A Dynamic Game under Ambiguity: Contracting for Delegated Experimentation

David Besanko\textsuperscript{1} \hspace{1cm} Jian Tong\textsuperscript{2}
Northwestern University \hspace{1cm} University of Southampton

Jianjun Wu\textsuperscript{3}
University of Arizona and Princeton Economics Group

First version: December 30, 2011
This version: June 25, 2012

\textsuperscript{1}Department of Management and Strategy, Kellogg School of Management, Northwestern University, Evanston, IL 60208. Email: d-besanko@kellogg.northwestern.edu
\textsuperscript{2}Economics Division, School of Social Sciences, University of Southampton, Southampton, SO17 1BJ, UK, Email: j.tong@soton.ac.uk
\textsuperscript{3}Department of Economics, University of Arizona, Tucson, AZ 85721-0108 and Princeton Economics Group. Email: jwu@eller.arizona.edu.
Abstract

Conventional (single-prior) Bayesian games of incomplete information are limited in their ability to capture the extent of informational asymmetry. In particular, they are not capable of representing complete ignorance of an uninformed player about an unknown parameter of the environment. Using a framework of contracting for delegated experimentation, we formulate and analyze a dynamic game of incomplete information that incorporates a multiple-prior belief system. Specifically, we consider a game with a principal contracting with an expert agent for his (observable) effort on a novel experiment — a Poisson process with unknown hazard rate. Although the expert agent has sufficient knowledge to form a single prior over the hazard rate, the principal initially has complete ignorance and her ambiguous beliefs are represented by the set of all plausible prior distributions over the hazard rate. We propose a new equilibrium concept — Perfect Objectivist Equilibrium — in which the principal, who has ambiguity aversion, draws inference about the agent’s prior from the observed history of the game via maximum likelihood updating. The new equilibrium concept thus also embodies a novel model of learning under ambiguity in the context of a dynamic game. Although the game is rich in its contractual space and strategic interactions, the unique (Markov) equilibrium outcome is a remarkably simple pooling contract with appealing economic properties. In addition, the underlying Markov Perfect Objectivist Equilibria are all belief-free. These are in sharp contrast with the set of Markov Perfect Bayesian Equilibria, which not only hinge on subjective pretence of knowledge, but also predict multiple continuum of equilibrium outcomes.
“Acknowledging what is known as known, what is not known as unknown, that is knowledge.”

– Analects of Confucius

“What has now appeared is that the mathematical concept of probability is inadequate to express our mental confidence or diffidence in making such inferences, and that the mathematical quantity which appears to be appropriate for measuring our order of preference among different possible populations does not in fact obey the laws of probability. To distinguish it from probability, I have used the term ‘likelihood’ to designate this quantity.”

– Sir R. A. Fisher (1925)

1 Introduction

Ad hoc subjective belief is commonplace in many economic models of rational behavior. Bayesian games, particularly, dynamic settings, often hinge on unexplained arbitrary prior beliefs about the economic environment. Defenders of the Savage-Bayesian paradigm would usually invoke the argument that inferences have to start from somewhere, so therefore an unexplained (unique) prior is a necessity. At a fundamental level, this line of argument has been exposed to be questionable in light of the recent literature on multiple priors and ambiguity (see Gilboa, Maccheroni, Marinacci and Schmeidler or GMMS, 2010; Manski, 2008), according to which, a multiple-prior belief system should be the rule and the unique-prior belief system is just an extreme exception. Epstein and Schneider (2007) further illustrate that in a multiple priors setting, some form of likelihood inference can suitably replace the familiar Bayesian inference.

In this paper, we use the example of ‘contracting for delegated experimentation’ to illustrate that the ad hoc subjective element can indeed be removed from the selection of prior distributions for a (dynamic) game of incomplete information. By allowing multiple prior distributions, arbitrary restriction on the set of priors, such as uniqueness, becomes unnecessary. As a consequence, the (iterative) selection of priors, in addition to updating beliefs based on them, becomes an ongoing process as the game is played. Learning (including the selection of priors and updating of beliefs) under ambiguity is based on observational facts and a likelihood function that is derived from the common knowledge of the game and suitable concept of equilibrium, of which ad hoc unique prior is no part.¹

¹If the observational data are not contaminated by random noises, as is the case in our model, the likelihood function degenerates and gives deterministic falsification of some subsets of hypotheses. The logical status of the inference degenerates to deductive inference.
The agent must continually decide how much to invest in solving the problem, and the investment of more funds can speed up the time of the breakthrough if the problem is indeed solvable. This basic setting admits to a number of specific economic applications, such as government subsidization of basic scientific research.

As an expert, the agent (a ‘Bayesian’ player) has sufficient knowledge to form a unique prior probability about whether the problem is solvable, and he updates his posterior probability according to the investment history in the absence of a breakthrough. Bayes Rule dictates that the posterior probability decreases over time conditional on non-zero investment and absence of a breakthrough. The principal (a ‘non-Bayesian’ player) — who does not know the agent’s prior belief (his “type”) but who can observe the investment history — relies on this history to infer the agent’s type. She must act on her ambiguous beliefs — represented by multiple priors and posteriors about the agent’s type — to determine how much to compensate the agent at each point in time. The principal’s deficiency of information about the true state of the world means that she cannot form complete pairwise rankings over alternative action plans (even without strategic uncertainty). In general, then, there may not exist any unanimous winners among all candidate action plans. How this potential indecision is resolved depends critically on the principal’s attitude toward ambiguity. We follow GMMS (2010) to represent this factor of decision making by the max-min expected utility criterion.23

Our paper makes five main contributions. First, we formulate and analyze a dynamic game of incomplete information with a multiple-prior belief system, as opposed to the conventional (single-prior) Bayesian game, a la Harsanyi (1967, 1968).4 The game we analyze is rich in strategic interactions, full of potential for multiplicity of equilibria, and would be expected to generate equilibria that would be sensitive to the choice of a prior. Yet, we find that the equilibrium outcome can be implemented by a remarkably simple (pure strategy) pooling contract in which the principal compensates the agent with a cost-reimbursement contract with a time invariant reimbursement rate. The information transmission in equilibrium is minimal, but it is adequate to ensure that the investment intensity and the cumulative level of investment are Pareto efficient (allowing for compensation transfer).5 Second, by developing a model that involves both learning (i.e., how a ‘non-Bayesian’ player learns) under ambiguity and strategic interactions between rational players, we add to and complement

---

2In the paper we also formally establish the equivalence between the maxmin expected utility criterion and the maxmin utility criterion for our setting, therefore we equivalently adopt the maxmin criterion originated by Abraham Wald (1950) in his “Statistical Decision Functions”.

3More precisely, the maxmin expected utility criterion is only a necessary condition for a most preferred action plan. To ensure sufficiency, a most preferred action plan must not be weakly dominated by any feasible alternative. This additional necessary condition was originally proposed by Manski (2008).

4Inspired by the title of Harsanyi’s seminal paper — “Games with Incomplete Information Played by ‘Bayesian’ Players” — A fitting alternative title of the current paper could be “A Dynamic Game with Incomplete Information Involving A ‘Non-Bayesian’ Player”.

5The notion of socially efficient allocation adopted in this paper is based on Pareto efficiency and the compensation principle, which states that the winner(s) must compensate the loser(s). The compensation transfer is explicitly modelled. Whenever Pareto efficiency is subject to informational constraints, the latter are fully reflected in the representation of preferences.
the emerging literature on learning under ambiguity (Epstein and Schneider, 2003, 2007). Third, to analyze our dynamic game, we present a new equilibrium concept — the Perfect Objectivist Equilibrium (POE) — which extends and contrasts the familiar Perfect Bayesian Equilibrium (PBE) that is the standard in dynamic (single-prior) Bayesian games. The term ‘objectivist’ emphasizes the importance of objective inference rules in the learning process. The new equilibrium concept also allows us to deal with out-of-equilibrium updating of beliefs in a systematic way. Fourth, to solve for the simplest equilibrium outcome of the model, we define Markov strategies of the players and characterize the Markov Perfect Objectivist Equilibria (MPOE), which we prove to exist and deliver the uniqueness of the (Markov) equilibrium outcome. The MPOE of the current dynamic game with incomplete information hinges on an (interactive learning-related) state variable, whose transition features a ratchet mechanism — it can only move in one direction — renown as the “ratchet effect” in dynamic mechanism design (with limited commitment) literature. Fifth, by solving for illustrative Markov Perfect Bayesian Equilibria (MPBE) and Belief-Free Equilibrium (BFE) of the model, we are able to compare them with the MPOE. The comparison reveals that the MPOE of the game is also a belief-free equilibrium, therefore is strongly robust with respect to the specification of subjective beliefs; in contrast, the MPBE of the game suffer a problem of multiplicity of equilibrium outcomes (for a given belief specification) as well as a compounding issue of “ad hocery” with respect to the specification of subjective beliefs.

The efficiency, uniqueness, robustness and simplicity of the (Markov) equilibrium outcome in our model are especially noteworthy. We start by noting that a conventional (single-prior) Bayesian game of incomplete information is limited in its ability to model the extent of informational asymmetry. Particularly, it is not capable of representing complete ignorance of an uninformed player about an unknown parameter of the environment, as is the case in our model. The reason is simple: in a single-prior Bayesian game the belief of an uninformed player is represented by a single probability distribution, which is always informative to some extent. This biases the extent of informational asymmetry downward. In contrast, this modelling bias is overcome by our multiple prior model, which can represent complete ignorance adequately. That is, the multiple-prior model allows more extreme asymmetry of information between the principal and agent than a conventional Bayesian game does. As a result, our analysis shows that ambiguity and ambiguity aversion weaken the principal’s ability to minimize the agent’s information rent. In effect, the principal becomes so passive in minimizing the agent’s information rent that the contract she offers does not sacrifice the efficiency of investment in a trade off to extract more expected surplus from trade. Consequently, there exists a unique Markov equilibrium outcome, which is also Pareto efficient (allowing for compensation transfer). The Markov equilibrium outcome is robust in the sense that it does not delicately depend on any ad hoc prior belief and out-of-equilibrium posterior belief. Due to Pareto efficiency it is also robust in that it is renegotiation-proof. The Markov equilibrium outcome is also very simple both structurally and dynamically. The main reason for this is the principal’s lack of ability to fully commit to predetermined contractual terms after new information is revealed by the agent, which is a reasonable assumption that we make. This lack of commitment ability by the principal makes the agent wary of and strate-
gically withholding revelation of sensitive information, which could be used by the principal against the agent’s interest in the future - the ratchet effect. With the information transmission from the agent to the principal being very limited (to avoid ratcheting), pooling contract arises and there is little incentive for the principal to alter the terms over time.

Our model of multiple-prior belief system is built upon the axiomatic foundation laid by GMMS (2010), which, in turn, is a synthesis of two main strands of multiple-prior (axiomatic decision-theoretic) models following the pioneering works of Bewley (1986, 2002) and Gilboa and Schmeidler (1989). In comparison with the axiomatic foundation of the subjective probability theory\(^6\) (notably, the popular version by Anscombe and Aumann (1963)), GMMS (2010) propose less restrictive sets of axioms for a pair of rational preference relations (representing objective rationality and subjective rationality respectively). Based on these more reasonable sets of rationality axioms, GMMS (2010) prove that the degree of rational belief of a decision maker, as revealed by hypothetical betting behavior, is represented by a set of (multiple) prior distributions (hence ambiguity), as opposed to the Savage-Bayesian unique-prior belief system. Since the multiple prior belief system makes the subjective expected utility theory unworkable, GMMS (2010) also provide an axiomatic foundation for the minimum expected utility theory, which is a substitute for the subjective expected utility theory.

Our model of learning under ambiguity is closely related to the seminal work of Epstein and Schneider (2007) that introduces the idea of using likelihood ratio test as a procedure for reevaluating prior distributions. Our paper adopts the likelihood inference as a generic replacement of Bayesian inference to avoid the ad hoc choice of unique prior distribution. We show that for learning under ambiguity, Bayesian updating (through application of Bayes’ Theorem) is duplicated by the iterative selection of priors (through maximum likelihood test). We also show, in our particular game-theoretic setting, the choice of critical value for the likelihood ratio test has no effect on inference. While Epstein and Schneider (2007) model learning about a memoryless (data generating) mechanism, we model a data generating mechanism that is history dependent. Therefore the two models are complementary. The multiple-prior likelihood inference we adopt can be seen as a novel synthesis of ideas from the three competing philosophies\(^7\) of statistics: Bayesian (e.g., Bayes’ Theorem), frequentist (e.g., hypothesis testing theory) and Fisherian (e.g., likelihood and sufficiency). A central theme of this new synthesis is the realization that a single prior (or posterior) probability distribution is not a sufficient statistic to summarize existent information\(^8\), while multiple prior distributions plus likelihoods more adequately represent knowledge (and ignorance\(^9\)).

---

\(^6\)The subjective probability theory was pioneered by Ramsey (1926), de Finetti (1937) and Savage (1954).

\(^7\)According to Efron (1998): “the development of modern statistical theory has been a three-sided tug of war between the Bayesian, frequentist and Fisherian viewpoints”. “In many ways the Bayesian and frequentist philosophies stand at opposite poles from each other, with Fisher’s ideas being somewhat of a compromise”. “The world of applied statistics seems to need an effective compromise between Bayesian and frequentist ideas.” While the “Fisherian synthesis” is expected to continue to do very well in the 21th century, a “new synthesis” may also emerge.

\(^8\)This statement is clearly Fisherian in spirit (see the quote from Fisher (1925).)

\(^9\)To adequately represent knowledge, it is important to indicate what is known, but also what is not
The ratchet effect is well-documented in the dynamic mechanism design (with limited commitment) literature (see Bolton and Dewatripont (2005) for a comprehensive survey). The fundamental approach underpinning this literature is the conventional Bayesian game framework originated by Harsanyi (1967, 1968). Perfect Bayesian Equilibrium (PBE) is the foundation of the characterization of the optimal mechanism. The majority of models in this literature has very simple dynamics – typically modelling only two periods. As a result, the concepts of Markov strategy and Markov equilibrium and the dynamic programming formulation of the problem are not explicitly used in the simplistic models. In general, the concept of Markov equilibrium, e.g., Markov Perfect Bayesian Equilibrium, is not well established in the Markov literature for dynamic games with incomplete information, even though, its counterpart for games with complete information – Markov Perfect Equilibrium – has been well established (see Maskin and Tirole (2001)). In our characterization of POE of the delegated experimentation model, we identify an (interactive epistemology-related) state variable and formalize its characteristic ratchet property. On any equilibrium path this state variable coincides with, but off equilibrium path departs from the (sufficient\(^\text{10}\)) statistic about the belief of the uninformed player (the principal). Due to the ratchet mechanism in the Markov strategies (similar to the renown trigger mechanism in the equilibrium strategies of the infinitely repeated games), a bootstrap (self-enforcing/self-sufficiency) property emerges in the MPOE. For our delegated experimentation model, this bootstrap property turns out to be so powerful that the optimality of the players’ equilibrium strategies have little reliance on the players’ beliefs about opponents’ private information – they virtually become “belief-free” in the sense how the solution concepts of belief-free games (with incomplete information) and belief-free equilibrium are defined (see Bergemann and Morris (2007); Hörner and Lovo (2009), Hörner, Lovo and Tomala (2011))). Our model of dynamic game with incomplete information and ambiguous beliefs and our solution concepts – POE and MPOE – share the same goal with the recent theoretical literature on belief-free games and belief-free equilibrium, which is the pursuit of equilibrium outcome that is robust to specific restrictions on beliefs. Against this goal, in the setting of the delegated experimentation model, the MPOE has performed exceedingly well – the unique MPOE outcome is also the unique belief-free equilibrium outcome. The procedure to solve for the MPOE also provides a practical procedure to identify the belief-free equilibrium. In contrast, there exist multiplicity of MPBE outcomes for our model, although one of them coincides with the belief-free equilibrium outcome, there are a continuum of other equilibrium outcomes that don’t. The concept of MPOE does appear to be very useful for filtering out the equilibria which can be rationalized by the concept of MPBE but nevertheless rely on arbitrary restrictions on the specification of beliefs – subjective pretence of knowledge – and therefore are not robust to such ad hocery.

Our paper is related to, but quite different from, Hörner and Samuelson’s (2009) analysis of incentives for experimenting agents. Like this paper, they consider repeated interaction known (see the quote from Analects of Confucius). It is now well known (see Edwards 1992, pp. 57-61) that a single probability distribution cannot accurately represent complete ignorance.

\(^{10}\)Sufficiency is defined in relation to the relevance to payoffs of the players.
between a principal and an agent to whom experimentation has been delegated. However, the emphasis in their paper is on the implications of hidden action by the agent in a setting in which the principal has a prior belief about the viability of the experiment. In this paper, the agent’s action is observable and contractible, and we emphasize the implications of ambiguity and ambiguity-aversion on the part of the principal.

The paper is organized in seven sections. Section 2 describes the model of delegated experimentation that forms the basis of our analysis. Section 3 explains formally what we mean by ambiguity, and lays out the basic structure of the game we analyze. Section 4 defines the Perfect Objectivist Equilibrium and lays the foundations for characterizing it in our model. Section 5 proceeds through a set of steps to characterize Markov Perfect Objectivist Equilibrium outcome for our model. We first characterize and establish the existence of a pooling Markov perfect objectivist equilibrium. We then show that this pooling equilibrium is Pareto efficient (allowing for compensation transfer). Finally we establish the uniqueness of Markov equilibrium outcome. Section 6 provides an extended discussion on the relationship with a variety of neighboring literature, including dynamic mechanism design and ratchet effect, Markov equilibrium, belief-free equilibrium, learning under ambiguity, and decision rules under ambiguity. Section 7 summarizes and concludes. Proofs of all results are in the Appendix.

2 Basic Model Formulation

We use the exponential bandit model of Keller, Rady, and Cripps (2005) to incorporate ambiguity into a dynamic game-theoretic setting with incomplete information. We assume that a principal hires an expert agent to solve a problem which may or may not be solvable. We refer to solving the problem as a “breakthrough.” If the agent achieves a breakthrough, the principal receives a “prize” equal to \( P > 0 \). The agent also receives a prize equal to \( A > 0 \).

Neither the principal nor the agent knows for certain that the problem is solvable. Specifically, let \( \omega \) be a binary variable, where \( \omega = 0 \) means a breakthrough can never occur, and \( \omega = 1 \) means a breakthrough is possible. The realization of \( \omega \) is unknown to both the principal and the agent unless a breakthrough occurs. However, the agent has expertise, grounded in his past experience and knowledge of scientific facts, that enables him to form a prior probability \( p_0 \equiv \Pr(\omega = 1) \in [0,1] \) that a solution exists. The experience and scientific facts that underpin \( p_0 \) are unknown to the principal and are thus private information to the agent. Therefore, \( p_0 \) can be interpreted as the agent’s unobservable type. In contrast to the conventional Bayesian approach in which the principal would have a given prior belief about \( p_0 \) (“a prior over the prior”), we assume that the principal in general has multiple prior beliefs over possible values of \( p_0 \), an assumption which has important implications for our analysis to be discussed below. The multiple prior assumption allows for the realistic possibility that the principal simply does not know what to believe about \( p_0 \), and it captures, as a special case, the possibility of complete ignorance in which the principal believes that any prior belief about \( p_0 \) is conceivable.
Conditional on the problem having a solution, the time at which the breakthrough occurs is random, but it can be influenced by the intensity of the agent’s investment in problem-solving activities. At each instant \( t \) in continuous time, the agent has at most one unit of a resource it can invest in problem solving. If the level of the agent’s investment is \( k_t \in [0, 1] \), then conditional on the problem being solvable, the hazard rate of a breakthrough is \( k_t \) \( dt \), where \( \lambda > 0 \) is a parameter whose reciprocal \( \frac{1}{\lambda} \) is the expected time to a breakthrough when the agent invests “flat out” \( (k = 1) \). The marginal cost of the resource is \( a > 0 \) per unit, so the agent’s total investment cost is \( C(k_t) = ak_t \).

The agent’s investment at each instant in time is assumed to be observable and contractible, so the contract by which the principal compensates the agent can be conditioned on \( k_t \). In setting up the model and presenting our analysis, we limit attention to linear contracts. Specifically, the payment from the principal to the agent for the period \([t, t+dt]\) is \( t\phi_tak_t \) and \( \phi_t \in [0, 1] \) is the reimbursement rate for the agent’s investment costs. For any posterior probability \( p_t \) that the problem is solvable, the agent’s (rate of flow of) instantaneous utility is thus

\[
v_t = \lambda k_t p_t \Pi_A - a (1 - \phi_t) k_t, \tag{1}\]

while the principal’s instantaneous utility is

\[
u_t = \lambda k_t p_t \Pi_P - \phi_t ak_t. \tag{2}\]

Throughout we maintain:

**Assumption 1** \( \lambda (\Pi_P + \Pi_A) \geq a. \)

Assumption 1 implies that the benefit-cost ratio for a problem known to be solvable is at least equal to 1.

As the agent invests over time and a breakthrough does not occur, the agent updates his posterior belief about the likelihood that a solution to the problem exists. To illustrate this updating process, let \( K_t = \int_0^t k_t \, d\tau \) denote the agent’s cumulative investment through time \( t \), and let the posterior belief as a function of cumulative investment be given by:

\[ p_t(K_t; p_0) = \Pr(\omega = 1|\text{no breakthrough by time } t). \]

The law of motion for \( p_t(K_t; p_0) \) is determined by Bayes’ Rule (applied to a Poisson process), and given by

\[
p_{t+dt}(K_{t+dt}; p_0) = \frac{p_t(K_t; p_0) e^{-\lambda k_t dt}}{(1 - p_t(K_t; p_0)) + p_t(K_t; p_0) e^{-\lambda k_t dt}},
\]

where \( e^{-\lambda k_t dt} = \Pr(\text{no breakthrough in } (t, t+dt) \mid \text{no breakthrough by } t, \omega = 1) \). The law of motion can be expressed as

\[
\frac{dp_t(K_t; p_0)}{dt} = \lim_{dt \to 0} \frac{p_t(K_t; p_0)(1-p_t(K_t; p_0))(e^{-\lambda k_t dt} - 1)}{(1 - p_t(K_t; p_0)) + p_t(K_t; p_0)e^{-\lambda k_t dt}} \]

\[ = -\lambda k_t p_t(K_t; p_0) (1 - p_t(K_t; p_0))^2. \]
which has the following closed-form solution:

\[ p_t (K_t; p_0) = \frac{1}{1 + \frac{1-p_0}{p_0} e^{\lambda K_t}}. \]

In the special case of constant investment \( k_\tau = 1 \) for \( \tau \in [0, t] \), \( K_t = t \), and \( p_t (t; p_0) = \frac{1}{1 + \frac{1-p_0}{p_0} e^{\lambda K_t}}. \) Interestingly, this Poisson process with uncertain binary hazard rates is equivalent to a Poisson process with a certain time-varying hazard rate given by \( \lambda k_t p_t (K_t; p_0) \).

### 2.1 Limiting Case: Investment without Contract

To shed light on the incentive issues in the model, we discuss the limiting case of investment in the absence of a contract between the principal and the agent. In this case, the principal would prefer that the agent invest no matter since solving the problem is a free option for the principal. By contrast, the agent’s most preferred investment profile \( k^A (p_t) \) is given by the following proposition.

**Proposition 1** In the absence of a contract between the principal and the agent, the agent’s most preferred investment profile is given by

\[
k^A (p_t) = \begin{cases} 
1 & \text{if } p_t \in (p^A, 1] \\
0 & \text{if } p_t \in [0, p^A] 
\end{cases},
\]

where \( p^A \) is given by

\[
p^A = \min \left\{ \frac{\alpha}{\lambda \Pi A}, 1 \right\} > 0.
\]

The agent’s most preferred investment profile specifies maximum investment intensity as long as the posterior belief exceeds a belief cutoff, and no investment otherwise. However, because \( p^A > 0 \) there will be a range of posteriors over which the agent invests less than the principal prefers.

