Dynamic Contracting with Moral Hazard Under Incomplete Information

Teddy Mekonnen *

November 12, 2017

Abstract

I study a continuous time principal-agent model in which an unknown parameter and the agent's hidden effort affect the distribution of observable outcomes. The principal and the agent learn about the parameter by observing past outcomes. The agent's current effort has an implicit long-term effect through the belief dynamics and a deviation in effort creates a persistent disparity between the principal's and the agent's beliefs. This disparity affects the rate of learning as well as how the two evaluate the expected distribution of future outcomes which in turn affects their evaluation of future payoffs. Placing minimal restrictions on how effort and the parameter interact, I derive necessary and sufficient conditions for incentive compatible contracts. In addition to the agent's promised utility, the covariance between the on-path posterior beliefs and the agent's total payoff serves as a second state variable capturing the marginal long-run effects of effort.

Keywords: Dynamic Contracting, principle-agent problem, continuous time, stochastic maximum principle.

*California Institute of Technology, SISL. Contact:mekonnen@caltech.edu

1 Introduction

A key component in every moral hazard problem is the production function - the mapping from the agent's hidden effort to a distribution of observable outcomes. The principal, unable to monitor effort, provides incentives by tying the agent's payoff to the observed outcome.[1] Thus, the structure of incentives depends on the structure of the production function. While the assumption of complete information about the production function is standard, in reality, outcomes may also depend on variables that are unobservable. For example, the school performance of a student depends not only on the amount of effort the student's teacher exerts but also on unobservable variables such as the student's ability. Hence, a principal designing incentive schemes for the teacher has to take into account the uncertainty about the student's ability.

When the agency problem is repeated multiple times, the principal and the agent learn about the unobserved variables by using past outcomes to update their beliefs. The posterior beliefs in turn affect the expected distribution of current and future outcomes. In this dynamic setting with learning, the hidden effort assumption plays a new role: the principal's posterior beliefs are based on a conjecture of the agent's strategy and the agent can manipulate the principal's beliefs by deviating away from the conjecture. Such deviations create a persistent gap between the principal's and the agent's posteriors. Furthermore, a deviation can create a disparity in the rate of learning between the agent and the principal.[2] Whether or not such a deviation is beneficial to the agent depends on the structure of the production function as well as the provision of dynamic incentives.

In this paper, I study a principal-agent contracting problem in which the distribution of outcomes depends on the agent's effort and an unknown parameter. In particular, I study a continuous time finite horizon problem where outcome is described by a Brownian motion with a drift. The drift plays the role of a production function and is controlled by the agent's effort and a fully persistent parameter. The principal, the residual claimant of the produced outcome, offers and commits to a history-dependent contract in order to provide incentives to the agent. The continuous time model allows for a tractable expression of the learning dynamics and simplifies the complexity of the problem through powerful optimization techniques such as

[1] I use feminine pronouns to refer to the principal and masculine ones to refer to the agent.

[2] Think of the distribution of outcomes for each effort level as an experiment. Suppose experiments become more informative as effort increases. Then an agent who deviates upwards from the principal's conjecture knows the quality of the experiment has improved. The agent will therefore update the posterior more aggressively than the principal.

the stochastic maximum principle (Bismut 1973, 1977; Zhou 1996).

I consider general production functions with minimal restrictions on how effort interacts with the parameter. The generality is helpful to understand how different assumptions on the production function, such as the complementarity between effort and the parameter, affect the learning process and the contracting problem. I provide general necessary and sufficient conditions for the agent's problem and show that incentives are based on two state variables: the agent's promised (continuation) utility and the covariance between the agent's payoffs and posterior beliefs. Under mild conditions on the production function, I also show that the agent is exposed to more risk under incomplete information than under complete information for some parameter value.

Since Abreu et al. (1986, 1990) and Spear and Srivastrava (1987), the use of the agent's promised utility as a state variable is well understood. More recently, in a continuous time setting with complete information about the production function, Sannikov (2008) shows that the promised utility alone is both necessary and sufficient for a recursive formulation of incentive compatible contracts. The principal provides incentives by controlling the sensitivity of the agent's promised utility to the random fluctuations in the observed outcome, thereby exposing the agent to some degree of risk. The optimal degree of risk exposure trades off the agent's marginal disutility of effort with the marginal productivity at each instant of time.

However, when there is incomplete information about the production function, a characterization of incentive compatible contracts may require additional state variables. The key issue that arises under incomplete information is that hidden action begets hidden information: the agent's belief dynamics depend on past effort choices that are unobserved by the principal. Thus, the agent's posterior beliefs become part of the agent's private information. I show that a second state variable, the discounted covariance between on-path beliefs and total payoffs, is necessary to account for the agent's private information.

The agent's private information affects payoffs through a single channel; future payoffs depend on the expected distribution of future outcomes which in turn depends on posterior beliefs. Thus, the second state variable must capture the interplay between the persistent effect of effort on beliefs and future incentives (the sensitivity of the agent's utility to outcome fluctuations). Under complete information, beliefs are degenerate and there is no persistent effect of effort. Similarly, if the agent is given constant utility, then there is no benefit to having private information about the parameter. In both cases, the covariance is a constant zero process and becomes a redundant state variable.

Another interpretation of the covariance state variable is possible by casting the contracting problem as a dynamic mechanism design problem. The agent's beliefs serve as "types" that are correlated over time. The principal designs a mechanism that affects the agent's payoff based on the history of "type reports". Misreporting today would require the agent to keep misreporting in the future in a manner that is consistent with how types evolve over time. Following the dynamic mechanism design approach of Pavan et al. (2014), the covariance state variable can be interpreted as the dynamic virtual surplus that accounts for the marginal effect of the agent's current type report on his payoffs as well as on his future reports.

The papers closest to this one are Prat and Jovanovic (2014), Demarzo and Sannikov (2016), and He et al. (2017), who characterize the optimal dynamic contract under incomplete information. The first paper considers the case in which the parameter is fully persistent and the latter two consider a parameter that evolves stochastically over time. However, in all three papers, effort and the parameter enter the production function additively, and as such, their conditions for incentive compatibility are a special case of the conditions presented in this paper. More importantly, an additively separable production function is not just a simplifying assumption; it implies the agent's current effort always has a persistent negative effect on his future payoffs that creates an incentive to shirk.[3]. This paper provides tractable conditions for incentive compatibility even when effort and the parameter interact more intricately, for example, when the production function is multiplicative as in the experimentation literature (Bolton and Harris, 1999).

This paper is also related to Sannikov (2014) who considers an agency problem where effort has explicit long-term consequences. An increase in the agent's current effort not only improves the current outcome but also raises the likelihood of good outcomes in the future. Thus, the agent's current effort always has a persistent positive effect on his future payoffs that lowers the incentives to shirk. In this paper, the agent's current effort has long-term effects only implicitly through learning. In Section 1.1.3, I use a simple two-period, two-outcome model to show that learning models exclude the case of unambiguously positive persistent effects.

My methodology borrows from Williams (2011, 2015) who is the first to apply the stochastic maximum principle to agency problems. Williams (2011) studies an agent who reports income or taste shocks to a principal in order to smooth his consumption. Williams (2015) characterizes the optimal contract in an agency model that has complete information about the production function but allows for hidden savings by the agent. The main difference to my paper is that any deviation by the agent (misreporting shocks or saving more than conjectured) has no impact on the production function. In contrast, the persistent effects of effort on beliefs, and thereby

[3]See Section 1.1.3.

the production function, is the crucial component of private information in my paper.

More broadly, this paper is related to the literature on dynamic long-term contracting in agency models that feature: complete information and lump sum payments (Holmstrom and Milgrom, 1987; Schattler and Sung, 1993; Sung, 1995; Cvitanić et al., 2009), explicit long term-effects of effort (Jarque, 2010; Hopenhayn and Jarque, 2010), funding experimentation (Bergemann and Hege, 1998, 2005; Horner and Samuelson, 2013; Kwon, 2017), and incomplete information with dogmatic beliefs (Adrian and Westerfield, 2009).

The rest of the paper is organized as follows: I first highlight how the interactions between effort and beliefs creates persistent effects through a simple two-period, two-outcome model in Section 1.1. In Section 2, I present the main continuous time model. I derive the necessary and sufficient conditions for incentive compatible contracts in Section 3. Section 4 concludes. All proofs are in the Appendix.

1.1 Two-Period Setting

Prior to the continuous time model, it is instructive to first explore two-period, two-outcome models. I consider three different principal-agent models: a standard moral hazard problem, a moral hazard problem with explicit long-run consequences, and a moral hazard problem with incomplete information. In all three models, the agent (he) chooses an action $a_t \in [0,1] \triangleq A$ for each period $t = 1, 2$. Given a wage $W \in \mathbb{R}$ and an action choice $a \in A$, the agent's payoff is given by

$$u(W) - c(a),$$

where $u'(\cdot), c'(\cdot) > 0$ and $u''(\cdot) < 0, c''(\cdot) \geq 0$. Future payoffs are discounted at rate $\delta \in (0,1)$.

The principal (she) cannot monitor the agent's action and only observes if the outcome in each period, y_t, is a success, S, or a failure, F. The principal-agent relationship begins with the principal offering a contract and committing to it. A contract, $\{a^r, w\}$, is a pair of action recommendations for each period,

$$a_1^r \in A, \text{ and } a_2^r : \{F, S\} \to A,$$

along with a pair of wage payments

$$w_1 : \{F, S\} \to \mathbb{R}, \text{ and } w_2 : \{F, S\}^2 \to \mathbb{R}.$$

A contract $\{a^r, w\}$ is incentive compatible if the agent is willing to follow the action recommendations. The following three models differ in how the agent's actions affect the distribution of observable outcomes.

1.1.1 Standard Moral Hazard Problem

Consider first a standard moral hazard model: in each period, $y_t = S$ with probability $f(a_t) \in [0,1]$ with $f'(\cdot) > 0$ so that a higher action leads to a more favorable distribution of outcomes. The distribution of outcomes in each period is independent of the action choices made in prior periods.

Fix a contract $\{a^r, w\}$. If the contract is incentive compatible, then given any first period outcome y_1, the recommendation in the second period $a_2^r(y_1)$ must be a solution to

$$\max_{a_2 \in A} f(a_2) \underbrace{\Big[u\big(w_2(y_1, S)\big) - u\big(w_2(y_1, F)\big)\Big]}_{\triangleq \zeta_2(y_1)} - c(a_2). \tag{1}$$

We can interpret $\zeta_2(y_1)$ as the bonus the agent gets for a successful outcome in period 2 conditional on already having produced y_1 in period 1. Similarly, the recommendation in the first period a_1^r must be a solution to

$$\max_{a_1 \in A} f(a_1) \underbrace{\Big[u\big(w_1(S)\big) + \delta V_2(S) - u\big(w_1(F)\big) - \delta V_2(F)\Big]}_{\triangleq \zeta_1} - c(a_1), \tag{2}$$

where

$$V_2(y_1) = f\big(a_2^r(y_1)\big)\zeta_2(y_1) + u\big(w_2(y_1, F)\big) - c\big(a_2^r(y_1)\big)$$

is the agent's continuation value (also referred to as promised utility) given outcome y_1. We can interpret ζ_1 as the agent's bonus for a successful outcome in period 1. The tuple $\{\zeta_1, \zeta_2(S), \zeta_2(F)\}$ only depend on the wages and recommendations specified in the contract and are therefore under the principal's control.

