# Incentive Compatibility with Endogenous States

Suehyun Kwon*

February 5, 2016

**Abstract**

This paper studies incentive compatibility with endogenous states. A principal hires an agent over the infinite horizon, and each period the agent takes an action unobservable to the principal. The principal observes an outcome and makes a payment. The key feature of the model is that the agent's action controls the transition probabilities of the state, which is unobservable to both parties. The paper first develops a dynamic programming method then applies it to study incentive compatibility. The necessary and sufficient condition for IC constraints consists of two IC constraints for each history on the equilibrium path. There is the usual local IC constraint, but the novel finding is the dynamic IC constraint. The agent's equilibrium strategy is linked across the infinite horizon, and it takes into account both direct effects and indirect effects of the agent's deviation. When the agent deviates, it leads to a different cost of action and different transition probabilities this period; but the marginal cost-benefit ratios of the agent's subsequent actions also change. Both the dynamic programming method and the necessary and sufficient condition for IC constraints can be applied to a more general class of models. The IC constraints highlight the difference between exogenous states and endogenous states. The dynamic IC constraint shows that local IC constraints are no longer sufficient in this environment.

## 1 Introduction

Consider an employee accumulating human capital. One can imagine that given the level of human capital the employee has, his effort together with the current human capital pins

*Kwon: University College London, suehyun.kwon@ucl.ac.uk, Department of Economics, University College London, 30 Gordon Street, London, WC1H 0AX United Kingdom. Some of the results were previously presented as "Recursive Formulation with Learning." I thank Dilip Abreu, Eduardo Azevedo, V. Bhaskar, Ken Binmore, Patrick Bolton, Antonio Cabrales, Sylvain Chassang, Yeon-Koo Che, Pierre-André Chiappori, Syngjoo Choi, Martin Cripps, Joyee Deb, Glenn Ellison, Amanda Friedenberg, Camilo García-Jimeno, Emir Kamenica, Philipp Kircher, Jihong Lee, David Levine, Lucas Maestri, George Mailath, Miltos Makris, Michael Ostrovsky, Sven Rady, Ariel Rubinstein, Yuliy Sannikov, Vasiliki Skreta, Bruno Strulovici, Balázs Szentes, Juuso Toikka, Juuso Välimäki and participants at UCL Economics Research Day, ESEM 2014, UCL-LSE Theory Workshop, UEA, Bristol, UCL Brown Bag for helpful comments.

down the distribution of his human capital in the following period. In a competitive market setting with a continuum of levels of human capital, this will be Holmström (1999) with endogenous states. If there are two levels of human capital, one can consider an employee acquiring and maintaining a skill. Once the employee has the skill, his effort determines the probability he maintains or loses the skill. If the employee doesn't have the skill, he can acquire the skill. In these settings, the employer doesn't always observe the employee's effort. Whether the employee has the skill depends on his effort, and it can be thought of as human capital that neither the employer nor the employee observes perfectly. This is a dynamic moral hazard problem where the state variable is endogenous and unobservable.

Little is known about dynamic moral hazard with endogenous states. We can ask which contract maximizes the principal's payoff, which tax policy maximizes welfare for the social planner, or in a competitive market setting, what types of dynamics we expect to observe. Board and Meyer-ter-Vehn (2013) study the competitive market setting, but to the best of my knowledge, there is no paper in the principal-agent setting. Furthermore, before we can study optimal contracts, we need to understand the constraints of the optimization problem, and in particular, we need to understand incentive compatibility.

Specifically, I study the following model. A principal hires an agent over the infinite horizon, and each period, the agent takes an action. The agent's action is unobservable to the principal, so there is moral hazard, and it controls the state transition, so the state is endogenous. The principal observes an outcome every period and makes a payment. The state is never observed, and both parties start with a common prior.

The focus of this paper is how incentive compatibility works in this environment. There are two main results: I develop a dynamic programming method that allows me to study this problem, and I use this method to characterize the necessary and sufficient condition for IC constraints on the equilibrium path.

The dynamic programming method uses a different way to express the continuation value of the agent than the standard method. In the standard dynamic programming, there will be one value function to express the agent's continuation value, but in this setting, because the agent's belief on the state matters for the continuation value, and the beliefs of the principal and the agent diverge after the agent deviates, we need to understand and have a way of expressing the agent's continuation value as a function of his belief. I use hypothetical continuation values to linearize the agent's continuation value. I decompose the continuation values in a way that the agent's belief enters his continuation value linearly, and this allows me to express the agent's continuation value both on and off the equilibrium path with the same set of hypothetical continuation values.

I then continue to characterize the necessary and sufficient condition for IC constraints. I show that two types of IC constraints are necessary after each history. This departs from usual models where local IC constraints are sufficient and there is only one IC constraint after each history. The first IC constraint is similar to usual local IC constraints in a sense that it captures the tradeoff when the agent deviates only once. I consider a deviation where the agent deviates for one period but he conforms to the principal's expectation from the following period on, and the local IC constraint equalizes the marginal cost and the marginal benefit of this deviation.

The novel finding is the second type of IC constraint. It is a dynamic IC constraint that takes into account the agent's equilibrium strategy over the infinite horizon. It is no longer sufficient to consider local deviations, and in order to make sure that a contract is incentive compatible, the dynamic IC constraint has to be satisfied. It takes into account both the direct benefits and indirect benefits of the agent's deviation. After every history, the difference in cost from a deviation is the same as usual. But when the agent deviates, it affects the transition probabilities of the state which is the direct benefit. It also has a lasting effect in subsequent periods through indirect benefits. When the agent deviates, the principal and the agent have different beliefs on the state, and in particular, earlier deviations of the agent change the marginal cost-benefit ratios of subsequent deviations. Even if it's not profitable to deviate on the equilibrium path after a particular history, if the agent has already deviated, then from his perspective, the marginal benefit-cost ratio is different, and he might find it profitable to deviate again. The agent won't find it profitable to deviate in the current period alone, but if he's going to deviate again in the next period, then two deviations together could prove to be profitable, and this is why we need to worry about dynamic incentives.

The necessity of the dynamic IC constraint is new and is a feature of endogenous states. Detailed comparison to existing literature is in section 5.1, and the difference between exogenous states and endogenous states are discussed in section 5.2. In short, my paper can be thought of as Holmström (1999) with endogenous states that provides an analogue of the impulse response function in Pavan, Segal and Toikka (2014). As a result, local IC constraints are no longer sufficient.

The necessary and sufficient condition I find holds for a general class of models. I discuss how the model could be extended in section 5.3, and a few cases to note are as follows. The condition holds regardless of the objective function. Whether the principal is maximizing his own profit or the social planner is maximizing welfare, the incentive compatibility has to be satisfied. It also holds for every incentive-compatible contracts; as long as it is incentive compatible or is an equilibrium, this condition has to be satisfied. It also has to be satisfied whether it's a principal-agent setting or a competitive market setting. It holds regardless of whether a state is absorbing or non-absorbing. It has to be satisfied for both risk-neutral agents and risk-averse agents. As long as we have continuity at infinity, it has to be satisfied regardless of the length of the horizon. In particular, it holds for both finite-horizon models and infinite-horizon models. In short, this is a characterization of incentive compatibility, and whenever we have incentive compatibility, the necessary and sufficient condition has to be satisfied.

The rest of the paper is organized as follows. Section 2 describes the model, and the dynamic programming method is explained in section 3. Section 4 applies the dynamic programming method to study IC constraints, and the results are further discussed in section 5. Section 6 concludes.

## 2 Model

A principal hires an agent over the infinite horizon, $t = 0, 1, 2, \cdots$. The principal and the agent share the common discount factor $\delta \in (0, 1)$. Each period, the principal offers a contract, and the agent decides whether to accept. If the agent accepts, the agent chooses an effort $p \in [0, 1)$ which is his private information. An outcome is realized and observed by both parties. The principal makes a payment, and they move to the next period. If the agent rejects, the parties receive their outside options and continue in the following period.

There are three state variables in each period: they are the underlying state, the agent's action and the outcome. There are two states, $\omega_1$ and $\omega_2$. The states are unobservable to both the principal and the agent, and the parties start with a common prior in the beginning of period 0. The agent's action is his private information which leads to moral hazard, and it controls the transition probabilities of the state. When the agent chooses $p$, the probability of going from $\omega_2$ this period to $\omega_1$ next period is $p$. The probability of staying in $\omega_1$ is $r(p) \in [0, 1]$, where $r(\cdot)$ is differentiable. Let $P(p)$ be the transition matrix for the states; $P_{ij}(p)$ is the probability of going from state $i$ to state $j$ next period. When the agent chooses $p$, $P(p)$ is given by

$$\begin{pmatrix} r(p) & 1 - r(p) \\ p & 1 - p \end{pmatrix}.$$

The cost of effort for the agent is $c(\cdot)$ which is differentiable, weakly convex, increasing and $c(0) = 0$. Since the transition probabilities are controlled by the agent's action, the underlying state is endogenous.

The last state variable is the outcome $y \in \mathcal{Y} \subset \mathbb{R}$. The probability of each outcome is pinned down by the underlying state. Let $f_1(y), f_2(y)$ be probabilities of outcome $y$ in states 1 and 2, respectively. I assume that there is no atom and that $f_1(y), f_2(y) > 0$ for all $y \in \mathcal{Y}$. The state transition depends on the current state and the agent's effort, and the outcome only depends on the current state. This is not a restriction which I'll comment

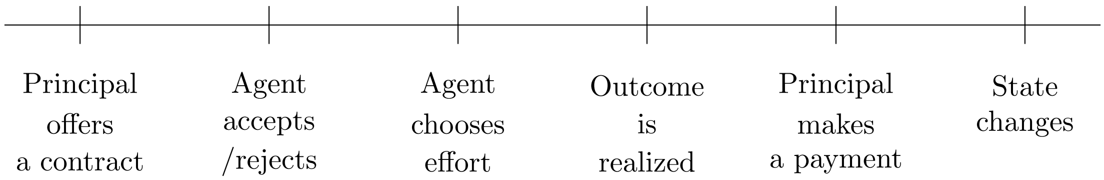


Figure 1: Timing within period $t$

on after describing the timing. The timing of the game is described in Figure 1. When the agent accepts, the state changes at the end of the period after the principal makes a payment. If the agent rejects, the state doesn't change. I chose this timing because it's closer to models with exogenous states, but an alternative timing, which allows the state to change first then allows the outcome to be realized, gives exactly the same results. The outcome distribution is further discussed in section 5.2.