This incentive problem can also be seen by analyzing the largest cumulative investment preferred by the agent, i.e., the total investment that would be accumulated if no breakthrough occurs until the agent reaches its belief cutoff \( p^A \). The largest cumulative investment level preferred by the agent, \( K^A (p_0) \), must satisfy:

\[
p^A = \frac{1}{1 + \frac{1-p_0}{p_0} e^{\lambda K^A (p_0)}},
\]

To see this, let \( NB = \text{“no breakthrough”} \), then

\[
\Pr (NB \text{ in } [t, t+dt] \text{ by time } t) = \frac{\Pr (NB \text{ by time } t+dt)}{\Pr (NB \text{ by time } t)} = \frac{p_0 e^{-\lambda K^A +dt} (1-p_0)}{p_0 e^{-\lambda K^A} + (1-p_0)}
\]

\[
= \frac{p_0 (1-p_0) e^{\lambda K^A +dt} e^{-\lambda K^A dt}}{p_0 + (1-p_0) e^{\lambda K^A}} e^{-\lambda K^A dt} = \frac{p_0}{p_{t+dt}} e^{-\lambda K^A dt} = e^{-\lambda \ln p_t} e^{-\lambda K^A dt} = e^{-\lambda p_t K^A dt}.
\]

The last equality uses \( d \ln p_t = -\lambda k_t (1-p_t) dt \) which follows from the law of motion \( \frac{dp_t}{dt} = -\lambda k_t p_t (1-p_t) \).
which implies

\[
K^A(p_0) = \frac{1}{\lambda} \ln \left( \frac{(1-p_0^A)}{p_0^A (1-p_0)} \right).
\]

(4)

Note that this is increasing in \(p_0\), so the more optimistic is the agent’s prior, the greater will be the agent’s optimal largest cumulative investment.

Because the principal prefers that the agent continue to invest indefinitely, the cumulative investment most preferred by the principal is effectively unbounded. The unwillingness of the agent to continue investing indefinitely highlights the need of a contract for the two players to achieve mutually favorable trade.

3 Ambiguity in a Dynamic (Differential) Game of Incomplete Information

3.1 Learning under Ambiguity

3.1.1 Observable Actions

We assume that the reimbursement rate \(\phi_t\) is determined through a series of short-term contracts. Thus, in the period \([t - dt, t)\), the principal announces the reimbursement rate \(\phi_t\), which is legally binding for the period \([t, t + dt)\). Because there is a time gap between the principal’s offer \(\phi_t\) and the agent investment action \(k_t\), these moves are sequential and therefore \(k_t\) can depend on \(\phi_t\).

To account for the beginning of the game, let \(0_{-2}\) and \(0_{-1}\) denote almost the same calendar time as \(t = 0\), but with the logical sequence: \(0_{-2} < 0_{-1} < 0\). At time \(t = 0_{-2}\) the principal offers a menu of initial contracts \(c \in \mathcal{C} = \{\Phi|\Phi(\hat{p}_0) \in [0, 1], \hat{p}_0 \in [0, 1]\}\), where \(\Phi(\hat{p}_0) = \phi_0\) and \(\hat{p}_0\) is the agent’s self-declared lower bound of type. An initial contract is a menu of short-term contracts in the form of \(\phi_0\). Given any menu \(\Phi \in \mathcal{C}\), at time \(t = 0_{-1}\), the agent of type \(p_0\) declares the lower bound of his type to be \(\hat{p}_0 \in [0, 1]\). There is a default lower bound which is \(\hat{p}_0 = 0\) if the agent does not declare otherwise. Given the realized pair \((c, \hat{p}_0) \in \mathcal{C} \times [0, 1]\), an initial contract \(\phi_0 \in [0, 1]\) is uniquely determined by \(c(\hat{p}_0)\).

Denote by \(\mathcal{H}^t\) the set of all possible histories for the period \(\{0_{-2}, 0_{-1}\} \cup [0, t]\), defined by

\[
\mathcal{H}^t = \{ h^t|h^t = (c, \hat{p}_0, \phi^t, k^t), c \in \mathcal{C}, \hat{p}_0 \in [0, 1], \phi^t = \{\phi_t|\tau \in [0, t]\}, k^t = \{k^t|\tau' \in [0, t]\}\}
\]

for \(t \in [0, \infty)\).\(^{12}\) In addition, define the histories \(\mathcal{H}^{0_{-2}}\) and \(\mathcal{H}^{0_{-1}}\) such that \(\mathcal{H}^{0_{-2}} = \{h^{0_{-2}}|h^{0_{-2}} = \emptyset\}\) and \(\mathcal{H}^{0_{-1}} = \{h^{0_{-1}}|h^{0_{-1}} = c \in \mathcal{C}\}\). The realized history of observable actions, \(h^t \in \mathcal{H}^t\) for \(t \in \{0_{-2}, 0_{-1}\} \cup [0, \infty)\), which includes information of the initial contract, is public information.

\(^{12}\) Throughout the paper, we use subscript \(t\) to refer to actions or contract terms at a given point in time \(t\), while we use superscript \(t\) to refer to histories of actions and contract terms up to time \(t\).
The principal’s action space $G_t$ for time $t \geq 0$ is given by $G_t = \{\phi_{t+dt} | \phi_{t+dt} \in [0,1]\}$. Let $G_{0-2} = \{g_{0-2} | g_{0-2} = c \in C\}$ and $G_{0-1} = \emptyset$, i.e., the principal does not move at $t = 0-1$. The agent’s action space $X_t$ for time $t \geq 0$ is given by $X_t = K_t = \{k_t | k_t \in [0,1]\}$. Let $X_{0-2} = \emptyset$ and $X_{0-1} = \{x_{0-1} | x_{0-1} = \hat{p}_0 \in [0,1]\}$. Thus, prior to $t = 0$, the principal and the agent move sequentially and the agent observes the principal action before he moves.

Define a pure strategy for the principal, denoted by $s_P$, as a map $s_P : H^t \rightarrow G_t$, $\forall t \in \{0-2, 0-1\} \cup [0,\infty)$. Define a pure strategy for the agent, denoted by $s_A$, as a map $s_A : H^t \rightarrow X_t$, $\forall t \in \{0-2, 0-1\} \cup [0,\infty)$. Denote by $S_P$ (respectively, $S_A$) the set of all possible pure strategies of the principal (respectively, agent). Thus, $s_P \in S_P$ and $s_A \in S_A$. In keeping with the standard assumptions in the differential games literature, we impose the following regularity condition.\(^\text{1}\)

**Assumption 2** $S_P$ and $S_A$ are such that $s_P (h^t)$ and $s_A (h^t, p_0)$ are continuous and differentiable with respect to $t$ almost everywhere.

Note that the history $h^t \in H^t$ for $t \geq 0$ contains information on $\phi_t$, which is the contract for period $[t, t + dt]$. In contrast, the history $h^t$ does not contain information on $k_t$. This asymmetry reflects the sequential relation between $\phi_t$ and $k_t$. The choice of $k_t$ can depend on $\phi_t$. The converse does not hold, i.e., the choice of $\phi_t$ cannot depend on $k_t$. So the principal at any point of time is a first mover or leader, and agent is a follower. The principal has the ability to commit (e.g., she can commit to $\phi_{t+dt}$ at time $t$), but the agent cannot commit to anything apart from playing the best responses. This asymmetry between the principal and the agent in the sequence of moves and commitment ability gives the principal a first mover advantage. The agent does have one potential advantage, though, which is the asymmetry of information.

A pure strategy profile $s = (s_P, s_A)$ is a map

$$s : \{(h^t, p_0) | h^t \in H^t, t \in \{0-2, 0-1\} \cup [0,\infty), p_0 \in [0,1]\} \rightarrow G_t \times X_t,$$

such that $s_P (h^t) \in G_t$, $s_A (h^t, p_0) \in X_t$ for all $t \in [0,\infty)$. Let $S$ denote the set of all almost everywhere continuous and differentiable pure strategy profiles, thus, $s \in S$. Note, the principal’s pure strategy $s_P (h^t)$ does not depend on parameter $p_0$, which is not observable to her. In contrast, the agent’s strategy is a function of $p_0$. At $t = 0-1$ the agent’s strategy defines a revelation function $R \in R = \{\hat{R} | \hat{R} (p_0; c) \in [0,1], c \in C\}$, such that $s_A (h^{0-1}, p_0) = R (p_0; c) \in X_{0-1}$ for all $p_0 \in [0,1]$. If

$$R (p_0; c) \leq p_0 \text{ for all } p_0 \in [0,1],$$

then the revelation is said to be truthful based on the given $c \in C$. If

$$R (p_0; c) = p_0 \text{ for all } p_0 \in [0,1],$$

\(^{13}\)See Fudenberg and Tirole (1991), Chapter 13.
then the revelation is said to be truthful and complete based on the given $c \in C$.

The learnable parameter space is $\Theta = \{ \theta = (\omega, p_0) \, | \, (\omega, p_0) \in [0, 1] \times [0, 1] \}$. The principal's learning about $\omega$ depends on observing a breakthrough. Her belief about $p_0$ is inferred from observational data on history $h^t$. If $\omega = 1$ and is revealed by a breakthrough, the game ends and the data generating process stops without fully resolving the ambiguity about $p_0$.

As noted earlier, the principal initially lacks information about $p_0$ and therefore holds multiple priors about it. She can learn about $p_0$ by observing the agent's action history from $h^t$.

The principal can logically infer the range of $p_0$ from the data of history $h^t$. The inference rule involves a judgement about the equilibrium of the game. In principle, there could be multiple inference rules if there were multiple equilibrium outcomes, or the data generating process involved off-equilibrium-path play. Although both multiplicity of priors and multiplicity of inference rules can cause ambiguity, in the current paper, the emphasis is on multiple priors. The justification for this emphasis is due to the assumption that the data-generating mechanism is the play of the game along the equilibrium paths of a commonly known pure strategy equilibrium, which determines the likelihood function uniquely.

### 3.1.2 The Principal's Priors Over $p_0$

We now formulate the principal's multiple priors over $p_0$. Each prior over $p_0$ is a function $\mu_0 : \Sigma \to [0, 1]$, where $\Sigma$ is the Borel $\sigma$-algebra of subsets of $[0, 1]$. The set of all priors of the principal denoted by $\mathcal{M}_0$, is defined as $\mathcal{M}_0 = \{ \mu_0 | \mu_0 : \Sigma \to [0, 1] \}$.

The principal's knowledge about the process determining realized histories of the game is summarized by a set of likelihood functions

$$\mathcal{L} = \{ l : \mathcal{H}^t \times \{ t \} \times S \times \{ p_0 | p_0 \in [0, 1] \} \to [0, 1] \, | \, t \in \{ 0\_2, 0\_1 \} \cup [0, \infty) \} ,$$

where $l(\cdot; t, s, p_0)$ is a conditional probability measure\(^{14}\) on $\mathcal{H}^t$. For example, $l(h^t; t, s, p_0)$ is the probability of history $h^t$ conditional on $(t, s, p_0)$. Each duple $(\mu_0, l) \in \mathcal{M}_0 \times \mathcal{L}$ represents a theory, where $l$ reflects the (structural) model specification and $\mu_0$ is a distribution on the value of the unknown parameter $p_0$. If $\mu_0$ is a Dirac measure (i.e., an indicator function) given by

$$\delta_{p_0}(A) = \begin{cases} 
1 & \text{if } p_0 \in A \\
0 & \text{otherwise} \end{cases} , \forall A \subseteq [0, 1] ,$$

then $\mu_0 = \delta_{p_0}$ represents a unique specification of the parameter value.

A likelihood function has the argument $s$, which is a strategy profile. The likelihood is plausible only if $s$ is plausible. In principle, the likelihood could reflect both "signal" (which is the equilibrium of the game) and "noise" (which could include measurement error or out-of-equilibrium play). In the current study, we abstract from all noise components, and as a result, $\mathcal{L}$ is a singleton and thus $l$ is unique for any given $(t, s, p_0)$. Multiplicity of

\(^{14}\)The likelihood $l(\cdot; t, s, p_0)$ is expressed as a probability over the observation space $\mathcal{H}^t$; it is NOT a probability over the parameter space $[0, 1]$. Thus, $l(h^t; t, s, p_0)$ does not obey the laws of probability when $h^t$ is fixed and $p_0$ is treated as a variable, e.g., the integration over $p_0$ does not add up to one. That is why we use the term "likelihood" as opposed to "probability."
likelihood only arises from multiplicity of \( s \in S \). If the data generating mechanism is a pure strategy equilibrium, i.e., \( s = (s_P, s_A) \), then \( s \) determines an equilibrium path for each given \( p_0 \in [0, 1] \); thus, \( s \) determines a unique conditional likelihood \( l(\cdot; s, \cdot) \). Thereby only a unique conditional likelihood function \( l(\cdot; s, \cdot) \) passes the likelihood ratio test (which will be specified presently) at each point of time.

Denote by \( \mathcal{L}_0(s) \) the restriction of \( \mathcal{L} \) to \( s \). \( \mathcal{L}_0(s) \) is a singleton of conditional likelihood, and \( l(\cdot; s, \cdot) \in \mathcal{L}_0(s) \) represents the unique pure strategy equilibrium that the players coordinate to play.

### 3.1.3 Likelihood Inference — Iterative Selection of Priors and Updating

At each point of time, each theory duple \((\mu_0, l) \in \mathcal{M}_0 \times \mathcal{L}_0(s) \) competes with other members to better explain the history data \( h^t \). In the absence of a breakthrough and the agent’s voluntary declaration, the principal forms a set of posteriors against each realized history \( h^t \). She (iteratively) selects the subset of prior distributions, each of which has to exceed a critical value of likelihood of generating the data set \( h^t \). Denote by \( \mu_t(\cdot|h^t; \mu_0, l) \) the conditional probability measure, which is calculated by Bayes’ Rule,

\[
d\mu_t(\cdot|h^t; \mu_0, l) = \frac{l(h^t; t, s, \cdot) d\mu_0(\cdot)}{\int_0^1 l(h^t; t, s, \bar{p}_0) d\mu_0(\bar{p}_0)}, \tag{5}
\]

where

\[
L(h^t; t, s, \mu_0, l) = \int_0^1 l(h^t; t, s, \tilde{p}_0) d\mu_0(\tilde{p}_0) \tag{6}
\]

is the (generalized) likelihood (over distribution \( \mu_0 \) as opposed to over parameter \( p_0 \)). It follows that

\[
\mu_t(A|h^t; \mu_0, l) = \frac{\int_{p_0 \in A} l(h^t; t, s, p_0) d\mu_0(p_0)}{L(h^t; t, s, \mu_0, l)} \tag{7}
\]

for all \( A \in \Sigma \). Note, the (variable) argument of the functional \( L(h^t; t, s, \mu_0, l) \) is \( \mu_0 \), while \( h^t \) is fixed. In general, a hypothesis is expressed as a distribution \( \mu_0 \) instead of a single parameter value \( p_0 \). In the special case that \( \mu_0(\tilde{p}_0) \) degenerates to a Dirac measure, i.e., \( \mu_0(\tilde{p}_0) = \delta_{p_0}(\tilde{p}_0) \), then

\[
L(h^t; t, s, \mu_0, l) = \int_0^1 l(h^t; t, s, \tilde{p}_0) d\delta_{p_0}(\tilde{p}_0) = l(h^t; t, s, p_0).
\]

From (7) and (6) it is apparent that conditional probability measure \( \mu_t(\cdot|h^t; \mu_0, l) \) is well defined only if \( L(h^t; t, s, \mu_0, l) > 0 \); i.e., the theory duple \((\mu_0, l) \) is not contradicted by the observed history \( h^t \). In principle, there can be a continuum of competing theories (hypotheses) that are candidate explanations for the observed history \( h^t \). Likelihood inference can play a useful role to quantitatively assess the relative merits of these competing hypotheses.\(^{15}\) Fixing the history \( h^t \) and underlying strategy profile \( s \), the likelihood ratio between

---

\(^{15}\)For an authoritative account on likelihood inference, see Edwards (1992).
hypotheses \((\mu_0, l)\) and \((\bar{\mu}_0, \hat{l})\) is given by

\[
\frac{L(h^t; t, s, \mu_0, l)}{L(h^t; t, s, \bar{\mu}_0, \hat{l})}.
\]

We say that \((\mu_0, l)\) has more support from the data than \((\bar{\mu}_0, \hat{l})\) if this ratio exceeds 1. Based on pair-wise likelihood ratios, the principal can discriminate among all priors in \(\mathcal{M}_0\). To formulate the likelihood inference procedure which is used by the principal, let \(\alpha \in [0, 1]\) be the critical value of the likelihood ratio test the principal uses. Let the set of all accepted posteriors against \(h^t\) be given by

\[
\mathcal{M}_t^\alpha \left( h^t \right) = \left\{ \mu_t \left( h^t; \mu_0, l \right) \mid \mu_0 \in \mathcal{M}_0, l \in \mathcal{L}_0(s) \right\}, \\
\mathcal{M}_t^{\text{FBU}} \left( h^t \right) = \left\{ \mu_t \left( h^t; \mu_0, l \right) \mid \mu_0 \in \mathcal{M}_0, l \in \mathcal{L}_0(s) \right\}, \\
\mathcal{M}_t^{\text{MLU}} \left( h^t \right) = \left\{ \mu_t \left( h^t; \mu_0, l \right) \mid \mu_0 \in \mathcal{M}_0, l \in \mathcal{L}_0(s), \\
L(h^t; t, s, \mu_0, l) = \max_{\bar{\mu}_0 \in \mathcal{M}_0} L(h^t; t, s, \bar{\mu}_0, \hat{l}) \right\}. 
\]

Intuitively, the principal admits a posterior probability measure \(\mu_t \left( h^t; \mu_0, l \right)\) if and only if \(\mu_t \left( h^t; \mu_0, l \right)\) is a Bayesian update of \(\mu_0 \in \mathcal{M}_0\) such that there exists \(l \in \mathcal{L}_0(s)\) and the likelihood of \((\mu_0, l)\), given by \(L(h^t; t, s, \mu_0, l)\), is at least a fraction \(\alpha\) of the possible maximum likelihood over all \((\bar{\mu}_0, \hat{l}) \in \mathcal{M}_0 \times \mathcal{L}_0(s)\). Note that the special case of \(\alpha = 0\) corresponds to Full Bayesian Updating (FBU):

\[
\mathcal{M}_t^{\text{FBU}} \left( h^t \right) = \left\{ \mu_t \left( h^t; \mu_0, l \right) \mid \mu_0 \in \mathcal{M}_0, l \in \mathcal{L}_0(s) \right\}.
\]

By contrast, the special case of \(\alpha = 1\) corresponds to Maximum Likelihood Updating (MLU)

\[
\mathcal{M}_t^{\text{MLU}} \left( h^t \right) = \left\{ \mu_t \left( h^t; \mu_0, l \right) \mid \mu_0 \in \mathcal{M}_0, l \in \mathcal{L}_0(s), \right. \\
L(h^t; t, s, \mu_0, l) = \max_{\bar{\mu}_0 \in \mathcal{M}_0} L(h^t; t, s, \bar{\mu}_0, \hat{l}) \left. \right\}. 
\]

In principle the choice of critical value \(\alpha \in [0, 1]\) may potentially introduce an ad hoc subjective element into the inference rule. This turns out not to be the case in our current game-theoretic context, since the inference is entirely deductive. As has been established previously, for \(\mu_t \left( h^t; \mu_0, l \right)\) to be well defined, \(L(h^t; t, s, \mu_0, l) > 0\) is necessary. Consequently, for \(\mathcal{M}_t^\alpha \left( h^t \right)\) to be well defined, \(\max_{\bar{\mu}_0 \in \mathcal{M}_0} L(h^t; t, s, \bar{\mu}_0, \hat{l}) > 0\) is necessary, that is, there must exist some \((\bar{\mu}_0, \hat{l}) \in \mathcal{M}_0 \times \mathcal{L}_0(s)\) that is not contradicted by the data \(h^t\).

We are interested in learning through inferences that are based on some notion of pure strategy equilibrium of the game. For a given pure strategy equilibrium, a unique likelihood

---

\^16 The terms FBU and MLU are borrowed from Gilboa and Marinacci (2011), which is the most recent comprehensive survey of the ambiguity literature.
function can be established because a unique equilibrium path is determined for each realization of the state variable \( p_0 \). That is, there exists a (forecast) map \( f : S \times \{ p_0 \} \rightarrow \mathcal{H}^\infty \) such that \( \forall (s, p_0) \in S \times \{ p_0 \}, f (s, p_0) = h^\infty \in \mathcal{H}^\infty \) predicts the entire equilibrium path of the game (conditional on the absence of a breakthrough). Let \( f^t (s, p_0) \) be the restriction of \( f (s, p_0) \) to the set \( \{ 0_{-2}, 0_{-1} \} \cup [0, t] \). Therefore, \( f^t (s, p_0) \) predicts the on-equilibrium-path history for the period \( \{ 0_{-2}, 0_{-1} \} \cup [0, t] \).

Define the (partial) identification correspondence, denoted by \( I_d \), such that
\[
I_d (h^t) = \{ p_0 \in [0, 1] | f^t (s, p_0) = h^t \}.
\] (11)

The correspondence \( I_d : \mathcal{H}^t \rightarrow \{ p_0 \} \) maps any given on-equilibrium-path history to a subset of agent types. Since agent type subset is rarely singleton, the identification is usually partial. The next theorem then states the updating of the principal’s belief through pure strategy equilibrium (strategy) profile-based likelihood inference.

**Theorem 1** Suppose \( s \) is a pure strategy equilibrium (strategy) profile. Then \( \mathcal{L}_0 (s) = \{ l \} \) is a singleton such that
\[
l (h^t; t, s, p_0) = \delta_{f^t(s,p_0)} (h^t) = \begin{cases} 
1 & \text{if } h^t = f^t (s, p_0) \\
0 & \text{otherwise}
\end{cases},
\]
where \( \delta_{f^t(s,p_0)} (h^t) \) is a Dirac measure. For all on-equilibrium-path history \( h^t \), for all \( \alpha \in [0, 1] \),
\[
\mathcal{M}^\alpha_t (h^t) = \mathcal{M}_t^{MLU} (h^t) = \{ \mu_0 \in \mathcal{M}_0 | \mu_0 (I_d (h^t)) = 1 \}.
\] (12)

It is interesting to note, by (8) \( \mathcal{M}^\alpha_t (h^t) \) is defined as a set of posterior distributions, which are the outcomes of applying Bayes’ Theorem. According to (12), \( \mathcal{M}^\alpha_t (h^t) \) appears to be a set of prior distributions, which is the subset of \( \mathcal{M}_0 \) that passes certain (iterative) selection. This comparison reveals the fact that the effect of Bayesian updating on beliefs is duplicated by the effect of selecting prior distributions by the maximum likelihood test. In this result, we see a seamless synthesis of the ideas from the Bayesian, frequentist and Fisherian schools of statistics, under the condition that the set of ambiguous priors \( \mathcal{M}_0 \) is sufficiently large (inclusive).

The fact that the endogenous critical value of the likelihood ratio test can be set to \( \alpha = 1 \) is because in a pure strategy equilibrium, in the absence of measurement error, the principal only makes deductive inferences which rule out all possibilities that are contradicted by the observational evidence, but do not rule out any possibility that is consistent with such evidence.\(^{18}\) In the language of classical hypothesis-testing theory, the probability of type I

---

\(^{17}\) The term partial identification is borrowed from Manski (1995). In the current context, the pure strategy equilibrium (strategy) profile-based inferential problem is abstracted from statistical inference problem, therefore is purely a problem of identification \( a la \) Manski (1995).

