion under the probability measure in which the agent reports truthfully, i.e.,

$$d\tilde{B}_t = \frac{dY_t - (\xi - \lambda Y_t)dt}{\sigma}.$$

If the agent does not report his cash flow truthfully, $\tilde{B}$ is no longer a martingale: plugging Equation (2) into the previous equation yields

$$d\tilde{B}_t = \frac{\lambda (Y_t - X_t)dt + L_t dt + \sigma dB_t}{\sigma}.$$ (12)

Intuitively, the cash flow shocks reported by the agent’s are biased upward if (i) the agent overreports his cash flow ($L_t > 0$) or (ii) the actual cash flow is lower than the reported cash flow ($X_t < Y_t$). Effect (i) is straightforward. Effect (ii) comes from mean-reversion: if $X_t < Y_t$, then $X_t$ has a higher mean than what the principal expects based on $Y_t$. Even if the agent does not produce any additional lie ($L_t = 0$), reported increments are positively biased.

It follows from (12) that when the agent lies, his promised utility evolves as

$$dW_t = (rW_t - u(C_t))dt + S_t(L_t dt + \lambda G_t dt + \sigma dB_t),$$ (13)

where $\{B_t\}_{t \geq 0}$ is the standard Brownian motion. The coefficient $S_t$ is the sensitivity of the agent’s promised utility to the agent’s report increment at time $t$. It will be treated as a choice variable of the principal in the recursive formulation of the problem.

**Agent’s Incentives:** Persistence of the agent’s private information implies that past lies can have a long term impact on incentives: if the agent has lied even for a short period before time $t$, he has affected the report history $Y^t = \{Y_s\}_{s \leq t}$ and therefore also affected his future consumption flow $C$ and his future incentives to report the truth. Intuitively, if the agent underreports his cash flow increment $dX_t$ at time $t$, he affects his utility through two channels. First, this reduces his promised utility, which depends on the report $dY_t$ by a sensitivity factor $S_t$, and thus reduces the consumption stream that the principal aims to give the agent. Second, to deliver a given level of promised utility, the principal must provide higher transfers to the agent if the agent generates lower cash flows. The first channel incentivizes the agent to report higher cash flows and the second channel incentivizes him to report lower cash flows. For the contract to be truthful, these two incentives must balance each other, at least on the equilibrium path. The agent’s reporting incentives are the subject of Section 7.

### 5.1 Structure of the Principal’s Payoff

Consider a truthful, renegotiation-proof contract $C$ for the transformation group $G$ constructed in Section 4. Since the contract is truthful, there is no difference on the equilibrium path between $X_t$
and $Y_t$ and between $V_t$ and $W_t$.

We will now use $(w, y)$ instead of $(v, x)$ to denote the state, to make these equalities clear. For any state $(w, y) \in (-\infty, 0) \times \mathbb{R}$, let $\Pi(w, y)$ denote the principal’s expected payoff for the transformation of $(C, 0)$ that corresponds to state $(w, y)$. From Proposition 1, $\Pi(w, y)$ is also the principal’s continuation payoff after any history leading to state $(w, y)$. We now derive the functional form of $\Pi(w, y)$.

First, consider the dependence on the cash flow level $y$. As noted in Section 4.1, the only difference in continuation payoffs for the principal between histories starting in state $(w, y)$ and $(w, y')$ concerns the expected discounted transfer from the agent $\Upsilon(y) = E[\int_0^\infty e^{-rt}Y_t dt | w, y]$, because the processes $C$ have the same distribution independently of $y$.

To compute $\Upsilon(y)$, note that since the contract is truthful, we have $Y_t \equiv X_t$. Moreover, the cash flow process $X_t$ has an explicit formula, which is computed in Appendix A (Equation (34)), the expected discounted transfer from the agent satisfies

$$\Upsilon(y) = \int_0^\infty e^{-rt} \left( e^{-\lambda t} y + E \left[ \int_t^\infty e^{\lambda(s-t)} \xi ds \right] \right) dt.$$  

After simplification, this yields

$$\Upsilon(y) = \frac{y}{r + \lambda} + \frac{\xi}{r (r + \lambda)}.$$  

This yields the following result:

$$\Pi(w, y) = \Pi(w, 0) + \frac{y}{r + \lambda}.$$  

Next, consider two different promised utility levels $w_0$ and $w_1 = \beta w_0$ for the agent, where $\beta > 0$. The translation performed in Section 4.2 implies that

$$\Pi(w_1, y) = \frac{\log(\beta)}{\theta r} + \Pi(w_0, y).$$

The previous analysis yields the following result.

**Proposition 8** For any truthful, renegotiation-proof contract $C$, the principal’s payoff satisfies the following relation:

$$\Pi(w', y') = \frac{(y' - y)}{r + \lambda} + \frac{\log(w'/w)}{r \theta} + \Pi(w, y).$$

In particular, the principal’s continuation payoff in state $(w, y)$ satisfies

$$\Pi(w, y) = \frac{y}{r + \lambda} + \frac{\log(-w)}{\theta r} + \Pi(-1, 0).$$

Proposition 8 has an immediate but important corollary:
Corollary 2 If $C$ is truthful and state consistent, the principal’s continuation payoff takes the form $\Pi_t = \Pi(W_t, Y_t)$ where $\Pi(\cdot, \cdot)$ is twice continuously differentiable.

5.2 Contractual Variables

We now derive the functional form of contractual variables $C_t, S_t$ for all renegotiation-proof contracts.

Proposition 9 For any truthful renegotiation-proof contract, there exist constants $c_1, \bar{s}$ such that

$$C_t = c_1 - \frac{\log(-W_t)}{\theta} \quad a.s. \quad (17)$$

$$S_t = -W_t \bar{s} \quad a.s. \quad (18)$$

for all $t \geq 0$.

The parameters $c_1$ and $\bar{s}$ are the consumption and sensitivity levels at time 0 of a contract transformation starting with promised utility $-1$ and any cash flow level. Proposition 9 shows that any truthful renegotiation-proof contract is fully determined by the parameters $c_1$ and $\bar{s}$. Theorem 3, below, will show that $c_1$ and $\bar{s}$ must satisfy an additional equation, which reduces the set of renegotiation-proof contracts to a one-parameter family. Proof. Given dynamic equation (11), the statement of Proposition 9 is equivalent to $W$ being a geometric Brownian motion or, equivalently, to the Itô process $Z_t = \log(-W_t)$ having constant drift and volatility.