The outside options for each party are normalized to 0. The principal is risk neutral, and the agent's utility function is $u(\cdot)$ which is increasing, concave and $u(0) = 0$. I can allow the agent to be risk neutral, and it is discussed in section 5.3.

No one observes the state, and the agent knows his effort and the outcome. The principal only observes the outcome each period, and he offers payments conditional on the history of all outcomes. I consider pure-strategy sequential equilibria of the game, but the strategy space is not restricted otherwise. I assume that the principal has no commitment power when he offers a contract; this can be generalized to allow different degrees of commitment power, and I discuss it in section 5.3.

Specifically, let $p_t, y_t, w_t$ be the agent's action, the outcome and the payment in period $t$. $d_t = 1(0)$ denotes that the agent accepts (rejects) the contract in period $t$. The public history, which coincides with the principal's history, at period $t$ is $h^t = (d_0, y_0, w_0, d_1, y_1, w_1, \cdots, d_t, y_t) \in \mathcal{H}^t$. If $d_t = 0$, then $y_t = \emptyset = w_t$. The set of public histories is $\mathcal{H} = \cup_{t=0}^{\infty} \mathcal{H}^t$. The agent's private history at period $t$ is $\tilde{h}^t = (d_0, p_0, y_0, w_0, d_1, p_1, y_1, w_1, \cdots, d_{t-1}, p_{t-1}, y_{t-1}, w_{t-1}) \in \tilde{\mathcal{H}}^t$. The set of agent's private histories is $\tilde{\mathcal{H}} = \cup_{t=0}^{\infty} \tilde{\mathcal{H}}^t$. The principal's strategy is $\sigma : \mathcal{H} \to \mathbb{R}$, and the agent's strategy is $\tilde{\sigma} : \tilde{\mathcal{H}} \to \{\emptyset\} \cup [0, 1)$. If the agent chooses $\emptyset$, $d_t = 0$. If the agent chooses $p \in [0, 1)$, then $d_t = 1$ and $p_t = p$. Throughout the paper, tilde denotes the agent's true private history/strategy, and hat denotes the agent's private history/strategy the principal believes is the true history/strategy.

I focus on undetectable deviations in this paper. I assume that after a deviation is detected, the principal and the agent repeat the static Nash equilibrium. The principal makes no payments, and anticipating no payment, the agent puts in zero effort. The parties are indifferent between taking their outside options and no effort. Since the contract offered by the principal, the agent's decision to accept or reject and the principal's payment are all observed by both parties, I focus on incentive compatibility of the agent's choice of effort.

This model can encompass many different settings. There are only two states, but $r(\cdot)$ is not restricted. For example, it can allow for both absorbing states and non-absorbing states. The set of outcomes $\mathcal{Y}$ can be any subset of $\mathbb{R}$, and in particular, it allows for a continuum of outcomes. The set of actions is also a continuum, $[0, 1)$. The agent can be risk neutral or risk averse. Section 5.3 further discusses how the model could be extended.

## 3 Decomposition of Continuation Values

This section describes the main building block for results in section 4. The results I derive in section 4 use an unconventional dynamic programming method, and I will describe how it's done in this section.

The key to my dynamic programming method is the decomposition idea. Instead of considering one continuation value for a given history of outcomes, I decompose it into a linear combination of hypothetical continuation values. The number of hypothetical continuation values I need depends on the number (cardinality) of the states and the outcomes.

Before going into detail, let me describe how standard dynamic programming would approach this problem, and why it doesn't give us the results in this model. In order to use dynamic programming, we need to be able to express the continuation value of the agent by some value function. But we also need an expression for the deviation payoff of the agent as we need to compare the two in the IC constraint. In this class of models, we can't use the same value function on and off the equilibrium path, that only depends on the principal's history. Given a history of outcomes, we know what's the payment the principal is going to make. However we also need to know what's the cost of action for the agent, and more importantly, we need to know the probability of reaching each history. Suppose we know what the agent is going to do in the continuation game. Given a continuation strategy of the agent, we know the cost of action, but we still need to know the probability of reaching each history. The value function needs the agent's belief on the state as one of its arguments. Furthermore, once the agent deviates, the principal and the agent have different beliefs on the state and the principal doesn't even know the agent's belief on the state.

Let me illustrate the last point a bit more. Suppose the belief on the state in the beginning of period $t$ is $\pi^t$. The agent's equilibrium strategy is to choose $p$. When an outcome $y$ is realized, the belief on the state is updated to

$$\pi^0 = \left( \frac{\pi_1^t f_1(y)}{\pi_1^t f_1(y) + \pi_2^t f_2(y)}, \frac{\pi_2^t f_2(y)}{\pi_1^t f_1(y) + \pi_2^t f_2(y)} \right).$$

If the agent takes action $p$, both the principal and the agent believe

$$\pi^{t+1} = \pi^0 P(p) = \left( \frac{\pi_1^t f_1(y)}{\pi_1^t f_1(y) + \pi_2^t f_2(y)}, \frac{\pi_2^t f_2(y)}{\pi_1^t f_1(y) + \pi_2^t f_2(y)} \right) \begin{pmatrix} r(p) & 1 - r(p) \\ p & 1 - p \end{pmatrix}$$

in the beginning of period $t+1$. However, if the agent deviates to $p' \neq p$, the principal still believes $\pi^{t+1}$, but the agent believes

$$\tilde{\pi}^{t+1} = \pi^0 P(p') = \left( \frac{\pi_1^t f_1(y)}{\pi_1^t f_1(y) + \pi_2^t f_2(y)}, \frac{\pi_2^t f_2(y)}{\pi_1^t f_1(y) + \pi_2^t f_2(y)} \right) \begin{pmatrix} r(p') & 1 - r(p') \\ p' & 1 - p' \end{pmatrix}.$$

In order to know the agent's deviation payoff, we need to understand how the value function depends on the belief on the state, and this problem becomes even more complicated when the agent deviates multiple periods. At the bottom line, we cannot use the same value function on and off the equilibrium path.

The decomposition idea circumvents this problem by representing the agent's continuation value, both on and off the equilibrium path, by a linear combination of a set of hypothetical continuation values. The agent's belief on the state enters the linear combination as weights on hypothetical continuation values. The agent's action also enters the linear combination as weights.

To give a concrete example, suppose there are two outcomes. The probability of a high outcome is $p_g \in (0, 1]$ in the good state and $0 \leq p_b < p_g$ in the bad state. The continuation payoff of the agent from period $t+1$ on depends on continuation strategies of the principal and the agent and the agent's belief $\pi^{t+1}$. The deviation payoff of the agent from period $t+1$ on depends on the continuation strategy of the principal conditional on observing outcome $y$, the continuation strategy of the agent after having deviated to $p' \neq p$ and the agent's belief $\tilde{\pi}^{t+1}$.

Now I can decompose the continuation values as follows. Suppose (i) an outcome $y$ is realized in period $t$, (ii) the state in period $t+1$ is $\omega_i$ and (iii) the agent's continuation strategy coincides with what the principal expects him to do. If the agent hasn't deviated up to period $t$, point (iii) just means that the agent follows his equilibrium strategy. If the agent has deviated at some point in the first $t+1$ periods, point (iii) could mean that the agent deviates from his strategy conditional on his private history; but conforming to the principal's expectation is an available strategy for the agent, and suppose for the moment that's the agent's continuation strategy. Furthermore, this strategy includes many of the local IC constraints discussed in the literature. If the agent diverts cash only for this instance and doesn't divert any more in the future, or if the agent misreports his type only this period and reports truthfully from next period on, the strategy I just described includes these strategies.

When the three conditions are met, we can express the agent's continuation value by some number $V_{y\omega_i}$. Since we fixed the strategies of the principal and the agent, we know the agent's action and the principal's payment after each history. If we know the agent's action each period and the state in period $t+1$, we know the probability of each history from period $t+1$ on and can find the continuation value of the agent. I call $V_{y\omega_i}$ a hypothetical continuation value for $(y, \omega_i)$ because it is conditional on getting outcome $y$ this period and going to state $\omega_i$ in the following period. Neither the principal nor the agent observes the state, and even if they happen to be in state $\omega_i$ in the following period, they will never know that the agent's true continuation value is $V_{y\omega_i}$. Nevertheless, we can compute the hypothetical continuation value for each pair of $(y, \omega_i)$.