\(^{18}\) Since the inferential problem pertinent to our model is confined to identification problem, it turns out that deductive inference is adequate for the task. Therefore our approach is in spirit close to Keynes’s logical approach to probability (as a measure of rational beliefs), which is much concerned with the legitimacy of inference (see Keynes 1921).
error is set to be zero. As a matter of fact the updating from \( M_0 \) to \( M_t(h_t) \) is essentially a truncation of the support of \( M_0 \) (which is \([0, 1]\)) to \( \mathcal{I}_d(h_t) \). \( M_t(h_t) \) admits all probability measures whose supports lie within \( \mathcal{I}_d(h_t) \). When \( \alpha = 1 \) is chosen, the probability of type II error is uniformly minimized but still positive. The fact (as revealed in the proof of Theorem 1) that \( [0, 1] \) is sufficient for representing deductive inference implies that inferential correctness is not sensitive to the choice of critical value \( \alpha \). Without loss of generality, in the remainder of the paper, we confine our analysis to the case in which \( \alpha = 1 \). And reference to the value of \( \alpha \) is omitted. We therefore adopt the simple notation \( M_t(h_t) \) such that
\[
M_t(h_t) = M_t^{MLU}(h_t) = M_t^1(h_t).
\]

So far we have only dealt with on-equilibrium-path histories. For any off-equilibrium-path history \( h_t \), by definition we must have \( \mathcal{I}_d(h_t) = \emptyset \), which implies \( \mu_0(\mathcal{I}_d(h_t)) = 0 \) for all \( \mu_0 \in M_0 \). That is, the pure strategy equilibrium (strategy) profile-based identification process must fail if the history is off equilibrium path. To complete the updating of beliefs for off equilibrium histories, we propose the following definition.

**Definition 1** If \( l(h_t; t, s, p_0) = 0 \) for all \( \mu_0 \in M_0 \) and \( l \in \mathcal{L}_0(s) \) (i.e., \( \mathcal{I}_d(h_t) = \emptyset \)), then \( M_t(h_t) \) is defined by
\[
M_t(h_t) = M_0.
\]

This definition captures the idea that the identification failure (i.e., \( \mathcal{I}_d(h_t) = \emptyset \)) indicates that the data \( h_t \) is generated by off equilibrium path behavior, which implies that the pure strategy equilibrium (strategy) profile-based likelihood function is not valid for explaining the data. As a result, learning based on the invalid likelihood function should be undone, and the beliefs should reverse to the initial priors. In this case, all competing priors are contradicted by data \( h_t \), and no prior prevails.

It follows from Theorem 1 and Definition 1 that the set of posterior beliefs is a subset of priors, i.e., \( M_t(h_t) \subseteq M_0 \). Therefore the set of all posteriors is essentially a set of all priors that pass the maximum likelihood test. A prior \( \mu_0 \) is rejected by \( h_t \) if and only if \( \mu_0 \in M_0 \backslash M_t(h_t) \). If \( \mu_0 \) is rejected, the posterior for \( \mu_0 \) is not well-defined, therefore discarded.

### 3.1.4 Initial Revelation

In a pure strategy equilibrium, \( s \) is common knowledge to all players. Thus, \( s(h^{t-1}, \cdot) = (c, R) \) are common knowledge. Also, \( h^0 = (c, \hat{p}_0) \) is known by \( t = 0 \). From these, the principal can deduce the set of plausible types of the agent to be
\[
\mathcal{I}_d(h^0) = \{ p_0 \in [0, 1] \mid f^0(s, p_0) = h^0 \} = \{ p_0 \in [0, 1] \mid R(p_0; c) = \hat{p}_0 \}.
\]

---

19 Equivalently, the test is uniformly-most-powerful (U.M.P.).
It follows that,
\[ \mathcal{M}_0(h^0) = \{ \mu_0 \in \mathcal{M}_0 | \mu_0(\mathcal{I}_d(h^0)) = 1 \}, \]
or equivalently, \( P(\mathcal{M}_0(h^0)) = \mathcal{I}_d(h^0) \), where \( P(A) \) denotes the support of the set of measures: \( A \), i.e., \( P(A) = \cup_{\mu \in A} \text{supp}(\mu) \).

We write \( h^0 = (c;R) \) to save space. The initial revelation based on \((c;R)\) is a separating outcome if and only if \( \mathcal{I}_d(h^0) \) is a singleton for all \( p_0 \in [0,1] \). The initial revelation based on \((c;R)\) is a pooling outcome if and only if \( \mathcal{I}_d(h^0) = [0,1] \) for all \( p_0 \in [0,1] \). For example, if no agent type opts out of the default lower bound, \( \hat{p}_0 = 0 \) — i.e., if \( R(p_0;c) = 0 \) for all \( p_0 \in [0,1] \) — then the initial revelation based on \((c,R)\) is a pooling outcome. The initial revelation based on \((c,R)\) is a semi-separating outcome if \( \mathcal{I}_d(h^0) \subset [0,1] \) and \( \mathcal{I}_d(h^0) \) is non-singleton for some \( p_0 \in [0,1] \).

To conclude this section on learning under ambiguity, note that our model of learning about the ambiguous parameter \( p_0 \) can be summarized by the tuple \((\Theta, \mathcal{M}_0, \mathcal{L}(s))\). This has marked similarity with Epstein and Schneider (2007); but the differences warrant some remarks. First, the likelihood function in our model is derived endogenously from the equilibrium of the game. In contrast, the likelihood function in Epstein and Schneider (2007) is exogenously given. Second, as a result, \( \mathcal{L}(s) \) is a singleton for a given pure strategy equilibrium (strategy) profile in our model. Third, in general \( \alpha \) can be interpreted as a preference parameter that reflects the decision maker’s aversion to type II error in hypothesis testing, i.e., \( \alpha = 1 \) means maximum aversion, and \( \alpha = 0 \) means maximum tolerance to type II errors. In principle, a preferred trade off between type I and II errors should be modelled as part of the overall (subjective) attitude towards ambiguity. This concern, however, does not arise in our analysis, since the value of \( \alpha \) (for \( \alpha \in [0,1] \)) does not affect the (deductive) inferential outcome in pure strategy equilibrium in our model, we can simply assign \( \alpha = 1 \) without loss of generality.

3.2 Principal’s Objective and Best Response

In Bayesian games in which a player has a prior belief about the environment, specification of the player’s objective function is straightforward: conditional on the history of play, the player forms an expectation of the function it seeks to maximize based on the posterior beliefs implied by that history. However, if the player does not have a unique prior, the formulation of the player’s objective is more complex. Here we describe the objective for a principal facing ambiguity following approaches inspired by Manski (2008) and GMMS (2010).

These papers deal with the problem of modelling decision making under uncertainty when a decision maker does not have enough information to quantify uncertainty using a single probability measure, which is the situation faced by the principal in this model. Manski (2008) formulates a two-step procedure in which the first step is to eliminate all weakly-dominated actions, and the second step is to maximize (over non-dominated actions) a not-uniquely specified utility function (which could be minimum, minimum-regret or expected utility function). GMMS (2010) axiomatize the problem of a decision maker who
has a pair of preference relations: objectively rational preferences and subjectively rational preferences. If the decision maker chooses based on objectively rational preferences, he can defend his choice to others; if the decision maker chooses based on subjectively rational preferences, he cannot be convinced by others that his choice was wrong. Objectively rational preferences generate a unanimous but incomplete ordering of actions, while subjectively rational preferences generate a complete ordering of actions that can be represented by a minimum expected utility function (with respect to all priors in the set of the decision maker’s possible priors). GMMS (2010) demonstrate that given two plausible conditions (consistency and caution), there exists a common set of priors (that can be justified by a given set of inference rules) that enable the decision maker’s choices to be represented either by objectively rational preferences or by subjectively rational preferences. This provides a foundation for decision making based on the maximum rule; basing decisions on a minimum expected utility function can be thought of as a way of completing an otherwise incomplete preference ordering based on objective rationality.

Our formulation can be seen as an extended application of GMMS (2010) to a dynamic game-theoretic setting. To develop this formulation, we define, derive, and characterize the principal’s minimum expected utility function — which we call the worst-case value function.

To begin, note that the principal’s instantaneous conditional expected utility (conditional on \((p_{t}(K_t;p_0), s, p_0))\) is given by

\[
u_t(p_{t}(K_t;p_0)|s, p_0) = \lambda k_{t+\tau}(p_0) e^{\tau} \int_0^{\tau} e^{-t} dt + e^{-\lambda t} k_{t+\tau} \int_0^{\tau} e^{-t} dt\]

where \(\lambda k_{t+\tau}(p_0)\) is the hazard rate of the Poisson (breakthrough) process.

Given \((p_{t}(K_t;p_0), s, p_0))\) the principal’s conditional value function is given by

\[
W_{t}(p_{t}(K_t;p_0)|s, p_0) = \int_t^{\infty} u_t(p_{t}(K_t;p_0)|s, p_0) e^{-\lambda t} k_{t+\tau} \int_0^{\tau} e^{-t} dt + e^{-\lambda t} k_{t+\tau} \int_0^{\tau} e^{-t} dt W_{t+\tau}(p_{t+\tau}(K_{t+\tau};p_0)|s, p_0),
\]

where the term \(e^{-\lambda t} k_{t+\tau} \int_0^{\tau} e^{-t} dt\) is the probability that no breakthrough occurs in the time interval \([t, \tau]\) for \(\tau \geq t\) conditional on no breakthrough by time \(t\). We can now define the principle’s worst-case value function:

**Definition 2** In the absence of a breakthrough prior to time \(t\), the principal’s worst-case value function \(U_t(h^t; s)\) is determined by the plausible posterior that, for a given pure strategy equilibrium \(s = (s_p, s_A)\), minimizes the principal’s conditional value; i.e.,

\[
U_t(h^t; s) = \min_{\mu_0 \in cl(\mathcal{M}_t(h^t))} \int_0^{\tau} W_{t}(p_{t}(K_t;p_0)|s, p_0) d\mu_0(p_0),
\]

where \(cl(X)\) is the closure of set \(X\), i.e., the smallest closed superset of \(X\), or the set of all limit points of \(X\); the definition of \(cl(\mathcal{M}_t(h^t))\) is based on convergence in distribution (CDF).
We can immediately establish the following equivalences:

**Theorem 2** The principal’s worst-case value function has the following equivalent expressions:

(i) \( U_t(h^t; s) = \min_{\mu_0 \in cl(\mathcal{M}_t(h^t))} \int_0^1 W_t(p_t(K_t; p_0) | s, p_0) d\mu_0(p_0) \);

(ii) \( U_t(h^t; s) = \min_{\hat{s}_0 \in cl(\mathcal{D}_t(h^t))} W_t(p_t(K_t; p_0) | s, p_0) \);

(iii) \( U_t(h^t; s) = \min_{p_0 \in cl(P(\mathcal{M}_t(h^t)))} W_t(p_t(K_t; p_0) | s, p_0) \),

where \( \mathcal{D}_t \) is the largest Dirac subset of \( \mathcal{M}_t(h^t) \), \( P(\mathcal{M}_t(h^t)) \) is the support of \( \mathcal{M}_t(h^t) \) and \( cl(X) \) is the closure of set \( X \); the definition of \( cl(\mathcal{D}_t(h^t)) \) is based on convergence in distribution (CDF).

To develop our notion of equilibrium below, we need to define the principal’s best response conditional on her set of accepted posteriors. We build toward that definition by first presenting:

**Definition 3** The conditional weak dominance relation over \( \mathcal{S}_P \times \mathcal{S}_A \) conditional on \( \mathcal{M}_t(h^t) \), denoted by \( \succ^{*}_{\mathcal{M}_t(h^t)} \), is defined such that, for all \( (s_p, s_A), (\hat{s}_p, \hat{s}_A) \in \mathcal{S}_P \times \mathcal{S}_A \), \( (s_p, s_A) \succ^{*}_{\mathcal{M}_t(h^t)} (\hat{s}_p, \hat{s}_A) \) iff

\[
\int_0^1 W_t(p_t(K_t; p_0) | s_p, s_A, p_0) d\mu_0(p_0) \geq \int_0^1 W_t(p_t(K_t; p_0) | \hat{s}_p, \hat{s}_A, p_0) d\mu_0(p_0),
\]

for all \( \mu_0 \in cl(\mathcal{M}_t(h^t)) \) and

\[
\int_0^1 W_t(p_t(K_t; p_0) | s_p, s_A, p_0) d\mu_0(p_0) > \int_0^1 W_t(p_t(K_t; p_0) | \hat{s}_p, \hat{s}_A, p_0) d\mu_0(p_0),
\]

for some \( \mu_0 \in cl(\mathcal{M}_t(h^t)) \).

To illustrate this definition, we can establish:

**Proposition 2** For given \( \mathcal{M}_t(h^t) \) and all \( (s_p, s_A), (\hat{s}_p, \hat{s}_A) \in \mathcal{S}_P \times \mathcal{S}_A \), the following three statements are equivalent:

(i) \( (s_p, s_A) \succ^{*}_{\mathcal{M}_t(h^t)} (\hat{s}_p, \hat{s}_A) \).

(ii) \( W_t(p_t(K_t; p_0) | s_p, s_A, p_0) \geq W_t(p_t(K_t; p_0) | \hat{s}_p, \hat{s}_A, p_0), \)

for all \( \delta_{p_0} \in cl(\mathcal{D}_t(h^t)) \), where \( \mathcal{D}_t(h^t) \) is the Dirac subset of \( \mathcal{M}_t(h^t) \), and

\( W_t(p_t(K_t; p_0) | s_p, s_A, p_0) > W_t(p_t(K_t; p_0) | \hat{s}_p, \hat{s}_A, p_0), \)

for some \( \delta_{p_0} \in cl(\mathcal{D}_t(h^t)), \)

(iii) \( W_t(p_t(K_t; p_0) | s_p, s_A, p_0) \geq W_t(p_t(K_t; p_0) | \hat{s}_p, \hat{s}_A, p_0), \)

for all \( p_0 \in cl(P(\mathcal{M}_t(h^t))) \) and

\( W_t(p_t(K_t; p_0) | s_p, s_A, p_0) > W_t(p_t(K_t; p_0) | \hat{s}_p, \hat{s}_A, p_0), \)

for some \( p_0 \in cl(P(\mathcal{M}_t(h^t))) \).
We can now state:

**Definition 4** For a given \( \mathcal{M}_t(h^t) \) and \( s_A \in \mathcal{S}_A \), the principal’s set of all conditional best responses (conditional on \( \mathcal{M}_t(h^t) \)) is given by

\[
\mathcal{S}_P^* (s_A; \mathcal{M}_t(h^t)) = \left\{ s_P \in \mathcal{S}_P \mid s_P \in \arg \max_{s'_P \in \mathcal{S}_P} U_t(h^t; s'_P, s_A) \text{ and } (s_P, s_A) \not\in s_P^*_{\mathcal{M}_t(h^t)}(s_P, s_A), \forall \hat{s}_P \in \mathcal{S}_P \right\}.
\]

This definition entails that, informally speaking, \( s_P \) should not be conditionally weakly dominated by any \( \forall \hat{s}_P \in \mathcal{S}_P \). Given the agent’s strategy \( s_A \in \mathcal{S}_A \) and the principal’s belief represented by \( \mathcal{M}_t(h^t) \), the principal’s objective is to play her best response \( s_P \in \mathcal{S}_P^* (s_A; \mathcal{M}_t(h^t)) \). The principal’s preferences over alternative strategies are determined by the worst-case-scenario performance of these strategies. A (continuation) strategy can be optimal only if it yields the best worst-case-scenario performance. If there are multiple strategies that have the best worst-case-scenario performance, only the non-weakly dominated strategies can be optimal.\(^{20}\)

Turning now to the agent, the agent’s instantaneous utility is given by

\[
v_{t_r}(p_t(K_t; p_0) \mid s, p_0) = \lambda k_{t_r} p_{t_r}(K_t; p_0) \Pi_A - a (1 - \phi_{t_r}) k_{t_r}.
\]

Allowing for arbitrary (as opposed to optimal) investment \( k_{t_r} \in [0, 1] \) the agent’s value function is given by

\[
V_t(p_t(K_t; p_0) \mid s, p_0) = \int_{t}^{\infty} v_{t_r}(p_t(K_t; p_0) \mid s, p_0) e^{-\lambda \int_{s}^{t} k_{t} p_{t} d\tau} e^{-r(t-\tau)} d\tau \\
= v_t(p_t(K_t; p_0) \mid s, p_0) dt + e^{-(\lambda k_t p_t + r)} dt V_{t+dt}(p_{t+dt}(K_{t+dt}; p_0) \mid s, p_0).
\]

(15)

### 4 Solution Concept

Our solution concept is an extension of the familiar perfect Bayesian equilibrium to a setting in which the uninformed player has multiple priors. We call it the *perfect objectivist equilibrium*. The term ‘objectivist’ is to emphasize the dominant role of objective inference, i.e., likelihood inference, in learning under ambiguity, as is modelled in the current paper, in contrast to Bayesian inference that underlines the perfect Bayesian equilibrium.

**Definition 5** A pure strategy perfect objectivist equilibrium (POE) of the game is a tuple \( (s, \mathcal{M}) \), where \( \mathcal{M} \) represents the beliefs of the principal; i.e.,

\[
\mathcal{M} = \{ \mathcal{M}_t(h^t) \mid h^t \in \mathcal{H}_t, t \in \{0, \infty\} \cup [0, \infty) \},
\]

and the pure strategy profile \( s = (s_P, s_A) \) is a map

\[
s : \{ (h^t, p_0) \mid h^t \in \mathcal{H}_t, t \in \{0, \infty\} \cup [0, \infty), p_0 \in [0, 1] \} \to \mathcal{G}_t \times \mathcal{X}_t,
\]

\(^{20}\)In the Appendix, we provide an example of a weakly dominated strategy. However, to understand the example, one must first understand the equilibrium analysis developed below. For this reason, the example is presented immediately after the proof of Theorem 4.
such that,

(i) for any time \( t \in [0, \infty) \) and realized history \( h^t \), given the conditional posteriors \( M_t(h^t) \), the continuation strategies derived from \( s_A \) and \( s_P \) are mutual best responses. That is, for all history \( h^t \in H^t \), all \( t \in \{0_-, 0_+\} \cup [0, \infty) \),

\[
V_t(p_t(K_t; p_0)|s_P, s_A, p_0) \geq V_t(p_t(K_t; p_0)|s_P, s'_A, p_0)
\]

for all \( s'_A \in S_A \) for all \( p_0 \in [0, 1] \), and

\[
s_P \in S_P^* (s_A; M_t(h^t)).
\]

(ii) The initial set of posteriors is given by \( M_{0^{-}} (h^{0^{-}}) = M_0 \). The belief updating is through maximum likelihood inference as described in Section 3.1.3. Specifically, (a) \( L_0(s) \), the set of likelihood functions restricted to strategy profile \( s \), is a singleton; (b) for all \( t \in \{0_-, 0_+\} \cup [0, \infty) \), the set of all accepted posteriors \( M_t(h^t) \) against any on-equilibrium-path history \( h^t \) is derived from maximum likelihood inference based on \( s \in L_0(s) \), and \( M_t(h^t) = M_0 \) off equilibrium path; The determination of \( M_0 (h^0) \) is as described in Section 3.1.4.

In Definition 5, condition (i) requires that given the posteriors for any given history (either on or off equilibrium path), all continuation strategies of the agent must be best responses to the principal’s continuation strategies, and subject to this constraints, the continuation strategies of the principal must be best response too. The above two conditions must also hold for the special cases: (a) At \( t = 0_{-} \), \( s_P (h^{0_{-}}, p_0) = c \in C \),

\[
s_P \in S_P^* (s_A; M_{0_{-}} (h^{0_{-}})).
\]

(b) At \( t = 0_{-1} \), \( s_A (h^{0_{-1}}, p_0) = R (p_0; c) \in [0, 1] \), and for all \( s'_A (h^{0_{-1}}, p_0) = R' (p_0; c) \in [0, 1] \)

\[
V_{0_{-1}} (p_0|s_P, s_A, p_0) \geq V_{0_{-1}} (p_0|s_P, s'_A, p_0).
\]

Condition (ii) requires that belief updating follows the “likelihood inference” described in Section 3.1.3 (as opposed to the familiar Bayesian inference). Particularly, it requires that for any history that is off the equilibrium path, every plausible equilibrium-based prediction of the play of the game is contradicted by the data and the set of all posteriors reverts to equal the set of all priors.

The perfect objectivist equilibrium could potentially take many forms. We define some of the forms it could take, and we then indicate what it means for a POE to be Pareto dominant.

### 4.1 Pooling Equilibrium

A POE is a pooling equilibrium if and only if the initial revelation is a pooling outcome, that is, if and only if the set of plausible types \( I_d(c, R (p_0, c)) = [0, 1] \) for all \( p_0 \in [0, 1] \). The condition reflects the fact that by time \( t = 0 \) no private information about \( p_0 \) has been (credibly) revealed by the agent through the initial contracting behavior.
4.2 Separating Equilibrium

A POE is a separating equilibrium if and only if the initial revelation is a separating outcome, i.e., if and only if \( \mathcal{I}_d(c, R(p_0, c)) \) is a singleton for all \( p_0 \in [0, 1] \). This is possible only if the principal’s mechanism, by time \( t = 0 \), can induce the agent to voluntarily reveal his type truthfully (perhaps subject to some one-to-one mapping) and completely.

4.3 Semi-Separating Equilibrium

A POE is a semi-separating equilibrium if and only if initial revelation is semi-separating, i.e., if \( \mathcal{I}_d(c, R(p_0, c)) \) is non-singleton for some \( p_0 \in [0, 1] \). This is possible only if the principal’s mechanism, by time \( t = 0 \), can induce the agent to voluntarily reveal his type truthfully but incompletely.

We emphasize that the concepts of pooling, separating, and semi-separating equilibria pertain to the initial revelation. The fact that the POE is a pooling equilibrium does not necessarily imply that the equilibrium investment behavior of one agent-type will correspond to the equilibrium investment behavior of another agent-type as the game evolves.

4.4 Closed-loop and Markov Perfect Objectivist Equilibrium

In general a closed-loop strategy may depend on the full history \( h^t \) at time \( t \). In a subset of POE of this model, for \( t \geq 0 \) the players’ strategies are Markovian. Furthermore, there exists an explicit set of (payoff-relevant) state variables that sufficiently summarize the payoff-relevant part of history \( h^t \). They include \( K_t, \phi_t \), which are known to both players. They also include a state variable \( q_t \), which summarizes the principal’s knowledge about the agent type. This state variable (which will be defined in section 5.1), is also known to both players. In addition, the agent knows \( p_0 \), which is his private information. Therefore, the players’ strategies can be explicitly expressed as Markov strategies. We call each POE in this subset a Markov perfect objectivist equilibrium (MPOE), i.e., a POE in which all players (and player types) use Markov strategies.

4.5 Pareto Dominant Equilibrium Outcome

**Definition 6** A Pareto dominant equilibrium outcome is an equilibrium outcome which Pareto dominates all other (different) equilibrium outcomes, that is, in comparison with any other (different) equilibrium outcome, it does not make the principal or any type of the agent worse off; and it makes either the principal or at least one type of the agent strictly better off. “To make the principal worse off” means, from the principal’s perspective, the outcome

---

21 There is no established generic definition of Markov strategy for dynamic games with incomplete information in the Markov literature yet. Our analysis of Markov strategy inevitably proceeds on a case-by-case basis.
is either weakly conditionally dominated by, or has lower worst-case value than, the outcome in comparison. “To make the principal strictly better off” means, from the principal’s perspective, the outcome weakly conditionally dominates the outcome in comparison.