To prove that the drift and volatility of $Z$ are constant, we will use an intermediate result. Let $(w, y)$ denote the initial state of the truthful renegotiation-proof contract $C$. We fix a time $T$ and history up to time $T$ and let $\tilde{C}$ denote the continuation contract following this history. Let $\hat{C}$ denote the transformation of $\tilde{C}$ that starts in state $(\tilde{w}, \tilde{y}) = g \circ (w, y)$ (which is always possible since $g \circ \mathcal{S} = \mathcal{S}$) we have $\hat{C} = \mathcal{G}_{g^{-1}}(\tilde{C})$. The contract $\tilde{C}$ is truthful as a continuation of $C$, which is truthful, and so is $\hat{C}$ because $\mathcal{G}$ preserves truthfulness (Propositions 5 and 6). The contract $\hat{C}$ may proposed by the principal at time 0, instead of contract $C$, given the initial state $(w, y)$. We now show that the stochastic processes $C = (C_t)_{t \geq 0}$ and $\hat{C} = (\hat{C}_t)_{t \geq 0}$ must in fact be identical functions of the realized cash-flow process $Y = (Y_s)_{s \geq 0}$. To state this lemma, which is proved in the Appendix, let $\mathbb{L}$ denote the vector space of processes $A$ such that $E[\int_0^\infty e^{-rt}|A_t|dt] < \infty$ endowed with the norm $\|A\|_\mathbb{L} = E[\int_0^\infty e^{-rt}|A_t|dt]$. By assumption, any contract $C$ belongs to $\mathbb{L}$ under truth-telling.

Lemma 1 The processes $C$ and $\hat{C}$ are identical in $\mathbb{L}$: $\|C - \hat{C}\|_\mathbb{L} = 0$. 

23
Let \( \{\hat{W}_t\}_{t \geq 0} \) denote the agent’s promised utility under contract \( \hat{C} \). \( \hat{W} \) is an \( \mathbb{R}^Y \)-adapted Itô process given by
\[
\hat{W}_t = E \left[ \int_t^\infty e^{-r(s-t)} u(\hat{C}_s) ds \mid Y_z : z \leq t \right].
\]
(19)
Lemma 1 implies that \( \hat{W} \) is identical to \( W \) and that the process \( \hat{Z} : \{\hat{Z}_t = \log(-\hat{W}_t)\}_{t \geq 0} \) is identical to the process \( Z : \{Z_t = \log(-W_t)\}_{t \geq 0} \).

Let \( \mu_t \) and \( \sigma_t \) denote the drift and volatility of \( Z \) at time \( t \). Our objective is to show that \( \mu_t \) and \( \sigma_t \) are constant. Since \( \hat{Z} \) and \( Z \) are identical, we have:
\[
d\hat{Z}_t = dZ_t = \mu_t dt + \sigma_t dB_t.
\]
By definition of \( \mathcal{G} \), the contract \( \hat{C} = \mathcal{G}_{y^{-1}}(\hat{C}) \) is obtained by translating the process \( \hat{C} = \{C_{T+t}\}_{t \geq 0} \) by the constant \( \frac{1}{2}(\log(-W_T) - \log(-w)) \) and making it independent of the initial condition \( Y_T \), as explained in Section 4.1. The latter condition is equivalent to the requirement that \( \hat{C} \) depends on the Brownian path \( \{B_t\}_{t \geq 0} \) in the same way that \( \hat{C} \) depends on the Brownian path \( \{B_{T+t} - B_T\}_{t \geq 0} \).\(^{26}\)

From (19) and the fact that \( u(c) = -\exp(-\theta c) \), this implies that the process \( \{\hat{Z}_t\}_{t \geq 0} \) is identical in law to the process \( \hat{Z} = \{Z_{T+t} - Z_T + Z_0\}_{t \geq 0} \) conditioned on \( \mathcal{F}_T \). The latter process has drift \( \{\mu_{T+t}\}_{t \geq 0} \) and volatility \( \{\sigma_{T+t}\}_{t \geq 0} \). Equality in law of \( \hat{Z} \) and \( \hat{Z} \) implies that the drift and volatility of \( \hat{Z} \) at time zero is equal to the drift and volatility of \( Z \) at time \( t = 0 \), i.e.,
\[
\mu_0 = \mu_T \quad \sigma_0 = \sigma_T.
\]
Since the time \( T \) and history up to time \( T \) were arbitrary, this shows that \( \mu_T \) and \( \sigma_T \) are constant and proves the proposition.

Proposition 9 immediately implies the following corollary:

**Corollary 3** *Any truthful renegotiation-proof contract is Markovian: there exist functions \( c(\cdot) \) and \( s(\cdot) \) such that \( C_t = c(W_t, Y_t) \) and \( S_t = s(W_t, Y_t) \) for all \( t \).*

In fact, Proposition 9 shows that these functions are independent of \( Y_t \): the agent’s consumption at any time depends only on his current promised utility, and so does the sensitivity \( S_t \).\(^{27}\)

\(^{26}\)With a truthful contract starting at cash flow \( y \), the reported cash flow process \( Y \) has an explicit formula, given by \( Y_t = e^{-\lambda t} y + \int_0^t e^{\lambda(s-t)} \xi ds + \int_0^t e^{\lambda(s-t)} \sigma dB_s \) (see Equation (34)). Therefore, the path of \( Y \) is entirely pinned down by the initial condition \( y \) and the path of \( B \). Independence of the contract with respect to the initial condition \( y \) then shows the claim. Remark: although easier to understand with the closed form of \( Y \), the claim holds as long as the cash flow process \( Y \) solves an SDE with noise \( B \) that has a unique solution (see, e.g., Cherny (2002, Theorem 3.2)).

\(^{27}\)For readers familiar with the concept of symmetry in physics and Noether’s theorem, this property may be viewed
Proposition 9 provides a necessary condition for the form of truthful renegotiation-proof contracts. This necessary condition is strong, as it reduces the family of such contracts to a two parameter family $c_1$ and $\bar{s}$. This reduction to two parameters is essentially due to state consistency and convexification.\textsuperscript{28}

Truthfulness itself imposes one more relationship between the variables $c_1$ and $\bar{s}$, which we will express in terms of utility. Let $u_1 = u(c_1)$. The following result is shown in Section 7.