If the recommended actions are in the interior of A, we can use the first order necessary conditions from (1) and (2) to derive the following expressions:

1. $\zeta_2(y_1) = \frac{c'\left(a_2^r(y_1)\right)}{f'\left(a_2^r(y_1)\right)}, \forall y_1 \in \{F, S\}$ and
2. $\zeta_1 = \frac{c'(a_1^r)}{f'(a_1^r)}$.

In each period, the bonus captures the trade-offs between the current marginal cost and the current marginal productivity of the recommended action. Past actions or outcomes have no direct effect.

In a continuous time setting, Sannikov (2008) shows that the principal can incentivize the agent to choose the desired action by controlling the sensitivity of his promised utility to the performance measure, i.e., the change in output. Furthermore, the sensitivity of the promised utility is simply the ratio of marginal cost to marginal productivity at the desired action level. Hence, the agent's per period bonus for successful outcomes in the two-period model is analogous to the promised utility sensitivity in a continuous time model.

1.1.2 Moral Hazard Problem with Long-run Consequences

A more general model of moral hazard accommodates for past actions and outcomes to have a persistent effect on the distribution of future outcomes. For example, a boxer who exerts himself early in a match improves his chances of winning the first few rounds but may not be as powerful in the latter rounds. In contrast, a student who studies hard early in the quarter not only improves his chances for passing the midterm but also the final exam.

Consider an alteration to the previous model such that in the second period, the outcome is $y_2 = S$ with probability $g(a_2|a_1, y_1) \in [0, 1]$. Assume that g is differentiable in both a_1 and a_2, and that $g_{a_2}(\cdot|a_1, y_1) > 0$ for all histories $\{a_1, y_1\}$.[4] Nothing changes in the first period.

Fix a contract $\{a^r, w\}$. Given any first period action choice a_1 and outcome y_1, let $a_2^*(a_1, y_1)$ be the action that maximizes the agent's period 2 payoff, i.e., $a_2^*(a_1, y_1)$ solves

$$\max_{a_2 \in A} g(a_2|a_1, y_1) \underbrace{\Big[u\big(w_2(y_1, S)\big) - u\big(w_2(y_1, F)\big)\Big]}_{\triangleq \eta_2(y_1)} - c(a_2)$$

From the perspective of period 2, nothing fundamental has changed between the standard model in Section 1.1.1 and the current one. As before, $\eta_2(\cdot)$ is the agent's bonus for a successful outcome in period 2. The principal chooses the bonus via the wages she offers while the agent affects the distribution of second period outcomes through his action choice. Let

$$V_2(a_1, y_1) = g\big(a_2^*(a_1, y_1)|a_1, y_1\big)\eta_2(y_1) + u\big(w_2(y_1, F)\big) - c\big(a_2^*(a_1, y_1)\big)$$

[4]For some subset $X \subseteq \mathbb{R}^n$ and some differentiable function $G : X \to \mathbb{R}$, I write $G_{x_i}(x)$ for $\partial G(x)/\partial x_i$ when $n > 1$ and $G'(x)$ when $n = 1$.

be the agent's continuation value for any arbitrary history $\{a_1, y_1\}$. In period 1, the agent solves

$$\max_{a_1 \in A} f(a_1)\Big[u\big(w_1(S)\big) + \delta V_2(a_1, S) - u\big(w_1(F)\big) - \delta V_2(a_1, F)\Big] + \delta V_2(a_1, F) - c(a_1). \tag{3}$$

The contract $\{a^r, w\}$ is incentive compatible if a_1^r solves (3) and $a_2^r(y_1) = a_2^*(a_1^r, y_1)$ for all y_1.

Notice the differences between (2) and (3): in the latter, the agent's first period action directly enters the continuation values because it has a persistent effect on the distribution of future outcomes. Hence, the agent's bonus in the first period endogenously depends on his first period action (unlike ζ_1 that is exogenously determined via the wages in the contract). Consequently, the principal in this altered setting can only control the on-path bonus, given by

$$\eta_1 = u\big(w_1(S)\big) + \delta V_2(a_1^r, S) - u\big(w_1(F)\big) - \delta V_2(a_1^r, F).$$

If the agent deviates to an action $a_1 \neq a_1^r$, he would face a different bonus.

If the recommended actions are in the interior of A, we can again use the necessary first order conditions to derive the following expressions for the bonuses:

1. $\eta_2(y_1) = \frac{c'\left(a_2^r(y_1)\right)}{g_{a_2}\left(a_2^r(y_1)|a_1^r, y_1\right)}$, $\forall y_1 \in \{F, S\}$, and

2. $\eta_1 = \frac{c'(a_1^r)}{f'(a_1^r)} - \frac{\delta}{f'(a_1^r)}\Big\{f(a_1^r) g_{a_1}\big(a_2^r(S)|a_1^r, S\big)\eta_2(S) + (1 - f(a_1^r)) g_{a_1}\big(a_2^r(F)|a_1^r, F\big)\eta_2(F)\Big\}$,

Similar to the first model, the agent's bonus in each period captures the trade-offs between the marginal cost of effort and the marginal productivity of effort. However, the first period bonus, η_1, is also inter-temporally linked to future bonuses due to the persistence of the first period action on future outcomes.

After a first period outcome y_1, the agent exerts effort in the second period only if the principal offers him a strictly positive bonus, i.e., $\eta_2(y_1) > 0$. If $g_{a_1} < 0$ (as in the example of the boxer), the agent faces a negative persistent effect; by exerting effort in the first period, the agent is less likely to earn a bonus in period 2. Hence, the principal must offer a more high-powered incentive scheme in the first period to compensate the agent for the additional trade-off he faces. In other words, the principal front loads incentives. On the other hand, if $g_{a_1} > 0$ (as in the example of the student), exerting effort in period 1 leads to higher chances of earning a bonus in both periods. Since, the bonus in period 2 amplifies the agent's incentives to work in period 1, the principal back loads incentives.

The term in curly braces above captures the marginal persistent effect that a deviation from the recommended period 1 action has on the agent's continuation value. When the first period action has no direct effect on the distribution of future outcomes, i.e., $g_{a_1} = 0$, we recover the expressions for ζ_1. However, if $g_{a_1} \leq 0$ ($g_{a_1} \geq 0$), then a principal who wishes to implement a recommended action a_1^r must offer a higher (lower) bonus in the current model than in the standard model of Section 1.1.1, i.e., $\eta_1 \geq \zeta_1$ ($\eta_1 \leq \zeta_1$).

A similar conclusion is reached by Sannikov (2014) in a continuous time setting that is analogous to the current two period model with

$$f(a_1) = a_1,$$
$$g(a_2|a_1, y_1) = (1-\beta)a_2 + \beta a_1, \forall y_1 \in \{F, S\}$$

where $\beta \in (0, 1/2)$. As the first period action has a positive persistent effect on the distribution of future outcomes ($g_{a_1} > 0$), incentives are back loaded: the sensitivity of the promised utility starts off lower than the ratio of marginal cost to marginal productivity of effort and is subsequently adjusted upwards as time progresses.

1.1.3 Moral Hazard Problem with Incomplete Information

Consider a second variant to the standard model: in each period $t = 1, 2$, $y_t = S$ with probability $h(\theta, a_t) \in [0, 1]$, where $\theta \in \{\theta_L, \theta_H\}$ is an unobserved parameter. Assume that $h_a(\theta, \cdot) > 0$ for all θ. The principal and the agent have a common prior $\Pr(\theta = \theta_H) = p_1 \in (0, 1)$. Given a history $\{a_1, y_1\}$, the agent updates his belief in period 2 to $p_2(a_1, y_1)$ using Bayes rule.[5] It is easy to check that

$$\frac{\partial p_2(a_1, S)}{\partial a} \geq 0 \Leftrightarrow \frac{h_a(\theta_H, a_1)}{h_a(\theta_L, a_1)} \geq \frac{h(\theta_H, a_1)}{h(\theta_L, a_1)}, \tag{4}$$

and

$$\frac{\partial p_2(a_1, F)}{\partial a} \geq 0 \Leftrightarrow \frac{1 - h(\theta_H, a_1)}{1 - h(\theta_L, a_1)} \geq \frac{h_a(\theta_H, a_1)}{h_a(\theta_L, a_1)}. \tag{5}$$

This third model is a special case of the second one in which past actions affect the expected distribution of future outcomes via the beliefs the agent holds. In particular, we can embed this model into the second one by setting

$$f(a_1) = p_1 h(\theta_H, a_1) + (1 - p_1) h(\theta_L, a_1)$$

[5] If the parameter θ is observable, or if there is no learning, then the third model is equivalent to the first one.

and

$$g(a_2|a_1, y_1) = p_2(a_1, y_1)h(\theta_H, a_2) + \big(1 - p_2(a_1, y_1)\big)h(\theta_L, a_2).$$

The type of persistent effect that arises from exerting effort in period 1 is tied to how beliefs evolve which in turn depends on the interaction between the parameter and actions. Given any history $\{a_1, y_1\}$,

$$g_{a_1}(a_2|a_1, y_1) = \frac{\partial p_2(a_1, y_1)}{\partial a} \underbrace{\Big(h(\theta_H, a_2) - h(\theta_L, a_2)\Big)}_{\triangleq \Delta h(a_2)},$$

and we can express the marginal effect of increasing the first period action a_1 by

$$\underbrace{-c'(a_1)}_{\text{marginal cost}} \quad + \quad \underbrace{\eta_1 f'(a_1)}_{\substack{\text{marginal instantaneous benefit} \\ \text{from increased productivity}}} \quad + \tag{6}$$

$$f(a_1)\underbrace{\eta_2(S)\Delta h\big(a_2^*(a_1, S)\big)\frac{\partial p_2(a_1, S)}{\partial a}}_{\text{persistent effect conditional on } y_1=S} + (1 - f(a_1))\underbrace{\eta_2(F)\Delta h\big(a_2^*(a_1, F)\big)\frac{\partial p_2(a_1, F)}{\partial a}}_{\text{persistent effect conditional on } y_1=F}.$$

Unlike the model in Section 1.1.2, actions in this model cannot have an unambiguous positive persistent effect: If $g_{a_1} > 0$, then $\partial p_2(\cdot, y_1)/\partial a$ and $\Delta h(a_2)$ must have the same sign for all a_2 and y_1. However, using (4) and (5),

$$\frac{\partial p_2(\cdot, y_1)}{\partial a} > (<)0, \forall y_1 \implies \frac{1 - h(\theta_H, a)}{1 - h(\theta_L, a)} > (<)\frac{h(\theta_H, a)}{h(\theta_L, a)}, \forall a \implies \Delta h(a) < (>)0, \forall a.$$

On the other hand, unambiguous negative persistent effects are accommodated by this setting.
Example 1: Let $h(\theta, a) = 0.5(a + \theta)$, with $1 > \theta_H > \theta_L > 0$.
As $\Delta h(a) = 0.5(\theta_H - \theta_L) > 0$, the sign of g_{a_1} is pinned down by the sign of $\partial p_2(a_1, y_1)/\partial a$. Using (4), $\partial p_2(\cdot, S)/\partial a < 0$ as

$$1 = \frac{h_a(\theta_H, a)}{h_a(\theta_L, a)} < \frac{h(\theta_H, a)}{h(\theta_L, a)} = \frac{\theta_H + a}{\theta_L + a}, \;\; \forall a \implies g_{a_1}(a_2|\cdot, S) < 0, \forall a_2.$$

Similarly, using (5), $\partial p_2(\cdot, F)/\partial a < 0$ as

$$1 = \frac{h_a(\theta_H, a)}{h_a(\theta_L, a)} > \frac{1 - h(\theta_H, a)}{1 - h(\theta_L, a)} = \frac{2 - (\theta_H + a)}{2 - (\theta_L + a)}, \;\; \forall a \implies g_{a_1}(a_2|\cdot, F) < 0, \forall a_2.$$

Hence, providing the agent with incentives to work in period 2 dampens his period 1 incentives. The agent's period 1 bonus is ultimately higher than the simple marginal effort to marginal productivity ratio he would earn under complete information.