The next step is to express the agent's continuation value from period $t$ on in terms of $V_{y\omega_i}$'s. Consider the following figure: There are two outcomes, $H$ or $L$ in period $t$, and there

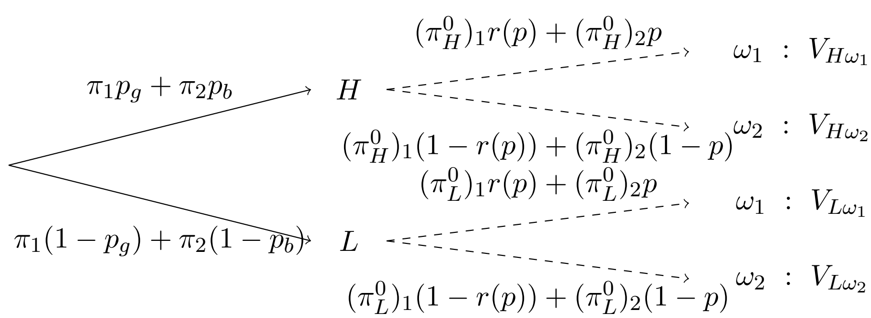


Figure 2: Decomposition with two outcomes

are two states we could be in period $t+1$, $\omega_1$ and $\omega_2$. We can use the above figure to write down the agent's continuation value with $V_{H\omega_1}, V_{H\omega_2}, V_{L\omega_1}, V_{L\omega_2}$, but we can decompose it one step further, and I think it helps with intuition.

Essentially, even though the principal and the agent don't know the state they're in, if they knew which state they are in, they know the exact probabilities of outcomes $H$ and $L$ and the transition probabilities of the state. This explains the probabilities of events in the figure. What needs a bit more attention is the continuation values from period $t+1$ on. There are eight cases to consider: there are two states and two outcomes in period $t$, and there are two states in period $t+1$. My claim is that the agent's hypothetical continuation value doesn't depend on the current state. The payment in period $t$ depends on the outcome, but not the actual state in period $t$. Once we factor out the effects of the state on probabilities in period $t$, the state in period $t$ no longer affects the continuation value from period $t+1$ on. The continuation value from period $t+1$ on only depends

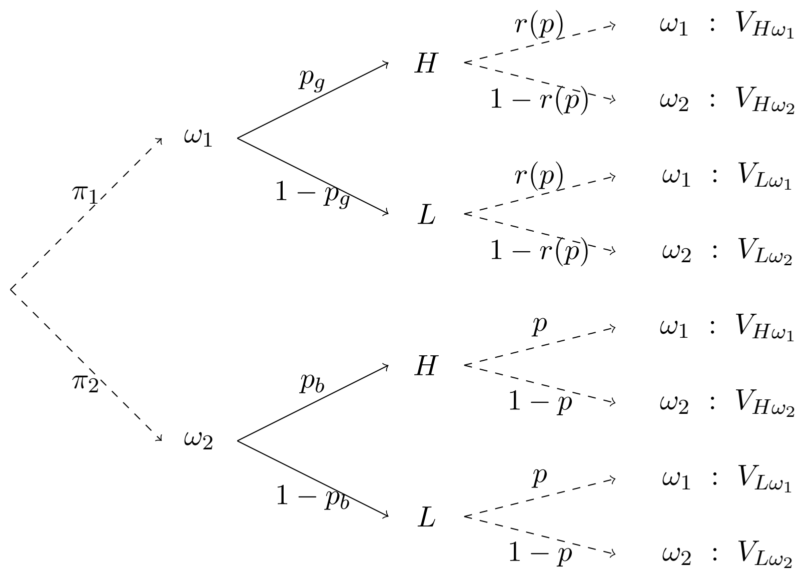


Figure 3: Decomposition with two outcomes (2)

on the continuation strategies and probabilities of each history from that point on; the continuation strategies depend on the history, not the state in period $t$, and once we consider hypothetical continuation values conditional on the state in period $t+1$, the current state no longer matters. Therefore, if the outcome is $y$ and the state in period $t+1$ is $\omega_i$, then the agent's continuation value is $V_{y\omega_i}$ regardless of the current state, and we can express the agent's continuation value as follows:

$$\begin{aligned}&-c(p)+\pi_1(p_g(u(w(H))+\delta(r(p)V_{H\omega_1}+(1-r(p))V_{H\omega_2}))\\&+(1-p_g)(u(w(L))+\delta(r(p)V_{L\omega_1}+(1-r(p))V_{L\omega_2})))\\&+\pi_2(p_b(u(w(H))+\delta(pV_{H\omega_1}+(1-p)V_{H\omega_2}))+(1-p_b)(u(w(L))+\delta(pV_{L\omega_1}+(1-p)V_{L\omega_2})))\end{aligned}$$

where $w(y)$ is the payment for outcome $y$ in period $t$.

This idea generalizes to more than two outcomes. We need one more notation for the agent's strategy. Given the principal's history $h^t$, there exists a private history of the agent $\hat{h}^t$ that the principal believes is the agent's true private history. The agent's deviation strategy I described above is formally $\sigma'_{|\tilde{h}^t,d_t,p'} = \tilde{\sigma}_{|\hat{h}^t,d_t,p_t}$.

**Proposition 1.** *Suppose in period $t$, the agent accepts the contract and chooses $p \in [0,1)$. Let $V_{y\omega_i}$ be the hypothetical continuation value of the agent from period $t+1$ on if (i) an outcome $y$ is realized in period $t$, (ii) the state in period $t+1$ is $\omega_i$ and (iii) the agent's continuation strategy coincides with what the principal expects him to do. ($\sigma'_{|\tilde{h}^t,d_t,p'} = \tilde{\sigma}_{|\hat{h}^t,d_t,p_t}$)*

*Given the agent's belief $\pi$ in the beginning of period $t$, the agent's continuation value from period $t$ on is given by*

$$-c(p)+\pi_1\int u(w(y))+\delta(r(p)V_{y\omega_1}+(1-r(p))V_{y\omega_2})dF_1+\pi_2\int u(w(y))+\delta(pV_{y\omega_1}+(1-p)V_{y\omega_2})dF_2.$$

*Hypothetical continuation values conditional on history up to period $t-1$ are*

$$V_1 = -c(p) + \int u(w(y)) + \delta(r(p)V_{y\omega_1} + (1-r(p))V_{y\omega_2})dF_1,$$
$$V_2 = -c(p) + \int u(w(y)) + \delta(pV_{y\omega_1} + (1-p)V_{y\omega_2})dF_2.$$

The above expression uses more than one hypothetical continuation values to express the agent's continuation value, but its advantage is that we can express the agent's deviation payoffs using the same hypothetical continuation values. Because both the agent's prior and the agent's action enters the continuation value of the agent as weights on hypothetical continuation values, we just need to change the weights to express deviation payoffs, and there's no longer any need for additional value functions. If the agent has a different prior $\tilde{\pi}$, the agent's continuation value from period $t$ on would be

$$-c(p)+\tilde{\pi}_1\int u(w(y))+\delta(r(p)V_{y\omega_1}+(1-r(p))V_{y\omega_2})dF_1+\tilde{\pi}_2\int u(w(y))+\delta(pV_{y\omega_1}+(1-p)V_{y\omega_2})dF_2.$$

If the agent deviates to $p' \neq p$ in period $t$, the agent's deviation payoff would be

$$-c(p')+\pi_1\int u(w(y))+\delta(r(p')V_{y\omega_1}+(1-r(p'))V_{y\omega_2})dF_1+\pi_2\int u(w(y))+\delta(p'V_{y\omega_1}+(1-p')V_{y\omega_2})dF_2.$$

Now we're ready to discuss the IC constraints.

# 4 Incentive-Compatibility Constraints

This section presents main results of the paper. I show that two types of IC constraints are necessary and sufficient to characterize the agent's equilibriums strategy. The first IC constraint is more standard and says that the marginal cost of a local deviation equals the marginal benefit. The second IC constraint is dynamic and new; it links the agent's actions over the infinite horizon. When the agent deviates, it changes the transition probabilities going into the next period, and it both changes the underlying environment and creates information asymmetry between the principal and the agent. In particular, earlier deviations change the marginal cost-benefit ratios of subsequent deviations. If we only consider the first type of IC constraints, there could be profitable multi-period deviations, and it is no longer sufficient to consider one local IC constraint after each history.

First consider the following deviation. The agent hasn't deviated in the first $t$ periods. He deviates in period $t$ but conforms to the principal's expectation from period $t+1$ on. Note that this is different from the usual one-step deviation. The agent's strategy might dictate that after a deviation, the continuation strategy of the agent is different from what the principal expects him to do. If the agent conforms to the principal's expectation from the following period on, this could mean that the agent deviates from his strategy infinitely many times. But in any case, if the agent follows this strategy, the agent's IC constraint looks like the following:

$$
\begin{aligned}
&- c(p) + \pi_1 \int u(w(y)) + \delta(r(p)V_{y\omega_1} + (1-r(p))V_{y\omega_2})dF_1 \\
&+ \pi_2 \int u(w(y)) + \delta(pV_{y\omega_1} + (1-p)V_{y\omega_2})dF_2 \\
\geq &- c(p') + \pi_1 \int u(w(y)) + \delta(r(p')V_{y\omega_1} + (1-r(p'))V_{y\omega_2})dF_1 \\
&+ \pi_2 \int u(w(y)) + \delta(p'V_{y\omega_1} + (1-p')V_{y\omega_2})dF_2. \qquad (1)
\end{aligned}
$$

Remember the agent follows a strategy $\sigma'_{|\tilde{h}^t,d_t,p'} = \tilde{\sigma}_{|\hat{h}^t,d_t,p_t}$. I've used Proposition 1 to express the agent's continuation values on and off the equilibrium path. Since we consider a deviation where the agent conforms to the principal's expectation from $t+1$ on, we can express the agent's payoffs in a tractable way. However, as I'll show later, IC constraints we get with these strategies turn out to be both necessary and sufficient. For now, we only know that this is a necessary condition. Because the agent is allowed to follow this strategy, this type of IC constraint has to be satisfied after every history.