**Theorem 3** The set of truthful renegotiation-proof contracts is characterized by Equations (17) and (18) where $u_1$ is a function of $\bar{s}$:

$$u_1(\bar{s}) = -\frac{\bar{s}\lambda}{\theta - \bar{s}}.$$  \hspace{1cm} (20)

Theorem 3 shows that any truthful renegotiation-proof contract is characterized by a single parameter $\bar{s}$. To optimize the principal’s payoff with respect to $\bar{s}$, the next proposition computes this payoff explicitly.

**Proposition 10** The principal’s payoff is given by:

$$\frac{1}{\theta r} \left[ \log(-u_1(\bar{s})) + \frac{1}{r} \left( r + u_1(\bar{s}) - \frac{1}{2}\sigma^2 \bar{s}^2 \right) + D(w, y) \right]$$  \hspace{1cm} (21)

where

$$D(w, y) = \log(-w) + \frac{\theta}{r + \lambda}(ry + \xi).$$

6 Comparative Statics

The following comparative statics obtain by maximizing (21) with respect to $\bar{s}$, ignoring the term $D(w, y)$, which is independent of $\bar{s}$.

6.1 Persistence

**Proposition 11** The optimal sensitivity $\bar{s}$ is decreasing in $\lambda$.

as a conservation law of the contract that stems from state consistency: state consistency with respect to the state $Y_t$ may be viewed as a symmetry of the contract expressed in terms of the left group action $C \mapsto G_{(1,\delta y)}(C)$ and invariance of the contract with respect to $Y_t$ is a consequence of this symmetry.\textsuperscript{28}From an algebraic perspective, each truthful Markovian contract generates an orbit in the set $\mathcal{T}$ of all truthful Markovian contracts through the left action of group $G$. Proposition 9 says that the quotient of $\mathcal{T}$ with respect to these orbits (or equivalence classes) is two-dimensional.
Intuitively, when the cash flow is more persistent ($\lambda$ is lower), the agent’s current report has more bearing on expected future cash flows. This raises the magnitude of the agency problem and requires a higher a sensitivity to induce truthful reporting.

The proof of Proposition 11 is straightforward. Ignoring terms and factors in the principal’s objective (21) that do not influence the choice of $\bar{s}$, the optimal choice of $\bar{s}$ is the solution to

$$\log(-u_1(\bar{s})) + \frac{1}{r} \left( r + u_1(\bar{s}) - \frac{1}{2} \sigma^2 \bar{s}^2 \right),$$

where

$$u_1(\bar{s}) = \frac{-\bar{s}\lambda}{\theta - \bar{s}}.$$ 

The function $u_1$ is submodular in $(\bar{s}, \lambda)$, as is easily checked. Therefore, the second term of (22) is submodular. The logarithmic term breaks up into separate functions of $\bar{s}$ and $\lambda$, and is therefore modular in $(\bar{s}, \lambda)$. Submodularity of the objective implies that $\bar{s}$ is decreasing in $\lambda$ (see Topkis 1978).

The intuition here can be back traced to the incentive compatibility condition (26):

$$\bar{s} = \theta \frac{-u_1}{\lambda - u_1}.$$ 

A lower $\lambda$ (higher persistence) implies a higher sensitivity $\bar{s}$.

### 6.2 Impact of Volatility and Discount Rate

It is easy to check from (21) that the principal’s payoff is decreasing in the volatility of the agent’s cash flow. The next result shows that the optimal sensitivity coefficient is also decreasing in the volatility parameter.

**Proposition 12** The optimal sensitivity $\bar{s}$ is decreasing in $\sigma$.

**Proof.** Expression (21) is submodular in $\sigma$ and $\bar{s}$. The result follows from Topkis (1978).

This result is intuitive: A higher volatility exposes the agent to more risk, other things equal. Since the agent is risk averse, it is optimal for the principal to offset this by reducing the agent’s exposure to exogenous cash flow fluctuations, by reducing $\bar{s}$. Perhaps more surprisingly, the sensitivity is decreasing in the patience of the players, as shown in the Appendix.

**Proposition 13** The optimal sensitivity $\bar{s}$ is increasing in $r$.

Intuitively, a less patient agent is more concerned about getting an immediate subsidy (or tax reduction) than about his future utility. Inducing truth-telling thus requires a higher sensitivity of the agent’s future utility to his current report.
6.3 Long-Run Properties

Bloedel, Krishna, and Strulovici (2020) show that the agent’s utility always has a negative drift under the optimal contract, but that the long-run properties of consumption and promised utility depend on the parameters of the agent’s cash flow process. Fixing the volatility parameter, the optimal contract induces immiserisation when persistence is low, but sends the agent to bliss when persistence is high. Fixing the degree of persistence, the optimal contract induces immiserisation when volatility is high, but sends the agent to bliss when volatility is low.

7 Necessary and Sufficient Conditions for Incentive Compatibility

This section proves Theorem 3: it shows that (20) is necessary and sufficient for the contracts of Proposition 9 to be incentive compatible (i.e., truthful). The argument used to prove sufficiency illustrates a strategy, potentially useful in other problems, to deal with an unbounded reporting domain by expending the strategy space of the agent.

7.1 Necessity

Consider any contract that has the form of Proposition 9, which is characterized by parameters \( c_1 \in \mathbb{R} \) and \( \bar{s} \geq 0 \), such that \( C_t = c(W_t) \) and \( S_t = s(W_t) \) for all \( t \), where the functions \( c(\cdot) \) and \( s(\cdot) \) are given by \( c(w) = c_1 - \log(-w)/\theta \) and \( s(w) = \bar{s}(-w) \). From (13), the agent’s promised utility satisfies

\[
dW_t = (r + u_1)W_t dt + s(W_t)(dG_t + \lambda G_t dt + \sigma dB_t)
\]

where \( G_t = Y_t - X_t \) where \( u_1 = u(c_1) \). Moreover, the agent’s actual consumption is \( c(W_t) - G_t \). Therefore, the agent cares about \( Y_t \) and \( X_t \) only through their difference \( G_t \). The agent’s optimization problem reduces to

\[
v(w, g) = \sup_L \mathbb{E} \left[ \int_0^\infty e^{-rt} \left( u(c(W_t) - G_t) \right) dt \right]
\]

subject to \( dG_t = L_t dt, G_0 = g \), and (23). The HJB equation for this problem is

\[
0 = \sup_l \left\{ u(c(w) - g) - rv(w, g) \right. \\
+ v_w (rw - u(c(w)) + s(w)(l + \lambda g)) + v_g l + \frac{1}{2} (s(w))^2 \sigma^2 v_{ww} \right\}.
\]