A Pareto dominant equilibrium outcome, if it exists, is a plausible equilibrium refinement. We return to this point later in the analysis.

5 Characterization of the POE

We start this section by establishing the following generic necessary conditions for a POE.

**Lemma 1** A necessary condition for a POE is that the following participation constraints must be satisfied: $U_t \geq 0$ for all $t \in \{0, \infty\}$ and $V_t \geq 0$ for all $t \in \{0, \infty\}$.

In the remainder of this section, we proceed as follows. First, we do some generic characterization of Markov perfect objectivist equilibria (MPOE) of $t = 0$ continuation games. Next, in Section 5.2 we study a particular pooling equilibrium, where the principal’s Markov strategy is a simple time-invariant reimbursement rate. Finally, in Section 5.3, we establish the uniqueness of Markov equilibrium outcome.

5.1 An MPOE of Continuation Games at $t = 0$

The Markov strategy of the agent: $k_t(K_t, \phi_t, q_t; p_0)$ is a function of state variables: $K_t$, $\phi_t$, $q_t$ and $p_0$, where $q_t$ is a payoff-relevant state variable (to be defined below) that summarizes the principal’s knowledge about the agent’s type. Notice, $\phi_t$ is a state variable for time $t$ as well as a choice variable of the principal at time $t - dt$. Recall that $K_t \equiv \int_0^t k_r dr$. It follows that $K_{t+dt} - K_t = k_t dt$.

It can be shown from the first-order approximation of (15) that $V_t(p_t(K_t; p_0)|s, p_0)$ satisfies the following recursion:

$$ rV_t = v_t - \lambda k_t p_t V_t + \frac{dV_t}{dt}, $$

where $v_t = [-a (1 - \phi_t) + \lambda p_t \Pi_A] k_t$. Since the agent plays a Markov strategy, $V_t$ can be written as $V(K_t, \phi_t, q_t; p_0)$, where $V(\cdot)$ does not directly depend on time $t$. Thus $\frac{dV_t}{dt} = \lim_{dt \to 0} \frac{V(K_{t+dt}, \phi_{t+dt}, q_{t+dt}; p_0) - V(K_t, \phi_t, q_t; p_0)}{dt}$ can be decomposed as

$$ \frac{dV_t}{dt} = \frac{\partial V}{\partial K_t} k_t + \frac{\partial V}{\partial q_t} dq_t + \frac{\partial V}{\partial \phi_t} d\phi_t. $$

Since $k_t$ is chosen after $K_t$ and $\phi_t$ have been determined, and simultaneously (and independently) chosen with $\phi_{t+dt}$, it has no effects on $K_t$, $\phi_t$ or $\frac{d\phi_t}{dt}$. The Bellman equation for the agent’s optimal investment strategy is thus given by

$$ rV = \frac{\partial V}{\partial \phi_t} d\phi_t + \max_{k_t \in [0,1]} \left\{ \left[ -a (1 - \phi_t) + \lambda p_t \Pi_A - \lambda p_t V + \frac{\partial V}{\partial K_t} \right] k_t + \frac{\partial V}{\partial q_t} dq_t \right\}. $$

(16)
If the term \( \frac{\partial V}{\partial q_t} \frac{dq_t}{dt} \) depends on action \( k_t \) (negatively), then there exists an informational strategic effect that affects the agent’s investment behavior. The agent may underinvest in order to mimic a lower type. Such informational strategic effect does not exist if

\[
\frac{\partial V}{\partial q_t} = 0 \quad \text{or} \quad \frac{dq_t}{dt} = 0;
\]

(17)

the first equality holds if \( V(K_t, \phi_t, q_t; p_0) \) does not depend on \( q_t \) in the neighborhood of \( q_t \); the second holds if \( \phi_t \) is such that the choice of \( k_t \in [0, 1] \) reveals no new information about the agent’s type. Condition (17) requires that the Markov strategy of the principal must not attempt to and cannot induce the agent to reveal new sensitive information about his type. In what follows we characterize MPOE such that condition (17) holds. First, we define the cutoff type \( \tilde{p}(K_t, \phi_t) \) such that

\[
-a (1 - \phi_t) + \lambda p_t \Pi_A - \lambda p_t V + \frac{\partial V}{\partial K_t} < 0
\]

for all \( p_0 \leq \tilde{p}(K_t, \phi_t) \) for all \( V \geq 0 \) and \( \frac{\partial V}{\partial K_t} \leq 0 \). More explicitly, we have

\[
p_t(K_t; \tilde{p}(K_t, \phi_t)) = \frac{a (1 - \phi_t)}{\lambda \Pi_A}
\]

or equivalently,

\[
\tilde{p}(K_t, \phi_t) = \frac{1}{1 + \left[ \frac{\lambda \Pi_A}{a (1 - \phi_t)} - 1 \right] e^{-\lambda K_t}}.
\]

(18)

Now we define the state variable \( q_t \) such that

\[
q_{t+dt} = \max_{\tau \in [0, t]} \{ \tilde{p}(K_{\tau}, \phi_{\tau}) \mathbb{1}_{k_\tau > 0}, q_0 \} ;
\]

\[ q_0 = \inf P_0, \]

(19)

where \( P_t \equiv P(\mathcal{M}_t(h^t)) \) is a simplified notation; \( 1_{k_r > 0} \) is an indicator function. Note, \( q_t \) has two important properties: (i) the Markov property

\[
q_{t+dt} = \begin{cases} 
\tilde{p}(K_t, \phi_t) \mathbb{1}_{k_t > 0} & \text{if } \tilde{p}(K_t, \phi_t) \mathbb{1}_{k_t > 0} > q_t, \\
q_t & \text{otherwise};
\end{cases}
\]

(20)

(ii) the ratchet property

\[
\frac{dq_t}{dt} \geq 0,
\]

(21)

i.e., \( q_t \) can only move in one direction.

The economic meaning of the state variable \( q_t \) is the following. It is a (biased) point estimate of the agent’s type, and the principal in formulating her strategy would treat the
agent as if his type is \( q_t \). Suppose, hypothetically, the agent behaved myopically, i.e., he ignored the continuation value of the experimental project. He would therefore choose \( k_t = 1 \) if the net flow benefit of the investment is positive and \( k_t = 0 \) otherwise. The cutoff myopic type at time \( \tau \) accepting \( \phi_\tau \) is \( \bar{p}(K_\tau, \phi_\tau) \). Thus \( q_t \) represents the max-min myopic type that is recorded based on the history \( h^t \in \mathcal{H}^t \) for \( t > 0 \). In a way, \( q_t \) maps into a record of the maximum concession the agent (controlling for \( K_t \)) has ever made in his history of bargaining with the principal over the reimbursement rate \( \phi_\tau \). The principal then would treat the agent as if he would be willing to accept the same concession (controlling for \( K_t \)) at any future time \( t' > t \). If she wants to induce the agent to invest at time \( t_0 \), she would not be willing to demand less concession (controlling for \( K_{t_0} \)). This ratchet mechanism then consolidates any temporary concession by the agent into a permanent concession (controlling for \( K_{t_0} \)). The implication is that if the agent deviates from an equilibrium level of concession by conceding more (e.g., behaving myopically), then this deviation will lead to reduction in his information rent as well as the value of his project.

To better understand the ratcheting behavior of the principal in case sensitive information about agent type is revealed to her, we turn to the contracting problem faced by a principal when she knows \( p_0 \). In this case, the players’ strategies do not depend on state variable \( q_t \). Instead, they take the forms of \( k_t(K_t, \phi_t; p_0) \) and \( \phi_{t+dt}(K_t; p_0) \). The agent’s Bellman equation for optimal investment strategy is simplified to:

\[
rv = \frac{\partial V}{\partial \phi_t} + \max_{k_t \in [0,1]} \left\{ -a(1-\phi_t) + \lambda p_t \Pi_A - \lambda p_t V + \frac{\partial V}{\partial K_t} \right\} k_t.
\]

(22)

Since the expression in the curly bracket in (22) is linear in \( k_t \), the optimal \( k_t \) must therefore satisfy the following condition:

\[
\tilde{k}_t^*(K_t, \phi_t; p_0) = \begin{cases} 
1 & \text{if } p_t \geq \frac{\frac{a(1-\phi_t)}{\lambda(\Pi_A - V)}}{rac{\partial V}{\partial K_t}} \\
0 & \text{otherwise},
\end{cases}
\]

i.e., the agent’s optimal investment intensity \( k_t \) only takes values: 1 or 0. We can reformulate the optimal investment strategy in terms of the cutoff type \( \tilde{p}(K_t, \phi_t) \) as follows.

\[
\tilde{k}_t^*(K_t, \phi_t; p_0) = \begin{cases} 
1 & \text{for } p_0 \geq \tilde{p}(K_t, \phi_t), \\
0 & \text{otherwise,}
\end{cases}
\]
or equivalently,

\[
\tilde{k}_t^*(K_t, \phi_t; p_0) = \begin{cases} 
1 & \text{for } K_t \leq \tilde{K}(\phi_t, p_0), \\
0 & \text{otherwise,}
\end{cases}
\]

(23)

\footnote{This state variable does not have to reflect the correct beliefs of the principal, which must be the product of equilibrium-based legitimate inference. This is because, off equilibrium path, the set of posterior beliefs equals the initial set of beliefs \( M_0 \) and becomes stationary. In contrast, the state variable \( q_t \) may still evolve over time and affect the principal’s action. The fact \( q_t \) is a biased estimate of the agent’s type and that it is used to guide the players’ strategies does not imply that the players are not (sequentially) rational, or use illegitimate inference to form beliefs in any way. The state variable \( q_t \) and the common knowledge about it are simply part of a coordination device used by the players.}
where
\[ \tilde{K}(\phi_t, p_0) = \frac{1}{\lambda} \ln \left[ \left( \frac{\lambda \Pi_A}{a(1 - \phi_t)} - 1 \right) \frac{p_0}{1 - p_0} \right]. \] (24)

It is optimal for the principal to refrain from offering compensation contract for \( K_t \in [0, K^A(p_0)] \), because the agent himself has a sufficient incentive to invest. If the principal wants to induce more cumulative investment, she has to offer positive reimbursement when \( K_t > K^A(p_0) \). Let \( \tilde{\phi}_t(K_t; p_0) \geq 0 \) be the minimum rate of reimbursement to keep \( \tilde{k}_t^*(K_t, \tilde{\phi}_t; p_0) = 1 \). Consider \( \tilde{\phi}_t(K_t; p_0) \) is such that
\[ \tilde{p}(K_t; \tilde{\phi}_t) = p_0. \]
That is, the cutoff type \( \tilde{p}(K_t; \tilde{\phi}_t) \) would equal the true type. Equivalently,
\[ \tilde{\phi}_t = 1 - \frac{\lambda \Pi_A}{a} p_t(K_t; p_0). \] (25)
It follows the true type will continue to invest as his incentive compatible constraint is binding.

To avoid negative value of \( \tilde{\phi}_t \), for \( K_t \in [0, K^A(p_0)] \) the value of \( \tilde{\phi}_t \) should be bounded by zero, hence, \( \tilde{\phi}_t(K_t; p_0) = 0 \). Equation (25) applies to the case of \( k_t \geq K^A(p_0) \); overall, it should be generalized to
\[ \tilde{\phi}_t(K_t; p_0) = \max \left\{ 1 - \frac{\lambda \Pi_A}{a} p_t(K_t; p_0), 0 \right\}. \] (26)

To find the principal’s optimal Markov strategy, we also need to know the principal’s preferred termination threshold value for \( K_t \) given \( \tilde{\phi}_t(K_t; p_0) \). The solution to this problem is full-information Pareto efficient (allowing for compensation transfer), taking into account the agent’s investment incentive compatibility constraint. The first-order approximation of (13) implies that the principal’s value function \( W_t = W(K_t, \tilde{\phi}_t; p_0) \) satisfies the following recursive relation:
\[ rW = \frac{\partial W}{\partial \phi_t} \frac{d\tilde{\phi}_t}{dt} - \tilde{\phi}_t a k_t + \lambda k_t p_t \Pi_P + \lambda p_t W k_t + \frac{\partial W}{\partial K_t} k_t. \]

Let \( \tilde{k}_t \) be the principal’s preferred investment intensity. Then \( \tilde{k}_t \) must satisfy the following Bellman equation:
\[ rW = \frac{\partial W}{\partial \phi_t} \frac{d\tilde{\phi}_t}{dt} + \max_{k_t \in \{0, 1\}} \left\{ -\tilde{\phi}_t a + \lambda p_t \Pi_P - \lambda p_t W + \frac{\partial W}{\partial K_t} \right\} \tilde{k}_t \]

The maximization problem has the following bang-bang solution:
\[ \tilde{k}_t(p_0) = \begin{cases} 1 & \text{if } p_t > \frac{a \tilde{\phi}_t - \frac{\partial W}{\partial K_t}}{\lambda (\Pi_P - W)}, \\ 0 & \text{otherwise}. \end{cases} \] (27)
The switching point must be characterized by $W = 0$ and $\frac{\partial W}{\partial K_t} = 0$ and hence asymptotically

$$\tilde{\phi}_t (K_t; p_0) |_{K_t = K^{**}(p_0)} = \phi^{**} \equiv \frac{\Pi_P}{\Pi_P + \Pi_A},$$

(28)

where

$$K^{**}(p_0) \equiv \tilde{K} (\phi^{**}, p_0)$$

(29)

is the termination threshold value of $K_t$. Let

$$p^{**} \equiv p_t (K^{**}(p_0); p_0) = \frac{a}{\lambda (\Pi_P + \Pi_A)},$$

(30)

where $p^{**}$ is the termination threshold value of $p_t (K_t; p_0)$. Note, $p^{**} = \bar{p} (0, \phi^{**})$.

Notice that $\phi^{**}$ has the property:

\text{Share of cost} = \text{Share of benefit}.

Intuitively, at the socially optimal termination point, the social surplus from trade is zero. Unless the share of cost equals the share of benefit, trade would make one party worse off than no trade.

**Lemma 2** Suppose (hypothetically) $p_0$ is known to the principal, and the Markov strategy of the agent is given by $\tilde{k}_t^* (K_t, \phi_t; p_0)$ as defined by (23). Then Markov strategy of the principal, $\tilde{\phi}^*_{t+dt} (K_t; p_0)$ for all $t \geq 0$, is optimal if and only if

$$\tilde{\phi}^*_{t+dt} (K_t; p_0) =
\begin{cases}
0 & \text{if } K_t \in [0, K^A (p_0)) \\
1 - \frac{\lambda A}{a} p_{t+dt} (K_t + dt; p_0) & \text{if } K_t \in [K^A (p_0), K^{**}(p_0)) \\
\leq \phi^{**} & \text{if } K_t \geq K^{**}(p_0).
\end{cases}$$

(31)

The strategy $\tilde{\phi}^*_{t+dt} (K_t; p_0)$, given $\tilde{k}_t^* (K_t, \phi; p_0)$ and conditional on the principal knowing $p_0$, is Pareto efficient (allowing for compensation transfer) and (subject to this) extracts the maximum surplus for the principal that is feasible using the short-term contract $\phi_{t+dt}$.

**Lemma 3** Suppose (hypothetically) that $p_0$ is common knowledge. Given appropriate compensation transfer, the following is a Pareto efficient investment policy:

$$k_t = \begin{cases}
1 & \text{if } K_t \leq K^{**}(p_0) \\
0 & \text{otherwise}.
\end{cases}$$

(32)

The set $[0, K^A (p_0))$ is empty if $K^A (p_0) < 0$, which is possible because from (4) it follows that $K^A (p_0) < 0 \Leftrightarrow p_0 < p^A$. The interpretation of a negative threshold $K^A (p_0)$ is that $p_0$ is too small for the agent to invest without compensation.
The problem is that, however, $p_0$ is not known to the principal. Lemma 2 helps explain why the agent has a good reason not to reveal accurate information about $p_0$. That is, if the agent did, then the principal would use this information to minimize $\phi_{t+dt}$ as well as the agent’s surplus from trade in the future – the ratchet effect.

**Proposition 3** If there exists an MPOE such that $\inf P_0 \leq p^{**} < \sup P_0$, and the principal’s Markov strategy takes the form $\phi_{t+dt} = \phi(K_t, q_t)$, then $\phi_t = \phi^{**}$ for all state $(K_t, q_t)$ that is on an equilibrium path. Consequently,

$$\frac{\partial V}{\partial q_t} = 0,$$

and

$$\frac{dq_t}{dt} > 0 \text{ if and only if } k_t > 0.$$

Intuitively, $\phi^{**}$ is the reservation price for the principal. In an MPOE the players’ strategies are Markov. The state variables $(K_t, q_t, \phi_t)$ continue to evolve only if $k_t > 0$. Once $k_t(K_t, q_t, \phi_t) = 0$ occurs, $(K_t, q_t, \phi_t)$ become stationary, and $k_t$ will also become stationary, i.e., $k_{t'} = 0$ for all $t' \geq t$. This means $k_t = 0$ triggers a termination of the trading relationship. For the principal, if $\phi_{t'} < \phi^{**}$, such a termination is suboptimal, which means the principal would have incentive to increase $\phi_{t'+dt}$, which contradicts the presumption that the principal’s Markov strategy is sequentially rational as required by the definition of POE. The implication is that to ensure that the principal’s offer $\phi_{t+dt} = \phi(K_t, q_t)$ is truly a take-it-or-leave-it offer (for which no amount of rejection by the agent can change it), the principal’s offer price must reach her reservation price $\phi^{**}$.

Now we turn to characterization of the off-equilibrium-path strategies of an MPOE. Since off-equilibrium-path strategies don’t affect the equilibrium outcome, they are only interesting for the purpose of ensuring that equilibrium Markov strategies of the players are sequentially rational (even off equilibrium path). This is an requirement of the definition of POE. We start by considering an MPOE such that in all off-equilibrium-path coarsest payoff-relevant states $(K_t, q_t)$,

$$\frac{dq_t}{dt} > 0 \text{ only if } \phi_t \geq \phi^{**}. \quad (33)$$

Intuitively, condition (33) can be explained by the bargaining strategy of the agent as follows: if he has never conceded from the following target before, he will insist on the target that the principal’s offer $\phi_t$ should be no less than her reservation price $\phi^{**}$. On the one hand, guided by this bargaining strategy, whenever the agent is required to “concede” or (equivalently) to “reveal sensitive information”, i.e., to accept $\frac{dq_t}{dt} > 0$, the agent will surely reject the request (by choosing $k_t = 0$) unless the offer $\phi_t$ is no less than the principal’s reservation price $\phi^{**}$. On the other hand, if the principal follows her Markov strategy, then the Markov strategy

---

24By the term “coarsest” we mean the list of state variables is as short as possible, so that the partition of history space is as coarse as possible.
must (i) be no higher than the reservation price, i.e., \( \phi_t = \phi (K_{t-dt}, q_{t-dt}) \leq \phi^* \). In such an MPOE, suppose at time \( t \) the payoff-relevant state \((K_t, q_t, \phi_t)\) is such that \( \frac{dt}{dt} > 0 \) for \( k_t > 0 \), then for any \( t' > t \) the payoff-relevant state \((K_{t'}, q_{t'}, \phi_{t'})\) in the equilibrium outcome of the continuation game supposing \((K_t, q_t, \phi_t)\) must be such that \( \phi_{t'+dt} = \phi (K_{t'}, q_{t'}) = \phi^* \). For any \( \tau \geq t + dt \), either \( \frac{dt}{dt} > k_t > 0 \) or \( \frac{dt}{dt} = 0 \). The former implies \( \phi (K_{t'+dt}, q_{t'+dt}) = \phi^* \) (following the same argument). The latter implies \((K_{t'+dt}, q_{t'+dt}) = (K_t, q_t)\) and hence \( \phi (K_{t'+dt}, q_{t'+dt}) = \phi (K_t, q_t) = \phi^* \). This means \( \phi_{t'} \) is time-invariant and not dependent on \( q_t \). Consequently, \( \frac{\partial V}{\partial q_t} = 0 \). In this case condition (17) holds, the expression in the curly bracket in (16) is linear in \( k_t \), the optimal \( k_t \) must therefore satisfy the following condition:

\[
k^*_t (K_t, \phi_t, q_t; p_0) = \begin{cases} 
1 & \text{if } p_t \geq \frac{a(1-\phi_t) - \frac{\partial V}{\partial K_t}}{\lambda (\Pi_A - V)} \\
0 & \text{otherwise},
\end{cases}
\]

i.e., the agent’s optimal investment intensity \( k_t \) only takes values: 1 or 0. At a switching point (if it exists), the following equation holds:

\[
p_t (K_t; p_0) = \frac{a (1 - \phi_t) - \frac{\partial V}{\partial K_t}}{\lambda (\Pi_A - V)}. \tag{34}
\]

We can reformulate the optimal investment strategy in terms of the cutoff type \( \bar{p} (K_t, \phi_t) \) and state the boundary conditions of the differential dynamic system as follows:

**Lemma 4** If there exists an MPOE that satisfies condition (33), then the agent’s investment strategy must be given by

\[
k^{**}_t (K_t, \phi_t, q_t; p_0) = \begin{cases} 
1 & \text{for } p_0 \geq \bar{p} (K_t, \phi_t) \text{ and } q_t \geq \bar{p} (K_t, \phi_t), \\
1 & \text{for } p_0 \geq \bar{p} (K_t, \phi_t), \text{ } q_t < \bar{p} (K_t, \phi_t) \text{ and } \phi_t \geq \phi^*, \\
0 & \text{otherwise},
\end{cases} \tag{35}
\]

and following boundary condition holds.

\[
V_t = 0 \text{ and } \frac{\partial V}{\partial K_t} = 0 \text{ for } p_0 \leq \bar{p} (K_t, \phi_t).
\]

There is an equivalent formulation of \( k^{**}_t (K_t, \phi_t, q_t; p_0) \) as follows:

\[
k^{**}_t (K_t, \phi_t, q_t; p_0) = \begin{cases} 
1 & \text{for } K_t \leq \bar{K} (\phi_t, p_0) \text{ and } q_t \geq \bar{p} (K_t, \phi_t), \\
1 & \text{for } K_t \leq \bar{K} (\phi_t, p_0), \text{ } q_t < \bar{p} (K_t, \phi_t) \text{ and } \phi_t \geq \phi^*, \\
0 & \text{otherwise},
\end{cases} \tag{36}
\]

\[25\]By the phrase “equilibrium outcome of the continuation game” we mean the players only play their equilibrium strategies that are assigned to the current continuation game, i.e., there is no deviation by any player.
Lemma 5 Suppose the principal’s Markov strategy is given by

\[
\phi_{t+dt}^{**}(K_t, q_t) = \begin{cases} 
0 & \text{if } K_t \in [0, K^A(q_t)) \\
1 - \frac{\Lambda A}{a} p_{t+dt}(K_t + dt; q_t) & \text{if } K_t \in [K^A(q_t), K^{**}(q_t)) \\
\phi^{**} & \text{if } K_t \geq K^{**}(q_t),
\end{cases}
\]  

(37)

and the agent’s investment strategy is given by \( k_t^{**}(K_t, \phi_t, q_t; p_0) \) as defined in (36). Then \( \phi_{t+dt}^{**}(K_t, q_t) \) and \( k_t^{**}(K_t, \phi_t, q_t; p_0) \) are mutual best responses following for all \( t \geq 0 \) histories. Given \( \phi_{t+dt}^{**}(K_t, q_t) \), the agent’s strategy \( k_t^{**}(K_t, \phi_t, q_t; p_0) \) can be equivalently formulated as

\[
k_t^{**}(K_t, \phi_t, q_t; p_0) = \begin{cases} 
1 & \text{if } \phi_t \geq \phi_{t-}^{**}(K_{t-}, q_{t-}) \text{ and } K_t \leq \bar{K}(\phi_t, p_0), \\
0 & \text{otherwise}.
\end{cases}
\]  

(38)

5.2 A Pooling Equilibrium

We now define a candidate pooling equilibrium.