This example is analogous to the continuous time setting of Prat and Jovanovic (2014) who show that when the agent's effort and the unobserved state are additively separable, incentives are front loaded: the sensitivity of the agent's promised utility is much higher in early periods. Unfortunately, even when we impose a lot of structure on $h(\theta, a)$, it is generally difficult to conclude the overall effect of a_1.

Example 2: Let $h(\theta, a) = 0.5 + 0.5a\theta$ with $1 > \theta_H > \theta_L > 0$.
As $\Delta h(a) = 0.5a(\theta_H - \theta_L) \geq 0$, the sign of g_{a_1} is pinned down by the sign of $\partial p_2(a_1, y_1)/\partial a$. Using (4), $\partial p_2(\cdot, S)/\partial a > 0$ as

$$\frac{\theta_H}{\theta_L} = \frac{h_a(\theta_H, a)}{h_a(\theta_L, a)} > \frac{h(\theta_H, a)}{h(\theta_L, a)} = \frac{1 + a\theta_H}{1 + a\theta_L}, \;\; \forall a \implies g_{a_1}(a_2|\cdot, S) > 0, \forall a_2 > 0.$$

Therefore, providing the agent with incentives to work in period 2 after $y_1 = S$ also amplifies his period 1 incentives. In contrast, using (5), $\partial p_2(\cdot, F)/\partial a < 0$ as

$$\frac{\theta_H}{\theta_L} = \frac{h_a(\theta_H, a)}{h_a(\theta_L, a)} > \frac{1 - h(\theta_H, a)}{1 - h(\theta_L, a)} = \frac{1 - a\theta_H}{1 - a\theta_L}, \;\; \forall a \implies g_{a_1}(a_2|\cdot, F) < 0, \forall a_2 > 0.$$

Therefore, providing the agent with incentives to work in period 2 after $y_1 = F$ dampens his period 1 incentives. Whether exerting effort in period 1 has an overall positive or negative persistent effect will depend on which of the two opposite effects above is stronger. For example, if the principal sets $\eta_2(F) = 0$, the agent has no incentives to work in the second period after $y_1 = F$. The only remaining persistent effect from (6) would then be the positive effect after $y_1 = S$. However, we would have to solve the full contracting problem to reach a conclusive answer in the two-period, two-outcome model. The next sections provides a general treatment of incentive compatible dynamic contracts in a continuous time setting where we can say more.

2 Model

A principal wishes to hire an agent over a finite horizon $[0, T]$. At each time t, the agent can take some action $a_t \in A \triangleq [0, \bar{a}]$ that is unobserved by the principal. Let $\boldsymbol{a} \triangleq (a_t)_{0 \leq t \leq T}$ represent an arbitrary action policy. The agent's actions affect the publicly observable output process

$\boldsymbol{Y} = (Y_t)_{0\leq t\leq T}$ which evolves according to

$$dY_t = f(\theta, a_t)dt + \sigma dB_t,$$

where $\boldsymbol{B} \triangleq (B_t)_{t\geq 0}$ is the standard one-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$, $\sigma > 0$ is a finite constant, and $\theta \in \Theta \triangleq \{\theta_0, \theta_1\}$ is an exogenous and unobservable parameter. The drift function $f : \Theta \times A \to \mathbb{R}$ captures how actions interact with the parameter to affect the average flow of output.

(A.1) $f : \Theta \times A \to R$ is continuously differentiable in a with $f_a(\theta, \cdot) > 0$ for all $\theta \in \Theta$.

Neither the principal nor the agent observe the parameter. At $t = 0$, they share a common prior that $\theta = \theta_1$ with probability $p_0 \in [0, 1]$. The agent knows the history of output as well as the history of his actions. Hence, the agent's information set at time t is given by the filtration $\mathcal{F}_t^a \triangleq \sigma(Y_s, a_s : 0 \leq s \leq t)$. An admissible action policy for the agent is any policy $\boldsymbol{a}$ such that $a_t \in A$ is $\mathcal{F}_t^a$-adapted.

Given an action policy $\boldsymbol{a}$, the agent updates his belief at time t from p_0 to p_t^a which evolves according to[6]

$$\begin{aligned} dp_t^a =& p_t^a(1 - p_t^a)\frac{\Delta f(a_t)}{\sigma}\left[\frac{dY_t - \bar{f}(p_t^a, a_t)dt}{\sigma}\right] \\ =& p_t^a(1 - p_t^a)\frac{\Delta f(a_t)}{\sigma} dZ_t^a. \end{aligned} \tag{7}$$

The function

$$\Delta f(a) \triangleq f(\theta_1, a) - f(\theta_0, a)$$

captures two pieces of information. First, when $\Delta f(a) \geq 0$, a large flow of output is evidence in favor of θ_1. Symmetrically, when $\Delta f(a) \leq 0$, a large flow of output is evidence in favor of θ_0. Second, the larger $|\Delta f(a)|$ is, the stronger the evidence. Given a posterior belief $p \in [0, 1]$ and action $a \in A$,

$$\bar{f}(p, a) \triangleq pf(\theta_1, a) + (1 - p)f(\theta_0, a)$$

represents the expected flow of output. The process $\boldsymbol{Z^a} \triangleq (Z_t^a)_{t\geq 0}$ is the standard Brownian motion under the measure Q^a - the probability measure induced by the action policy $\boldsymbol{a}$. The

[6]See Lipster and Shiryaev, Chapter 9.

stochastic process dZ_t^a represents the discrepancy at time t between the observed change in output dY_t and the expected change in output $\bar{f}(p_t^a, a_t)dt$.[7]

The belief process (7) has an intuitive description. For example, assume $\Delta f(a_t) > 0$ so that a high output flow is evidence of θ_1. Whenever the realized flow of output dY_t exceeds the agent's expectation $\bar{f}(p_t^a, a_t)dt$, he updates his belief upwards. Similarly, when the realized flow of output falls short of the agent's expectation, he updates his belief downwards. Moreover, beliefs are updated significantly when the uncertainty, $p_t^a(1-p_t^a)$, is high or when the evidence, $|\Delta f(a_t)|$, is strong.

Since the principal cannot observe the agent's actions, her information set at time t is given by the filtration $\mathcal{F}_t^Y \triangleq \sigma(Y_s : 0 \leq s \leq t)$. At $t = 0$, the principal offers and commits to a contract. An admissible contract is a tuple $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T \rangle$, where $(\boldsymbol{a^r}, \boldsymbol{c}) \triangleq (a_t^r, c_t)_{0 \leq t \leq T}$ is a pair of recommendation and compensation policies such that $(a_t^r, c_t) \in A \times \mathbb{R}$ is $\mathcal{F}_t^Y$-adapted, and $C_T \in \mathbb{R}$ is an $\mathcal{F}_T^Y$-measurable terminal payment. Thus, recommendations and payments at time t can depend in an arbitrary way on the history of output but they cannot depend on what the principal does not know such as future realizations of output, the parameter, or the agent's past actions.

At time t, the principal updates her belief based on the history of output and action recommendations from p_0 to p_t^r, which evolves according to

$$dp_t^r = p_t^r(1 - p_t^r)\frac{\Delta f(a_t^r)}{\sigma} dZ_t^{a^r},$$

where the process $\boldsymbol{Z}^{\boldsymbol{a^r}} \triangleq (Z_t^{a^r})_{t \geq 0}$ is the standard Brownian motion under Q^{a^r} - the probability measure induced when the agent follows the recommended policy $\boldsymbol{a^r}$.[8] If the agent follows the recommendations, the agent's and the principal's beliefs (and filtrations) also coincide.

For an action-wage pair $(a, c) \in A \times \mathbb{R}$, the agent's flow payoff is given by $u(a, c)$ which he discounts at a constant rate $r > 0$. Furthermore, given a terminal payment of $c \in \mathbb{R}$, his terminal utility is given by $U(c)$.

(A.2) $u : A \times \mathbb{R} \to \mathbb{R}$ is continuously differentiable in a with $u_a(\cdot, c) < 0$ for all c.

Given an admissible contract $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T \rangle$ and action policy $\boldsymbol{a}$, the agent's payoff at time t

[7] In non-linear filtering problems, the process Z^a is called the innovation process. See Kallianpur et al. (1972).
[8] $dZ_t^{a^r} = (dY_t - \bar{f}(p_t^r, a_t^r)dt)/\sigma$. See proof of Theorem 1 for more details.

is given by

$$V_t(\boldsymbol{a}) \triangleq \int_0^t e^{-rs} u(a_s, c_s) ds + E_t^a \left[\int_t^T e^{-rs} u(a_s, c_s) ds + e^{-rT} U(C_T) \right],$$

where E_t^a is a shorthand for $E[\cdot|\mathcal{F}_t^a]$. Specifically, it is the expectation operator with respect to the measure Q^a at time t. I further assume the following integrability conditions hold at each period t for any admissible contract $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T \rangle$ and action policy $\boldsymbol{a}$:

(A.3) $E_t^a \left[\left(\int_t^T e^{-r(s-t)} u(a_s, c_s) ds + e^{-r(T-t)} U(C_T) \right)^2 \right] < \infty,$

(A.4) $E_t^a \left[\left(\int_t^T e^{-r(s-t)} u_a(a_s, c_s) ds \right)^2 \right] < \infty.$

If the agent follows the recommendations, his expected utility at time t is a Q^{a^r}-martingale. Specifically, for all $s > t$, $E_t^{a^r}[V_s(\boldsymbol{a^r})] = V_t(\boldsymbol{a^r})$. Therefore, applying the martingale representation theorem, there exists a unique square-integrable and $\mathcal{F}_t^Y$-adapted stochastic process $(e^{-rt}\sigma\zeta_t)_{t\geq 0}$ such that

$$V_t(\boldsymbol{a^r}) = V_0(\boldsymbol{a^r}) + \int_0^t e^{-rs} \sigma \zeta_s dZ_s^{a^r}.$$

Define the agent's time t promised utility (continuation payoff from following the recommendation) by

$$W_t(\boldsymbol{a^r}) \triangleq E_t^{a^r} \left[\int_t^T e^{-r(s-t)} u(a_s^r, c_s) ds + e^{-r(T-t)} U(C_T) \right],$$

with terminal condition $W_T(\boldsymbol{a^r}) = U(C_T)$. It evolves according to

$$\begin{aligned} dW_t(\boldsymbol{a^r}) =& \Big(rW_t(\boldsymbol{a^r}) - u(a_t^r, c_t) \Big) dt + \sigma \zeta_t dZ_t^{a^r} \qquad (8) \\ =& \Big(rW_t(\boldsymbol{a^r}) - u(a_t^r, c_t) \Big) dt + \zeta_t \Big(dY_t - \bar{f}(p_t^r, a_t^r) dt \Big). \end{aligned}$$

The term ζ_t captures how sensitive the agent's continuation utility is to the observed outcome. For each additional unit of output above the principal's expectation, $\bar{f}(p_t^r, a_t^r)dt$, the agent's promised utility is increased by ζ_t utils. Notice that ζ_t serves an analogous purpose to the bonuses in the two-period, two outcome setting; to provide the agent with incentives to work.