It is easy to show by applying the same controls starting from different values of \( w \) that the value function has the form \( v(w, g) = wf(g) \) for some function \( f \). Incentive compatibility is characterized
by the condition \( f(0) = 1 \), which means that the agent cannot do better than get his promised utility \( w \) if he has not lied in the past. The Bellman equation becomes, after simplification and dividing throughout by \((-w)\),

\[
0 = \sup_{\ell} \{ u_1 \exp(\theta g) + f(g)(-u_1 + \bar{s}(\ell + \lambda g)) - f'(g)\ell \}. \tag{25}
\]

A priori, the function \( f \) defining the agent’s value function need not be everywhere differentiable, but it is a viscosity solution of (25). Moreover, \( v(w, g) \) is clearly decreasing in the gap \( g \), which implies that \( f \) is increasing and, hence, a.e. differentiable and everywhere left- and right-differentiable. Incentive compatibility implies that when \( g = 0 \) it is optimal for the agent not to increase the gap above zero. From (25), this requires that

\[
\bar{s} - f_r(0) \leq 0.
\]

Likewise, when \( g = 0 \), it must be optimal not to decrease the gap below zero, which requires that

\[
\bar{s} - f_l(0) \geq 0.
\]

Therefore, incentive compatibility implies that \( f_l(0) \leq \bar{s} \leq f_r(0) \).

We now prove that these inequalities are tight. Suppose by contradiction that one of these inequalities is strict, for example, \( \bar{s} < f_r(0) \). This implies that there exists a right neighborhood of 0 such that for \( g \in (0, \eta) \) it is strictly optimal for the agent to reduce the gap to zero, at an infinitely negative rate. We approximate the agent’s optimal strategy by considering an arbitrarily negative lying rate \(-K\) over some small time window, and taking the limit as \( K \to +\infty \). Specifically, suppose that at time \( t \) the state is \((w, g)\) for some fixed \( g \in (0, \eta) \). The agent expected utility at time \( t \) is \( v(w, g) \). Suppose that the agent lies at an arbitrarily large rate \(-K\) between times \( t \) and \( t_K = t + g/K \), which brings the gap to 0 at time \( t_K \). The agent’s promised utility over this time interval obeys the dynamic equation

\[
dW_t = W_t(r + u_1 + K\bar{s})dt + \bar{s}(-W_t)\sigma dB_t,
\]

which integrates to \( W_{t_K} = \hat{W}_{t_K} \exp(\bar{s}K) \) where \( \hat{W}_{t_K} \) is the agent’s promised utility at time \( t_K \) if he doesn’t lie during this period. The agent’s expected utility at time \( t_K \) is therefore \( v(W_{t_K}, G_{t_K}) = \hat{W}_{t_K} \exp(\bar{s}g)\exp(0) \) since by construction \( G_{t_K} = 0 \). As \( K \to +\infty \), we have \( t_K \to t, \hat{W}_{t_K} \to w \), and \( v(W_{t_K}, G_{t_K}) \leq v(w, g) \) since lying at the arbitrarily large rate \(-K\) is only one of many possible strategies for the agent. Moreover, incentive compatibility requires that \( f(0) = 1 \). Taken together, these observations yield \( wf(g) \geq w \exp(\bar{s}g) \). Dividing by \( w < 0 \) and noting that the argument works for all \( g \in (0, \eta) \), we get

\[
f(g) \leq \exp(\bar{s}g)
\]
for all $g \in (0, \eta)$. Since the functions on both sides of the equation are equal to 1 for $g = 0$, we get
\[
\frac{d[\exp(\bar{s}g)]}{dg} \bigg|_{g=0} = \bar{s},
\]
which yields the desired contradiction. By a similar argument, $f_t(0) = \bar{s}$.

In conclusion, $f$ is differentiable at 0, with derivative $\bar{s}$. This implies that the value function $\bar{v}(w, x, y) = v(w, y - x)$ of the agent is differentiable with respect to $x$ and to $y$ whenever $x = y$. In the Appendix (Lemma 2), this differentiability is used in an envelope argument to show that, necessarily,
\[
\bar{s} = f'(0) = \bar{v}_x(w, x, x) = -\bar{v}_y(w, x, x) = \int_0^\infty e^{-(r+\lambda)t}u'(C_t + X_t - Y_t)dt.
\]
The right-hand side is computed explicitly in the Appendix (Lemma 3), yielding
\[
\bar{s} = \frac{\theta(-u_1)}{\lambda - u_1} \in (0, \theta). \tag{26}
\]
Notice that $\bar{s} < \theta$ for all $u_1 < 0$ and $\lambda > 0$. Intuitively, if $\bar{s}$ were higher than $\theta$, the agent would want to exaggerate the cash flow in order to artificially increase his promised utility. The effect of earning less than the reported cash flow would reduce the agent’s flow utility at rate $\theta$, which would be dominated by the increase in promised utility governed by the sensitivity parameter $\bar{s}$. Incentive compatibility rules this case out.

7.2 Sufficiency

We have established that (20) is necessary for incentive compatibility. We now show that it is sufficient. The objective (25) is linear in $\ell$, which has unbounded domain. If the contract is not truthful, the agent therefore wants to lie at an infinite rate. To accommodate for this, we allow the agent to report jumps in his cash flows, which expands the agent’s strategy space. Admissibility of the agent’s strategy is defined by extending the transversality condition (4) to this space.

Let $W_{t+}(\Delta L)$ denote the agent’s promised utility if he reports a jump $\Delta L$ in his cash flow at time $t$. For contracts with a fixed sensitivity parameter, as considered here, a natural closure of the contract to report jumps is to stipulate that
\[
W_{t+}(\Delta L) = \exp(-\bar{s}\Delta L)W_t. \tag{27}
\]
To see this, notice that if the agent lies at an arbitrarily large rate $K$ between times $t$ and $t + \varepsilon$, his promised utility satisfies, ignoring second-order effects (see the analysis of the necessary condition, above, for a formal argument), the dynamic equation $dW_t = \bar{s}(-W_t)Kdt$. This yields $W_{t+\varepsilon} = \ldots$
\[ \exp(-K\bar{s}\varepsilon)W_t, \] and results in a gap change \[ G_{t+\varepsilon} = G_t + K\varepsilon. \] Combining the last two equations yields \[ W_{t+\varepsilon} = \exp(-\bar{s}(G_{t+\varepsilon} - G_t))W_t, \] which explains (27).