Definition 7 Let the tuple \((s^{**}, \mathcal{M}^{**})\) be such that

(a) at \( t = 0 \),

\[s_{p}^{**}(h^{0}) = \phi^{**} \text{ for all } \hat{p}_0 \in [0, 1];\]

(b) at \( t = 0 \), \( s_{A}^{**} \) is a best response to \( s_{p}^{**} \) for all \( h^{0} \), specifically,

\[s_{A}^{**}(h^{0}, p_0) = \inf \arg \max_{\hat{p}_0 \in [0,1]} \Phi(\hat{p}_0);\]

thus \( s_{A}^{**}(h^{0}, p_0) = 0 \) with \( h^{0} = \{\Phi|\Phi(\hat{p}_0) = \phi^{**}\} \);

(c) for all \( t \geq 0 \),

\[s_{p}^{**}(h^t) = \phi_{t+dt}^{**}(K_t, q_t),\]

\[s_{A}^{**}(h^t, p_0) = k_t^{**}(K_t, \phi_t, q_t; p_0) = \begin{cases} 
1 & \text{if } \phi_t \geq \phi_{t-}^{**}(K_{t-}, q_{t-}) \text{ and } K_t \leq \bar{K}(\phi_t, p_0), \\
0 & \text{otherwise}.
\end{cases}\]

(39) (40)

\( \mathcal{L}_0(s^{**}) \) is derived from \( s^{**} \); the derivation of \( \mathcal{M}_t(h^t) \in \mathcal{M}^{**} \) is as described in Section 3.1.3, \( \mathcal{M}_0(h^0) = \mathcal{M}_0 \).

The agent’s strategy \( s_{A}^{**}(h^t, p_0) \) given by (40) is characterized by the feature that for \( t > 0 \) all higher types \( p_0 \in (\bar{p}(K_t, \phi^{**}), 1] \) mimic the behavior of the marginal type \( \bar{p}(K_t, \phi^{**}) \). If the marginal type’s participation constraint is violated (i.e., \( \phi_t < \phi^{**} \)) then all higher types stop investing. So \( s_{A}^{**}(h^t, p_0) \) is a partial pooling strategy. This strategy essentially sets a reservation rate for \( \phi_t \) (at \( \phi^{**} \)), below which the agent would reject the offer through suspending investment. By such strategy, the agent also withholds sensitive information about his true type.

We now establish that the candidate pooling equilibrium in Definition 7 is indeed an MPOE.
Theorem 3  The tuple \((s^{**}, \mathcal{M}^{**})\) is a pooling pure strategy MPOE.

In what follows we illustrate how likelihood inference works using the pooling equilibrium example. It can be shown that (40) is equivalent to

\[
k_t^{**}(K_t, \phi_t, q_t; p_0) = \begin{cases} 
1 & \text{if } \phi_t \geq \phi_t^{**}(K_{t-dt}, q_{t-dt}) \text{ and } p_0 \leq \bar{p}(K_t, \phi^{**}), \\
0 & \text{otherwise}.
\end{cases}
\]

where \(\bar{p}(K_t, \phi^{**})\) is the cutoff type defined by (18). In the pooling equilibrium with \(\phi_t = \phi^{**}\) for all \(t \geq 0\), \(P_0 = [0, 1]\).

If \(k_0 = 1\) is observed at \(t = dt\), then the principal can infer \(p_0 \geq \frac{1}{\exp(A)} = p^*\) and that \(p^*\) is the (inclusive) maximum lower bound of \(p_0\), i.e., \(p^* = \inf P_t = \min P_t\) and \(p^* \notin P_{dt}\). Also, 1 can be inferred as the (inclusive) least upper bound of \(p_0\), i.e., \(p_0 \leq 1 = \sup P_{dt} = \max P_{dt}\) and \(1 \in P_{dt}\). If the agent terminates investment at time \(t\), then it can be inferred that the true type of the agent is given by \(p_0 \geq \bar{p}(K_t, \phi^{**})\); otherwise, \(\bar{p}(K_t, \phi^{**})\) can be inferred, at time \(t\), as the maximum (inclusive) lower bound of \(p_0\), and \(1\) remains the (inclusive) least upper bound. That is, \(p_0 \in [\bar{p}(K_t, \phi^{**}), 1]\).

If \(k_0 = 0\) is observed at \(t = dt\), then it can be inferred that \(p_0 < p^*\) and \(p^*\) the (non-inclusive) least upper bound of \(p_0\) and \(0\) is the (inclusive) maximum lower bound, i.e., \(p_0 \in [0, p^*]\).

Formally, in the pooling equilibrium with \(\phi_t = \phi^{**}\) for all \(t \geq 0\), \(P_0 = [0, 1]\), if \(k_0 = 1\) is observed at time \(dt\) then \(\inf P_{dt} = \min P_{dt} = p^*\) and \(\sup P_{dt} = \max P_{dt} = 1\), and all \(p_0 \in [p^*, 1]\) are plausible. At time \(t + dt\), if \(k_t = 1\) is observed, then \(\inf P_{t+dt} = \min P_{t+dt} = \bar{p}(K_t, \phi^{**})\), \(\sup P_{t+dt} = \max P_{t+dt} = 1\), and every \(p_0 \in [\bar{p}(K_t, \phi^{**}), 1]\) is plausible. As the agent continues to invest in the absence of a breakthrough, \(\inf P_t = \min P_t\) gradually increases from \(p^*\) to \(\sup P_t = \max P_t = 1\) in a Logistic (diffusion) process. Since

\[
\frac{d}{dt} \frac{1}{\inf P_t} = \frac{\lambda}{1 - p^*} \left( \sup P_t - \inf P_t \right),
\]

ambiguity is reduced smoothly in this case. \(P_t\) is a closed interval and \(P_t' \subset P_t\) for \(t' > t\). All probability distributions (including all Dirac measures) over the closed plausible set \(P_t\) belong to the closed plausible set \(\mathcal{M}_t(h')\).

Along the equilibrium path, if investment is terminated at time \(t\), then (in the limit) \(\inf P_{t+dt} = \min P_{t+dt} = \sup P_{t+dt} = \max P_{t+dt} = \bar{p}(K_t, \phi^{**})\) and only \(p_0 = \bar{p}(K_t, \phi^{**})\) is plausible. At this point, ambiguity is resolved abruptly, and

\[
\inf P_{t'} = \min P_{t'} = \sup P_{t'} = \max P_{t'} = \bar{p}(K_t, \phi^{**})
\]

for all \(t' > t\). That is, the true value of \(p_0\) becomes known to the principal after time \(t\).

Even though the principal does not know \(p_0\), the cumulative investment level delivered by a type-\(q_0\) agent in the POE described in Theorem 3 is the full-information Pareto efficient cumulative investment (allowing for compensation transfer). Thus, even though the information transmission in equilibrium is minimal, the allocational properties of the equilibrium are as if the principal actually knew the agent’s prior belief.
Theorem 4 The pooling equilibrium given by Definition 7 generates a Pareto efficient outcome, where the investment policy is given by

$$k_t = \begin{cases} 1 & \text{if } K_t \leq K^{**} (p_0) \\ 0 & \text{otherwise.} \end{cases}$$

The principal has $\phi^{**}$ as her reservation reimbursement rate. In the pooling equilibrium given by Definition 7, this reservation price is always reached along any equilibrium path. Let $V^{**}(K_t, \phi^{**}, p_0)$ denote the value function of the agent along an equilibrium path in this equilibrium.

Given that the investment intensity in this equilibrium is socially efficient at each point of time and the surplus given to the agent (through the expected benefit of breakthrough and compensation transfer) is maximized, $V^{**}(K_t, \phi^{**}, p_0)$ must set an upper bound to the value function of the agent for all MPOE. Furthermore, as we establish in Proposition 4, this equilibrium generates a Pareto dominant equilibrium outcome in case there exist multiple equilibrium outcomes. For each possible type of the agent, the MPOE presented by Definition 7 gives the agent its largest payment, while the principal is indifferent between the payoff it receives under this equilibrium and the outcomes under other Markov equilibria if they exist.

Proposition 4 The pooling equilibrium given by Definition 7 generates the Pareto dominant MPOE outcome if there exist multiple MPOE outcomes.

Let $V^A(K_t, 0, p_0)$ denote the value function of the agent in the absence of compensation contract with the principal. Define $B(K_t, p_0) \equiv V^{**}(K_t, \phi^{**}, p_0) - V^A(K_t, 0, p_0)$, which is the least upper bound of (expected) information rent the agent can potentially extract from the contracting with the principal.

Is it possible for any type of the agent to earn less information rent in any MPOE outcome than in the current pooling equilibrium? Presumably, each agent type who can potentially earn positive information rent should have an incentive to protect his information rent (a sort of natural property right) by not revealing sensitive private information in the contracting process. Can this kind of motivation effectively rule out the existence of (allocationally and distributively) different MPOE outcomes? This is the question to which we turn next.

### 5.3 Uniqueness of MPOE Outcome and Information Transmission

We confine the solution concept of our analysis to Markov Perfect Objectivist Equilibrium (MPOE). So far we have confined our analysis to a specific class of MPOE which involves state variable $q_t$. To derive general results about all MPOE, we either need to establish that there is no other class of MPOE, or to show the results apply to all other classes of MPOE if they (ever) exist. In this section we will do the latter. First, the next lemma extends the result: $\phi_t = \phi^{**}$, stated in Proposition 3.
Lemma 6 For any equilibrium path of an MPOE such that \( \inf P_0 \leq p^{**} < \sup P_0 \), the allocational and distributional outcome is given by

\[
\phi_t = \phi^{**},
\]

and

\[
k_t = \begin{cases} 
1 & \text{if } K_t \leq K^{**} (p_0) \\
0 & \text{otherwise},
\end{cases}
\]

for all \( t \geq 0 \).

So far we have identified a particular MPOE that exists. Are there other MPOEs? We can rule out separating equilibria in general.

Proposition 5 There exists no separating equilibrium.

To understand why separating equilibria are not possible, note that because the principal seeks to maximize the worst-case value function, she focuses on minimizing the possibility of funding agent types with low priors. That, coupled with the principal’s inability to commit to a reimbursement rate, deprives her the ability to offer information rents to high types in exchange for their truth-telling. As a result, a high type has a strong incentive to misrepresent his prior, which renders separating equilibria impossible.

The nonexistence of separating equilibrium is one instance of the general rule that the agent (who can potentially earn positive information rent) has an incentive to limit the revelation of his true type in order to protect his potential information rent. The information about the agent’s type is transmitted through two channels: first by the initial revelation and then by the observable flow of investment \( k_t \) (and reimbursement claim) in response to the stream of reimbursement rate \( \phi_t \). If the initial revelation is pooling, then no information is transmitted by it. When initial revelation is semi-separating, it is important to know whether the information transmitted is sensitive pertaining to information rent. For example, if the agent has no potential to earn information rent (i.e., \( p_0 < p^{**} \)), then any level of revelation about his type is trivial - the information has no effect on his information rent, which is zero anyway. Formally, we make a distinction between trivial and non-trivial semi-separating equilibria.

Definition 8 The initial revelation based on \( h_0 = (c, R) \) is a non-trivial semi-separating outcome if the set of all plausible types \( \mathcal{I}_d (h_0) \subset [0, 1] \), \( \mathcal{I}_d (h_0) \) is non-singleton for some \( p_0 \in [0, 1] \), and \( \inf \mathcal{I}_d (h_0) > p^{**} \) for some \( p_0 \in [0, 1] \). The initial revelation based on \((c, R)\) is a trivial semi-separating outcome if \( \mathcal{I}_d (h_0) \subset [0, 1] \), \( \mathcal{I}_d (h_0) \) is non-singleton for some \( p_0 \in [0, 1] \), and \( \inf \mathcal{I}_d (h_0) \leq p^{**} \) for all \( p_0 \in [0, 1] \). A POE is called a non-trivial semi-separating equilibrium if and only if initial revelation is a non-trivial semi-separating outcome. A POE is called a trivial semi-separating equilibrium if and only if initial revelation is a trivial semi-separating outcome.
There are many trivial semi-separating equilibria that give rise to the same allocational and distributional outcome as the pooling equilibrium defined in Definition 7 and established in Theorem 3.

**Proposition 6** There exists an infinite set of trivial semi-separating MPOE. Their (allocative and distributional) outcomes are identical to the equilibrium defined in Definition 7 and established in Theorem 3.

The result follows immediately from Theorem 3 and the following example: \((\bar{s}, \mathcal{M})\) such that

\[
\bar{s}_p (h^{0-2}) = c (\hat{p}_0) = \phi^{**} \text{ for all } \hat{p}_0 \in [0, 1],
\]

\[
\bar{s}_A (h^{0-1}, p_0) = \bar{R} (p_0; c) = \begin{cases} p_0 & \text{if } p_0 < p^{**} \\ p^{**} & \text{otherwise} \end{cases},
\]

and for \(\forall t \geq 0\) and \(\forall h^t \in \mathcal{H}^t\),

\[
\bar{s}_p (h^t) = \bar{\phi}_{t+dt} = \phi_{t+dt}^* (K_t, q_t),
\]

and

\[
\bar{s}_A (h^t; p_0) = k_t^* (K_t, \phi_t, q_t; p_0) = \begin{cases} 1 & \text{if } \phi_t \geq \phi_t^{**} (K_{t\text{-}dt}, q_{t\text{-}dt}) \text{ and } K_t \leq K (\phi_t, p_0) \\ 0 & \text{otherwise}. \end{cases}
\]

Obviously, in any of the trivial semi-separating POE described by Lemma 6, the agent’s information rent reaches its upper bound, just as in the pooling POE established in Theorem 3. By Proposition 3 this property holds for all MPOE such that \(\phi_{t+dt} = \phi (q_t, K_t)\) is the principal’s Markov strategy. Intuitively, the reason why the agent’s information rent is maximized is because he never lets the principal know unambiguously that the reimbursement rate \(\phi^{**}\) is too “generous” for the agent. Notice, \(p_t (K_t; \inf P_t) = p^{**}\) (and equivalently \(K_t = K^{**} (\inf P_t)\)) holds for all \(t \geq 0\). The ambiguity-averse principal therefore can never justify a reduction of \(\phi_t\) to below \(\phi^{**}\) (which is the upper bound of \(\phi_t\)). If the principal wanted to control the agent’s information rent, she would have to be able to induce the agent to reveal sensitive information such as \(p_t (K_t; \inf P_t) > p^{**}\) (when that is the case). In all MPOE such that \(\phi_{t+dt} = \phi (q_t, K_t)\), the agent would always refuse such attempt by the principal.

Intuitively the principal faces a trade off between inducing the marginal low type to invest (socially) efficiently and reducing the information rents of the higher types. The extreme asymmetry of information (assumed by our model) between the principal and agent tilts the balance toward trading with the marginal low type efficiently and tolerating the information rents earned by the higher types. In equilibrium, a higher type can anticipate this outcome, and sees the maximal information rent as his “natural entitlement”. An offer of \(\phi_t < \phi^{**}\) would be seen as an off-equilibrium path behavior by the principal and would be “rejected” (i.e., responding by \(k_t = 0\)).

\[\text{Note that } \mathcal{M} \text{ represents the beliefs of the principal corresponding to this specific equilibrium.}\]
Similarly, the agent has an incentive to protect sensitive information at the initial revelation stage. Intuitively, this force may prevent the transmission of non-trivial information about the agent’s type. Indeed, the next lemma rules out the existence of non-trivial semi-separating MPOE.

**Proposition 7** There exists no non-trivial semi-separating MPOE.

The key insight of the proof of the above proposition is that due to the initial extreme information asymmetry, the principal cannot credibly force the high types of agent to reveal sensitive information above his true type. This means that all high types of agent can credibly resist the principal’s any (counter factual) attempt to induce the agent to reveal sensitive information. Although the principal has a first mover advantage, it is entirely offset by the disadvantage of extreme informational asymmetry.

**Theorem 5** There exists a unique MPOE (allocational and distributional) outcome. For all \( t \geq 0 \), in the absence of a breakthrough, the reimbursement rate at which compensation transfer occurs is

\[
\phi_t = \phi^{**},
\]

and the investment is given by

\[
k_t = \begin{cases} 
1 & \text{if } K_t \leq K^{**} (p_0) \\ 
0 & \text{otherwise.}
\end{cases}
\]

As already established by Theorem 4, this outcome is also socially efficient (i.e., Pareto efficient, allowing compensation transfer).

**Proposition 8** For all MPOE such that the principal’s Markov strategy takes the form \( \phi_{t+dt} = \phi (K_t, q_t) \), there is a close relationship between the state variable \( q_t \) and the principal’s beliefs about the agent’s type such that for all \( t \geq 0 \) with \( K_t > 0 \)

\[
\inf P_t = \begin{cases} 
q_t & \text{on an equilibrium path,} \\
0 & \text{off equilibrium path.}
\end{cases}
\]

It follows that for any state along an equilibrium path such that \( t \geq 0 \) with \( K_t > 0 \),

\[
\inf P_t = q_t,
\]

\[
\inf P_{t+dt} = \begin{cases} 
\bar{p} (K_t, \phi^{**}) & \text{if } k_t > 0, \\
\inf P_t & \text{if } k_t = 0,
\end{cases}
\]

and

\[
\frac{d \inf P_t}{dt} = \begin{cases} 
\lim_{dt \to 0} \frac{\bar{p}(K_t, \phi^{**}) - \inf P_t}{dt} & \text{if } k_t > 0, \\
0 & \text{if } k_t = 0.
\end{cases}
\]
6 Relationship to the Literature

In this section, we discuss how the model analyzed here relates to a number of important literatures, including dynamic mechanism design with limited commitment and belief-free equilibrium. We also discuss other decision rules besides max-min that have been considered in the ambiguity literature.

6.1 Conventional Bayesian Dynamic Mechanism Design with Limited Commitment

To conserve space, we do not undertake a full review of the literature on conventional Bayesian dynamic mechanism design. This literature is voluminous (see Bolton and Dewatripont (2005)), and it is undergoing a new spurt of growth. (See, for example, Pavan et al (2011), or Doepke and Townsend (2006).) Here we focus on the problem of lack of commitment (in bilateral contracting under dynamic adverse selection), because this is the reason why an optimal long term contract cannot simply be the repetition of an optimal static contract (when the agent’s hidden type is time-invariant). Suppose that each type of the agent truthfully reveals his true type in the first period of a multiple-period model. Without commitment, there will be re-contracting between the principal and the agent in each period (due to unilateral violation of the extended optimal static contract). This then implies that the agent will not be able to enjoy any information rent after the initial period. Taking this prospect into account, the agent will not be induced by the optimal static contract to fully truthfully reveal his type. In general, the optimal dynamic contract will be affected by what is well known in the literature as the “ratchet effect”. Particularly, some degree of “pooling” may arise from equilibrium dynamic contracting.

Should the ratchet effect not arise (hypothetically), the optimal dynamic contract could have been the repetition of an optimal static contract leading to a separating equilibrium. In the presence of a ratchet effect, it becomes harder for the principal to profitably induce each type of the agent to fully truthfully reveal his true type. Depending on the strength of the ratchet effect, it may bias the equilibrium of the principal’s optimal mechanism towards more or less pooling. If one could rank the possible types of equilibrium of the model—from separating, to semi-separating (partial pooling), to pooling—then one could measure in clearly-defined terms the strength of the ratchet effect. Hence, a separating equilibrium indicates the weakest and a pooling equilibrium represents the strongest ratchet effect. Given the structure of a dynamic mechanism design (with fixed hidden agent’s type) problem, it is interesting to know how the prior belief of the principal affects the strength of the ratchet effect. Particularly, does less informative prior belief lead to stronger ratchet effect? As far as we know, this question has not been explicitly and systematically addressed in the literature. In what follows we approach this question in a particular way that will provide a partial answer. Given that our model assumes the least informative prior beliefs and induces a unique MPOE that involves initial pooling, we ask whether it is possible, if we replace the belief of the principal by a particular single-prior belief system (holding all else fixed),
that there exists a pure strategy Markov perfect Bayesian equilibrium (MPBE) such that
the initial revelation is (at least) non-trivial semi-separating. The answer is affirmative and
stated in the following proposition.

**Proposition 9** Consider the delegated experimentation model in which the principal has a
prior belief that is a uniform probability density \( \frac{1}{p^{**} + 1 - p} \) over closed intervals \([0, p^{**}]\) and
\([\bar{p}, 1]\), and zero probability (a hole) over open interval \((p^{**}, \bar{p})\). Then there exists a non-
trivial semi-separating MPBE for each of a continuum of games indexed by the parameter
\( \bar{p} \in [p^{**}, 1] \) where the equilibrium \((\bar{s}, \bar{M})\) is such that:\(^{27}\)

\[
\bar{M}_0 = \begin{cases} 
\mu | d\mu (p_0) = \begin{cases} 
\frac{1}{p^{**} + 1 - p} dp_0 & \text{for } p_0 \in [0, p^{**}] \\
0 & \text{for } p_0 \in (p^{**}, \bar{p}) \\
\frac{1}{p^{**} + 1 - p} dp_0 & \text{for } p_0 \in [\bar{p}, 1]
\end{cases}.
\end{cases}
\]

\[
\bar{s}_P (h^{0-2}) = c (\hat{p}_0) = \phi_0 \text{ for all } \hat{p}_0 \in [0, 1], \text{ where } \phi_0 = \max \left( 1 - \frac{\lambda \Pi_A}{a} \bar{p}, 0 \right),
\]

\[
\bar{s}_A (h^{0-1}, p_0) = \tilde{R} (p_0; c) = \begin{cases} 
0 & \text{if } p_0 < \bar{p} \\
\frac{1}{\bar{p}} & \text{otherwise}
\end{cases}
\]

That is, all types of agent with \( p_0 \in [\bar{p}, 1] \) pool together for initial revelation and all types of
agent with \( p_0 \in [0, \bar{p}) \) also pool together for initial revelation. And \( \forall t \geq 0 \) and \( \forall h^t \in \mathcal{H}^t \),

\[
\bar{s}_P (h^t) = \tilde{\phi}_{t+dt} (K_t, q_t) = \begin{cases} 
0 & \text{for } K_t \in [0, K^A (\max (\bar{p}, q_t))] \\
1 - \frac{\lambda \Pi_A}{a} p_{t+dt} (K_t + dt; \max (\bar{p}, q_t)) & \text{for } K_t \in \left( K^A (\max (\bar{p}, q_t)), K^{**} (\max (\bar{p}, q_t)) \right) \\
\phi^{**} & \text{for } K_t \geq K^{**} (\max (\bar{p}, q_t))
\end{cases}.
\]

(Where \( \mathcal{M}_t (h^t) \) is a singleton) and

\[
\bar{s}_A (h^t; p_0) = \tilde{k}_t (K_t, \phi_t, q_t; p_0) = \begin{cases} 
1 & \text{if } \phi_t \geq \tilde{\phi}_t \text{ and } K_t \leq \tilde{K} (\phi_t, p_0) \\
0 & \text{otherwise}.
\end{cases}
\]

All types in \([\bar{p}, 1]\) pool together initially and it becomes suboptimal for the principal to
violate the participation constraint of the (initial) marginal type, i.e., \( p_0 = \bar{p} \). In general
all higher types understand that by mimicking the behavior of the (current) marginal type
(including the initial one and each new one) they can maximize their information rents.