When there is no uncertainty, i.e. $p_0 \in \{0,1\}$, the process given in (8) is independent of the agent's past actions. This property reduces the problem of writing incentive compatible contracts to a Markov-problem where the promised utility is a state variable. In contrast, when there is uncertainty with $p_0 \in (0,1)$, the entire history of past actions affects the agent's posterior, which in turn affects the process $\boldsymbol{Z^a}$. Thus, the evolution of the agent's promised utility at time t depends on the agent's period t belief. Since the agent's beliefs are his private information, the principal cannot directly use them as another state variable in the contracting problem. Additionally, if the agent deviated in the past, his continuation payoff may evolve differently from (8).

3 Incentive Compatible Contracts

For a proposed contract $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T \rangle$, the agent's problem is to choose an admissible action policy that maximizes his expected payoff,

$$\max_{\boldsymbol{a}} V_0(\boldsymbol{a}) \qquad \text{s.t.} \tag{AP}$$

$$dY_t = \bar{f}(p_t^a, a_t)dt + \sigma dZ_t^a,$$

$$dp_t^a = p_t^a(1-p_t^a)\frac{\Delta f(a_t)}{\sigma}dZ_t^a.$$

A contract $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T \rangle$ is incentive compatible if the recommendation policy $\boldsymbol{a^r}$ is a solution to the agent's problem (AP).

3.1 Necessary Conditions

The formulation of the agent's problem in (AP) is straightforward. The agent affects his payoff through three channels: (i) he incurs a cost associated with his action choice, (ii) he affects the average flow of output dY_t which then affects his compensation, and (iii) he affects the evolution of his posterior belief as well as the probability measure he uses to evaluate his continuation payoffs. Nevertheless, solving (AP) directly is difficult as the proposed contract can depend on the entire output path making the problem non-Markovian.

Another, now standard, formulation of the agent's problem is to consider the entire output path on $[0, T]$ as a random variable.[9] The agent's action changes the probability measure over the different possible output paths that can be realized. I use this latter formulation which is closer to the two-period, two-outcome setting in which the agent's action changes the probability of realizing a successful outcome.

Let $\boldsymbol{Z^0} \triangleq (Z_t^0)_{t \geq 0}$ be a standard Brownian motion under some probability measure Q^0 and let $\mathcal{Z}_t^0$ be the natural filtration generated by Z_t^0. As $\sigma > 0$ is a constant, the tuple (Y, Z^0, Q^0) is a weak solution to the SDE

$$dY_t = \sigma dZ_t^0.$$

That is, the process Y_t has no drift under the measure Q^0.

For any admissible action policy $\boldsymbol{a}$, define Γ_t^a to be a $\mathcal{Z}_t^0$-measurable and square-integrable stochastic exponential process given by

$$\Gamma_t^a = \exp\left(\int_0^t \frac{\bar{f}(p_s^a, a_s)}{\sigma} dZ_s^0 - \frac{1}{2} \int_0^t \left| \frac{\bar{f}(p_s^a, a_s)}{\sigma} \right|^2 ds \right),$$

with $\Gamma_0^a = 1$ as an initial condition. Given Assumption $(A.1)$, Novikov's condition holds and Γ_t^a is a Q^0-martingale.[10] Therefore, Γ_t^a is the Radon-Nikodym derivative capturing the change in measure $\frac{dQ^a}{dQ^0}|_{\mathcal{Z}_t^0}$. Using the Girsanov theorem, we can define the standard Brownian motion under Q^a as the process $\boldsymbol{Z^a}$ such that

$$Z_t^a = Z_t^0 - \int_0^t \frac{\bar{f}(p_s^a, a_s)}{\sigma} ds.$$

Furthermore, the tuple (Y, Z^a, Q^a) is a weak solution to the SDE[11]

$$dY_t = \bar{f}(p_t^a, a_t)dt + \sigma dZ_t^a.$$

Notice that under the measure Q^a, we recover the setup where Y_t has a drift of $\int_0^t \bar{f}(p_s^a, a_s)ds$. We can apply the change of measure to transform the agent's expected payoff from

$$E_0^a \left[\int_0^T e^{-rt} u(a_t, c_t)dt + e^{-rT} U(C_T) \right]$$

[9]This approach was first suggested by Mirrlees (1974) and developed by Holmstrom (1979). For examples in discrete time, see Abreu et al. (1986, 1990), Spear and Srivastravа (1987), and Green (1987). For examples in continuous time, see Sannikov (2008), Cvitanić et al. (2009), and Williams (2009).

[10]Specifically, $|\bar{f}(p, a)| \leq \max_{\theta \in \Theta} \max\{|f(\theta, 0)|, |f(\theta, \bar{a})|\} < \infty$. See Lipster and Shiryaev, Chapter 6.

[11]From $(A.1)$, the function f satisfies the linear growth and Lipschitz conditions for a solution.

to

$$E_0^0\left[\int_0^T e^{-rt}\Gamma_t^a u(a_t,c_t)dt+e^{-rT}\Gamma_T^a U(C_T)\right]$$

where E_t^0 is the expectation operator with respect to the measure Q^0 in period t.

Let $X_t^a \triangleq p_t^a \Gamma_t^a$ be the process that captures the covariation between beliefs and the change in measure.[12] Using Itô's lemma, X_t^a evolves according to

$$dX_t^a = X_t^a \frac{f(\theta_1,a_t)}{\sigma} dZ_t^0.$$

We can now reformulate the agent's problem given in (AP). Fix a contract $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T\rangle$. The agent's problem is to choose an admissible action policy that maximizes his expected payoff,

$$\max_{\boldsymbol{a}} E_0^0\left[\int_0^T e^{-rt}\Gamma_t^a u(a_t,c_t)dt+e^{-rT}\Gamma_T^a U(C_T)\right] \tag{RAP}$$

$$\text{s.t.} \qquad \Gamma_t^a = \Gamma_t^a \frac{\bar{f}\left(\frac{X_t^a}{\Gamma_t^a},a_t\right)}{\sigma} dZ_t^0, \qquad \Gamma_0^a = 1$$

$$dX_t^a = X_t^a \frac{f(\theta_1,a_t)}{\sigma} dZ_t^0, \qquad X_0^a = p_0.$$

Since the measure Q^0 is now independent of the agent's action, the reformulated agent's problem (RAP) is an optimal control problem with (Γ_t^a, X_t^a) as the state variables and a_t as the control variable. We can directly apply the stochastic maximum principle (Bismut 1973, 1978, Zhou 1996) to identify necessary conditions for incentive compatibility.

Theorem 1 *Suppose assumption A.1-A.4 hold. If a contract $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T\rangle$ is incentive compatible, there exist $\mathcal{F}_t^Y$-adapted, square-integrable co-states $(\gamma_t, \sigma\eta_t)$ and $(\lambda_t, \sigma\varphi_t)$ such that*

i. γ_t evolves according to

$$d\gamma_t = \Big(r\gamma_t - u(a_t^r,c_t)+\eta_t p_t^r \Delta f(a_t^r)\Big)dt + \sigma\eta_t dZ_t^{a^r} \tag{9}$$

with $\gamma_T = U(C_T)$,

[12] The transformation of p_t^a to X_t^a is a convenient rescaling that simplifies the expression for the Hamiltonian.

ii. λ_t *evolves according to*

$$d\lambda_t = \Big(r\lambda_t - \Delta f(a_t^r)\big(\eta_t + \varphi_t(1-p_t^r)\big)\Big)dt + \sigma\varphi_t dZ_t^{a^r}, \tag{10}$$

with $\lambda_T = 0$*, and*

iii. For each t*,* a_t^r *solves*

$$\max_{a_t \in A} u(a_t, c_t) + \eta_t \bar{f}(p_t^r, a_t) + \varphi_t p_t^r f(\theta_1, a_t) \tag{11}$$

almost surely.

3.2 Interpretation of Co-States

In the two-period, two-outcome models of Section 1.1, the principal provides incentives for the agent through a bonus. In particular, the bonus captures the difference in the agent's contemporaneous and continuation payoffs when he generates a successful outcome versus a failed one. Furthermore, when the agent's actions have long run consequences, as in Section 1.1.2 and 1.1.3, the principal must account for how bonuses and actions in the second period affect incentives in the first period.

An analogous characterization of incentive compatibility is possible in the continuous time setting. This section reformulates the necessary conditions for incentive compatibility in Theorem 1 by using two state variables: $W_t(\boldsymbol{a}^r)$, the agent's promised utility, and R_t, to be defined further below.

Proposition 1 *For each* $t \in [0, T]$*, the process* $\gamma_t + p_t^r \lambda_t = W_t(\boldsymbol{a^r})$*, almost surely.*

To build some intuition, suppose both the agent and the principal know that $\theta = \theta_0$, i.e., $p_0 = p_t^a = 0$ for all $t \in [0, T]$ and all admissible action policy $\boldsymbol{a}$. The evolution of γ_t given in (9) simplifies to

$$d\gamma_t = \big(r\gamma_t - u(a_t^r, c_t)\big)dt + \eta_t\big(dY_t - f(\theta_0, a_t^r)dt\big)$$

with $\gamma_T = U(C_T)$. Thus, the process γ_t would be equivalent to the agent's promised utility in a complete information setting with $\theta = \theta_0$. In particular, γ_t would be equivalent to

$$W_t(\boldsymbol{a^r}; \theta_0) \triangleq E_t^{a^r, \theta_0}\left[\int_t^T e^{-r(s-t)} u(a_s^r, c_s) ds + e^{-r(T-t)} U(C_T)\right],$$

where $E_t^{a^r,\theta_0}$ is the expectation operator given the filtration $\sigma(Y_s, a_s^r, \theta_0 : 0 \le s \le t)$.