Report jumps amount to impulse controls on the part of the agent (see for example, Øksendal and Sulem (2004)). The HJB equation (25) becomes

\[ 0 = \max \left\{ \sup_{\ell \in \mathbb{R}} \left\{ u_1 \exp(\theta g) + f(g)\left(-u_1 + \bar{s}(\ell + \lambda g)\right) - f'(g)\ell \right\}, \right. \]
\[ \left. \sup_{\Delta L \in \mathbb{R}\setminus\{0\}} \left\{ \exp(-\bar{s}\Delta L)f(g + \Delta L) - f(g) \right\} \right\}. \quad (28) \]

We now show that the function \( f(g) = \exp(\bar{s}g) \) solves this equation. With this value for \( f \), the second term of the equation is always equal to zero. Therefore, it suffices to show that

\[ \sup_{\ell} \left\{ u_1 \exp(\theta g) + \exp(\bar{s}g)\left(-u_1 + \bar{s}(\ell + \lambda g)\right) - \bar{s}\exp(\bar{s}g)\ell \right\} \leq 0 \quad (29) \]

for all \( g \). The terms involving \( \ell \) cancel out and (29) reduces to

\[ u_1 \exp(\theta g) + \exp(\bar{s}g)\left(-u_1 + \bar{s}\lambda g\right) \]

Convexity of the exponential function implies that for all \( g \neq 0 \),

\[ \exp(\theta g) > \exp(\bar{s}g) + \exp(\bar{s}g)(\theta - \bar{s})g. \]

Since \( u_1 < 0 \), this implies that (29) is satisfied if

\[ \exp(\bar{s}g)\left(u_1(\theta - \bar{s})g + \bar{s}\lambda g\right) \leq 0. \]

The second factor is zero, from (20), which concludes the proof.

An optimal control associated with the Bellman equation is to set \( \Delta L = -g \) if \( g \neq 0 \) and \( \Delta L = 0 \) otherwise, and \( \ell \) always equal to zero. It is optimal for the agent to (i) always report truthfully if he has been truthful in the past, and (ii) immediately correct any existing gap between real and reported cash flows. In particular, if the principal did not know the initial cash flow, the contract would still be incentive compatible. This optimal control is essentially unique: between two impulse controls, the first term of the Bellman equation must be equal to zero, which holds only if \( g = 0 \).

8 Discussion

8.1 Summary: Concept of Renegotiation

This paper introduces a concept of renegotiation for stochastic games that strengthens internal consistency by exploiting comparisons across states. These comparisons are weaker than those
imposed by strongly renegotiation-proof equilibria (Farrell and Maskin (1989)) and could be applied to study stable “social norms” (Asheim (1991)), with the following cognitive interpretation: players may recognize relations between equilibria across different states, just as they recognize, in a repeated game, whether continuation payoffs across different histories are Pareto ranked. This requires players to think by analogy rather than come up with radically new equilibria. From a modeling perspective, economic analysis often relies on specific payoff or dynamic structures (e.g., CARA or CRRA preferences, linear or geometric growth). For such models, the present approach provides a natural way to model renegotiation.

The concept is well-defined in any of the following situations: (i) there is no private information, (even when time is discrete), (ii) there is private information but we focus on truthful contracts (as in Strulovici (2011)), or (iii) we use a diffusion model, as in the main application of this paper. With private information, the analysis has relied on a diffusion model to produce an Observability Theorem and a Revelation Principle that addresses informational asymmetries.29

8.2 Robustness and Extensions

Agent Effort to Generate Higher Cash Flows: The agent could affect his real cash flow by putting some privately observed effort. This extension is realistic in many settings (e.g., to study labor incentives and taxation) and connects the model to numerous papers on dynamic moral hazard, such as Sannikov (2008, 2014). An earlier version of this paper (Strulovici (2011)) studies this extension: the agent privately chooses some effort that affects his cash flow: $dX_t = (\xi - \lambda X_t + A_t)dt + \sigma dB_t$, where $A_t = \{A_t\}_{t \geq 0}$ is the agent’s effort process and the agent incurs an additively separable cost $\phi(A_t)$ at time $t$ from effort $A_t$. Although they are more complicated, all the arguments and results of this paper continue to go through:30 assuming that $\phi$ takes the exponential form $\phi(a) = \tilde{\phi}\exp(\chi a)$ for some positive parameters $\tilde{\phi}, \chi$, the definition of state consistency and the form of renegotiation-proof contacts (consumption and sensitivity) are unchanged and the

29While the focus on adapted strategies is quasi-universal in economic models with diffusion processes (for instance, the Martingale Representation Theorem relies on this assumption), this focus is generally not innocuous. See Bernard (2016) for a theory of arbitrarily mixed strategies in continuous time. As explained in the Introduction, renegotiation tends to reveal any private information that is required to implement efficient allocations as communication becomes arbitrarily frequent (Maestri (2017), Strulovici (2017)) which suggests that a richer diffusion model in which renegotiation is modeled with an explicit protocol would also lead to a truthful contracts. Such a model seems challenging to formalize but would be interesting to explore.

30The only significant change concerns the convexification argument. Convexifying the cost of effort across two contracts leads to different paths for the cash flow process and affects consumption under each path. This issue can be addressed by backing out the underlying Brownian path from the agent’s report process and expressing the contracts in these terms.
agent’s cost of effort is now proportional to the negative of the agent’s utility: \( \phi(A_t) = \phi_1(-W_t) \)
where \( \phi_1 = \bar{s}/\chi > 0 \) and the incentive compatibility condition (20) is replaced by the condition:
\[ u_1(\bar{s}) = \bar{s}(\chi \lambda + \bar{s})/(\chi(\theta - \bar{s})). \]
The comparative statics also go through (some are harder to prove—see Strulovici (2011)) and new results and comparative statics obtain with respect to the agent’s cost of effort.

**Reporting Constraints** Truthful renegotiation-proof contracts clearly remain truthful if the agent faces additional constraints on cash flow reports and transfers. For example, the agent could be unable from over-reporting his cash flows beyond some upper bound. Such constraints reduce the set of possible deviations and facilitate truth-telling.

**Self-Insurance and Private Savings** Bloedel, Krishna, and Strulovici (2020) consider an alternative implementation of the model, in which the principal allows the agent to self insure by investing in some asset whose return is predetermined by the principal. They show that equilibria of the self insurance problem are equivalent to the renegotiation-proof equilibria studied here. They also study the possibility of private savings and this addition pins down the sensitivity parameter of the contract and yields the contract studied by Williams (2011).