Note, the unique prior distribution is subjective. The principal does not have to justify
it on any objective basis (which does not have to exist). It is possible that the principal’s
subjective prior belief is wrong. Such an error, however, can never be proven by the observa-
tional evidence generated by the equilibrium described above. A subjective Bayesian model

\(^{27}\)Note that \( \mathcal{M} \), whose members are all singletons, represents the single-distribution beliefs of the principal
corresponding to this specific equilibrium. Coupling this belief system with POE allows POE to nest perfect
Bayesian equilibrium (PBE).
of dynamic game with incomplete information allows this. Features of such non-falsifiable subjective prior belief, e.g., parameter \( \bar{p} \), can play a crucial role in affecting the strength of the ratchet effect, as is illustrated by the following proposition.

Let the parameter \( \bar{p} \in [p^{**}, 1] \) represent the informativeness of the principal’s single-prior belief system. It also represents the degree of nontrivial semi-separation of the MPBE outcome. The variable \( (1 - \bar{p}) \in [0, 1 - p^{**}] \) represents the strength of the ratchet effect.

**Proposition 10** The degree of nontrivial semi-separation of the MPBE outcome is strictly increasing in the informativeness of the principal’s single-prior belief system, represented by \( \bar{p} \), and the strength of the ratchet effect is strictly decreasing in \( \bar{p} \). In the limiting case: \( \bar{p} = p^{**} \), the least informative single-prior belief system results in the largest ratchet effect: the initial revelation is pooling or trivial semi-separating, and the same as the result of the least informative multiple-prior belief system. Ceteris Paribus, \( \phi_t \) is non-increasing in \( \bar{p} \).

The above result shows that a player’s subjective pretense of knowledge (i.e., \( \bar{p} > p^{**} \)), if publicly known, can possibly influence the strategy of an opponent and hence the equilibrium outcome. Such subjective pretense of information can be false but not observationally falsifiable in an equilibrium. While such subjective belief may benefit the believer in a bargaining situation (if it happens to be compatible with the true state) by providing a means of commitment, if it is false (i.e., \( p_0 \in (p^{**}, \bar{p}) \)), such commitment becomes a source of Pareto inefficient outcome and harms the believer as well.

There is another problem with the subjective probabilistic beliefs described above—they make the bargaining problem nontrivial and lead to multiplicity of MPBE. In the unique MPOE outcome we have established, the ambiguity-averse principal’s decisions are influenced by the pessimistic expectation of zero value of surplus, making the bargaining problem trivial. As a result, a unique MPOE outcome arises. By contrast, subjective probabilistic beliefs allow the principal to entertain arbitrary expected values of surplus from trade. A Bayesian principal may expect a strictly positive value of surplus to be divided between herself and agent. As a result, the bargaining problem is not trivial and multiplicity of MPBE arises. The following proposition gives some concrete examples.

**Proposition 11** In the delegated experimentation model with a Bayesian principal, there exists a continuum of MPBE indexed by parameter \( \hat{\phi} \in (\phi^{**}, 1] \), where the equilibrium \( (\hat{s}, \hat{\mathcal{M}}) \) is such that

\[
\hat{\mathcal{M}}_0 = \left\{ \mu \left| \frac{d\mu(p_0)}{dp_0} = 1 \right. \text{ for } p_0 \in [0, 1] \right\}
\]

i.e., the set of prior beliefs is a singleton uniform distribution,

\[
\hat{s}_p \left( \hat{h}^{0,-2} \right) = c(\hat{p}_0) = \phi_0 \text{ for all } \hat{p}_0 \in [0, 1], \text{ where } \phi_0 = \hat{\phi},
\]

\[
\hat{s}_A \left( \hat{h}^{0,-1}, p_0 \right) = \hat{R}(p_0; c) = \bar{p} \left( 0, \hat{\phi} \right)
\]
(that is, the initial revelation is a pooling outcome), and for $\forall t \geq 0$ and $\forall h_t \in \mathcal{H}^t$,
\[
\hat{s}_P (h_t) = \hat{\phi}_{t+dt} (w_t) = (1 - w_t) \hat{\phi} + w_t \phi^{**}
\]
where $\hat{\phi}$ is such that
\[
u_0 \left( \frac{\hat{p} (0, \hat{\phi}) + 1}{2} | \hat{\phi}, 1, p_0 \right) \geq 0, \tag{46} \]
\[
w_t = 1 (\hat{\phi} - \phi_r)_{k_t > 0, \exists r \in [0, t]} , \tag{47} \]
and
\[
\hat{s}_A (h_t; p_0) = \hat{k}_t (K_t, \phi_t, w_t; p_0) = \begin{cases} 
1 & \text{if } \phi_t \geq \hat{\phi}, w_t = 0 \text{ and } K_t \leq \bar{K} (\phi_t, p_0) \\
1 & w_t = 1 \text{ and } K_t \leq \bar{K} (\phi_t, p_0) \\
0 & \text{otherwise.} 
\end{cases}
\]

The inequality (46) is the principal’s participation constraint at $t = 0$, which ensures that (conditional on $k_0 = 1$) the principal’s expected surplus from trade is non-negative. Note, the principal’s participation constraint (conditional on $k_t = 1$) at $t = 0$ is given by
\[
E_t \left[ u_t \left( \left( K_t, \hat{\phi} \right) \right) \right] = \int_{\hat{p} (K_t, \hat{\phi})}^{1} \frac{u_t (p_t (K_t; p_0) \hat{\phi}, 1, p_0)}{1 - \hat{p} (K_t, \hat{\phi})} dp_0,
\]
\[
= \frac{u_t (E_t [p_t (K_t; p_0) \hat{p} (K_t, \hat{\phi})] | \hat{\phi}, 1, p_0)}{1 - \hat{p} (K_t, \hat{\phi})} \frac{u_t (\frac{\hat{p} (0, \hat{\phi}) + 1}{2} | \hat{\phi}, 1, p_0)}{1 - \hat{p} (K_t, \hat{\phi})} \tag{48} \]
\[
\geq 0 \text{ if and only if } u_0 \left( \frac{\hat{p} (0, \hat{\phi}) + 1}{2} | \hat{\phi}, 1, p_0 \right) \geq 0.
\]

For $t = 0$, if $\hat{\phi} \to \phi^{**}$ then $\hat{p} (K_t, \hat{\phi}) \to p^{**}$ the left hand side of (46) must be strictly positive. Then there must exist $\hat{\phi} \in (\phi^{**}, 1]$ such that for all $\hat{\phi} \in \left( \phi^{**}, \hat{\phi} \right]$ inequality (46) is satisfied. The state variable $w_t$ is the “weakness” indicator. $w_t = 1$ means there exists historical evidence to show the agent has in the history violated the threat of rejecting any offer if $\phi_t < \hat{\phi}$. That is, the reputation for being tough on one’s bargaining position has been damaged. So $w_t$ is also a “credibility damage” indicator. The properties of $w_t$ include: irreversibility, i.e.,
\[
w_{t'} = 1 \text{ if } w_t = 1 \text{ for all } t' > t;
\]
and the Markov property. Note,
\[
w_{t+dt} = \max \left( w_t, 1 \phi_t < \phi, k_t > 0 \right)
\]
$w_{t+dt}$ only depends on $(w_t, \phi_t, k_t)$, conditional on which $w_{t+dt}$ does not depend on past history. In this MPOE, there is a ratchet effect with respect to credibility. The credibility (state variable) is not so closely linked to the type of the agent. It is essentially related to the agent’s
determination/resolution to select a particular equilibrium out of a continuum of equilibria. If the alleged resolution has been compromised in history, then it loses its power to influence the opponent’s incentive. Once the agent damages his credibility, the principal will ratchet up her own expected surplus from trade, and hence reduce the agent’s surplus by a finite positive value. For the agent, it is always worth to refrain from current investment in order to protect his credibility. The strategies of both the principal and agent are trigger strategies: the first instance (of accepting an lower offer) that damaged the agent’s credibility would trigger the principal to become very tough and the agent very soft afterwards. This strategies have the bootstrap (self sufficiency) property that makes them mutual best responses.

**Proposition 12** For $\hat{p} = p^{**}$, we have

$$\hat{\phi}_{t+dt}(w_t) \geq \phi^{**} = \hat{\phi}_{t+dt}(K_t, q_t)$$

for all $t \geq 0$; and the inequality is strict for some $t \geq 0$, particularly,

$$\hat{\phi}_0 > \phi^{**} = \tilde{\phi}_0.$$

The analysis in this section clearly demonstrates two weaknesses of PBE as a solution concept. First, it is prone to rationalizing epistemologically absurd outcomes, with beliefs that are wrong but never falsifiable in the equilibrium, as is shown by Proposition 9. Second, it is prone to multiplicity of equilibrium, as is illustrated by Proposition 11.

The comparison between the MPBE and MPOE is very revealing. While the MPOE outcome is unique, there exist a continuum of MPBE for the Bayesian version of the model with uniform priors. In practice, the uniform distribution is often treated as uninformative, at least approximately (though, in general, this is recognized to be incorrect; see, for example, Edwards 1992). The analysis above demonstrates that the set of MPBE based on (unambiguous) uniformly distributed prior belief is very different from the MPOE based on truly uninformative ambiguous beliefs. In this instance, the approximate representation of ignorance by the uniform distribution turns out to represent an “informative” (or rather a “misinformative”) pretence of knowledge that may bias the prediction of the outcome of the game. This example vindicates our claim that Bayesian model is inadequate to represent complete ignorance by some player about an opponent’s private information, and hence understates informational asymmetry.

### 6.2 Markov Perfect Equilibrium for Games with Incomplete Information

In the Markov literature, there is no established definition of Markov perfect equilibrium (MPE) for games with incomplete information. The closest related work is Maskin and Tirole (2001), which provides definitions of Markov strategy and MPE for games with observable actions (i.e., perfect or almost perfect information). According to Maskin and Tirole, “Informally, a Markov strategy depends only on payoff-relevant past events. More precisely,
it is measurable with respect to the coarsest partition of histories for which, if all other players use measurable strategies, each player’s decision-problem is also measurable.” For games with incomplete information, the evaluation of payoff depends on (posterior) belief, the inference about which depends on equilibrium. Therefore the judgement about whether a past event is payoff-relevant can also depend on equilibrium. Without specification of an equilibrium it is not clear what is meant by the “coarsest” partition of histories. Therefore the definition provided by Maskin and Tirole (2001) does not apply to games with incomplete information in general, and dynamic games (with incomplete information) under ambiguity in particular.

Our application of the concepts Markov strategy, Markov perfect objective equilibrium (MPOE) and Markov perfect Bayesian equilibrium (MPBE) proceed on a case-by-case basis, without the guidance of a generic definition. Our usage of the concept of Markov strategy implicitly requires that it must be measurable with respect to the coarsest partition of histories. For MPOE, it is impossible to reduce the list of state variables any further. The only potential candidate for reduction is $q_t$. If $q_t$ were removed, then the agent could safely behave myopically without fearing revealing sensitive information because the principal’s Markov strategy did not depend on such information. Given the agent’s myopic investment strategy, it would be optimal for the principal to experiment with reduced reimbursement rates and learn about the agent’s true type and benefit from this sensitive information. That means an MPOE must have a state variable like $q_t$, which summarizes the payoff-relevant knowledge of uninformed principal about the private information of her opponent. This state variable can capture the informational strategic effect (e.g., ratchet effect) on the players’ behavior. One potential benefit of focusing the analysis on Markov equilibrium is to avoid multiplicity of equilibrium outcomes. This is achieved by the MPOE of the delegated experimentation model (see Theorem 5). In contrast, the same cannot be said about the MPBE, as Proposition 11 demonstrates, there exist a continuum of trigger mechanisms represented by state variable $w_t$. The two state variables $q_t$ and $w_t$, representing the ratchet and trigger mechanisms respectively, are both as simple as possible. The associated state space are both the coarsest.

### 6.3 Belief-Free Equilibrium

There is a recent literature on belief-free equilibrium, which aims at characterizing equilibria of games with incomplete information that are robust to specification of beliefs (see Hörner and Lovo (2009), Hörner, Lovo and Tomala (2011)). Hörner, Lovo and Tomala (2011) define a belief-free equilibrium as a strategy profile such that, “after every history, every player’s continuation strategy is optimal, given her information, and independently of the information held by the other players. That is, it must be a subgame-perfect equilibrium for every game of complete information that is consistent with the player’s information.” Our discussion below is based on this definition. Belief-free equilibrium also relates to the concept of ex post (Nash) equilibrium (see Crémer and McLean (1985), Kalai (2004)). Bergemann and Morris (2007) introduce the notion of (static) belief-free incomplete information game, compare and
look into the epistemic foundations of three belief-free solution concepts (including belief-free equilibrium). In general, the solution concept of belief-free equilibrium applies to belief-free games with incomplete information, for which no priors are specified and keeping track of beliefs is not required.

Since our study of perfect objectivist equilibrium also pursues the goal of characterizing equilibria of games with incomplete information that are robust to specific restrictions on beliefs, it is natural to relate POE to belief-free equilibrium. It is clear that (confining to pure strategy profile in game) POE is a weaker solution concept than belief-free equilibrium. POE represents weaker ambiguity aversion by the uninformed player than in a belief-free equilibrium. The former is based on max-min plus non-weak-dominance criterion while the latter is based on the criterion of unanimous preferability (with respect to the multiple distribution belief system).

As a result, if there exists a pure strategy belief-free equilibrium for a belief-free game with incomplete information that corresponds to a dynamic game under ambiguity, then belief-free equilibrium pure strategy profile must coincide with strategy profile of a pure strategy POE. Therefore, if a pure strategy belief-free equilibrium exists, than looking for the corresponding pure strategy POE can be a useful way to approach it. As an example, the strategy profile of the pooling pure strategy MPOE in our model happens to be a pure strategy belief-free equilibrium of the corresponding belief-free game with incomplete information. The converse may not hold, i.e., the existence of a pure strategy POE may not ensure the existence of a pure strategy belief-free equilibrium.

Applying the notion of belief-free equilibrium to our model suggests that the belief-free equilibrium may be driven by the bootstrap property: the principal offers her reservation rate $\phi^{**}$ because every (sufficiently high) type of the agent always demands this reservation rate along an equilibrium path. If a (sufficiently high type) agent did not insist on this demand, he would be punished by the principal (and his future self), thereby reducing the reimbursement rate and his information rent.

This belief-free equilibrium strategy profile also coincides with the same pure strategy profile of an MPBE in a model with a Bayesian principal with all possible full-support priors. Given this result, one naturally wonders if imposing the restriction that “probabilistic beliefs should have full support” is sufficient to ensure all PBE deliver the belief-free equilib-

---

28 This belief-free equilibrium strategy profile also coincides with the same pure strategy profile of an MPBE of the strip-down model played by a Bayesian principal with all possible full-support priors. These do not include the priors discussed in Section 6.1. For the games mentioned by Proposition 9, should the combination of the strategy profile of the belief-free equilibrium and the subjective prior constitute a PBE, then zero-probability event could be observed on an equilibrium path and lead to contradiction with the prior itself, which is absurd.

For models with objectively formed ambiguous beliefs about opponents’ private information, a belief-free equilibrium (if exists) is always compatible with a POE.

29 These do not include the priors discussed in Section 6.1 with $\bar{p} > p^{**}$. For the games mentioned by Proposition 9 (conditional on $\bar{p} > p^{**}$), should the combination of the strategy profile of the belief-free equilibrium and the subjective prior constitute a PBE, then zero-probability event could be observed on an equilibrium path and lead to contradiction with the prior itself, which is absurd.
rium outcome. The answer is negative – it can be shown that even with full support, there exist other MPBE outcomes that are different from the belief-free equilibrium outcome.

6.4 Learning under Ambiguity and Dynamic Consistency

In a recent authoritative survey of the ambiguity literature, Gilboa and Marinacci (2011) suggest that Bayes’s rule can be extended to non-Bayesian beliefs in more than one ways. Two typical examples are full Bayesian updating (FBU) and maximum likelihood updating (MLU). If beliefs are given by a set of priors \( C \), and event \( B \) is known to have occurred, then the FBU set of posteriors (on \( B \)) is given by

\[
C_B = \{ p(\cdot|B) | p \in C \}.
\]

The MLU set of posteriors is given by

\[
C_B^M = \left\{ p(\cdot|B) | p \in \arg\max_{q \in C} q(B) \right\}.
\]

By definition, \( C_B^M \subseteq C_B \) for all \( B \). A sufficient condition for \( C_B^M = C_B \) for all \( B \) is that the set of priors \( C \) is large enough. More precisely, we have the following definition and proposition.

**Definition 9** The ambiguous beliefs represented by the set of priors \( C \) is sufficiently modest if for any \( B \) (which is possible to be proven to have occurred), (i) \( \max_{p \in C} p(B) = 1 \) and (ii) \( C_B \subseteq C \).

**Theorem 6** If \( C \) represents sufficiently modest beliefs, then \( C_B = C_B^M \) for all \( B \).

Condition (i) says that \( C \) is sufficiently large such that for any \( B \) (which is possible to be proven to have occurred) there must exist \( p \in C \) such that \( p(B) = 1 \). Condition (ii) says that \( C \) is sufficiently large such that the Bayesian update of any prior on any conditioning event must coincide with a prior (this can be itself or another one) belonging to \( C \).

The idea for proof is the following: If condition (i) is satisfied then for any \( B \), we have

\[
C_B^M = \{ p(\cdot|B) | p(B) = 1, p \in C \} = \{ q \in C | q(B) = 1 \}.
\]

If condition (ii) is also satisfied then for any \( B \) and any \( p(\cdot|B) \in C_B \), we have \( p(\cdot|B) \in C \), \( p(B|B) = 1 \), and hence \( p(\cdot|B) \in \{ q \in C | q(B) = 1 \} \). It follows that \( C_B \subseteq \{ q \in C | q(B) = 1 \} = C_B^M \). Since \( C_B \supseteq C_B^M \) for all \( B \), we must have

\[
C_B = C_B^M \text{ for all } B.
\]

In the modelling of learning under ambiguity in the current paper, we assume the set of the priors is \( M_0 \), which is already the largest feasible set of priors. Therefore, the condition for sufficiently modest beliefs is satisfied and we establish that FBU and MLU are equivalent.
for our model. There is another general criterion for saying \( C \) is large enough, the concept of rectangularity proposed by Epstein and Schneider (2003). The set of priors \( C \) is rectangular if “it can be decomposed into a set of current-period beliefs, coupled with next-period conditional beliefs, in such a way that any combination of the former and the latter is in \( C \)” (Gilboa and Marinacci, 2011).\(^{30}\) It is an open question as to the relationship between the sufficient modesty condition stated above and the rectangularity condition, particularly, whether inclusion relation applies one way or the other, or both.

The concept of rectangularity has another important application; i.e., it is closely related to the axiom of dynamic consistency (Epstein and Schneider, 2003). It has been recognized in the literature of dynamic models of ambiguous beliefs that the axiom of dynamic consistency is necessary for Bayes’s rule to be respected in updating of beliefs. In our model, rectangularity is satisfied and Bayes’s rule is respected in the sense that it is never violated, but also becomes redundant (or trivial) when we use maximum likelihood updating for likelihood inference. The Bayesian updates of the priors that generate the maximum likelihood are simply themselves.

### 6.5 Decision Rules under Ambiguity Other than Maxmin

It could be argued that the max-min criterion for decision under ambiguity may be too cautious or pessimistic. There is a rich variety of models of decision making under ambiguity that consider weaker representations of aversion to ambiguity. (See Gilboa and Marinacci (2011) for a survey.) In the current formulation, our definition of perfect objectivist equilibrium (POE) may seem to be tied up with the max-min criterion. However, in principle there should be no serious difficulty to extend the definition of POE to allow it to be compatible to some other well founded criterion for decision under ambiguity. For games with more than one players who have ambiguous beliefs, a general definition of POE needs to accommodate different attitudes toward ambiguity across players anyway. If there exists an associated belief-free equilibrium, then the specification of the attitudes toward ambiguity will not matter in the POE that corresponds to the belief-free equilibrium.

The definition of POE should always be tied up with the criterion that an optimal solution for player must not be weakly dominated with respect to updated beliefs. This when coupled with the max-min criterion is stronger than the max-min criterion. This additional criterion does have force: In Section A.13 of the Appendix, we give an example of a weakly dominated strategy that satisfies the max-min criterion. This criterion was originally proposed by Manski (2008). It is consistent with the axiomatic foundation for the max-min criterion provided by GMMS (2010). It is easy to see that the max-min expected utility function does not constitute a complete representation of the objective preferences of the decision maker. For instance, when the strict objectively rational preferences relation \( (\succ^*) \) applies to a pair of alternatives, the second alternative is weakly dominated by the first. This instance of the preference relation, however, cannot be adequately expressed by

\(^{30}\)We have not yet been able to identify any formal proof of the claim suggested by Gilboa and Marinacci (2011) that \( C^M_B = C_B \) if \( C \) is rectangular.
any max-min expected utility function because the latter assigns the same value to the two alternatives. It is compelling to argue that strict objectively rational preference relations should matter (for pairwise choices) whenever they are well-defined. Therefore, even when we adopt the max-min criterion, we treat it only as a necessary condition for rational decision making.

6.6 Can the POE Nest the PBE?

Technically, in the extreme case that the initial set of priors is a singleton (e.g., there exists a single objective probability), the POE degenerates to a perfect Bayesian equilibrium. In this sense, the POE can nest the PBE as a limiting case. This nesting result shows the generality of the definition of POE. POE therefore can be seen as an extension of the PBE to accommodate ambiguous beliefs. This extension is the most needed when an uninformed player lacks proof of or confidence in any single-prior belief. POE allows the solution concept not to hinge on the any subjective single-prior belief, and therefore to be robust to subjective belief specification. The POE, however, differs from the belief-free equilibrium discussed in Section 6.3 in that it keeps track of beliefs and demands equilibrium beliefs and strategies are mutually consistent after any history while the latter does not.

7 Summary and Conclusions

In this paper, we study a dynamic game of incomplete information in a model of delegated experimentation. The principal seeks a solution to a problem and contracts with an agent to find it, but a solution to the problem may not exist. Neither the principal nor the agent knows for certain whether the problem can be solved, but the agent’s expertise allows him to formulate a prior probability $p_0$ that the problem is solvable. This prior belief is private information to the agent. Unlike standard models of contracting with asymmetric information, we do not assume that the principal has a unique prior distribution over the agent’s prior. Instead, we allow the principal to hold multiple prior distributions over $p_0$. This gives rise to a process by which the principal iteratively selects beliefs over $p_0$ based on observational facts and a likelihood function that is derived from the common knowledge about the game and the concept of equilibrium. Following the work of Manski (2008) and GMMS (2010), the principal is assumed to maximize a worst-case value function that reflects an aversion to ambiguity. As a solution concept for the game, we propose an extension of perfect Bayesian equilibrium to a setting in which the uninformed principal has multiple priors. We call this extension the perfect objectivist equilibrium, or POE. We fully characterize the Markov perfect objectivist equilibrium or MPOE for the game of delegated experimentation, and show that there is a unique MPOE outcome, in which the principal offers the agent a simple linear contract that involves a time-invariant reimbursement of the agent’s costs of investment in solving the problem. This induces the agent to choose the investment profile and cumulative investment that is Pareto efficient (allowing for compensation). The reimbursement rate simply equates the share of cost to the share of benefit.
The idea that the principal may not be able to form a single prior over private information has potentially broad applicability. In principle, it allows better representation of ignorance in a player’s belief and hence more adequate account of the extent and scope of informational asymmetry between players. This is a breakthrough relative to the conventional Bayesian mechanism design literature. For example, in mechanism design models of optimal regulation (Baron and Myerson, 1982; Laffont and Tirole, 1986) or of non-linear pricing (Maskin and Riley, 1984), the principal’s (single) prior belief plays a critical role in shaping the nature of the mechanism. We intend to explore how the possibility multiple priors held by an ambiguity-averse principal changes the implications of these important benchmark models. In preliminary work (Besanko, Tong, Wu, 2011), we find that the Baron and Myerson mechanism does not emerge as a POE when the principal can hold multiple priors.