Now suppose that agent and the principal know that $\theta = \theta_1$, i.e., $p_0 = p_t^a = 1$ for all $t \in [0, T]$ and all admissible action policy $\boldsymbol{a}$. The evolution of $\gamma_t + \lambda_t$ simplifies to

$$d\gamma_t + d\lambda_t = \Big(r(\gamma_t + \lambda_t) - u(a_t^r, c_t)\Big)dt + (\eta_t + \varphi_t)\Big(dY_t - f(\theta_1, a_t^r)dt\Big)$$

with $\gamma_T + \lambda_T = U(C_T)$ which would represent the evolution of the agent's promised utility in a complete information setting with $\theta = \theta_1$. In particular, the process $\gamma_t + \lambda_t$ would be equivalent to

$$W_t(\boldsymbol{a^r}; \theta_1) \triangleq E_t^{a^r,\theta_1}\left[\int_t^T e^{-r(s-t)}u(a_s^r, c_s)ds + e^{-r(T-t)}U(C_T)\right],$$

where $E_t^{a^r,\theta_1}$ is the expectation operator given the filtration $\sigma(Y_s, a_s^r, \theta_1 : 0 \le s \le t)$.

Given the above intuition for the co-states under full information, it seems that we could interpret γ_t as the agent's promised utility conditional on $\theta = \theta_0$ and $\gamma_t + \lambda_t$ as the agent's promised utility conditional on $\theta = \theta_1$. Furthermore, the agent's promised utility under incomplete information would then be expressed as

$$W_t(\boldsymbol{a^r}) = p_t^r W_t(\boldsymbol{a^r}; \theta_1) + (1 - p_t^r)W_t(\boldsymbol{a^r}; \theta_0) = \gamma_t + p_t^r\lambda_t.$$

However, the above interpretation is problematic: when there is uncertainty about θ, the filtration $\sigma(Y_s, a_s^r, \theta : 0 \le s \le t)$ is richer than $\mathcal{F}_t^{a^r} \equiv \mathcal{F}_t^Y$. As γ_t and λ_t are $\mathcal{F}_t^Y$-adapted, such an interpretation would be inaccurate. Nevertheless, Proposition 1 shows that the intuition still holds.

The other key variable accounts for the long term effects of the recommended action choices on the agent's subsequent incentives. Let

$$R_t \triangleq E_t^{a^r}\left[\int_t^T e^{-r(s-t)}\zeta_s p_s^r(1 - p_s^r)\Delta f(a_s^r)ds\right],$$

with terminal condition $R_T = 0$. The agent always has (weakly) more information than the principal at any given time; he chooses the action policy that generates the data used to update beliefs. Any deviation at time t from the recommendation will generate asymmetric information through the posterior beliefs for the remainder of the contract.

To be more concrete, suppose the agent makes a "one-shot deviation" at time t from the recommendation a_t^r to $\tilde{a}_t \in A$. His belief process for the remainder of the contract is $(p_s^{\tilde{a}})_{t<s\le T}$. The principal, unaware of the deviation, continues to hold the belief process $(p_s^r)_{t<s\le T}$. Thus,

the deviation generates a persistent difference of $(p_s^{\tilde{a}} - p_s^r)_{t<s\leq T}$ between the agent's and the principal's beliefs.

In order to provide incentives to the agent at time $k > t$, the principal increases the agent's promised utility by $\zeta_k\big(dY_k - f(p_k^r, a_k^r)dk\big)$ which has an expected value of zero at time t under the probability measure Q^{a^r}.[13] However, from the agent's perspective, his incentives at time k have changed to

$$\underbrace{\zeta_k\big(dY_k - f(p_k^{\tilde{a}}, a_k^r)dk\big)}_{\text{expected value of zero under } Q^{\tilde{a}}} + \zeta_k \underbrace{\big(f(p_k^{\tilde{a}}, a_k^r) - f(p_k^r, a_k^r)\big)}_{=\Delta f(a_k^r)(p_k^{\tilde{a}} - p_k^r)} dk.$$

Aggregating the change in incentives over the remaining time $(t, T]$ yields the expected payoff difference from the deviation

$$E_t^{\tilde{a}}\left[\int_t^T e^{-r(s-t)}\zeta_s \Delta f(a_s^r)(p_s^{\tilde{a}} - p_s^r)ds\right].$$

Notice the similarity to the persistent marginal effects in (6) from the two-period , two outcome model of Section 1.1.3. The process R_t captures the discounted expected marginal benefit of such a deviation from the recommended action.

Following the dynamic mechanism design approach of Pavan et al. (2014), R_t can further be understood as the dynamic virtual surplus associated with the agent's "hidden type" (his true posterior belief). It accounts for the marginal effect of the agent's current "type report" on his payoffs as well as on his future "reports".[14] As I show below, R_t is equivalent to the covariance between the agent's total payoff and his on-path beliefs.

Proposition 2 *There exists a square-integrable and $\mathcal{F}_t^Y$-adapted process $\sigma\beta_t$ such that*

$$dR_t = \Big(rR_t - \zeta_t p_t^r(1 - p_t^r)\Delta f(a_t^r)\Big)dt + \sigma\beta_t dZ_t^{a^r}, \text{ with } R_T = 0. \tag{17}$$

Furthermore, R_t is equivalent to

i. $e^{rt}\text{cov}_t\big(V_T(\boldsymbol{a^r}), p_T^r\big)$, *and*

ii. $\lambda_t p_t^r(1 - p_t^r)$,

[13] More precisely, $E_t^{a^r}[\zeta_k\big(dY_k - f(p_k^r, a_k^r)dk\big)] = E_t^{a^r}[\sigma\zeta_k dZ_k^{a^r}] = 0$, for $k > t$.
[14] The latter effect is called the impulse response function.

almost surely.

The necessary conditions for incentive compatibility can now be expressed using $W_t(\boldsymbol{a}^r)$ and R_t as the key state variables. Theorem 2 generalizes the necessary conditions for incentive compatibility in the existing literature to a setting with both incomplete information, and general production and utility functions.

Theorem 2 *Suppose assumption A.1-A.4 hold. If a contract $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T \rangle$ is incentive compatible, there exist $\mathcal{F}_t^Y$-adapted and square-integrable processes $(W_t(\boldsymbol{a^r}), \sigma\zeta_t)$ and $(R_t, \sigma\beta_t)$ such that*

i. $W_t(\boldsymbol{a^r})$ evolves according to (8) with $W_T(\boldsymbol{a^r}) = U(C_T)$,

ii. R_t evolves according to (17) with $R_T = 0$, and

iii. For each t, a_t^r solves

$$\max_{a_t \in A} u(a_t, c_t) + \zeta_t \bar{f}(p_t^r, a_t) + \beta_t \Delta f(a_t) - R_t \frac{\Delta f(a_t^r)}{\sigma^2} \Big(f(\theta_1, a_t) - p_t^r \Delta f(a_t) \Big) \tag{18}$$

almost surely.

Consider first the case of complete information with $p_0 \in \{0, 1\}$ and $p_t^a = p_0$ for any t and any action policy $\boldsymbol{a}$. Then $R_t \equiv \lambda_t p_0(1 - p_0) = 0$ at all time t. As R_t is a constant process, the sensitivity process β_t is also equivalent to zero for all t. Hence, the necessary condition for incentive compatibility given by (18) reduces to

$$a_t^r \in \arg\max_{a_t \in A} u(a_t, c_t) + \zeta_t \underbrace{\bar{f}(p_t^r, a_t)}_{\bar{f}(p_0, a_t)}.$$

At each time t, the agent's promised utility increases by ζ_t per additional unit of output. A marginal increase in the action a_t would increase the average flow of output by $\bar{f}_a(p_t^r, a_t)$. Consequently, the marginal benefit to the agent is given by $\zeta_t \bar{f}_a(p_t^r, a_t)$ whereas the marginal cost from an increase in the action is $u_a(a_t, c_t)$. An incentive compatible recommendation a_t^r will balance these two opposing forces. In particular, if a_t^r is in the interior of A, then

$$\zeta_t = -\frac{u_a(a_t^r, c_t)}{\bar{f}_a(p_t^r, a_t^r)},$$

which is a similar necessary condition to the two-period, two-outcome model discussed in Section 1.1.1. Moreover, the necessary conditions from Sannikov (2008) are a special case of (18) with $f(\theta, a) = a$ and an additively separable utility function $u(a, c)$.

Consider next a marginal increase in the agent's action a_t^r under incomplete information, i.e., $p_0 \in (0, 1)$.[15] The instantaneous marginal effects still take the form

$$\underbrace{u_a(a_t^r, c_t)}_{\text{marginal cost}} \qquad + \qquad \underbrace{\zeta_t \bar{f}_a(p_t^r, a_t^r)}_{\substack{\text{marginal benefit from} \\ \text{raising the average flow of output}}} \qquad .$$

In addition to the instantaneous marginal effects, there is an effect on how the agent's subsequent beliefs evolve, and ultimately, how he responds to incentives over the remaining time period.

Suppose the agent makes a "one-shot deviation" at time t from the recommendation a_t^r to $\tilde{a}_t \in A$. As discussed earlier, the deviation strategy changes the agent's time $k > t$ incentives by $\zeta_k \Delta f(a_k^r)(p_k^{\tilde{a}} - p_k^r)dk$. Dividing this term by $\tilde{a}_t - a_t^r$ and taking the limit as the difference goes to zero gives

$$\zeta_k \Delta f(a_k^r) \frac{dp_k^r}{da_t^r} dk$$

which is the persistent marginal effect of a_t^r on the agent's time k incentives. Unfortunately, deriving an expression for dp_k^r/da_t^r proves difficult for general production functions.

To simplify, let $f(\theta, a) = \theta + a$ which implies $\Delta f(a) = \theta_1 - \theta_0 \triangleq \Delta_\theta$ is a constant. The necessary condition for incentive compatibility given by (18) reduces to

$$a_t^r \in \arg\max_{a_t \in A} u(a_t, c_t) + \zeta_t a_t - R_t \frac{\Delta_\theta}{\sigma^2} a_t. \tag{19}$$

Let $A_k^r = \int_0^k a_s^r ds$ be the agent's cumulative action at time k when he follows the recommendation. Conditional on parameter θ, the total output at time k is normally distributed with

$$Y_k \sim \mathcal{N}(A_k^r + \theta k, \sigma^2 k).$$

[15] It is no longer the case that $p_t^a = p_0$ for all t and all action policy $\boldsymbol{a}$.