**Permanent Outside Option** The agent may be able to leave the contract for some outside option at any time, or the principal may wish to prevent the agent’s continuation utility from becoming too low (e.g., to avoid immiserisation). This extension is discussed in Appendix B.

**Appendix A: Proofs**

**Proof of Theorem 1**

To prove strong uniqueness, we can analyze the problem under any probability measure that is absolutely continuous with respect to \( P \). Consider the probability measure \( Q^Y \) under which \( \frac{1}{\sigma} Y \) is the standard Brownian motion. This measure is absolutely continuous because \( Y \) has constant quadratic variation equal to \( \sigma \), and the drift of \( Y \) satisfies the standard integrability conditions.\(^{31}\) Equation (6) is an SDE with constant volatility coefficient and functional drift \(-L(t, \cdot)\).

Part 1. The boundedness assumed in the premise of Theorem 1 implies that condition (4.169) in Liptser and Shiryaev (2001) is satisfied. Theorem 4.13 of Liptser and Shiryaev (2001) then implies that (6) has a unique (in law) weak solution. This, together with Theorem 3.2 in Cherny (2002) implies pathwise uniqueness, which implies that \( X \) is adapted to \( Y \) and the existence of a unique

---

\(^{31}\)The drift of \( Y \) depends on two components: the drift of \( X \), which is linear in \( X \), and the lying process \( L \) which is assumed to be locally bounded.
Part 2. Equation (6) satisfies assumptions (a1)–(a4) of Huang (2018): Assumption (a1) and (a2) trivially hold because volatility is constant, assumption (a4) holds because the term \( b \) in Huang’s paper is identically equal to 0, and assumption (a3) is explicitly imposed in the premise of Theorem 1 and the premise of Part 2. Theorem 1.1 of Huang (2018) then implies pathwise uniqueness—which implies that \( X \) is adapted to \( Y \)—and the existence of a unique strong solution.\(^{33}\)

Part 3. Equation (6) satisfies conditions (C1)–(C3) of Bachmann (2020): Condition (C1) comes from the boundedness assumed in the premise of Theorem 1. Condition (C2) holds trivially because the volatility coefficient is constant and positive, and Condition (C3) is explicitly imposed in the premise of Part 2. Bachmann’s Theorem 1.5 then implies that pathwise uniqueness holds. This implies that \( X \) is adapted to \( Y \) and the existence of a strong solution. \( \blacksquare \)

Proof of Proposition 1

State consistency implies that, after any history leading up to some state \((\tilde{v}, \tilde{x}) = g \circ (v, x)\) and continuation equilibrium \((\tilde{C}, \tilde{L})\), we have

\[
\Pi(\tilde{C}, \tilde{L}) \geq \Pi(G_g(C, L)).
\]

Suppose by contradiction that this inequality is strict. Monotonicity of \( G \) applied this time to \( g^{-1} \) implies that

\[
\Pi(G_{g^{-1}}(\tilde{C}, \tilde{L})) > \Pi(G_{g^{-1}}(G_g(C, L))) = \Pi(G_{g^{-1}}(C, L)) = \Pi(C, L),
\]

which contradicts the state consistency of \( C \). \( \blacksquare \)

Proof of Proposition 4

Consider any reporting strategy \( L \) and let \( V^\lambda(L) \) denote the agent’s expected utility when using strategy \( L \) and facing contract \( C^\lambda \). Since the agent consumes \( C^\lambda_t - G_t \) at time \( t \) and the agent has utility \( u(c) = -\exp(-\theta c) \), we have

\[
V^\lambda(L) = E \left[ \int_0^\infty e^{-rt} U^\lambda_t e^{\theta G_L^t} dt | L \right]
\]

where \( U^\lambda_t = u(C^\lambda_t) \) and \( G^L_t = \int_0^t L_s ds \) is the gap. By construction of \( C^\lambda \), we have

\[
U^\lambda_t = \lambda U^1_t + (1 - \lambda) U^2_t
\]

\(^{32}\)This principle was stated and proved by Yamada and Watanabe (1971) for standard (as opposed to functional) SDEs. However, their proof extends without any change to functional SDEs, a fact that is well-known and widely used in the literature on functional SDEs. See, e.g., Cherny (2002, Figure 1) and Huang (2018, Theorem 2.8). I am grateful to Nikolai Krylov and Stefan Bachmann for pointing this out.

\(^{33}\)See Theorem 2.8 in Huang (2018).
for all $t$ and report histories $\{Y_s : s \leq t\}$, where $U_i^t = u(C_i^t)$ for $i \in \{1, 2\}$. This implies that

$$V^\lambda(L) = \lambda V^1(L) + (1 - \lambda) V^2(L)$$  \hspace{1cm} (30)

where $V^i(L)$ is the agent’s expected utility with strategy $L$ under contract $C^i$. By assumption, the contracts $C^1$ and $C^2$ are truthful, so for $i \in \{1, 2\}$, we have $V^i(L) \leq v$ for all $L$ and $V^i(L \equiv 0) = v$. Combining these observations with (30) yields $V^\lambda(L) \leq v$, and the inequality is tight if $L \equiv 0$. ■

**Proof of Proposition 5**

We prove that if a strategy $L$ gives a higher expected utility than another strategy $L'$ under the initial contract, then it also does so under the new contract $\hat{C}$. Suppose that $L$ gives higher expected utility than $L'$ under $C$. This means that

$$E \left[ \int_0^\infty e^{-rt} \left( u \left( C(Y_s : s \leq t) - \int_0^t L_s ds \right) \right) dt \right] \geq E \left[ \int_0^\infty e^{-rt} \left( u \left( C(Y'_s : s \leq t) - \int_0^t L'_s ds \right) \right) dt \right] \hspace{1cm} (31)$$

where

$$dY_t = \left[ L_t + \left( \xi - \lambda \left( Y_t - \int_0^t L_s ds \right) \right) \right] dt + \sigma dB_t$$

and

$$dY'_t = \left[ L'_t + \left( \xi - \lambda \left( Y'_t - \int_0^t L'_s ds \right) \right) \right] dt + \sigma dB_t$$

and the initial conditions $Y_0 = Y'_0 = x$.