If we look at the IC constraint more closely, one can see that this equalizes the marginal cost and the marginal benefit. (1) is equivalent to

$$
\delta(\pi_1(r(p) - r(p')) \int V_{y\omega_1} - V_{y\omega_2} dF_1 + \pi_2(p-p') \int V_{y\omega_1} - V_{y\omega_2} dF_2) \geq c(p) - c(p'). \qquad (2)
$$

Since the agent can deviate to any $p' \in [0,1]$, we need to consider both the cases $p' > p$ and $p' < p$. If (2) is satisfied for all $p'$, then

$$
\delta(\pi_1 r'(p) \int V_{y\omega_1} - V_{y\omega_2} dF_1 + \pi_2 \int V_{y\omega_1} - V_{y\omega_2} dF_2) = c'(p) \qquad (3)
$$

must hold for $p \neq 0$. Let $\int V_{y\omega_1} - V_{y\omega_2} dF_i = W_i$. In (3), $V_{y\omega_1} - V_{y\omega_2}$ is the marginal benefit of going to $\omega_1$ rather than $\omega_2$ in $t+1$. Therefore, $W_1$ is the expected benefit of going to $\omega_1$ when the outcome in $t$ hasn't been realized yet and they're in $\omega_1$ in $t$. $W_2$ is the analogue for $\omega_2$ in $t$. When the agent deviates to $p' \neq p$, the cost of effort is different, but it also changes the transition probabilities into the next period. The left-hand side of (3) is the marginal benefit of increasing $p$, and the right-hand side is the marginal cost of effort. If the agent conforms to the principal's expectation from $t+1$ on, the IC constraint for a local deviation equalizes the marginal benefit and the marginal cost.

**Proposition 2.** *Suppose the agent's prior in the beginning of period $t$ is $\pi$. Any incentive-compatible contract must satisfy the local IC constraint after every history on the equilibrium path:*

$$\delta(\pi_1 r'(p) W_1 + \pi_2 W_2) = c'(p), \ \forall p \neq 0,$$
$$\delta(\pi_1 r'(0) W_1 + \pi_2 W_2) \leq c'(0).$$

Now consider the following deviation. The agent hasn't deviated in the first $t$ periods. He then deviates two periods in a row, in $t$ and $t+1$, and conforms to the principal's expectation from $t+2$ on. In $t$, the agent deviates to $p' \neq p$, and in $t+1$, the agent deviates to $p'(y) \neq p(y)$. Consider expanding Figure 3 one more period allowing for the continuum of outcomes.

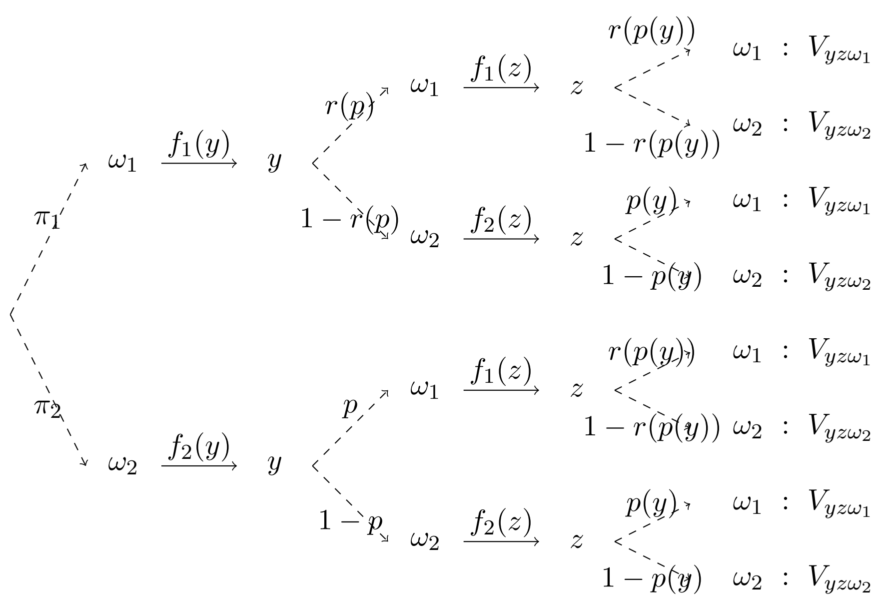


Figure 4: Decomposition for two periods

With a slight abuse of notation, let $V_{yzi}$ be the agent's continuation value if outcomes in period $t$ and $t+1$ are $y$ and $z$, respectively, and the state in $t+2$ is $\omega_i$. Let $W_{yi} = \int V_{yz1} - V_{yz2} dF_i(z)$. If we expand the continuation value of the agent from Proposition 1 for one more period, this is what we get:

$$
\begin{aligned}
&-c(p)+\pi_1\int u(w(y))+\delta(r(p)(-c(p(y))+\int u(w(z))+\delta(r(p(y))V_{yz\omega_1}+(1-r(p(y)))V_{yz\omega_2})dF_1\\
&+(1-r(p))(-c(p(y))+\int u(w(z))+\delta(p(y)V_{yz\omega_1}+(1-p(y))V_{yz\omega_2})dF_2)dF_1\\
&+\pi_2\int u(w(y))+\delta(p(-c(p(y))+\int u(w(z))+\delta(r(p(y))V_{yz\omega_1}+(1-r(p(y)))V_{yz\omega_2})dF_1\\
&+(1-p)(-c(p(y))+\int u(w(z))+\delta(p(y)V_{yz\omega_1}+(1-p(y))V_{yz\omega_2})dF_2)dF_2.
\end{aligned}
$$

If the agent deviates two periods in a row, the IC constraint looks like

$$
\begin{aligned}
&-c(p)+\pi_1\int u(w(y))+\delta(r(p)(-c(p(y))+\int u(w(z))+\delta(r(p(y))V_{yz\omega_1}+(1-r(p(y)))V_{yz\omega_2})dF_1\\
&+(1-r(p))(-c(p(y))+\int u(w(z))+\delta(p(y)V_{yz\omega_1}+(1-p(y))V_{yz\omega_2})dF_2)dF_1\\
&+\pi_2\int u(w(y))+\delta(p(-c(p(y))+\int u(w(z))+\delta(r(p(y))V_{yz\omega_1}+(1-r(p(y)))V_{yz\omega_2})dF_1)\\
&+(1-p)(-c(p(y))+\int u(w(z))+\delta(p(y)V_{yz\omega_1}+(1-p(y))V_{yz\omega_2})dF_2))dF_2\\
\geq&-c(p')+\pi_1\int u(w(y))+\delta(r(p')(-c(p'(y))+\int u(w(z))+\delta(r(p'(y))V_{yz\omega_1}+(1-r(p'(y)))V_{yz\omega_2})dF_1\\
&+(1-r(p'))(-c(p'(y))+\int u(w(z))+\delta(p'(y)V_{yz\omega_1}+(1-p'(y))V_{yz\omega_2})dF_2)dF_1\\
&+\pi_2\int u(w(y))+\delta(p'(-c(p'(y))+\int u(w(z))+\delta(r(p'(y))V_{yz\omega_1}+(1-r(p'(y)))V_{yz\omega_2})dF_1)\\
&+(1-p')(-c(p'(y))+\int u(w(z))+\delta(p'(y)V_{yz\omega_1}+(1-p'(y))V_{yz\omega_2})dF_2))dF_2. \qquad (4)
\end{aligned}
$$

(4) looks complicated, but it can be simplified in the following way.

$$
\begin{aligned}
&\delta(\pi_1(r(p)-r(p'))W_1+\pi_2(p-p')W_2)\\
&+\delta^2\int(r(p(y))-r(p'(y)))W_{y1}(\pi_1 r(p')dF_1+\pi_2 p' dF_2)\\
&+\delta^2\int(p(y)-p'(y))W_{y2}(\pi_1(1-r(p'))dF_1+\pi_2(1-p')dF_2)\\
\geq& c(p)-c(p')+\delta\int c(p(y))-c(p'(y))(\pi_1 dF_1+\pi_2 dF_2). \qquad (5)
\end{aligned}
$$

(5) tells us that the first deviation changes the marginal benefit of the second deviation. When (3) is satisfied, the agent's deviation in period $t$ alone is not profitable. However it changes the marginal benefit of the period-$(t+1)$ deviation through $\pi_1 r(p')dF_1+\pi_2 p'dF_2$ and $\pi_1(1-r(p'))dF_1+\pi_2(1-p')dF_2$, and it might become profitable to deviate in $t+1$.

Looking at (5) more closely, one can see that $W_1$ is the marginal benefit of increasing the probability of being in $\omega_1$ in the following period, conditional on being in $\omega_1$ this period. $W_2$ is the marginal benefit of increasing the probability of being in $\omega_1$ in the following period if they're in $\omega_2$ this period. $W_{y1}$ is the marginal benefit of increasing the probability of being in $\omega_1$ in $t+2$, conditional on the history up to $t$ and being in $\omega_1$ in $t+1$. When the agent deviates to $p'(y) \neq p(y)$, it changes the probability of being in $\omega_1$ in $t+2$, and $(r(p(y)) - r(p'(y)))W_{y1}$ is the benefit of deviating to $p'(y)$ if they're in $\omega_1$ in $t+1$. Furthermore, if the agent deviated to $p' \neq p$ in period $t$, $\pi_1 r(p')dF_1 + \pi_2 p' dF_2$ is the probability of having $y$ in period $t$ and being in $\omega_1$ in $t+1$. Therefore, from period-$t$ perspective, the benefit of period-$(t+1)$ deviation conditional on being in $\omega_1$ in $t+1$ is $\delta^2 \int (r(p(y)) - r(p'(y)))W_{y1}(\pi_1 r(p')dF_1 + \pi_2 p' dF_2)$, which depends on both the period-$t$ deviation and the period-$(t+1)$ deviation. Similarly, $W_{y2}$ is the marginal benefit of increasing the probability of being in $\omega_1$ in $t+2$, conditional on the history up to $t$ and being in $\omega_2$ in $t+1$. From period-$t$ perspective, the benefit of period-$(t+1)$ deviation conditional on being in $\omega_2$ in $t+1$ is $\delta^2 \int (p(y) - p'(y))W_{y2}(\pi_1(1-r(p'))dF_1 + \pi_2(1-p')dF_2)$.