Using Bayes rule, the agent's on-path posterior at time k is

$$p_k^r = \frac{p_0 \exp\left(-(Y_k - A_k^r - \theta_1 k)^2/2\sigma^2 k\right)}{p_0 \exp\left(-(Y_k - A_k^r - \theta_1 k)^2/2\sigma^2 k\right) + (1-p_0)\exp\left(-(Y_k - A_k^r - \theta_0 k)^2/2\sigma^2 k\right)}$$

$$= \frac{p_0 \exp\left(\frac{\Delta_\theta}{2\sigma^2}\left(2(Y_k - A_k^r) - (\theta_0 + \theta_1)k\right)\right)}{p_0 \exp\left(\frac{\Delta_\theta}{2\sigma^2}\left(2(Y_k - A_k^r) - (\theta_0 + \theta_1)k\right)\right) + 1 - p_0}.$$

By taking the derivative of the above expression with respect to a_t^r, the marginal effect of a_t^r on time k incentives can be expressed as

$$\zeta_k \Delta f(a_k^r) \frac{dp_k^r}{da_t^r} = -\zeta_k p_k^r (1 - p_k^r) \frac{\Delta_\theta^2}{\sigma^2} \frac{dA_k^r}{da_t^r} = -\zeta_k p_k^r (1 - p_k^r) \frac{\Delta_\theta^2}{\sigma^2}.$$

Collecting all such terms for $k \in (t, T]$, discounting them, and taking expectations yields

$$-E_t^{a^r}\left[\int_t^T e^{-r(s-t)} \zeta_k p_k^r (1 - p_k^r) \frac{\Delta_\theta^2}{\sigma^2}\right] = -R_t \frac{\Delta_\theta}{\sigma^2}.$$

Hence, the total marginal effect of the agent's action a_t^r is given by

$$u_a(a_t^r, c_t) + \zeta_t - R_t \frac{\Delta_\theta}{\sigma^2}$$

which corresponds to the maximization problem given in (19).[16] Thus, an incentive compatible contract balances (i) the instantaneous marginal cost of action, (ii) the instantaneous marginal increase in promised utility from increasing the average output flow, and (iii) the discounted long-run marginal effect of the recommended effort on future incentives captured through the covariance of payoffs and beliefs.

Theorem 2 generalizes the rough intuition presented in the linear additive case to more general production functions. Recall that $|\Delta f(a)|$ measures the informativeness of the observed output flow. Specifically, a large and positive (negative) $\Delta f(a)$ is a strong indication in favor of θ_1 (θ_0). In the general case, $\Delta f(\cdot)$ is not a constant which implies the agent's actions affect the quality of information. Hence, a deviation at time t not only creates a persistent gap between

[16] A similar necessary condition is derived by Prat and Jovanovic (2014) and Demarzo and Sannikov (2016) for linear additive production functions with a normally distributed parameter.

the agent's and the principal's beliefs but also affects the rate of instantaneous learning. For example, if $\Delta f'(a) > 0$, the agent can increase the quality of information by increasing his action choices, thereby increasing the variance in his belief process.[17] The term $\beta_t \Delta f'(a_t^r)$ in (18) accounts for this marginal effect of the agent's action on the sensitivity of the payoff-belief covariance.

I conclude this section by comparing the sensitivity of the agent's promised utility under complete and incomplete information. This corresponds to the amount of risk the agent needs to be exposed to in order to incentivize him.

Corollary 1 *Suppose assumptions A.1-A.4 hold,* $\text{sign}(\Delta f(a)) = \text{sign}(\Delta f(a')), \forall a, a' \in A$, *and* $\Delta f(a)$ *is monotone (increasing or decreasing). Given an incentive compatible contract* $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T \rangle$, *if*

i. The processes (η_t, φ_t) *from Theorem 1 are non-negative, and*

ii. $a_t^r \in Int(A)$,

then $\zeta_t \geq \min_{\theta \in \Theta} -u_a(a_t^r, c_t)/f_a(\theta, a_t^r)$.

Suppose the sensitivities of the co-states, (η_t, φ_t), are non-negative.[18] Assume that one of the parameters is more productive than the other, for example, $f(\theta_1, a) \geq f(\theta_0, a)$ for all $a \in A$. If higher actions lead to more informative signals, i.e., $\Delta f'(\cdot) \geq 0$, then, a principal who wishes to induce a particular level of effort must expose the agent to more risk under incomplete information than she would have under complete information with $\theta = \theta_1$. On the other hand, if higher actions lead to less informative signals, i.e., $\Delta f'(\cdot) \leq 0$, then the principal must expose the agent to more risk under incomplete information than she would have under complete information with $\theta = \theta_0$. As a consequence, if $f(\theta, a)$ is additively separable, then the agent is exposed to more risk under incomplete information.

3.3 Sufficient Conditions

In continuous time models of moral hazard with complete information, the necessary conditions for incentive compatibility are also sufficient.[19] In contrast, characterizing general and

[17] The variance increases because the more informative signal allows the agent to updates his posteriors more aggressively.

[18] While the sign of (η_t, φ_t) is an outcome of the principal's problem, heuristically, negative sensitivities serve to only reduce the agent's incentive to work at time t.

[19] See Sannikov (2008).

tractable sufficient conditions for models of moral hazard with private information is still an open question. Most papers instead use the first-order approach: solve a relaxed version of the principal's problem using the necessary conditions and check ex-post that the contract is incentive compatible. While the focus of this paper is on identifying necessary conditions for incentive compatibility under incomplete information, I provide a general, albeit intractable, sufficient condition.

Theorem 3 *Suppose assumptions A.1-A.4 hold. A contract $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T \rangle$ is incentive compatible if there exist $\mathcal{F}_t^Y$-adapted, square-integrable co-states $(\gamma_t, \sigma\eta_t)$ and $(\lambda_t, \sigma\varphi_t)$ such that*

i. γ_t evolves according to (9) with $\gamma_T = U(C_T)$,

ii. λ_t evolves according to (10) with $\lambda_T = 0$,

iii. For each t, a_t^r solves (11) almost surely, and

iv. For any arbitrary admissible policy $\boldsymbol{a}$ and associated belief process $(p_t^a)_{0\leq t\leq T}$,

$$E_0^a\left[\int_0^T e^{-rt}(p_t^a - p_t^r)\left(\eta_t\Big(\Delta f(a_t) - \Delta f(a_t^r)\Big) + \varphi_t\Big(f(\theta_1, a_t) - f(\theta_1, a_t^r)\Big)\right)dt\right] \leq 0. \qquad (20)$$

Theorem 3 states that the necessary conditions along with the additional condition (20) are sufficient for incentive compatibility. Under complete information, $p_t^a = p_t^r = p_0$ for all t, which implies that (20) always holds with equality. Thus, the necessary conditions in Theorem 1 are also sufficient. Unfortunately, when there is incomplete information, it is generally difficult to check if condition (20) holds without imposing a lot of structure on the production function.

Corollary 2 *Suppose assumptions A.1-A.4 hold. Further assume that (A) the production function $f(\theta, a)$ is additively separable with $\Delta f(\cdot) \geq 0$, and (B) for each t, the agent's action space is $A_t \triangleq [0, a_t^r]$. A contract $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T \rangle$ is incentive compatible if there exist $\mathcal{F}_t^Y$-adapted and square-integrable co-states $(\gamma_t, \sigma\eta_t)$ and $(\lambda_t, \sigma\varphi_t)$ such that*

i. γ_t evolves according to (9) with $\gamma_T = U(C_T)$,

ii. λ_t evolves according to (10) with $\lambda_T = 0$,

iii. For each t, a_t^r solves

$$\max_{a_t \in A_t} u(a_t, c_t) + \eta_t \bar{f}(p_t^r, a_t) + \varphi_t p_t^r f(\theta_1, a_t)$$

almost surely, and

iv. For each t, $\varphi_t \geq 0$.

The corollary applies to "shirking models" in which the agent can choose an action level up to the recommendation but not more.[20] For example, let the recommendation a_t^r be the amount of an input resource (raw materials, capital investment, or time) the principal allocates for production. A contract is incentive compatible if the agent chooses to use all of the allocated input resource as intended instead of diverting part of it for his personal use.

When the production function is additively separable and θ_1 is the more productive parameter, then any deviation (shirking) strategy generates a belief process that is more optimistic than that of the principal's. To see why, let $f(\theta, a) = \theta + a$. Suppose the agent shirks at time t and chooses $a_t < a_t^r$. The principal attributes only $dY_t - a_t^r dt$ of the total observed output flow to the parameter whereas the agent attributes $(a_t^r - a_t)dt > 0$ more to the parameter. Hence, the agent believes the parameter is more productive than the principal does which implies $p_s^a \geq p_s^r$ for all $s \geq t$. Given the additional structure imposed in Corollary 2, the expression in (20) simplifies to

$$E_0^a \left[\int_0^T e^{-rt} \varphi_t \underbrace{(p_t^a - p_t^r)}_{\geq 0} \underbrace{(a_t - a_t^r)}_{\leq 0} dt \right].$$

Thus, the necessary conditions along with a non-negative process $(\varphi_t)_{t\geq 0}$ are sufficient for a contract to be incentive compatible.

4 Conclusion

In this paper, I consider agency models under incomplete information to study the implicit persistent effects of effort on future incentive. I show the principal must keep track of the agent's promised utility, his on-path belief, and the covariance between beliefs and total payoffs. The latter captures the long-run effects from a marginal increase in current effort.

Since the production function places minimal restrictions on how effort and the parameter interact, the necessary conditions I derive can be used to explore dynamic agency problems using the first-order approach in models that have yet been considered. For example, the model can be used to study an extension of Bolton and Harris (1999) in which the principal can choose to either invest in a safe status-quo project or to finance a risky project managed by an agent.

[20] While the action space so far has been exogenously given and stationary, all the results can be directly extended to the case where $A_t \triangleq [0, a_t^r]$ by treating the action space as part of the contract specifications.

5 Appendix

Proof. Theorem 1:

Let (γ_t, η_t) and (λ_t, φ_t) be the $\mathcal{Z}_t^0$-adapted, square-integrable co-states associated with the state variables Γ_t^a and X_t^a respectively. The Hamiltonian for the control problem (RAP) is given by

$$\mathcal{H}(a, c, \Gamma, X, \gamma, \eta, \lambda, \varphi) \triangleq \Gamma u(a, c) + \bar{f}\left(\frac{X}{\Gamma}, a\right)\Gamma\eta + X\varphi f(\theta_1, a). \tag{12}$$

If the contract $\langle \boldsymbol{a^r}, \boldsymbol{c}, C_T\rangle$ is incentive compatible, then $\boldsymbol{a^r}$ must be a solution to (RAP). Therefore, at each t, a_t^r must solve

$$\max_{a_t \in A} \mathcal{H}(a_t, c_t, \Gamma_t^{a^r}, X_t^{a^r}, \gamma_t, \eta_t, \lambda_t, \varphi_t). \tag{13}$$

Furthermore, the co-states evolve according to:

$$d\gamma_t = \left(r\gamma_t - \mathcal{H}_\Gamma(a_t, c_t, \Gamma_t^{a^r}, X_t^{a^r}, \gamma_t, \eta_t, \lambda_t, \varphi_t)\right) dt + \eta_t dZ_t^0$$

$$= \left(r\gamma_t - u(a_t^r, c_t) - \sigma\eta_t\left(\frac{\bar{f}\left(\frac{X_t^{a^r}}{\Gamma_t^{a^r}}, a_t^r\right)}{\sigma} - \frac{\Delta f(a_t^r)}{\sigma}\frac{X_t^{a^r}}{\Gamma_t^{a^r}}\right)\right) dt + \sigma\eta_t dZ_t^0, \tag{14}$$

with terminal condition $\gamma_T = \frac{\partial \Gamma_T^{a^r} U(C_T)}{\partial \Gamma} = U(C_T)$, and

$$d\lambda_t = \left(r\lambda_t - \mathcal{H}_X(a_t, c_t, \Gamma_t^{a^r}, X_t^{a^r}, \gamma_t, \eta_t, \lambda_t, \varphi_t)\right) dt + \sigma\varphi_t dZ_t^0$$

$$= \left(r\lambda_t - \sigma\eta_t\frac{\Delta f(a_t^r)}{\sigma} - \sigma\varphi_t\frac{f(\theta_1, a_t^r)}{\sigma}\right) dt + \sigma\varphi_t dZ_t^0, \tag{15}$$

with terminal condition $\lambda_T = \frac{\partial \Gamma_T^{a^r} U(C_T)}{\partial X} = 0$.