A number of extensions of the delegated experimentation model presented here are also possible, perhaps especially in the context of the R&D subsidy application. As indicated in Section 6, it would be useful to explore conditions under which standard R&D funding mechanisms, such as time-limited restricted grants, can be rationalized using the framework developed here. More generally, we hope that the concept of multiple priors and the POE can provide a valuable framework for analyzing delegated experimentation with even richer structures than the setting explored here. Among the interesting extensions in this direction would be exploration of general contract spaces, unobservable action profiles that give rise to noisy signals, noise due to measurement errors, and the possibility that a given agent may be replaced by other agents if performance in achieving a breakthrough seems unsatisfactory.

A Appendix

A.1 Proof of Proposition 1

The proof is essentially identical to that of Proposition 3.1 in Keller, Rady and Cripps (2005) and is thus omitted.

A.2 Proof of Theorem 1

Given the pure strategy equilibrium with strategy profile s, for each realization of p_o the equilibrium path is given by f^t (s, p_o) and the likelihood function is uniquely given by l (h^t; t, s, p_o) = δ_f^t (s, p_o) (h^t). Given the realized on-equilibrium-history h^t it can be inferred that all conditional plausible states belong to I_d (h^t). That is, I_d (h^t) is the conditioning event for Bayesian updating. As a result, for all μ_0 ∈ M_0 such that μ_0 (I_d (h^t)) > 0,

31 More precisely we refer to the counterpart of “perfect objectivist equilibrium” in a static game, which is called “objectivist Nash equilibrium” (ONE), as an extension to conventional “Bayesian Nash equilibrium”.

45
the posterior $\mu_t (h^t; \mu_0, l)$ is well-defined and for all $A \subseteq \mathcal{I}_d (h^t)$ we have

$$
\mu_t (A | h^t; \mu_0, l) = \frac{\int_{p_0 \in A} \delta_{f^t(s,p_0)} (h^t) \, d\mu_0 (p_0)}{\int_0^1 \delta_{f^t(s,p_0)} (h^t) \, d\mu_0 (\tilde{p}_0)} = \frac{\int_{p_0 \in A} \delta_{f^t(s,p_0)} (h^t) \, d\mu_0 (p_0)}{\mu_0 (\mathcal{I}_d (h^t))},
$$

which implies

$$
\mu_t (\mathcal{I}_d (h^t) | h^t; \mu_0, l) = \frac{\int_{p_0 \in \mathcal{I}_d (h^t)} \delta_{f^t(s,p_0)} (h^t) \, d\mu_0 (p_0)}{\mu_0 (\mathcal{I}_d (h^t))} = 1.
$$

Since $\mathcal{M}_0$ includes all plausible probability distribution over $\Sigma$, i.e., $\mathcal{M}_0$ is sufficiently large, we must have $\mu_t (h^t; \mu_0, l) \in \mathcal{M}_0$. This implies for all $\alpha \in [0, 1]$

$$
\mathcal{M}_t^\alpha (h^t) \subseteq \{ \mu_0 \in \mathcal{M}_0 | \mu_0 (\mathcal{I}_d (h^t)) = 1 \}.
$$

For any on-equilibrium-path history $h^t$, we have

$$
\max_{\tilde{p}_0 \in \mathcal{M}_0} L \left( h^t; t, s, \tilde{p}_0, \tilde{l} \right) = \max_{\tilde{p}_0 \in \mathcal{M}_0} \int_0^1 \delta_{f^t(s,p_0)} (h^t) \, d\tilde{p}_0 (p_0')
$$

$$= \int_0^1 \delta_{f^t(s,p_0)} (h^t) \, d\tilde{p}_0 (p_0') \bigg|_{p_0 \in \mathcal{I}_d (h^t)} = \delta_{f^t(s,p_0)} (h^t) \bigg|_{p_0 \in \mathcal{I}_d (h^t)} = \delta_{h^t} (h^t) = 1.
$$

As a result, for all $\mu_0 \in \mathcal{M}_0$ such that $\mu_0 (\mathcal{I}_d (h^t)) = 1$, we have

$$
\mu_t (h^t; \mu_0, l) = \mu_0,
$$

for $l \in \mathcal{L}_0 (s)$, and

$$
L \left( h^t; t, s, \mu_0, l \right) = 1 \geq \alpha \max_{\tilde{p}_0 \in \mathcal{M}_0 \atop l \in \mathcal{L}_0 (s)} L \left( h^t; t, s, \tilde{p}_0, \tilde{l} \right) = \alpha.
$$

It follows that

$$
\mathcal{M}_t^\alpha (h^t) \supseteq \{ \mu_0 \in \mathcal{M}_0 | \mu_0 (\mathcal{I}_d (h^t)) = 1 \}.
$$

Overall, we have for all $\alpha \in [0, 1]$

$$
\mathcal{M}_t^\alpha (h^t) = \{ \mu_0 \in \mathcal{M}_0 | \mu_0 (\mathcal{I}_d (h^t)) = 1 \}.
$$

### A.3 Proof of Theorem 2

Let $\mu \in \mathcal{A}$, by definition

$$P (\mathcal{A}) = \bigcup_{\mu \in \mathcal{A}} \{ x \in [0, 1] | \mu (x) > 0 \}$$

and $cl (\mathcal{A})$ includes all $\mu \in \mathcal{A}$ and all limit points $\mu'$ of $\mathcal{A}$. That is $\mu' = \lim_{n \to \infty} \mu_n$ where $\mu_n \in \mathcal{A}$. Hence

$$P (\mu') = \left\{ x \in [0, 1] | \lim_{n \to \infty} \mu_n (x) > 0 \right\}.$$
As a preliminary step, we establish that if \( \mathcal{A} \) includes the set of all Dirac measures over support \( P(\mathcal{A}) \), then
\[
P(\text{cl} (\mathcal{A})) = \text{cl} (P(\mathcal{A}))
\]

First, we show \( \text{cl} (P(\mathcal{A})) \subseteq P(\text{cl} (\mathcal{A})) \).

Let \( x \in \text{cl} (P(\mathcal{A})) \). If \( x \in P(\mathcal{A}) \), then there exists \( \mu \in \mathcal{A} \) such that \( \mu (x) > 0 \). Because \( \mu \in \text{cl} (\mathcal{A}) \), we must have \( x \in P(\text{cl} (\mathcal{A})) \).

Suppose \( x \in \text{cl} (P(\mathcal{A})) \setminus P(\mathcal{A}) \), i.e., \( x \) is a boundary point of \( P(\mathcal{A}) \), then there exists \( x_N \in P(\mathcal{A}) \) and \( x_N \to x \). Define \( \mu_N = \delta_{x_N} \in \mathcal{A} \), we must have \( \mu_N \to \mu = \delta_x \) (convergence by distribution), which implies \( \mu \in \text{cl} (\mathcal{A}) \) and thus \( x \in P(\text{cl} (\mathcal{A})) \).

Second, we show \( P(\text{cl} (\mathcal{A})) \subseteq \text{cl} (P(\mathcal{A})) \).

Let \( x \in P(\text{cl} (\mathcal{A})) \), then there exists \( \mu \in \text{cl} (\mathcal{A}) \) such that \( \mu (x) > 0 \). If \( \mu \in \mathcal{A} \), then we must have \( x \in P(\mathcal{A}) \subseteq \text{cl} (P(\mathcal{A})) \). Now suppose \( \mu \notin \mathcal{A} \), then \( \mu \) is a boundary point of \( \mathcal{A} \) and there exists \( \mu_N \in \mathcal{A} \) such that \( \mu_N \to \mu \) (convergence by distribution). If \( \mu_N (x) > 0 \), then \( x \in P(\mathcal{A}) \subseteq \text{cl} (P(\mathcal{A})) \). If \( \mu_N (x) = 0 \) for all \( N \), then we must have \( \lim_{N \to \infty} \mu_N (x) = 0 \). Then \( \mu_N \to \mu \) implies \( x \) is a mass point for \( \mu \).

Suppose \( x \notin \text{cl} (P(\mathcal{A})) \), that is, \( x \) is an exterior point of \( P(\mathcal{A}) \). Then there must exist \( \varepsilon > 0 \) such that \( (x - \varepsilon, x + \varepsilon) \notin P(\mathcal{A}) \) and \( F_\mu \) is continuous at \( (x - \varepsilon, x) \) or \( (x, x + \varepsilon) \), where the CDF \( F_\mu \) is defined by
\[
F_\mu (x') = \int_0^{x'} d\mu (z)
\]
for all \( x' \in [0, 1] \).

We distinguish three cases: (a) \( x = 0 \); (b) \( x \in (0, 1) \); (c) \( x = 1 \).

(a) \( x = 0 \): There must exist \( \varepsilon > 0 \) such that \( (0, \varepsilon) \notin P(\mathcal{A}) \) and \( F_\mu (x') \) is continuous at \( x' \) for all \( x' \in (0, \varepsilon) \). On the one hand,
\[
\lim_{N \to \infty} F_{\mu_N} (x') = F_\mu (x') > 0.
\]
On the other hand, since \( F_{\mu_N} (x') = F_{\mu_N} (0) = 0 \) for all \( N \in \mathbb{N} \), we must have
\[
\lim_{N \to \infty} F_{\mu_N} (x') = 0.
\]
The above is a contradiction, which implies case (a) is empty.

(b) \( x \in (0, 1) \): There must exist \( \varepsilon > 0 \) such that \( (x - \varepsilon, x + \varepsilon) \notin P(\mathcal{A}) \), \( F_\mu (x') \) is continuous at \( x' \) for all \( x' \in (x - \varepsilon, x) \) and \( F_\mu (x'') \) is continuous at \( x'' \) for all \( x'' \in (x, x + \varepsilon) \). On the one hand,
\[
\lim_{N \to \infty} F_{\mu_N} (x') = F_\mu (x') < F_\mu (x'') = \lim_{N \to \infty} F_{\mu_N} (x'').
\]
On the other hand, since \( F_{\mu_N} (x) = F_{\mu_N} (x') = F_{\mu_N} (x'') \), we must have
\[
\lim_{N \to \infty} F_{\mu_N} (x) = \lim_{N \to \infty} F_{\mu_N} (x') = \lim_{N \to \infty} F_{\mu_N} (x'').
\]
We have therefore derived a contradiction which implies that case (b) is also empty.

(c) $x = 1$: There must exists $\varepsilon > 0$ such that $(1 - \varepsilon, 1) \not\subseteq P(A)$ and $F_{\mu} (x')$ is continuous at $x'$ for all $x' \in (1 - \varepsilon, 1)$. On the one hand,

$$\lim_{N \to \infty} F_{\mu_N} (x') = F_{\mu} (x') < 1.$$ 

On the other hand, since $F_{\mu_N} (x') = F_{\mu_N} (1) = 1$, we must have

$$\lim_{N \to \infty} F_{\mu_N} (x') = \lim_{N \to \infty} F_{\mu_N} (1) = 1.$$ 

We have therefore derived a contradiction which implies that case (c) is empty too.

We therefore have established that $x \in cl (P(A))$.

Note, By Theorem 1, $\mathcal{M}_t (h^t)$ is the set of all probability measures over the set $\mathcal{I}_d (h^t)$, where $\mathcal{I}_d (h^t)$ is the set of $p_0$ such that $h^t$ is predicted by $f^t (s; p_0)$. This implies

$$P (\mathcal{M}_t (h^t)) = \mathcal{I}_d (h^t),$$

$$P (\mathcal{D}_t (h^t)) = \mathcal{I}_d (h^t),$$

and that $\mathcal{D}_t (h^t)$ is the set of all Dirac measures over the set $\mathcal{I}_d (h^t)$.

Let $A = \mathcal{M}_t (h^t)$ and $A = \mathcal{D}_t (h^t)$ respectively, we therefore establish respectively that

$$P (cl (\mathcal{M}_t (h^t))) = cl (P (\mathcal{M}_t (h^t))),$$

$$P (cl (\mathcal{D}_t (h^t))) = cl (P (\mathcal{D}_t (h^t))).$$

Next we note that $cl (\mathcal{D}_t (h^t))$ is the set of all Dirac measures over the set $cl (P (\mathcal{D}_t (h^t)))$. This is obvious because by definition $cl (\mathcal{D}_t (h^t))$ is the set of all Dirac measures over the set $P (cl (\mathcal{D}_t (h^t)))$.

Now we prove the main claims. First, (i)$\Leftrightarrow$(ii) is due to the fact that $cl (\mathcal{M}_t (h^t))$ includes $cl (\mathcal{D}_t (h^t))$, which contains all Dirac measures over $cl (P (\mathcal{M}_t (h^t)))$. All members of $cl (\mathcal{M}_t (h^t))$, including all solutions of the minimization problem in expression (i), are a convex combination of some Dirac measures over $cl (P (\mathcal{M}_t (h^t)))$. There must exist a corner solution to the minimization problem in expression (i), that is a Dirac measure that belongs to $cl (\mathcal{D}_t (h^t))$ and also minimizes $W_t (p_t (K_t; p_0) | s, p_0)$. To see this, let $\tilde{\mu}_0 \in cl (\mathcal{M}_t (h^t))$ be a worst-case distribution. Let $P (\tilde{\mu}_0)$ be the support of $\tilde{\mu}_0$. Then there must exist $\tilde{p}_0 \in P (\tilde{\mu}_0)$ such that $\delta_{\tilde{p}_0}$ is also a worst-case distribution. Suppose the opposite, that is,

$$W_t (p_t (K_t; \tilde{p}_0) | s, \tilde{p}_0) > \int_0^1 W_t (p_t (K_t; p_0) | s, p_0) d\tilde{\mu}_0 (p_0)$$

for all $\tilde{p}_0 \in P (\tilde{\mu}_0)$. Integrating both sides of the above inequality w.r.t. $\tilde{\mu}_0 (\tilde{p}_0)$ over $P (\tilde{\mu}_0)$, we have

$$\int_{\tilde{p}_0 \in P (\tilde{\mu}_0)} W_t (p_t (K_t; \tilde{p}_0) | s, \tilde{p}_0) d\tilde{\mu}_0 (\tilde{p}_0) > \int_0^1 W_t (p_t (K_t; p_0) | s, p_0) d\tilde{\mu}_0 (p_0) \Leftrightarrow \int_0^1 W_t (p_t (K_t; p_0) | s, p_0) d\tilde{\mu}_0 (p_0) > \int_0^1 W_t (p_t (K_t; p_0) | s, p_0) d\tilde{\mu}_0 (p_0)$$

48
which is a contradiction. Since \( \delta_{\tilde{p}_0} \in cl(\mathcal{D}_t(h^t)) \subseteq cl(\mathcal{M}_t(h^t)) \), \( \delta_{\tilde{p}_0} \) must also solve the minimization problem in expression (ii). As a result, (i) ⇔ (ii).

Finally, (ii) ⇔ (iii) is obvious because \( cl(\mathcal{D}_t(h^t)) \) is the set of all Dirac measures over the set \( cl(\mathcal{M}_t(h^t)) \).

### A.4 Proof of Proposition 2

The proof of Theorem 1 has already established the following:

\[
P(cl(\mathcal{M}_t(h^t))) = cl(P(\mathcal{M}_t(h^t)))
\]
\[
P(\mathcal{M}_t(h^t)) = P(\mathcal{D}_t(h^t))
\]
\[
P(cl(\mathcal{D}_t(h^t))) = cl(P(\mathcal{D}_t(h^t)))
\]

The rest of the result follows immediately from Definition 3.

### A.5 Proof of Lemma 1

The result follows immediately from Definition 5.

### A.6 Proof of Lemma 2

Given the agent’s strategy \( \tilde{k}_t^* (K_t; \phi_t; p_0) \), the principal chooses \( \tilde{\phi}_{t+dt}^* (K_t; p_0) \) and induces the agent to follow the socially efficient investment strategy specified by (27). Hence \( \tilde{\phi}_{t+dt}^* (K_t; p_0) \) must be optimal because it allows the principal to extract the maximum surplus.

To show the necessity, we note that for \( K_t \in [0, K^A(p_0)] \), the principal must set \( \tilde{\phi}_t^* (K_t; p_0) = 0 \) otherwise she can increase her payoff by lowering \( \tilde{\phi}_t^* (K_t; p_0) \) without changing the agent’s investment level. Further more, we know

\[
\lim_{K_t \to K^{**}(p_0)} \tilde{\phi}_t^* (K_t; p_0) = \phi^{**}
\]

where \( K^{**}(p_0) \) is the optimal termination threshold value of \( K_t \). This implies for \( K_t \in (K^A(p_0), K^{**}(p_0)) \), we also have \( \tilde{\phi}_t^* (K_t; p_0) = \tilde{\phi}_t (K_t; p_0) \). Finally, for \( K_t \geq K^{**}(p_0) \), we need to have \( \tilde{\phi}_t^* (K_t; p_0) \leq \phi^{**} \) otherwise the agent will overinvest. As a result,

\[
\tilde{\phi}_{t+dt}^* (K_t; p_0) = \begin{cases} 
0 & \text{if } K_t \in [0, K^A(p_0)] \\
1 - \frac{\lambda t}{a} p_{t+dt}(K_t + dt; p_0) & \text{if } K_t \in (K^A(p_0), K^{**}(p_0)) \\
\leq \phi^{**} & \text{if } K_t \geq K^{**}(p_0)
\end{cases}
\]

is a necessary condition for \( \tilde{\phi}_{t+dt}^* (K_t; p_0) \) to be optimal.
A.7 Proof of Lemma 3

Lemma 2 implies that the compensation strategy \( \tilde{\phi}_{t+dt}^* (K_t; p_0) \) can induce the investment policy given by (32), keep the agent’s value the same as if there is no compensation transfer, and strictly improve and maximize the principal’s value for all \( K_t \geq 0 \) in the absence of a breakthrough. The outcome is obviously Pareto efficient for all \( K_t \geq 0 \) in the absence of a breakthrough.

A.8 Proof of Proposition 3

Preliminary claim (i): If there exists an MPOE such that \( \inf P_0 \leq p^{**} < \sup P_0 \), and the principal’s Markov strategy takes the form \( \phi_{t+dt} = \phi(K_t, q_t) \), then \( \phi_0 = \phi(0, q_0) = \phi^{**} \) for all state \((0, q_0)\) that is on an equilibrium path.

Suppose \( \phi_{t+dt} = \phi(K_t, q_t) \) and \( k_t = k(K_t, q_t; p_0) \) are the equilibrium Markov strategies of the principal and agent respectively. If the state (event) \((K_{t-dt}, q_{t-dt})\) is on an equilibrium path, then the offer \( \phi_t = \phi(K_{t-dt}, q_{t-dt}) \) must be a credible take-it-or-leave-it offer, i.e., it is genuinely not renegotiable. Formally, \( k_t = 0 \) implies \( dq_t/dt = 0 \) and \( dK_t/dt = 0 \), and it follows \( \frac{d\phi_t}{dt} = 0 \). In this case, \( (K_{t'}, \phi_{t'}, q_{t'}) \) becomes stationary for \( t' \geq t \). In this case, the agent would reject the offer only if the net flow benefit is negative. Therefore, if \( k_t = 0 \) then it must be legitimate to infer that

\[ p_0 < \bar{p}(K_t, \phi_t) \]

and hence

\[ \sup_{P_{t+dt}} P_0 = \bar{p}(K_t, \phi_t). \]

Suppose \( \phi_0 < \phi^{**} \) in an MPOE, then \( k_0 = 0 \) implies \( dq_0/dt = 0 \) and \( dK_0/dt = 0 \), and it follows \( \frac{d\phi_0}{dt} = 0 \). \((K_t, \phi_t, q_t)\) becomes stationary for \( t > 0 \). This strategy is not sequentially rational given that \( p_0 < \bar{p}(\phi_0, 0) \) because \( \phi_{t+dt} = \phi^{**} \) weakly dominates \( \phi_{2dt} = \phi_0 \) for \( p_0 \in [0, \bar{p}(\phi_0, 0)) \). Note the former strictly dominates the latter for \( p_0 \in [p^{**}, \bar{p}(\phi_0, 0)) \), while they are equivalent for \( p_0 \in [0, p^{**}] \). Thus the Markov strategy \( \phi(0, q_0) = \phi_0 \) cannot be an equilibrium strategy.

Now suppose \( \phi_0 > \phi^{**} \), then the type \( p_0' = p^{**} - \varepsilon \) would have incentive to invest, but the social surplus is negative and the type \( p_0' \) agent at least breaks even, the expected loss must be borne by the principal. The worst case must be even worse, and the worst case value function for the principal must be negative. For all types \( p_0 \geq p^{**}, \phi_0 = \phi^{**} \) is sufficient to induce \( k_0 = 1 \) there \( \phi_0 > \phi^{**} \) is unnecessarily too high. In conclusion, \( \phi_0 > \phi^{**} \) must be suboptimal.

Preliminary claim (ii): If there exists an MPOE such that \( \inf P_0 \leq p^{**} < \sup P_0 \), and the principal’s Markov strategy takes the form \( \phi_{t+dt} = \phi(K_t, q_t) \), then \( \phi_{t+dt} = \phi(K_t, q_t) = \phi^{**} \) for all state \((K_t, q_t)\) that is on an equilibrium path.

Suppose \( \phi_t < \phi^{**} \) for some \( t > 0 \), and \( \phi_{\tau} = \phi^{**} \) and \( k_\tau > 1 \) for all \( \tau < t \). Then \( \phi(t - dt, \bar{p}(\phi^{**}, t - dt)) < \phi^{**} \). Then \( k_t = 0 \) implies \( dq_t/dt = 0 \) and \( dK_t/dt = 0 \), and it follows \( \frac{d\phi_t}{dt} = 0 \). As a result, \((K_{t'}, \phi_{t'}, q_{t'})\) becomes stationary for \( t' \geq t \). This strategy is not sequentially rational given that \( p_0 < \bar{p}(\phi_t, t) \) because \( \phi_{t+2dt} = \phi^{**} \) weakly dominates \( \phi_{t+2dt} \).
for \( p_0 < \bar{p}(\phi_t, t) \). Thus the Markov strategy \( \phi(t, \bar{p}(\phi^**, t)) < \phi^** \) cannot be an equilibrium strategy.

Now suppose \( \phi_t > \phi^** \), then the worst case value function for the principal is negative, which is suboptimal. Also, for all types such that \( p_t(K_t; p_0) \geq p^** \), \( \phi_t = \phi^** \) is sufficient to induce \( k_t = 1 \) there \( \phi_t > \phi^** \) is suboptimally too high.

Proof of the main claim: The first part has been established by preliminary claims (i) and (ii). The second part is due to the fact that \( k_t = 1 \) means \( k_t \) is time-invariant and not dependent on \( q_t \). Consequently, \( \frac{\partial V}{\partial q_t} = 0 \). The third part is due to the fact that if and only \( k_t > 0 \) we have

\[
\bar{p}(K_t, \phi_t) = \bar{p}(K_{t-d} + k_t dt, \phi^**) > \bar{p}(K_{t-dt}, \phi^**) = q_t.
\]

### A.9 Proof of Lemma 4

Trivial.

### A.10 Proof of Lemma 5

Claim: \( \dot{\phi}_{t+dt}^**(K_t, q_t) \) defined by (37) satisfies the condition: \( \frac{d\nu}{dt} > 0 \) only if \( \phi(K_t, q_t) \geq \phi^** \) conditional on that the principal never deviates from strategy \( \phi^*_\tau(K_{\tau-dt}, q_{\tau-dt}) \) for \( \tau \geq t \).