Now consider another initial condition $\hat{x}$. The reporting processes under strategies $L$ and $L'$ are respectively

$$d\hat{Y}_t = \left[ L_t + \left( \xi - \lambda \left( \hat{Y}_t - \int_0^t L_s ds \right) \right) \right] dt + \sigma dB_t$$

and

$$d\hat{Y}'_t = \left[ L'_t + \left( \xi - \lambda \left( \hat{Y}'_t - \int_0^t L'_s ds \right) \right) \right] dt + \sigma dB_t$$

subject to the initial condition $\hat{Y}_0 = \hat{Y}'_0 = \hat{x}$.

By construction, the virtual cash flow processes follow the equations

$$d\tilde{Y}_t = d\hat{Y}_t - (\xi - \lambda \hat{Y}_t) dt + (\xi - \lambda \hat{Y}_t) dt$$

$$d\tilde{Y}'_t = d\hat{Y}'_t - (\xi - \lambda \hat{Y}'_t) dt + (\xi - \lambda \hat{Y}'_t) dt$$

subject to $\tilde{Y}_0 = \tilde{Y}'_0 = y$. Combining the last four equations pairwise yields

$$d\tilde{Y}_t = \left[ L_t + \left( \xi - \lambda \left( \hat{Y}_t - \int_0^t L_s ds \right) \right) \right] dt + \sigma dB_t$$

34
\[ d\tilde{Y}'_t = \left[ L'_t + \left( \xi - \lambda \left( \tilde{Y}'_t - \int_0^t L'_s ds \right) \right) \right] dt + \sigma dB_t \]

subject to the conditions \( \tilde{Y}_0 = \tilde{Y}'_0 = y \). By definition of the contract \( \hat{C} \), the agent’s expected utilities using strategies \( L \) and \( L' \) when facing contract \( \hat{C} \) are identical the expected utilities appearing in (31) and we have just shown that \( \tilde{Y} \) and \( \tilde{Y}' \) are subject to the same dynamic equations and initial condition as \( Y \) and \( Y' \). This prove that strategy \( L \) also dominates \( L' \) under contract \( \hat{C} \). ■

Proof of Lemma 1

Since \( C \) is truthful and state consistent, \( \hat{C} \) is also truthful and we have \( \Pi(C) = \Pi(\hat{C}) = \Pi(w, y) \) by Proposition 8. Fix any \( \lambda \in (0, 1) \). From Proposition 4, the convexification \( C^\lambda \) of \( C \) and \( \hat{C} \) is also truthful and gives the same promised utility \( w \) to the agent given \( y \). Moreover, strict concavity of the utility function \( u \) together with the definition of \( C^\lambda \) (Equation (9)) implies that \( C^\lambda_t \leq \lambda C_t + (1 - \lambda) \hat{C}_t \) for all \( t \) with a strict inequality whenever \( C_t \neq \hat{C}_t \). This implies that \( \Pi(C^\lambda) < \lambda \Pi(C^1) + (1 - \lambda) \Pi(C^2) = \Pi(w, y) \) and violates renegotiation-proofness of \( C \) (Equation (10)) unless \( \| C - \hat{C} \|_L = 0 \). ■

Necessary Conditions for Incentive Compatibility

**Lemma 2** For any truthful, renegotiation-proof contract, we have

\[ \emptyset_y(w, x, x) = -\emptyset_x(w, x, x) = -\int_0^\infty e^{-(r+\lambda)t} u'(C_t + X_t - Y_t) dt. \]  

(32)

**Proof.** The value function of the agent satisfies the optimization problem

\[ \emptyset(w, y, x) = \sup_{\emptyset} E \left[ \int_0^\infty e^{-rt} (u(C_t(Y_s : s \leq t) + X_t - Y_t)) dt \right], \]

where \( C_t(\cdot) \) is, for each \( t \), a functional that determines the consumption provided to the agent at time \( t \) given past reports \( \{Y_s : s \leq t\} \). If the initial cash flow is increased by \( \varepsilon \), this affects the distribution of future cash flows and, keeping the lying process fixed, of future reports. However, by a change of variable, one can control the path of the report process \( Y_t \), and make it independent from the initial cash flow change. Recall that

\[ dY_t = [(\xi - \lambda X_t) + L_t] dt + \sigma dB_t. \]

Making the change of variable \( \tilde{L}_t = L_t + (\xi - \lambda X_t) - (\xi - \lambda Y_t) \), we get

\[ dY_t = [(\xi - \lambda Y_t) + \tilde{L}_t] dt + \sigma dB_t. \]  

(33)

The agent’s strategy can be restated as choosing \( \tilde{L} \), rather than \( L \):

\[ \emptyset(w, y, x) = \sup_{\emptyset} E \left[ \int_0^\infty e^{-rt} (u(C_t(Y_s : s \leq t) + X_t - Y_t)) \right]. \]
subject to $Y_0 = y$, $X_0 = x$, (33), and

$$dX_t = (\xi - \lambda X_t)dt + \sigma dB_t.$$  

$$dW_t = (rW_t - u(C_t(Y_s : s \leq t)))dt + S_t (dY_t - (\xi - \lambda Y_t)dt).$$

If the contract is incentive compatible, it is optimal to set $\bar{L}_t = 0$ whenever initial conditions are such that $y = x$. Section 7 established that the value function is differentiable with respect to $x$ and $y$ whenever $x = y$. An application of the Envelope Theorem (Milgrom and Segal, 2002, Theorem 1) then implies that $\bar{v}_x(w, x, x)$ can be computed by evaluating the objective function at $\bar{L}_t \equiv 0$ or, equivalently, under the report process $Y_t$ starting from $y_0 = x_0$. Under this approach, $W$, $Y$, and $C$ are independent of the initial condition $x$, and

$$\bar{v}_x(w, x, x) = \int_0^\infty e^{-rt} \frac{d}{dx} E[u(X_t - Y_t + C_t(Y_s : s \leq t))] dt.$$  

Since the distribution of $\{Y_s\}_{s \leq t}$ is independent from the initial condition $x$, the inner derivative is equal to

$$E\left[u'(X_t - Y_t + C_t(Y_s : s \leq t)) \frac{dX_t}{dx}\right].$$

The process $X$ defined by the dynamic equation (1) is a generalization of Ornstein-Uhlenbeck processes, which can be explicitly integrated:

$$X_t = e^{-\lambda t} x + \int_0^t e^{\lambda(s-t)} \xi ds + \int_0^t e^{\lambda(s-t)} \sigma dB_s$$  

(34)

This implies that $\frac{dX_t}{dx} = e^{-\lambda t}$ and yields the formula of Lemma 2.