Using the fact that $\Gamma_t^{a^r} > 0$ and $\frac{X_t^{a^r}}{\Gamma_t^{a^r}} = p_t^r$, we can simplify (13) to (11). Additionally, using

$$dZ_t^{a^r} = dZ_t^0 - \frac{\bar{f}(p_t^r, a_t^r)}{\sigma}dt = dZ_t^0 - \frac{\bar{f}\left(\frac{X_t^{a^r}}{\Gamma_t^{a^r}}, a_t^r\right)}{\sigma}dt,$$

we can further rewrite (14) as (9) and (15) as (10). Finally, notice that if the agent follows the recommendations, he does not have any more information than the principal on-path. Hence, the filtrations $\mathcal{F}^Y, \mathcal{F}^a$, and $\mathcal{Z}^0$ coincide. ■

Proof. Proposition 1:

Applying Itô's lemma shows that $\gamma_t + p_t^r \lambda_t$ evolves according to

$$\Big(r(\gamma_t + p_t^r \lambda_t) - u(a_t^r, c_t)\Big)dt + \sigma \left(\eta_t + p_t^r \varphi_t + \lambda_t p_t^r (1-p_t)^r \frac{\Delta f(a_t^r)}{\sigma^2}\right) dZ_t^{a^r},$$

with the terminal conditions for the co-states satisfying $U(C_T) = \gamma_T + p_T^r \lambda_T = W_T(\boldsymbol{a^r})$. From (8),

$$U(C_T) = W_t(\boldsymbol{a^r}) + \int_t^T \big\{rW_s(\boldsymbol{a^r}) - u(a_s^r, c_s)\big\}ds + \int_t^T \sigma \zeta_s dZ_s^{a^r}$$

$$= \gamma_t + p_t^r \lambda_t + \int_t^T \Big\{r(\gamma_s + p_s^r \lambda_s) - u(a_s^r, c_s)\Big\}ds + \int_t^T \sigma \left(\eta_s + p_s^r \varphi_s + \lambda_s p_s^r (1-p_s)^r \frac{\Delta f(a_s^r)}{\sigma^2}\right) dZ_s^{a^r}$$

for each t. As such a decomposition is almost surely unique, we have $\gamma_t + p_t^r \lambda_t = W_t(\boldsymbol{a^r})$ and

$$\zeta_t = \eta_t + p_t^r \varphi_t + \lambda_t p_t^r (1-p_t)^r \frac{\Delta f(a_t^r)}{\sigma^2}$$

for all $t \leq T$ almost surely. ■

Proof. Proposition 2: Let

$$\hat{R}_t \triangleq E_t^{a^r}\left[\int_0^T e^{-rs} \zeta_s p_s^r (1-p_s^r) \Delta f(a_s^r) ds\right],$$

which is an $\mathcal{F}_t^Y$-martingale. Hence, there exists a square-integrable and $\mathcal{F}_t^Y$-adapted process $e^{-rt}\sigma\beta_t$ such that $d\hat{R}_t = e^{-rt}\sigma\beta_t dZ_t^{a^r}$. Since,

$$R_t = e^{rt}\hat{R}_t - \int_0^t e^{-r(s-t)} \zeta_s p_s^r (1-p_s^r) \Delta f(a_s^r) ds,$$

an application of Itô's lemma yields (17).

To see that $R_t = e^{rt}\text{cov}_t\big(V_T(\boldsymbol{a^r}), p_T^r\big)$, note that the covariation between the agent's expected utility and his on-path beliefs is governed by

$$d\langle V(\boldsymbol{a^r}), p^r\rangle_t = e^{-rt}\zeta_t p_t^r(1-p_t^r)\Delta f(a_t^r)dt$$

implying

$$\begin{aligned}
R_t =& e^{rt}E_t^{a^r}\left[\int_t^T d\langle V(\boldsymbol{a^r}), p^r\rangle_s\right] \\
=& e^{rt}\left\{E_t^{a^r}\left[V_T(\boldsymbol{a^r})p_T^r\right] - \underbrace{V_t(\boldsymbol{a^r})p_t^r}_{=E_t^{a^r}[V_T(\boldsymbol{a^r})]E_t^{a^r}\left[p_T^r\right]} - \underbrace{E_t^{a^r}\left[\int_t^T V_s(\boldsymbol{a^r})dp_s^r\right]}_{=0} - \underbrace{E_t^{a^r}\left[\int_t^T p_s^r dV_s(\boldsymbol{a^r})\right]}_{=0}\right\} \\
=& e^{rt}\,\mathrm{cov}_t(V_T(\boldsymbol{a^r}), p_T^r).
\end{aligned}$$

Finally, note that $R_T = \lambda_T p_T^r(1-p_T^r) = 0$ because $\lambda_T = 0$ by Theorem 1. Hence,

$$\begin{aligned}
0 =& R_t + \int_t^T \Big(rR_s - \zeta_s p_s^r(1-p_s^r)\Delta f(a_s^r)\Big)ds + \int_t^T \sigma\beta_s dZ_s^{a^r} \\
=& \lambda_t p_t^r(1-p_t^r) + \int_t^T\left(r\lambda_s p_s^r(1-p_s^r) - \underbrace{\left\{\eta_s + \varphi_s p_s^r + \lambda_s p_s^r(1-p_s^r)\frac{\Delta f(a_s^r)}{\sigma^2}\right\}}_{\equiv\zeta_s} p_s^r(1-p_s^r)\Delta f(a_s^r)\right)ds \\
&+ \int_t^T \sigma\left(\lambda_s p_s^r(1-p_s^r)(1-2p_s^r)\frac{\Delta f(a_s^r)}{\sigma^2} + \varphi_s p_s^r(1-p_s^r)\right)dZ_s^{a^r}.
\end{aligned}$$

A similar argument to Proposition 1 establishes that $R_t = \lambda_t p_t^r(1-p_t^r)$ and

$$\beta_t = R_t(1-2p_t^r)\frac{\Delta f(a_t^r)}{\sigma^2} + \varphi_t p_t^r(1-p_t^r)$$

for each $t \leq T$ almost surely. ■

Proof. Theorem 2: From Theorem 1, a necessary condition for incentive compatibility is that the recommendation a_t^r solves (11). We can rewrite (11) as (18) by substituting η_t by

$$\zeta_t - p_t^r \varphi_t - \lambda_t p_t^r (1 - p_t^r) \frac{\Delta f(a_t^r)}{\sigma^2},$$

$\lambda_t p_t^r (1 - p_t^r)$ by R_t, and $\varphi_t p_t^r (1 - p_t^r)$ by

$$\beta_t - R_t (1 - 2p_t^r) \frac{\Delta f(a_t^r)}{\sigma^2}.$$

■

Proof. Corollary 1: If $a_t^r \in Int(A)$, then from the first order conditions associated with (11),

$$u_a(a_t^r, c_t) + \eta_t \bar{f}_a(p_t^r, a_t^r) + p_t^r \varphi_t f_a(\theta_1, a_t^r) = 0. \tag{21}$$

From Proposition 1,

$$\zeta_t = (\eta_t + \varphi_t) p_t^r + \eta_t (1 - p_t^r) + R_t \frac{\Delta f(a_t^r)}{\sigma^2}. \tag{22}$$

Assume that $\Delta f(\cdot)$ does not change sign and the processes $(\eta_t, \varphi_t)_{0 \le t \le T}$ are non-negative. I first show that $R_t \frac{\Delta f(a_t^r)}{\sigma^2} \ge 0$. Substituting for ζ_t in (17), we get

$$dR_t = \left(R_t \left(r - p_t^r (1 - p_t^r) \left(\frac{\Delta f(a_t^r)}{\sigma} \right)^2 \right) - (\eta_t + p_t^r \varphi_t) p_t^r (1 - p_t^r) \Delta f(a_t^r) \right) dt + \sigma \beta_t dZ_t^{a^r}.$$

Hence,

$$R_t = E_t^{a^r} \left[\int_t^T e^{-r(s-t) + \int_t^s p_k^r (1 - p_k^r) \left(\frac{\Delta f(a_k^r)}{\sigma} \right)^2 dk} (\eta_s + p_s^r \varphi_s) p_s^r (1 - p_s^r) \Delta f(a_s^r) ds \right]$$

and the sign of R_t is the same as the sign of Δf.

If $\Delta f'(\cdot) \ge 0$, then $-\frac{u_a(a_t^r, c_t)}{f_a(\theta_0, a_t^r)} \ge -\frac{u_a(a_t^r, c_t)}{f_a(\theta_1, a_t^r)}$. Using the FOC (21) above,

$$(\eta_t + \varphi_t) p_t^r = -\frac{u_a(a_t^r, c_t)}{\bar{f}_a(\theta_1, a_t^r)} - \eta_t \frac{(1 - p_t^r)}{\bar{f}_a(\theta_1, a_t^r)}.$$

Substituting into (22), we have

$$\zeta_t = -\frac{u_a(a_t^r, c_t)}{f(\theta_1, a_t^r)} + \eta_t \frac{(1 - p_t^r) \Delta f'(a_t^r)}{f(\theta_1, a_t^r)} + R_t \frac{\Delta f(a_t^r)}{\sigma^2} \ge -\frac{u_a(a_t^r, c_t)}{f(\theta_1, a_t^r)}.$$

Similarly, if $\Delta f'(\cdot) \leq 0$, then $-\frac{u_a(a_t^r, c_t)}{f_a(\theta_1, a_t^r)} \geq -\frac{u_a(a_t^r, c_t)}{f_a(\theta_0, a_t^r)}$. Using the FOC (21),

$$\eta_t(1 - p_t^r) = -\frac{u_a(a_t^r, c_t)}{\bar{f}_a(\theta_0, a_t^r)} - (\eta_t + \varphi_t)\frac{(p_t^r)}{\bar{f}_a(\theta_0, a_t^r)}.$$