If the agent only deviates in $t+1$, it wouldn't be profitable, but because of period $t$ deviation, now it might become profitable. Even if (3) for all histories are satisfied, it doesn't automatically imply

$$\begin{aligned}
&\delta \int (r(p(y)) - r(p'(y)))W_{y1}(\pi_1 r(p')dF_1 + \pi_2 p' dF_2) \\
&+ \delta \int (p(y) - p'(y))W_{y2}(\pi_1(1 - r(p'))dF_1 + \pi_2(1 - p')dF_2) \\
\geq &\int c(p(y)) - c(p'(y))(\pi_1 dF_1 + \pi_2 dF_2)
\end{aligned}$$

if the agent deviates to $p' \neq p$ in period $t$. This is why these dynamic IC constraints are also necessary, and if these are not satisfied, (3) alone is not sufficient.

(5) generalizes to any $N$-period deviation:

$$\begin{aligned}
&\sum_{n=t}^{t+N-1} \delta^{n-t} \int \delta(\tilde{\pi}_1^n (r(p_n(\hat{h}^n)) - r(p'_n(\tilde{h}^n)))W_1(\hat{h}^{n+1}) \\
&+ \tilde{\pi}_2^n (p_n(\hat{h}^n) - p'_n(\tilde{h}^n))W_2(\hat{h}^{n+1})) - (c(p_n(\hat{h}^n)) - c(p'_n(\tilde{h}^n)))dG \geq 0.
\end{aligned}$$

where $\tilde{\pi}^n$ is the agent's belief given his private history in the beginning of period $n$ and $G$ is the CDF of the outcome given the agent's private history. $\tilde{h}^n$ is the agent's private history at the beginning of period $n$, and $\hat{h}^n$ is what the principal believes is the agent's private history. $p_n$ is the agent's equilibrium strategy, and $p'_n$ is the deviation strategy. We also know that the $N$-period IC constraint is sufficient for the $(N-1)$-period IC constraint, and if we take the limit as $N \to \infty$, then we get

$$\sum_{n=t}^{\infty} \delta^{n-t} \int \delta(\tilde{\pi}_1^n (r(p_n(\hat{h}^n)) - r(p'_n(\tilde{h}^n)))W_1(\hat{h}^{n+1})$$
$$+ \tilde{\pi}_2^n (p_n(\hat{h}^n) - p'_n(\tilde{h}^n))W_2(\hat{h}^{n+1})) - (c(p_n(\hat{h}^n)) - c(p'_n(\tilde{h}^n)))dG \geq 0. \tag{6}$$

**Proposition 3.** *Suppose the agent's prior in the beginning of period $t$ is $\pi$. Any incentive-compatible contract must satisfy the dynamic IC constraint after every history on the equilibrium path:*

$$\sum_{n=t}^{\infty} \delta^{n-t} \int \delta(\tilde{\pi}_1^n (r(p_n(\hat{h}^n)) - r(p'_n(\tilde{h}^n)))W_1(\hat{h}^{n+1})$$
$$+ \tilde{\pi}_2^n (p_n(\hat{h}^n) - p'_n(\tilde{h}^n))W_2(\hat{h}^{n+1})) - (c(p_n(\hat{h}^n)) - c(p'_n(\tilde{h}^n)))dG \geq 0.$$

Before analyzing the implication of the multi-period IC constraint, I'll first show that the local IC constraint and the dynamic IC constraint are necessary and sufficient conditions for all IC constraints on the equilibrium path.

**Proposition 4.** *Suppose the agent's prior in the beginning of period $t$ is $\pi$. The following IC constraints are necessary and sufficient conditions for all IC constraints on the equilibrium path:*

$$\delta(\pi_1 r'(p_t)W_1 + \pi_2 W_2) = c'(p_t), \ \forall p \neq 0$$
$$\delta(\pi_1 r'(0)W_1 + \pi_2 W_2) \leq c'(0),$$
$$\sum_{n=t}^{\infty} \delta^{n-t} \int \delta(\tilde{\pi}_1^n (r(p_n(\hat{h}^n)) - r(p'_n(\tilde{h}^n)))W_1(\hat{h}^{n+1})$$
$$+ \tilde{\pi}_2^n (p_n(\hat{h}^n) - p'_n(\tilde{h}^n))W_2(\hat{h}^{n+1})) - (c(p_n(\hat{h}^n)) - c(p'_n(\tilde{h}^n)))dG \geq 0.$$

Since the agent is always allowed to deviate to the strategies described above, we already know that these IC constraints are necessary. The main idea of the proof for sufficiency is somewhat related to the proof of one-step deviation principle. Suppose there is a profitable deviation $\sigma'$ for the agent that satisfies both of these IC constraints. Suppose $\sigma'$ gives an $\epsilon$-higher payoff to the agent than his equilibrium strategy. In a pure-strategy sequential equilibrium, we have continuity at infinity, and there exist $N$ sufficiently large and another deviation strategy $\sigma''$ such that the agent's payoff from $\sigma''$ is at least $\epsilon/2$ higher than his equilibrium payoff and

$$\begin{aligned}\sigma''(\tilde{h}^t, \tilde{h}^k) =& \sigma'(\tilde{h}^t, \tilde{h}^k) \ for \ k = 0, \cdots, N-1, \\ \sigma''_{|\tilde{h}^t, \tilde{h}^N} =& \tilde{\sigma}_{|\hat{h}^t, \hat{h}^N}\end{aligned}$$

In $\sigma''$, the agent deviates for $N$ periods as in $\sigma'$, but after $N$ periods, the agent conforms to the principal's expectation for the rest of the infinite horizon. However, this contradicts (6), and there does not exist a deviation strategy $\sigma'$ that gives at least $\epsilon > 0$ more to the agent than his equilibrium payoff.

I've shown so far that the two IC constraints for each history are necessary and sufficient conditions to characterize the agent's equilibrium strategy. There are a few things to note here. First, the dynamic IC constraint is new. To the best of my knowledge, the only other paper that looks at unobservable actions with endogenous states and ex-ante symmetric uncertainty is Board and Meyer-ter-Vehn. They consider Markov strategies while I allow for fully history-contingent strategies. I'll also compare to papers on exogenous states in section 5.1, but I would like to highlight that the dynamic IC constraint doesn't exist in the literature. The literature on dynamic moral hazard with ex-ante symmetric uncertainty has focused on models where local IC constraints are sufficient for all IC constraints. My result shows that once we move to endogenous states, we naturally have to consider dynamic incentives and multi-period deviations. This is related to the sufficiency of local IC constraints which I will discuss further in section 5.1.

Each term in the dynamic IC constraint takes into account the difference in costs and the direct benefit of affecting the transition probabilities into the next period. If the agent deviates one period and conforms to the principal's expectation from next period on, the agent's continuation value conditional on the state in the following period is the same as what the principal thinks he provides the agent with. The only difference is that the agent's deviation can affect the probability of being in each state in the following period, and the effect through transition probabilities this period is captured in each term.

However there is also an indirect effect because $\tilde{\pi}_1 dG$ depends on all past efforts, and a deviation in period $t$ has lasting effects. More specifically, the direct effect in the above paragraph are the differences $(r(p) - r(p'))W_{y1}$ and $(p - p')W_{y2}$. Hypothetical continuation values are the same as on the equilibrium path, but the agent affects state transition probabilities. We can also see in the expression that these direct benefits are integrated with respect to the cdf $G$, which depends on the sequence of actions the agent has taken, and the priors $\tilde{\pi}_1$ and $\tilde{\pi}_2$ are also multiplied to these terms. These priors are the agent's priors conditional on his private history. Therefore, if the agent deviates in period $t$, the agent affects the transition probabilities only in period $t$, but his action changes $\tilde{\pi}_1, \tilde{\pi}_2, dG$

in all subsequent periods. Different $\tilde{\pi}_1, \tilde{\pi}_2$ imply that earlier deviations change the marginal cost-benefit ratio of the agent; how these effects aggregate all together depends on $dG$.

Therefore, with endogenous states, we have a new type of IC constraint that hasn't existed in the literature. The dynamic IC constraint captures both direct effects and indirect effects of agent's deviations. I'll comment more on how these relate to papers on exogenous states and other papers with endogenous states in section 5.

Before comparing these IC constraints to the existing literature, let me comment on one more aspect of the dynamic IC constraint. One might think that the dynamic IC constraint takes into account the sequence of agent's equilibrium actions against a sequence of agent's deviations. But this dynamic IC constraint is stronger than that. For example, if we assume that the good state is absorbing and the cost function is quadratic, $c(p) = \frac{1}{2}p^2$, we can simplify the two-period IC constraint as follows:

$$\frac{1}{2}(1-p)^2 \geq \delta \int c(p(y))(\pi_1 dF_1 + \pi_2 dF_2).$$

This has to be satisfied after every history on the equilibrium path.

**Corollary 1.** *Suppose $r(p) = 1$ for all $p \in [0, 1)$ and $c(p) = \frac{1}{2}p^2$. After every history on the equilibrium path, if $p, p(y) \neq 0$, the following two-period IC constraint must hold:*

$$(1-p)^2 \geq \delta \int p(y)^2 (\pi_1 dF_1 + \pi_2 dF_2).$$

Note that this IC constraint is completely in terms of the agent's equilibrium strategy. We already saw from the local IC constraint that it only depends on the agent's equilibrium strategy, but it still takes into account the hypothetical continuation values. The dynamic IC constraint can be simplified to expressions only using equilibrium strategies, and the two-period IC constraint above is only in terms of the agent's actions in two subsequent periods. This shows that there are natural dynamic considerations that one needs to take into account.