To verify this, note that for \( K_t \in [K^A(q_t), K^*(q_t)] \) and \( k_t = 1 \)

\[
q_t = \bar{p}(K_t + dt, \phi^*_t) \geq \bar{p}(K_t + dt, \phi^**).
\]

The equality follows from the definition of strategy \( \dot{\phi}_{t+dt}^** \). The inequality follows from the fact \( K_{t+dt} = K_t + k_t dt \leq K_t + dt \). In the limit we have

\[
q_t \geq \bar{p}(K_t, \phi^**).
\]

It then follows that \( \frac{d\nu}{dt} = 0 \) (even if \( k_t = 1 \)).

For \( K_t \in [0, K^A(q_t)] \), we have

\[
K_t < K^A(q_t)
\]

which implies

\[
\bar{p}(K_t, 0) < \bar{p}(K^A(q_t), 0) = q_t,
\]

which implies \( \frac{d\nu}{dt} = 0 \) for \( K_t \geq K^*(q_t) \), \( \phi(K_t, q_t) = \phi^** \).

In conclusion, we have either \( \frac{d\nu}{dt} = 0 \) or \( \phi(K_t, q_t) = \phi^** \). As a result, \( \frac{d\nu}{dt} > 0 \) only if \( \phi(K_t, q_t) \geq \phi^** \).

Next we verify that \( \dot{\phi}_{t+dt}^**(K_t, q_t) \) and \( k_t^*(K_t, \phi_t, q_t; p_0) \) are mutual best responses following for all \( t \geq 0 \) histories. We will do this in two steps: (i) we verify this under strong condition that the principal never deviates from strategy \( \phi^*_\tau(K_{\tau-dt}, q_{\tau-dt}) \) for \( \tau \geq t \). (ii) we show this under the weak condition that the principal never deviates from strategy \( \phi^*_\tau(K_{\tau-dt}, q_{\tau-dt}) \) for \( \tau > t \), but temporary deviation at \( \tau = t \) occurs, i.e., \( \phi_t \neq \phi_t^*(K_{t-dt}, q_{t-dt}) \).

51
Step (i). Since the condition: \( \frac{dq}{dt} > 0 \) only if \( \phi(K_t, q_t) \geq \phi^{**} \) is satisfied under the strong condition, it then follows from Lemma 4 that \( k_t^{**}(K_t, \phi_t, q_t; p_0) \) as defined in (36) is the agent’s best response.

To verify that \( \phi_{t+dt}^{**}(K_t, q_t) \) is a best response to \( k_t^{**}(K_t, \phi_t, q_t; p_0) \) as defined in (36), note that \( \phi_{t+dt}^{**}(K_t, q_t) \) minimizes \( \phi_{t+dt} \) subject to the constraint: \( \frac{dq}{dt} = 0 \) or \( \frac{dq}{dt} > 0 \) and \( \phi_{t+dt} \geq \phi^{**} \). Along an equilibrium path, \( \phi_{t+dt}^{**}(K_t, q_t) = \phi^{**} \) and the strategy \( \phi_{t+dt}^{**}(K_t, q_t) \) induces the agent to invest in the socially efficient way (allowing for compensation transfer).

The strategy thus maximizes social surplus and subject to this and that the agent accepts the offer, it also minimizes the agent’s surplus, and hence is optimal for the principal. Off equilibrium path, if \( \phi_{t+dt}^{**}(K_t, q_t) = \phi^{**} \), then by the same argument, we can claim that the strategy is is optimal for the principal. If \( \phi_{t+dt}^{**}(K_t, q_t) < \phi^{**} \), then the strategy \( \phi_{t+dt}^{**}(K_t, q_t) \) minimizes \( \phi_{t+dt} \) subject to that the agent only behaves myopically. Since \( \inf P_t = 0 \) off equilibrium path, it is easy to see that the strategy \( \phi_{t+dt}^{**}(K_t, q_t) \) maximizes the worst case value for the principal and is not weakly dominated. Thus, it is optimal for the principal off equilibrium path as well.

Under the strong condition, given the principal’s strategy \( \phi_{t+dt}^{**}(K_t, q_t) \), (36) and (38) are equivalent. To verify this, we only need to show that both formulations of \( k_t^{**}(K_t, \phi_t, q_t; p_0) \) imply the identical outcome following any feasible payoff relevant state under the strong condition. This is exactly the case because both formulations imply that the agent would invest as if he were myopic given \( \phi_t = \phi_t^{**}(K_{t-dt}, q_{t-dt}) \), as well as identical compensation transfer from the principal to the agent - minimizing \( \phi_{t+dt} \) subject to the constraint that \( \frac{dq}{dt} > 0 \) only if \( \phi(K_t, q_t) \geq \phi^{**} \).

Step (ii). Suppose the principal never deviates from strategy \( \phi_{t+dt}^{**}(K_t, q_t) \) for \( \tau > t \), but for \( \tau = t \), \( \phi_t \neq \phi_{t+dt}^{**}(K_{t-dt}, q_{t-dt}) \) (the deference must be finite\(^{32}\)).

The definition of \( \phi_{t+dt}^{**}(K_{t-dt}, q_{t-dt}) \) entails that \( \phi_{t+dt}^{**}(K_{t-dt}, q_{t-dt}) \leq \phi^{**} \).

First, consider \( \phi_t < \phi_{t+dt}^{**}(K_{t-dt}, q_{t-dt}) \). It follows

\[
\bar{p}(K_t, \phi_t) > \bar{p}(K_t, \phi_{t+dt}^{**}) = q_t
\]

and thus \( \frac{dq}{dt} \to \infty \) if \( k_t > 0 \). That is, if \( k_t > 0 \) the state variable \( q_t \) would increase by a finite positive real number for a time period of finite positive length (in the absence of a breakthrough). This would mean a reduction of the agent’s information rent by a finite positive real number. The alternative choice by the agent is \( k_t = 0 \), which would keep \( (K_t, q_t) \) unchanged between time \( t \) and \( t+dt \), consequently no reduction of information rent. The only loss of the alternative choice is due discounting of the project value by the factor \( e^{-\tau dt} \). The loss is approximately by a fraction of \( rdt \to 0 \), clearly outweighed by the finite reduction of project value following the choice \( k_t > 0 \). The optimal choice then must be \( k_t = 0 \), either \( K_t > \bar{K}(\phi_t, p_0) \) or \( K_t \leq \bar{K}(\phi_t, p_0) \).

Second, consider \( \phi_t < \phi_{t+dt}^{**}(K_{t-dt}, q_{t-dt}) \geq 0 \). For \( K_t \in [0, K^A(q_t)] \), we have

\[
K_t < K^A(q_t)
\]

\(^{32}\)Deviation by infinitesimal quantity is not considered here based on the following argument: since an infinitesimal deviation is not well-defined and hence the offer \( \phi_{t+dt} \) is not well-defined. The deviation needs to be uniformly defined with respect to the not well-defined \( dt \).
which implies

$$q_t > \bar{p}(K_t, 0) > \bar{p}(K_t, \phi_t),$$

and hence $\frac{dn}{dt} = 0$. Then $k_t = 0$ is optimal if and only if $K_t > \bar{K}(\phi_t, p_0)$.

For $K_t \in [K^A(q_t), K^{**}(q_t)]$, we have

$$q_t = \bar{p}(K_t, \phi_t^{**}) > \bar{p}(K_t, \phi_t)$$

which implies $\frac{dn}{dt} = 0$. Then $k_t = 0$ is optimal if and only if $K_t > \bar{K}(\phi_t, p_0)$.

For $K_t \geq K^{**}(q_t)$, we have

$$q_t = \bar{p}(K_t, \phi_t^{**}) = \bar{p}(K_t, \phi_t^{**}) > \bar{p}(K_t, \phi_t),$$

which implies $\frac{dn}{dt} = 0$. Then $k_t = 0$ is optimal if and only if $K_t > \bar{K}(\phi_t, p_0)$.

In summary, conditional on $\phi_t \neq \phi_t^{**}(K_{t-dt}, q_{t-dt})$ the following statement holds:

(a) $k_t = 0$ is optimal if and only if $[K_t > \bar{K}(\phi_t, p_0), \text{ or } K_t \leq \bar{K}(\phi_t, p_0) \text{ and } \phi_t < \phi_t^{**}(K_{t-dt}, q_{t-dt})]$. (a) is equivalent to statement

(b): $k_t = 1$ is optimal if and only if $[K_t \leq \bar{K}(\phi_t, p_0) \text{ and } \phi_t > \phi_t^{**}(K_{t-dt}, q_{t-dt})]$. Thus, conditional on $\phi_t \neq \phi_t^{**}(K_{t-dt}, q_{t-dt})$, (b) is equivalent to (c): the agent’s strategy as defined by (38) is optimal.

Condition $\phi_t \neq \phi_t^{**}(K_{t-dt}, q_{t-dt})$ is equivalent to $[\phi_t > \phi_t^{**}(K_{t-dt}, q_{t-dt}) \text{ or } \phi_t < \phi_t^{**}(K_{t-dt}, q_{t-dt})]$, which is equivalent to

$$[[\phi_t > \phi_t^{**}(K_{t-dt}, q_{t-dt}) \text{ and } q_t \geq \bar{p}(K_t, \phi_t)] \text{ or } [\phi_t < \phi_t^{**}(K_{t-dt}, q_{t-dt}) \text{ and } q_t < \bar{p}(K_t, \phi_t)]]].$$

Thus, under the weak condition, given the principal’s strategy $\phi_t^{**}(K_t, q_t)$, (36) and (38) are equivalent, which hence implies that (36) is optimal as well.

In conclusion, given the principal’s strategy $\phi_t^{**}(K_t, q_t)$, (36) and (38) are equivalent; and $k_t^{**}(K_t, \phi_t, q_t; p_0)$, as defined by (36), is sequentially rational following any history of any $t \geq 0$, i.e., under both strong and weak conditions.

Given the agent’s strategy $k_t^{**}(K_t, \phi_t, q_t; p_0)$ as defined by (38), it is obvious that the principal’s strategy $\phi_t^{**}(K_t, q_t)$ minimizes $\phi_{t+dt}(K_t, q_t)$ subject to that the agent only behaves myopically. Since $\inf P_t = 0$ off equilibrium path, it is easy to see that the strategy $\phi_t^{**}(K_t, q_t)$ maximizes the worst case value for the principal and is not weakly dominated. Thus, it is optimal for the principal off equilibrium path under the weak condition.

A.11 Proof of Theorem 3

To establish $(s^{**}, M^{**})$ constitutes a POE, we need to establish part (i) and (ii) of Definition 5. Part (ii) has been established by discussion in Section 3.1.3 and Section 3.1.4 and Definition 7. Hence we are left to prove part (i), which requires us to establish $s_t^{**}$ and $s_t^{***}$ are mutual best responses for all $t \in \{0_2, 0_1\}$ and $[0, \infty)$. We will proceed in three steps.

Claim (i): $s_t^{**}$ and $s_t^{***}$ are mutual best responses for all $t \geq 0$ and all $h' \in \mathcal{H}'$.

This claim is an immediate implication of Lemma 5.

Claim (ii): $s_t^{**} (h^{0_1}, p_0) = \inf \arg \max_{\hat{p}_0 \in [0,1]} \Phi (\hat{p}_0)$. 

53
It is optimal for each type of the agent to choose the highest $\phi_0$ in the menu $\Phi$. If all types do so, the initial revelation is a pooling outcome and $q_0 = \inf P_0 = 0$. It is optimal to have $q_0 \leq p^{**}$ (in order to avoid ratcheting up $q_t$ for $t \geq 0$), and $s_A^{**}(h^{0-1}, p_0)$ satisfies this inequality.

Claim (iii): For all $t = 0_{-2}$, $s_A^{**}$ is a best response to $s_A^{**}$.

Note,
\[
\min_{p_0 \in [0,1]} W_{0-2} (p_0 | s_P^{**}, s_A^{**}, p_0) = 0 = \max_{s_P \in S_P} U_{0-2} (h^{0-2} | s_P, s_A^{**}) = \max_{s_P \in S_P} \min_{p_0 \in [0,1]} W_0 (p_0 | s_P, s_A^{**}, p_0).
\] (49)

Also, we claim given $s_A^{**}$, $s_P^{**}$ is never conditionally weakly dominated by any $s_P \in S_P$. To see this, suppose the opposite, that is, there exists $s_P \in S_P$, such that
\[
(s_P, s_A^{**}) \succ^*_{M_{0-2}} (s_P^{**}, s_A^{**}).
\]

Then, given that $s_A^{**}$ is (by definition) a best response to $s_P^{**}$ for all $h^{0-1}$ at $t = 0_{-1}$, $s_A^{**}(h^{0-1}, p_0) = \inf \arg \max_{p_0 \in [0,1]} \Phi(\hat{p}_0)$ applies. The $t = 0$ continuation game will always be based on a pooling initial revelation. Then $s_P^{**}(h^t) = \phi^{**}$ is unanimously optimal, which implies $s_P^{**}(h^t) = \phi^{**}$ cannot be conditionally weakly dominated by $s_P$, and thus a contradiction is delivered.

A.12 Proof of Theorem 4

For $\phi = \phi^{**}$, the termination investment level is given by $K(\phi, p_0) = K^{**}(p_0)$. The pooling equilibrium given by Definition 7 therefore generates the same investment policy (i.e., given by (32)) that is induced by compensation strategy $\tilde{\phi}_{t+dt} (K_t, p_0)$ under full information. Since the investment policy is a Pareto efficient allocation of resource, the difference between the two outcomes in question is purely distributional and due to the difference between the two compensation transfers. Both outcomes are Pareto efficient because in each case it is impossible to make one player better off without making the other worse off.

A.13 An Example of Weakly Dominated Strategy Which Satisfies the $\max - \min$ Criterion

Consider the (artificial) set of agent types: $[0, p^A]$. We argue that the following strategy of the principal, which is to offer no compensation to the agent, is a weakly dominated strategy which satisfies the max-min criterion. To see this, note that the aforementioned strategy is weakly dominated by the time-invariant pooling reimbursement strategy $\phi_t = \phi^{**}$, conditional on the agent’s investment strategy being given by (??). For all $p_t \in [0, p^{**}]$ and $t \geq 0$ both strategies give the same value of $W_t$, which is zero and also the minimum value. For all $p_t \in (p^{**}, p^A]$ and all $t \geq 0$, strategy $\phi_t = \phi^{**}$ gives strictly larger value of $W_t$ than the strategy of no compensation.
A.14 Proof of Proposition 4

For all equilibrium outcomes (if there exist multiple of them), \( U_{0-2} = 0 \) and none of them is weakly dominated by any other equilibrium outcome. At \( t = 0-2 \), the principal is “indifferent” between them. Among all the equilibrium outcomes, the agent gets the highest reimbursement rate and hence the largest amount of payment from the principal from the pooling equilibrium given by Definition 7 or its outcome-equivalent equilibria. This outcome makes each type of the agent best off. In any other equilibrium outcome (if it exists), some type of the agent is strictly worse off than in the pooling equilibrium given by Definition 7. Therefore the pooling equilibrium given by Definition 7 generates the Pareto dominant equilibrium outcome.

A.15 Proof of Lemma 6

Proposition 3 has already established that on any equilibrium path of an MPOE such that such that \( \inf P_0 \leq p^{**} < \sup P_0 \) and the state variable \( q_t \) is an argument of Markov strategies, \( \phi_t = \phi^{**} \). What remains to be shown is that if there exists any other MPOE, the result \( \phi_t = \phi^{**} \) should still hold on any equilibrium path. Without lose of generality, suppose there exists another MPOE with state variable \( \tilde{q}_t \) as argument of Markov strategies. In general \( \tilde{q}_t \) must have the Markov property. Our definition of Markov strategy\(^{33}\) requires that \( \tilde{q}_t \) be as simple as possible. Therefore, \( \tilde{q}_t \) must be as simple as \( q_t \) in the sense that it changes only if \( k_t > 0 \) (i.e., \( \tilde{q}_{t+dt} = \tilde{q}_t \) if \( k_t = 0 \)). Given the similarity between \( \tilde{q}_t \) and \( q_t \), the argument in proof of Proposition 3 applies to establishing \( t = t \) for the case of \( \tilde{q}_t \) as well.

A.16 Proof of Proposition 5

Suppose the opposite, that is, there exists a separating equilibrium, such that \( \mathcal{I}_d(c, R(p_0, c)) \) is a singleton for all \( p_0 \in [0, 1] \). Then \( \inf P_t = \inf \mathcal{I}_d(c, R(p_0, c)) = p_0 \). Let \( \phi_{t+dt}(K_t, \inf P_t) = s^P(h^t, t) \). For \( s^P \) to be optimal, it must provide no payment to any agent type \( p_0 \in [0, p^{**}) \), which entails for all \( p_0 \in [0, p^{**}) \), \( \phi_{t+dt}(K_t, \inf P_t) \leq \phi_{t+dt}^*(K_t, \inf P_t) \) as defined in (31). Also, for \( s^P \) to be optimal, it must be such that for all \( p_0 \in (p^{**}, 1] \), \( \phi_{t+dt}(K_t, \inf P_t) = \phi_{t+dt}^*(K_t, \inf P_t) \). This reimbursement rate, which maximizes social surplus and leaves no information rent for the agent, however, cannot be incentive compatible (i.e., optimal) for the \( p_0 = 1 \) type agent to voluntarily choose. To see this, note that if the agent \( p_0 = 1 \) reveals his type, he will be offered a contract specified in Lemma 2. We distinguish two cases.

Case (i): \( \frac{a}{\lambda A} \leq 1 \) which implies \( K_A(1) = \infty \). In this case the agent receives no payment from the principal. However, pretending to be any \( p_0 \in (p^{**}, 1) \) type agent, his investment cost will be covered partially before he stops investing and thus earns him some positive information rent. Thus, we reach a contradiction.

Case (ii): \( \frac{a}{\lambda A} > 1 \) which implies \( K_A(1) < 0 \). In this case the agent receives reimbursement at the constant rate \( \tilde{\phi}_{t+dt}^*(K_t, 1) = 1 - \frac{\lambda A}{a} \), which gives him no information rent. \(^{33}\)See Section 6.2 for further discussion of the definition of Markov strategy and Markov equilibrium.
However, pretending to be any \( p_0 \in (p^{**}, 1) \) type agent, he will get a higher reimbursement rate thus earns some positive information rent. Thus, we reach a contradiction.

A.17 Proof of Proposition 6

The result follows immediately from Theorem 3 and the example \((\bar{s}, \bar{M})\).

A.18 Proof of Proposition 7

Suppose the opposite, that is, there exists at least one non-trivial-semi-separating MPOE. Let \((c, R)\) represent such an equilibrium. Thus, \( \mathcal{Q}(p^{**} + \varepsilon) \equiv \{ p_0 | R(c, p_0) = R(c, p^{**} + \varepsilon) \} \) is the set of agent types which are indistinguishable from the marginal low type \( p^{**} + \varepsilon \) by the initial revelation. By Proposition 6 the principal must offer \( \phi_t = \phi^{**} \) to this group of types for all \( t \geq 0 \). There must exist at least another group of types such that \( \inf P_0 \geq \sup \mathcal{Q}(p^{**} + \varepsilon) > p^{**} \). For the latter group, it is optimal for the principal to make a sequence of take-it-or-leave-it offers \( \phi_0 = \max (1 - \frac{\Delta \mathcal{I}_a}{\mathcal{I}_a} \inf P_0, 0) < \phi^{**} \) and \( \phi_{t+dt} = \min (\max (1 - \frac{\Delta \mathcal{I}_a}{\mathcal{I}_a} P_{t+dt}(K_t + dt; \inf P_0), 0), \phi^{**}) \leq \phi^{**} \), which every type in the latter group will accept if \( K_t \leq \bar{K} (\phi_t, P_0) \). To see this, note that \( q_0 = \inf P_0 > p^{**} \). Condition (33) is satisfied and the result is the implication of Lemma 4. However, each type from this group can misrepresent his type to be in the former group and get the offer \( \phi_t = \phi^{**} \). This profitable deviation implies non-trivial-semi-separating MPOE cannot exist.

A.19 Proof of Theorem 5

Lemma 6 establishes the result for the case: \( \inf P_0 \leq p^{**} < \sup P_0 \). Proposition 7 establishes that \( \inf P_0 > p^{**} \) cannot happen in any MPOE. Therefore the result holds for the more general case \( \sup P_0 > p^{**} \). For the case \( \sup P_0 \leq p^{**} \), we have \( p_0 < p^{**} \), investment does not occur in equilibrium, therefore the result holds trivially.

A.20 Proof of Proposition 8

We first consider the case: \( \inf P_0 \leq p^{**} < \sup P_0 \). By Proposition 3 \( \frac{\partial V}{\partial q_t} = 0 \) on an equilibrium path, then Bellman equation (16) implies that for all \( p_0 \leq \bar{p}(K_t, \phi^{**}) \) choosing \( k_t = 0 \) is optimal for the agent. If \( k_t > 0 \) occurs, it can be legitimately inferred that the true type \( p_0 \geq \bar{p}(K_t, \phi^{**}) \) and hence \( \inf P_{t+dt} = \bar{p}(K_t, \phi^{**}) \). We also know that \( q_{t+dt} = \bar{p}(K_t, \phi^{**}) \) if \( k_t > 0 \).

If \( k_t = 0 \), but \( K_t > 0 \), then we must have \( k_0 > 0 \); otherwise, \( k_0 = 0 \) would imply an stationary state such that \( (q_t, K_t', \phi^{**}) = (0, 0, \phi_0) \) for all \( t' \geq 0 \), which deliver a contradiction with the assumption \( K_t > 0 \). It follows from \( k_0 > 0 \) that \( \inf P_{dt} = \bar{d}_{dt} = \bar{p}(0, \phi^{**}) = p^{**} \) and \( \inf P_t = q_t \geq p^{**} \) for \( t > 0 \) with \( K_t > 0 \). For any moment of time \( \tau > 0 \), if \( k_{\tau} > 0 \) both \( q_\tau \) and \( \inf P_\tau \) will ratchet up in an identical way; if \( k_{\tau} = 0 \) both \( q_\tau \) and \( \inf P_\tau \) will remain
unchanged and stay at the identical level for ever. This implies for all \( t \geq 0 \) with \( K_t > 0 \)

\[
\inf P_{t+dt} = q_{t+dt} \text{ on an equilibrium path.}
\]

Since \( \inf P_0 = q_0 \), we must also have

\[
\inf P_t = q_t \text{ on an equilibrium path.}
\]

Now we consider the case: \( \sup P_0 \leq p^* \). Thereby, \( p_0 < p^* \) holds and the principal has no interest to induce the agent to invest and \( k_t = 0 \) holds for all \( t \geq 0 \) on equilibrium path. As a result, \( q_t = q_0 = \inf P_0 = \inf P_t \). The first and last equalities are because both \( q_t \) and the principal’s beliefs have become stationary.

Last we consider the case: \( \inf P_0 > p^* \). By Proposition 7 this case is off equilibrium path, which means \( \inf P_t = \inf P_0 = 0 \).

The rest of the proof is trivial.

A.21 Proof of Proposition 9

The proof is similar to the proof of Theorem 3 (apart from that given the agent’s strategy the principal’s strategy is always unanimously optimal), therefore is omitted.

A.22 Proof of Proposition 10

It is an immediate implication of Proposition 9.

A.23 Proof of Proposition 11

The proof of existence of \( \phi \) has been given in the main text following the proposition, which also establishes the trigger strategy component and the bootstrap property. The rest of the proof is similar to the proof of Theorem 3 (apart from that given the agent’s strategy the principal’s strategy is always unanimously optimal), therefore is omitted.

A.24 Proof of Proposition 12

Trivial.

A.25 Proof of Theorem 6

The idea for proof has been given in the main text.
References