Substituting into (22), we have

$$\zeta_t = -\frac{u_a(a_t^r, c_t)}{f(\theta_0, a_t^r)} - (\eta_t + \varphi_t)\frac{p_t^r \Delta f'(a_t^r)}{f(\theta_0, a_t^r)} + R_t \frac{\Delta f(a_t^r)}{\sigma^2} \geq -\frac{u_a(a_t^r, c_t)}{f(\theta_0, a_t^r)}.$$

■

Proof. Theorem 3: Take any arbitrary admissible action policy $\boldsymbol{a}$, the belief process it generates $(p_t^a)_{0 \leq t \leq T}$, and the belief process generated by the recommendations $(p_t^r)_{0 \leq t \leq T}$. The agent's total payoff is given by

$$V_T(\boldsymbol{a}) = \int_0^T e^{-rs} u(\tilde{a}_s, c_s) ds + e^{-rT} U(C_T).$$

Using the terminal conditions that $\gamma_T = U(C_T)$ and $\lambda_T = 0$, we have

$$V_T(\boldsymbol{a}) = \int_0^T e^{-rs} u(\tilde{a}_s, c_s) ds + e^{-rT} \Big(\gamma_T + \lambda_T p_T^a\Big).$$

From (7), (9), and (10), we can derive the evolution of $e^{-rt}(\gamma_t + \lambda_t p_t^a)$ as

$$e^{-rt}\Bigg(- u(a_t^r, c_t) + \eta_t \Delta f(a_t^r)(p_t^r - p_t^a) + p_t^a \varphi_t \Big(\Delta f(a_t)(1 - p_t^a) - \Delta f(a_t^r)(1 - p_t^r)\Big)\Bigg) dt +$$

$$e^{-rt}\left(\lambda_t p_t^a (1 - p_t^a) \frac{\Delta f(a_t)}{\sigma}\right) dZ_t^a + e^{-rt}\sigma\left(\eta_t + p_t^a \varphi_t\right) dZ_t^{a^r}$$

$$= e^{-rt}\Bigg(- u(a_t^r, c_t) + \eta_t \Big(\bar{f}(p_t^a, a_t) - \bar{f}(p_t^r, a_t^r)\Big) + p_t^a \varphi_t f(\theta_1, a_t) - p_t^r \varphi_t f(\theta_1, a_t^r)\Bigg) dt +$$

$$e^{-rt}\Big(\eta_t \Delta f(a_t^r) + \varphi_t f(\theta_1, a_t^r)\Big)(p_t^r - p_t^a) dt + \underbrace{e^{-rt}\sigma\left(\lambda_t p_t^a (1 - p_t^a) \frac{\Delta f(a_t)}{\sigma^2} + \eta_t + p_t^a \varphi_t\right)}_{\triangleq \mathcal{X}_t} dZ_t^a,$$

where the equality follows by the change of measure from Q^{a^r} to Q^a such that

$$dZ_t^{a^r} = dZ_t^a + \frac{\bar{f}(p_t^a, a_t) - \bar{f}(p_t^r, a_t^r)}{\sigma} dt.$$

Let $H(a_t, p_t) \triangleq u(a_t, c_t) + \eta_t \bar{f}(p_t, a_t) + p_t \varphi_t f(\theta_1, a_t)$. Combining the expression for $V_T(\boldsymbol{a})$ with the evolution of $e^{-rt}(\gamma_t + \lambda_t p_t^a)$, we get

$$V_T(\boldsymbol{a}) = \underbrace{\gamma_0 + \lambda_0 p_0}_{=W_0(\boldsymbol{a^r}) \equiv V_0(\boldsymbol{a^r})} + \int_0^T e^{-rt} u(\tilde{a}_t, c_t) dt + \int_0^T d\Big(e^{-rt}(\gamma_t + \lambda_t p_t^a)\Big)$$

$$= V_0(\boldsymbol{a^r}) + \int_0^T e^{-rt}\Big(H(a_t, p_t^a) - H(a_t^r, p_t^r) - H_p(a_t^r, p_t^r)(p_t^a - p_t^r)\Big) dt + \int_0^T \mathcal{X}_t dZ_t^a.$$

As $V_t(\boldsymbol{a})$ is a Q^a-martingale, $E_0^a\left[V_T(\boldsymbol{a})\right] = V_0(\boldsymbol{a})$. Taking expectations of the right hand side under Q^a yields

$$\begin{aligned}
V_0(\boldsymbol{a}) =& V_0(\boldsymbol{a^r}) + E_0^a\left[\int_0^T e^{-rt}\Big\{H(a_t, p_t^a) - H(a_t^r, p_t^r) - H_p(a_t^r, p_t^r)(p_t^a - p_t^r)\Big\} dt\right] \\
=& V_0(\boldsymbol{a^r}) + E_0^a\left[\int_0^T e^{-rt}\left\{H(a_t, p_t^r) - H(a_t^r, p_t^r) + \Big(H_p(a_t, p_t^r) - H_p(a_t^r, p_t^r)\Big)(p_t^a - p_t^r)\right\} dt\right] \\
\leq& V_0(\boldsymbol{a^r}) + E_0^a\left[\int_0^T e^{-rt}\Big(H_p(a_t, p_t^r) - H_p(a_t^r, p_t^r)\Big)(p_t^a - p_t^r) dt\right] \\
=& V_0(\boldsymbol{a^r}) + E_0^a\left[\int_0^T e^{-rt}(p_t^a - p_t^r)\bigg(\eta_t\Big(\Delta f(a_t) - \Delta f(a_t^r)\Big) + \varphi_t\Big(f(\theta_1, a_t) - f(\theta_1, a_t^r)\Big)\bigg) dt\right],
\end{aligned}$$

where the second equality follows from the linearity of $H(a_t, p_t)$ in p_t, and the inequality follows from the fact that for all $t \in [0, T]$, $a_t^r \in \arg\max_{a \in A} H(a, p_t^r)$ almost surely. Hence, if the second term in the last line is non-positive, we have $V_0(\boldsymbol{a}) \leq V_0(\boldsymbol{a^r})$ for any arbitrary admissible policy $\boldsymbol{a}$.[21] ■

[21] Note that if $H(a_t, p_t)$ was concave in (a_t, p_t), then the necessary conditions would also be sufficient. However, the Hessian of $H(a_t, p_t)$ is not negative semi-definite.

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti. "Optimal cartel equilibria with imperfect monitoring." *Journal of Economic Theory* 39, no. 1 (1986): 251-269.

——. "Toward a theory of discounted repeated games with imperfect monitoring." *Econometrica: Journal of the Econometric Society* (1990): 1041-1063.

Adrian, Tobias, and Mark M. Westerfield. "Disagreement and learning in a dynamic contracting model." *The Review of Financial Studies* 22, no. 10 (2009): 3873-3906.

Bergemann, Dirk, and Ulrich Hege. "Venture capital financing, moral hazard, and learning." *Journal of Banking & Finance* 22, no. 6 (1998): 703-735.

——. "The financing of innovation: Learning and stopping." *RAND Journal of Economics* (2005): 719-752.

Bismut, Jean-Michel. "Conjugate convex functions in optimal stochastic control." *Journal of Mathematical Analysis and Applications* 44, no. 2 (1973): 384-404.

——. "An introductory approach to duality in optimal stochastic control." *SIAM review* 20, no. 1 (1978): 62-78.

Bolton, Patrick, and Christopher Harris. "Strategic experimentation." *Econometrica* 67, no. 2 (1999): 349-374.

Cvitani, Jakša, Xuhu Wan, and Jianfeng Zhang. "Optimal compensation with hidden action and lump-sum payment in a continuous-time model." *Applied Mathematics & Optimization* 59, no. 1 (2009): 99-146.

Demarzo, Peter M., and Yuliy Sannikov. "Learning, termination, and payout policy in dynamic incentive contracts." *The Review of Economic Studies* 84, no. 1 (2016): 182-236.

Green, Edward J. "Lending and the smoothing of uninsurable income." *Contractual arrangements for intertemporal trade* 1 (1987): 3-25.

He, Zhiguo, Bin Wei, Jianfeng Yu, and Feng Gao. "Optimal long-term contracting with learning." *The Review of Financial Studies* 30, no. 6 (2017): 2006-2065.

Hlmstrom, Bengt. "Moral hazard and observability." *The Bell journal of economics* (1979): 74-91.

Holmstrom, Bengt, and Paul Milgrom. "Aggregation and linearity in the provision of intertemporal incentives." *Econometrica: Journal of the Econometric Society* (1987): 303-328.

Hopenhayn, Hugo, and Arantxa Jarque. "Unobservable persistent productivity and long term contracts." *Review of Economic Dynamics* 13, no. 2 (2010): 333-349.

H orner, Johannes, and Larry Samuelson. "Incentives for experimenting agents." *The RAND Journal of Economics* 44, no. 4 (2013): 632-663.

Jarque, Arantxa. "Repeated moral hazard with effort persistence." *Journal of Economic Theory* 145, no. 6 (2010): 2412-2423.

Kallianpur, G., Masatoshi Fujisaki, and Hiroshi Kunita. "Stochastic differential equations for the non linear filtering problem." *Osaka Journal of Mathematics* 9, no. 1 (1972): 19-40.

Kwon, Suehyun. "Dynamic moral hazard with persistent states". *mimeo*, 2017.

Liptser, Robert, and Albert N. Shiryaev. *Statistics of random Processes: I. general Theory.* Vol. 5. Springer Science & Business Media, 2013.
Mirrlees, James. "Notes on welfare economics, information and uncertainty." *Essays on economic behavior under uncertainty* (1974): 243-261.

Pavan, Alessandro, Ilya Segal, and Juuso Toikka. "Dynamic mechanism design: A myersonian approach." *Econometrica* 82, no. 2 (2014): 601-653.

Prat, Julien, and Boyan Jovanovic. "Dynamic contracts when the agent's quality is unknown." *Theoretical Economics* 9, no. 3 (2014): 865-914.

Sannikov, Yuliy. “A continuous-time version of the principal-agent problem.” *The Review of Economic Studies* 75, no. 3 (2008): 957-984.

——. “Moral hazard and long-run incentives.” *Unpublished working paper*, Princeton University (2014).

Sch attler, Heinz, and Jaeyoung Sung. “The first-order approach to the continuous-time principalagent problem with exponential utility.” *Journal of Economic Theory* 61, no. 2 (1993): 331-371.

Spear, Stephen E., and Sanjay Srivastava. “On repeated moral hazard with discounting.” *The Review of Economic Studies* 54, no. 4 (1987): 599-617.

Sung, Jaeyoung. “Linearity with project selection and controllable diffusion rate in continuous-time principal-agent problems.” *The RAND Journal of Economics* (1995): 720-743.

Williams, Noah. “Persistent private information.” *Econometrica* 79, no. 4 (2011): 1233-1275.

——. “A solvable continuous time dynamic principalagent model.” *Journal of Economic Theory* 159 (2015): 989-1015.

Zhou, Xun Yu. ”Sufficient conditions of optimality for stochastic systems with controllable diffusions.” *IEEE Transactions on Automatic Control* 41, no. 8 (1996): 1176-1179.