**Lemma 3** The following equality holds for any truthful renegotiation-proof contract:

$$E\left[\int_0^\infty e^{-(r+\lambda)t} u'(C_t)dt\right] = (-w) \frac{(-u_1)}{\lambda - u_1}.$$  

**Proof.** Since $u'(C_t) = -\theta u(C_t)$,

$$E\left[\int_0^\infty e^{-(r+\lambda)t} u'(C_t)dt\right] = E\left[\int_0^\infty e^{-(r+\lambda)t} \theta u(C_t)dt\right] = \theta u_1 \int_0^\infty e^{-(r+\lambda)t} E[W_t]dt.$$  

Moreover,

$$dW_t = [rW_t - u(C_t)]dt + S_t \sigma dB_t.$$  

With the exponential utility specification, we have $u(C_t) = -W_t u_1$. Letting $\vartheta(t) = E[W_t] (\vartheta(0) = w)$, this implies that

$$\frac{d\vartheta}{dt}(t) = (r + u_1) \vartheta(t).$$

36
and hence that
\[ E[W_t] = e^{(r+u_1)t}w. \] (35)
Integrating this expression over time yields the result.

**Proof of Proposition 10**

The principal’s objective is to maximize \( \Pi(w, x) \) with respect to \( u_1 \). From (14), and neglecting for now terms that are not affected by the contract choice, this is equivalent to solving
\[
\max_{u_1} \int_0^\infty e^{-rt} \left( \frac{\alpha_t}{r + \lambda} - E[C_t] \right) dt
\]
where
\[
\alpha_t = \frac{E[\log(-W_t)]}{\chi} \quad \text{and} \quad E[C_t] = c_1 - \frac{E[\log(-W_t)]}{\theta}.
\]
Letting \( Z_t = \log(-W_t) \), Itô’s formula implies that
\[
dZ_t = (r + u_1)dt - \frac{1}{2}\sigma^2s^2dt - s\sigma dB_t.
\]
Therefore,
\[
E\log(-W_t) = \log(-w) + (r + u_1)t - \frac{1}{2}\sigma^2s^2t.
\]
The principal’s objective is to maximize
\[
-\frac{c_1}{r} + \frac{1}{\theta} \int_0^\infty e^{-rt}E\log(-W_t)dt.
\]
Replacing \( c_1 \) and \( E\log(-W_t) \) by their formulas in terms of \( u_1, t \), and the parameters of the model, and integrating the last term proves the proposition.

**Proof of Proposition 13**

The optimal value of \( \bar{s} \) is unchanged if we multiply (21) by \( r \) and get rid of the term \( r \) in the factor \((r + u_1 - 1/2\sigma^2s^2)\). The resulting objective function is equal to
\[
r \log(-u_1(\bar{s})) + \left( u_1(\bar{s}) - \frac{1}{2}\sigma^2s^2 \right).
\]
The function \( u_1(\bar{s}) \) is increasing in \( \bar{s} \). Therefore, the objective is supermodular in \((\bar{s}, r)\) and the result follows (see, e.g., Topkis (1978)).

**Appendix B: Permanent Outside Option**

Suppose that the agent is allowed to quite at any time and receive continuation utility \( \bar{w} \). This imposes a new individual rationality constraint \( W_t \geq \bar{w} \). An alternative interpretation is that the
principal wishes to guarantee a continuation utility level $w$ to the agent at all times, for example to avoid immiserisation. How does this affect renegotiation? Some earlier arguments go through. First, any internally-consistent contract has continuation payoffs that only depend on the current state $(w, y)$. Second, comparing cash-flow levels, the argument of Section 4.1 goes through, so that contractual variables should depend only on promised utility, not on the cash-flow level. Moreover, the principal’s payoff function should vary across cash flow levels according to Equation (14).

However, the constraint reduces the ability to compare contracts across promised utility levels: given a truthful contract $C$ starting in state $(w_1, y)$ for $w_1 > w$, there is no guarantee that, for $w_2 \in (w, w_1)$, the transformation of $C$ that starts in state $(w_2, y)$ satisfies the individual constraint. This reduces the set of challengers that can be constructed through the transformation and prevents us from exploiting earlier comparisons to derive the closed-form formulas obtained in Section 4.2.\(^{34}\)

Conceptually, however, the problem is similar to the unconstrained case. At one extreme, for $w$ far above $\bar{w}$, the optimal renegotiation-proof contract should be very similar to optimal contract of the unconstrained case, and the payoff and contractual functions should be well approximated by the closed-form functions derived for that case. At the other extreme, if $w = \bar{w}$, the principal has very few options to keep the agent in the relationship. Indeed, the only contracts that guarantee that the constraint is not violated are those for which (i) the sensitivity parameter of the promised utility is exactly zero (for otherwise the promised utility of the agent might drop below $\bar{w}$), and (ii) the drift is positive (to push $W_t$ higher away from $\bar{w}$), for example by providing a low utility flow. Such extreme contractual characteristics are clearly not required for $w$ high above $\bar{w}$.

The cross-state comparison imposes restrictions on the principal’s continuation payoff. One direction of the unconstrained analysis carries over to the constrained case, yielding a continuum of inequalities satisfied by the principal payoffs. Suppose that $w_2 > w_1$. The $(w_2, y)$-transformation of $C$ satisfies the constraint if $C$ did, suggesting the following conjecture. For any individually-rational, truthful, renegotiation-proof contract $C$, let $\Pi(w, y)$ denote the principal’s payoff under any continuation of $C$ starting in state $(w, y)$.

**Conjecture 1** For any states $(w, y), (w', y')$ such that $w' \geq w$,

$$
\Pi(w', y') \geq \frac{(y' - y)}{\tau + \lambda} + \frac{\log(w'/w)}{r} \left( \frac{1}{\theta} + \frac{1}{\chi(r + \lambda)} \right) + \Pi(w, y). \tag{36}
$$

We may conjecture that for the optimal renegotiation-proof contract subject to the constraint, the sensitivity factor is increasing in $w$ (instead of constant, in the unconstrained case) from 0 for $w = \bar{w}$ to the unconstrained optimum $\bar{s}$ (i.e., the maximizer of (21)), as $w$ gets arbitrary large.

\(^{34}\)Note that the convexification argument is unaffected by the restriction, since if both $W_t^1$ and $W_t^2$ exceed $\bar{w}$, so does any convex combination.
References


39