Lastly, I'll also comment more on this in the following section, but we can compare the dynamic IC constraint with the impulse response function in Pavan-Segal-Toikka (2014). Their model is about dynamic adverse selection, and what I capture with the dynamic IC constraint here is how to think about dynamic moral hazard. This is not an orthogonalization because the effect of agent's past actions are not orthogonalized; it works through the agent's beliefs and probabilities of each history, but we can think of direct benefits versus indirect benefits.

# 5 Discussion

This section discusses the IC constraints derived in section 4. I will first compare the results to existing literature and then compare the difference between exogenous states and endogenous states. It helps to discuss the existing literature first to understand exactly what changes when the states are endogenous.

I'll then discuss how my results extend to a more general class of models. Because I characterize the necessary and sufficient condition for IC constraints, it has to be satisfied for every incentive-compatible contract or an equilibrium. It also holds for any objective function of the principal, and it holds for competitive-market settings. As long as I have continuity at infinity, I can consider any degree of commitment power, and the results also hold for finite horizon models. Whether the state is absorbing or non-absorbing doesn't matter for my results, and the risk attitude also doesn't matter.

## 5.1 Related Literature

The key aspects of my model are ex-ante symmetric uncertainty, moral hazard and endogenous states. My results hold for both principal-agent settings and competitive-market settings. I will elaborate on related papers on endogenous states in two paragraphs, but first of all, ex-ante symmetric uncertainty and moral hazard are related to career concerns. The competitive-market version of my model can be thought of as Holmström (1999) with endogenous states. The principal-agent version of my model provides an analogue of the impulse response function in Pavan-Segal-Toikka (2014) for dynamic moral hazard with ex-ante symmetric uncertainty.

The analogue with exogenous states in the principal-agent setting would be the class of models where a principal hires an agent and there's ex-ante symmetric uncertainty and moral hazard. This literature includes Prat and Jovanovic (2014) and DeMarzo and Sannikov (2015).[1] The common finding in these papers is that the principal has to leave informational rent to the agent, and the second-best contract focuses on the tradeoff between rent and efficiency. The principal and the agent start with a common prior on the state, and when the agent deviates, their posteriors on the state diverge. The information asymmetry after an agent's deviation persists over time, and the agent collects informational rent because of the potential information asymmetry off the equilibrium path. The analogue with exogenous states is also related to experimentation, including Hörner and Samuelson (2013). Section 5.2 discusses the difference between endogenous states and exogenous states in more detail.

There are few papers with endogenous states when the agent's action is unobservable.

[1]See also Bhaskar (2014), He et al (2014) and Kwon (2015).

Board and Meyer-ter-Vehn (2013) study a competitive-market setting; the firm exerts effort in continuous time, and with a poisson arrival, the quality of the firm changes. The transition probability when the firm gets a chance depends on the firm's effort at the time, and in that sense, the state (quality) is endogenous. The market (consumers) doesn't observe the firm's effort. The difference from my model is that it's a competitive-market setting and the firm knows its own quality. They characterize Markov perfect equilibria while I focus on incentive compatibility with fully history-contingent strategies in this paper. Fershtman and Pakes (2012) study experience-based equilibrium where the space of strategies allowed is different from my model.

In my model, both the underlying state and the agent's action are unobservable to the principal. The underlying state is also unobservable to the agent. The literature on taxation of human capital accumulation mostly considers settings where the state and/or the agent's action is observable to the social planner. This literature allows for endogenous states in the principal-agent setup, but there is no ex-ante symmetric uncertainty or moral hazard. The relevant papers include Farhi and Werning (2013) and Stantcheva (2015). On the other hand, Kapička (2015) and Kapička and Neira (2015) consider settings where both the agent's ability and the agent's action are unobservable to the planner. But the agent knows his own ability in their papers. Makris and Pavan (2016) study the setting where the agent's productivity depends on his productivity and income in the previous period.

Pavan, Segal and Toikka (2014) provide a necessary condition for the agents' IC constraints for dynamic mechanism design in a general quasilinear environment. They also provide sufficient conditions when the state is Markov. My results are the analogue of their impulse response function to dynamic moral hazard and ex-ante symmetric uncertainty. In my model, the agent's effort affects the transition probabilities only within a given period, and the state is both Markov and endogenous. The impulse response function in Pavan, Segal and Toikka orthogonalize the effect of agent's report; in my model, it is not an orthogonalization, because the indirect effects work through the agent's belief and the probability of each history, and the effect of past actions can't be separated out individually. But it shows that the agent's action can be decomposed into residual effects in every period, and essentially, the effect of the agent's actions in a particular period is summarized by one expression. However, there is a crucial difference between their impulse response function and my result. Their necessary condition is the first-order condition for the IC constraints, and when they show the sufficient conditions for Markovian states (which is what I have with endogenous states) the first-order conditions are necessary and sufficient under regularity conditions. I show that the dynamic IC constraint is necessary with moral hazard and endogenous states. The first-order condition or the local IC constraint people typically consider is necessary and it is one of many conditions implied by my condition, but it is never sufficient. We must satisfy the dynamic IC constraint in order for a contract to be incentive compatible. The key difference from their model, and the reason for the necessity of the dynamic IC constraint is that the agent never knows the state in my model and the state is endogenous. The principal and the agent start with a common prior on the state, but because the agent never learns the state and his actions affects transition probabilities, all of his past actions matter for his belief on the state. Whereas in the Markovian environment in Pavan, Segal and Toikka, the agent knows his type, and the mechanism designer knows all of the past allocations, so the relevant uncertainty is the current type.

Because the necessary and sufficient condition for IC constraints requires the dynamic IC constraint, my result is also related to the literature on sufficiency of local IC constraints. As mentioned above, it is related to Pavan, Segal and Toikka (2014), but another relevant paper is Carroll (2012).

## 5.2 Endogenous States versus Exogenous States

To focus on the difference between endogenous states and exogenous states, one should compare my results to papers on dynamic moral hazard with ex-ante symmetric uncertainty. Papers in the principal-agent setup are listed in section 5.1. These papers typically show the sufficiency of local IC constraints and characterize the properties of the second-best contracts. Since my results focus on the necessary and sufficient condition for IC constraints, there are two main differences. First, it is no longer sufficient to consider local IC constraints. I have shown that the dynamic IC constraint, in addition to the usual local IC constraint, has to be satisfied after every history on the equilibrium path. Second, the agent's rent is no longer purely informational. With exogenous states, the agent can create information asymmetry between him and the principal off the equilibrium path, and that's the source of his rent; the second-best contract focuses on the tradeoff between the informational rent and efficiency. My results show that if the state is endogenous, then the agent's rent is more than informational.

I will focus on the nature of agent's rent in this section; it also gives intuition for the necessity of the dynamic IC constraint. Consider the following model with exogenous states. A principal hires an agent over the infinite horizon, and each period, the agent takes an action. An outcome is realized, and the principal makes a payment. The principal only observes the outcomes, and the agent knows his own action and the outcome. The state is never observed, and the principal and the agent start with a common prior. The state follows an exogenous Markov chain, and the distribution of outcomes in a given period is determined by the state and the agent's action.

In this class of models, when the agent deviates, it leads to a different distribution of outcomes. The agent knows his own action and updates his belief accordingly. However, the principal doesn't know that the agent deviated and updates his belief using the equilibrium action of the agent. The principal and the agent have different beliefs on the state after the agent deviates. Once the beliefs diverge, the difference in beliefs persist over time, and the agent collects rent in subsequent periods. Since the state is exogenous, the underlying environment/technology doesn't change, and the agent's rent is purely informational.

On the other hand, when the agent deviates in my model, it changes the transition probabilities of the state in the next period. The agent updates his belief using the action he took, while the principal updates his belief using the agent's equilibrium action. In this regard, the information asymmetry after a deviation is common between endogenous states and exogenous states. The difference in beliefs persists over time with endogenous states as well.

However, the agent's rent is more than informational with endogenous states. When the transition probabilities of the state changes, the agent's deviation has a direct benefit by changing the production technology. It also has indirect benefits by changing the marginal cost-benefit ratio of subsequent periods. With exogenous states, because the rent is informational, one can order the agent's beliefs in terms of tightness of the IC constraint. The sufficiency of local IC constraints implies that the IC constraint is the tightest when the agent has never deviated before, and that's why it is sufficient to consider the local IC constraints. If the agent doesn't want to deviate when he has never deviated, then the agent doesn't want to deviate after a few deviations. On the contrary, the dynamic IC constraint in Proposition 4 shows that local IC constraints are not sufficient with endogenous states. The marginal cost-benefit ratios in subsequent periods change after a deviation, one can no longer order the agent's beliefs in the same way as with exogenous states. Dynamic incentives and multi-period deviations have to be taken into account.

The last comment on the difference between endogenous states and exogenous states is about the outcome distribution in my model. I assumed that the transition probabilities depend on the current state and the agent's action, and the outcome distribution only depends on the current state. This isolates the effect of endogenous states. As mentioned earlier, an alternative timing one can consider is (i) the principal offers a contract (ii) the agent accepts/rejects (iii) the agent exerts effort (iv) the state changes (v) the outcome is realized and (vi) the principal makes a payment. This will have an interpretation that the outcome this period is a signal of the agent's action this period. I get exactly the same results with this timing. I kept the timing in section 2 because it's closer to the timing with exogenous states, but it doesn't affect the results in any way. One can also ask what if the agent's action affects the outcome distribution in the current period as well. In this case, the agent's action affects both the outcome distribution this period and the state transition into the next period. The principal's inference on the state is more complicated than what

I have, but the the principal still infers about one object, the state, from one observation, the outcome, and the main intuition will go through.

### 5.3 Extensions

I derived Propositions 1 to 4 for the model described in section 2, but the results hold for a more general class of models. I would first like to emphasize that this is a necessary and sufficient condition for incentive compatibility. Any incentive-compatible contract or any equilibrium has to satisfy these two IC constraints. This has two main implications. The first implication is that any optimal contract or best equilibrium has to satisfy these IC constraints. The second implication is that they have to be satisfied regardless of the objective; they have to be satisfied in a competitive-market setting, but in case of a principal-agent setting, it has to be satisfied when the principal maximizes his own profit and when the social planner maximizes welfare.

In terms of commitment power, we essentially need continuity at infinity. As long as we have continuity at infinity, the results extend to models where the principal has within-period commitment power, full commitment power or any other variations. If the time horizon is finite, we need to rule out the trivial case where the agent exerts no effort by backward induction, but for instance, if the principal has full commitment power over a finite horizon, my results still hold.

I set up the model with a risk-averse agent, but the results hold with a risk-neutral agent as well. In case of a risk-neutral agent, it would be good to rule out the case where the principal sells the firm to the agent, but we can introduce limited liability with risk neutrality, and we can also consider a competitive-market setting with a risk-neutral agent.

I can also allow the transition probability $r(\cdot)$ to take any form, and in particular, I can allow the state to be absorbing or non-absorbing. The model can also be generalized to allow for a continuum of states; checking both IC constraints will be a bit more complicated, but the main result, that we need both the local IC constraint and the dynamic IC constraint, holds.

In the principal-agent setting, results are based on the payments the agent receives after each history. In a competitive-market setting, one needs to make assumptions about the agent's per-period payoff based on the public history the market observes. The agent's hypothetical continuation values will then be constructed using these per-period payoffs, and the IC constraints resulting from the hypothetical continuation values are the necessary and sufficient condition for the competitive-market setting.

## 6 Conclusion

In this paper, I considered dynamic moral hazard with endogenous states and characterized the necessary and sufficient condition for IC constraints on the equilibrium path. To study incentive compatibility with endogenous states, I first developed a dynamic programming method. It uses the decomposition idea so that the agent's continuation value after a given history is expressed as a linear combination of hypothetical continuation values. It allows us to express the agent's deviation payoffs with the same set of hypothetical continuation values, and the only difference between the continuation values on and off the equilibrium path are the agent's belief and the actions which enter the linear combination as weights. The decomposition itself can be applied to other settings as well. I used it to study endogenous states, but it can be applied to exogenous states with moral hazard and ex-ante symmetric uncertainty. Kwon (2014) has results on exogenous states.

The necessary and sufficient condition for incentive compatibility consists of two types of IC constraints. The local IC constraint is similar to the usual local IC constraints, and it equalizes the marginal cost and the marginal benefit of the agent's deviation if he were to deviate for one period and conforms to the principal's expectation from the following period on. The dynamic IC constraint is the new feature of endogenous states that hasn't existed in the literature. It links the agent's equilibrium actions over the course of infinite horizon and captures both direct effects and indirect effects of the agent's deviation. When the agent deviates, it changes the transition probabilities of the state in the following period, which is the direct benefit. But it changes the agent's belief on the state and also the probability of each history in all subsequent periods. Even though the agent's action only affects the state transition for one period, it changes the marginal cost-benefit ratios of the agent's actions in the future. When the local IC constraint is satisfied, a single deviation is not profitable, but if the agent has already deviated in the past, then an additional deviation might become profitable. We no longer have the sufficiency of local deviations, and the dynamic IC constraint is a necessary condition for incentive compatibility.

My results provide an analogue of Pavan, Segal and Toikka (2014) for dynamic moral hazard with ex-ante symmetric uncertainty. It's not an orthogonalization as in their impulse response function, and the agent's past actions affect a later period through the agent's belief and the probability of reaching a particular history. However, local IC constraints are sufficient with Markov states in dynamic adverse selection while the dynamic IC constraint is necessary in my model. The dynamic IC constraint also highlights the difference between endogenous states and exogenous states. The agent's rent is more than informational when the state is endogenous.

I've already discussed how the necessary and sufficient condition for incentive compatibility holds for a more general class of models, but incentive compatibility is the first step towards understanding endogenous states. Going forward, I could apply the characterization of incentive compatibility to study competitive market settings with general information structure, or I could also study the second-best contract in the principal-agent setup. Taxation of human capital, when it accumulates as a result of agent's action and the agent's action is unobservable, would be a good application for the latter.

## A Proofs

*Proof of Proposition 1.* When the agent conforms to the principal's expection from period $t+1$ on, the principal and the agent have correct beliefs about the continuation game. Suppose the principal follows his equilibrium strategy given $h^t$ in period $t$, we know the strategies of both parties from period $t+1$ on. Furthermore, conditional on the state in $t+1$, we know the probability of each history. Therefore, there exist $V_{y\omega_i}$ for each pair of $(y, \omega_i)$ such that the agent's continuation value from period $t+1$ on conditional on being in $\omega_i$ in period $t+1$ is $V_{y\omega_i}$. In the beginning of period $t$, if the agent accepts the contract, he has to choose $p \in [0, 1)$. The agent's action affects the probability of $(y, \omega_i)$, and the agent's expected payoff is

$$-c(p)+\pi_1 \int u(w(y))+\delta(r(p)V_{y\omega_1}+(1-r(p))V_{y\omega_2})dF_1+\pi_2 \int u(w(y))+\delta(pV_{y\omega_1}+(1-p)V_{y\omega_2})dF_2.$$

Conditional on the state in $t$, the agent's hypothetical continuation values are

$$V_1 = -c(p) + \int u(w(y)) + \delta(r(p)V_{y\omega_1} + (1 - r(p))V_{y\omega_2})dF_1,$$
$$V_2 = -c(p) + \int u(w(y)) + \delta(pV_{y\omega_1} + (1 - p)V_{y\omega_2})dF_2.$$

□

*Proof of Proposition 2.* Suppose the agent's prior at the beginning of period $t$ is $\pi$. The agent's expected payoff from choosing $p$ is

$$-c(p)+\pi_1 \int u(w(y))+\delta(r(p)V_{y\omega_1}+(1-r(p))V_{y\omega_2})dF_1+\pi_2 \int u(w(y))+\delta(pV_{y\omega_1}+(1-p)V_{y\omega_2})dF_2.$$

If the agent deviates to $p' \neq p$ but conforms to the principal's expection from period $t+1$ on, his expected payoff is

$$-c(p')+\pi_1 \int u(w(y))+\delta(r(p')V_{y\omega_1}+(1-r(p'))V_{y\omega_2})dF_1+\pi_2 \int u(w(y))+\delta(p'V_{y\omega_1}+(1-p')V_{y\omega_2})dF_2.$$

Since the agent's strategy from period $t+1$ on coincides with the principal's expectation from period $t+1$ on, the agent's continuation value conditional on the state in $t+1$ is the same as on the equilibrium path. The agent's deviation only affects the transition probabilities in period $t$. Therefore, the local IC constraint is

$$\begin{aligned}
&-c(p)+\pi_1\int u(w(y))+\delta(r(p)V_{y\omega_1}+(1-r(p))V_{y\omega_2})dF_1\\
&+\pi_2\int u(w(y))+\delta(pV_{y\omega_1}+(1-p)V_{y\omega_2})dF_2\\
\geq&-c(p')+\pi_1\int u(w(y))+\delta(r(p')V_{y\omega_1}+(1-r(p'))V_{y\omega_2})dF_1\\
&+\pi_2\int u(w(y))+\delta(p'V_{y\omega_1}+(1-p')V_{y\omega_2})dF_2\\
\Leftrightarrow&\ \delta(\pi_1\int(r(p)-r(p'))(V_{y\omega_1}-V_{y\omega_2})dF_1+\pi_2\int(p-p')(V_{y\omega_1}-V_{y\omega_2})dF_2)\geq c(p)-c(p').
\end{aligned}$$

Let $W_i=\int V_{y\omega_1}-V_{y\omega_2}dF_i$. Since the IC constraint has to hold for all $p'$, it has to hold for both the left limit and the right limit as $p'\to p$, and we have

$$\delta(\pi_1 r'(p)W_1+\pi_2W_2)=c'(p),\ \forall p\neq 0.$$

When $p=0$, we only have the right limit, and we get

$$\delta(\pi_1 r'(0)W_1+\pi_2W_2)\leq c'(0).$$

□

*Proof of Proposition 3.* Suppose the agent's prior in the beginning of period $t$ is $\pi$ and the agent deviates $N$ periods. After $N$ periods, the agent conforms to the principal's expectation. Let $p(h^n)$ denote the agent's equilibrium action and $p'(h^n)$ denote the deviation. Conditional on history $\tilde{h}^t$, having no detectable deviations and the agent conforming to the principal's expectation from period $t+N$ on, one can express the agent's expected payoff as a function of his strategy from period $t$ to $t+N-1$, $X(\tilde{p}_t,\cdots,\tilde{p}_{t+N-1})$. The agent's IC constraint is

$$X(p_t,\cdots,p_{t+N-1})-X(p'_t,\cdots,p'_{t+N-1})\geq 0,$$

which is equivalent to

$$\begin{aligned}
& X(p_t,\cdots,p_{t+N-1})-X(p'_t,p_{t+1},\cdots,p_{t+N-1})\\
& +X(p'_t,p_{t+1},\cdots,p_{t+N-1})-X(p'_t,p'_{t+1},p_{t+2},\cdots,p_{t+N-1})\\
& +\cdots\\
& +X(p'_t,\cdots,p'_{t+N-2},p_{t+N-1})-X(p'_t,\cdots,p'_{t+N-1})\\
\geq\ & 0.
\end{aligned}$$

Let $\tilde{\pi}^n$ be the agent's belief in period $n$ given his private history where $n\in\{t,\cdots,t+N-1\}$. If the agent conforms to the principal's expectation from period $n+1$ on, the net loss from the local deviation is

$$X(p'_t,\cdots,p'_{n-1},p_n,\cdots,p_{t+N-1})-X(p'_t,\cdots,p'_n,p_{n+1},\cdots,p_{t+N-1}).$$

From Proposition 2, we can rewrite the net loss from the local deviation given history $\tilde{h}^n$ as

$$\begin{aligned}
&\delta(\tilde{\pi}_1^n\int(r(p_n(\hat{h}^n))-r(p'_n(\tilde{h}^n)))(V_1(\hat{h}^{n+1})-V_2(\hat{h}^{n+1}))dF_1\\
&+\tilde{\pi}_2^n\int(p_n(\hat{h}^n)-p'_n(\tilde{h}^n))(V_1(\hat{h}^{n+1})-V_2(\hat{h}^{n+1}))dF_2)\\
&-(c(p_n(\hat{h}^n))-c(p'_n(\tilde{h}^n)))
\end{aligned}$$

where $V_i(\hat{h}^n)$ is the hypothetical continuation value of the agent conditional on history $\hat{h}^n$ and being in state $\omega_i$ in period $n$. Let $\int V_1(\hat{h}^n)-V_2(\hat{h}^n)dF_i=W_i(\hat{h}^n)$. Note that there are two histories $\tilde{h}^n$ and $\hat{h}^n$. I consider a particular deviation strategy such that after one more deviation, the agent conforms to the principal's expectation. $\hat{h}^n$ is the agent's private history the principal believes with probability 1. $\tilde{h}^n$ is the agent's true private history. Furthermore, the agent reaches $\tilde{h}^n$ with the pdf generated by his true private history. Therefore, from period$-t$ perspective, the net loss of one more deviation in period $n$ is

$$\delta^{n-t}\int\delta(\tilde{\pi}_1^n(r(p_n(\hat{h}^n))-r(p'_n(\tilde{h}^n)))W_1(\hat{h}^{n+1})+\tilde{\pi}_2^n(p_n(\hat{h}^n)-p'_n(\tilde{h}^n))W_2(\hat{h}^{n+1}))-(c(p_n(\hat{h}^n))-c(p'_n(\tilde{h}^n)))dG.$$

The $N-$period IC constraint is

$$\begin{aligned}
&\sum_{n=t}^{t+N-1}\delta^{n-t}\int\delta(\tilde{\pi}_1(r(p_n(\hat{h}^n))-r(p'_n(\tilde{h}^n)))W_1(\hat{h}^n)\\
&+\tilde{\pi}_2(p_n(\hat{h}^n)-p'_n(\tilde{h}^n))W_2(\hat{h}^n))-(c(p_n(\hat{h}^n))-c(p'_n(\tilde{h}^n)))dG\geq 0.
\end{aligned}$$

This dynamic IC constraint has to be satisfied for any $N$. Furthermore, given the $N-$period IC, we can always let $p'_{t+N-1} = p_{t+N-1}$, and the $(N-1)-$period IC is implied by the $N-$period IC. Therefore, after every history $h^t$ on the equilibrium path, the following IC constraint must hold:

$$\sum_{n=t}^{\infty} \delta^{n-t} \int \delta(\tilde{\pi}_1^n(r(p_n(\hat{h}^n)) - r(p'_n(\tilde{h}^n)))W_1(\hat{h}^{n+1}) + \tilde{\pi}_2^n(p_n(\hat{h}^n) - p'_n(\tilde{h}^n))W_2(\hat{h}^{n+1})) - (c(p_n(\hat{h}^n)) - c(p'_n(\tilde{h}^n)))dG.$$

□

*Proof of Proposition 4.* The necessity of the IC constraints follows from Propositions 2 and 3. I'll next show the sufficiency of these two IC constraints. Since the dynamic IC constraint implies the IC constraint for any $N-$period deviations, it is sufficient to show that the agent cannot improve his payoff without violating one of these IC constraints.

Suppose the agent didn't deviate in the first $t$ periods and has a profitable deviation strategy $\sigma'_{\tilde{h}^t}$ from period $t$ on. Suppose $\sigma'$ provides $\epsilon$ more to the agent than his equilibrium strategy $\sigma$. By continuity at infinity, there exists $N$ sufficiently large and another deviation strategy $\sigma''_{\tilde{h}^t}$ such that the agent's payoff from $\sigma''_{\tilde{h}^t}$ is at least $\frac{\epsilon}{2}$ higher than his equilibrium payoff and

$$\begin{aligned} \sigma''(h^t, h^k) =& \sigma'(h^t, h^k) \ for \ k = 0, \cdots, N-1, \\ \sigma''_{|h^t, \tilde{h}^N} =& \tilde{\sigma}_{|h^t, \hat{h}^N}. \end{aligned}$$

(**??**) shows that the agent has another deviation strategy $\sigma''$ that coincides with $\sigma'$ for the first $N$ periods, coincides with the principal's expectation from period $t+N$ on and gives the agent $\frac{\epsilon}{2}$ more than his equilibrium payoff. But this is a contradiction to the assumption that there exists no profitable $N-$period deviation. Therefore, there exists no profitable deviation strategy. □

*Proof of Corollary 1.* From (5), we have

$$\begin{aligned} &\delta(\pi_1(r(p) - r(p'))W_1 + \pi_2(p - p')W_2) \\ &+ \delta^2 \int (r(p(y)) - r(p'(y)))W_{y1}(\pi_1 r(p')dF_1 + \pi_2 p' dF_2) \\ &+ \delta^2 \int (p(y) - p'(y))W_{y2}(\pi_1(1 - r(p'))dF_1 + \pi_2(1 - p')dF_2) \\ \geq& c(p) - c(p') + \delta \int c(p(y)) - c(p'(y))(\pi_1 dF_1 + \pi_2 dF_2). \end{aligned}$$

When $r(p) = 1,\ \forall p$, and $c(p) = \frac{1}{2}p^2$, this simplifies to

$$\delta\pi_2(p - p')W_2 + \delta^2 \int (p(y) - p'(y))W_{y2}\pi_2(1 - p')dF_2$$
$$- \frac{1}{2}(p^2 - p'^2) - \delta \int \frac{1}{2}(p(y)^2 - p'(y)^2)(\pi_1 dF_1 + \pi_2 dF_2) \geq 0.$$

From (3), we have

$$\delta\pi_2 W_2 = c'(p) = p,$$
$$\delta\frac{\pi_2 f_2(y)(1 - p)}{\pi_1 f_1(y) + \pi_2 f_2(y)}W_{y2} = c'(p(y)) = p(y).$$

Therefore, the two-period IC constraint becomes

$$\frac{1}{2}(p - p')^2 + \delta \int (\frac{1 - p'}{1 - p}p(y)(p(y) - p'(y)) - \frac{1}{2}(p(y) - p'(y))^2)(\pi_1 dF_1 + \pi_2 dF_2) \geq 0.$$

Since the left-hand side is minimized when

$$\frac{p(y)}{1 - p} = \frac{p'(y)}{1 - p'},$$

the two-period IC constraint is satisfied if and only if

$$(1 - p)^2 \geq \delta \int p(y)^2(\pi_1 dF_1 + \pi_2 dF_2).$$

□

# References


[1] Bhaskar, V. "The Ratchet Effect Re-examined: A Learning Perspective." (2014) Working paper.

[2] Board, Simon, and Moritz Meyer-ter-Vehn. "Reputation for quality." *Econometrica* 81.6 (2013): 2381-2462.

[3] Carroll, Gabriel. "When are local incentive constraints sufficient?" *Econometrica* 80.2 (2012): 661-686.

[4] DeMarzo, Peter M. and Yuliy Sannikov. "Learning, Termination, and Payout Policy in Dynamic Incentive Contracts." (2015) Working paper.

[5] Farhi, Emmanuel, and Iván Werning. "Insurance and taxation over the life cycle." *Review of Economic Studies* 80.2 (2013): 596-635.

[6] Fershtman, Chaim, and Ariel Pakes. "Dynamic games with asymmetric information: A framework for empirical work*." *Quarterly Journal of Economics* 127.4 (2012): 1611-1661.

[7] Harris, Milton, and Bengt Holmström. "A theory of wage dynamics." *Review of Economic Studies* (1982): 315-333.

[8] He, Zhiguo, Bin Wei, Jianfeng Yu, and Feng Gao. "Optimal Long-term Contracting with Learning." (2014) Working paper.

[9] Holmström, Bengt. "Managerial incentive problems: A dynamic perspective." *Review of Economic Studies* 66.1 (1999): 169-182.

[10] Hörner, Johannes, and Larry Samuelson. "Incentives for experimenting agents." *RAND Journal of Economics* 44.4 (2013): 632-663.

[11] Kapička, Marek. "Pareto Efficient Taxation with Learning by Doing." (2015) Working paper.

[12] Kapička, Marek and Julian Neira. "Optimal Taxation with Risky Human Capital." (2015) Working paper.

[13] Kwon, Suehyun. "Recursive Formulation with Learning." (2014) Working paper.

[14] Kwon, Suehyun. "Dynamic Moral Hazard with Persistent States." (2015) Working paper.

[15] Makris, Miltiadis, and Alessandro Pavan. "Incentives for Endogenous Types: Taxation under Learning-by-Doing." (2016) Working paper.

[16] Pavan, Alessandro, Ilya Segal, and Juuso Toikka. "Dynamic Mechanism Design: A Myersonia Approach." *Econometrica* 82.2 (2014): 601-653.

[17] Prat, Julien, and Boyan Jovanovic. "Dynamic contracts when the agent's quality is unknown." *Theoretical Economics* 9.3 (2014): 865-914.

[18] Stantcheva, Stefanie. "Optimal Taxation and Human Capital Policies over the Life Cycle." (2015) Working paper.