CESifo WORKING PAPERS

6621
2017

August 2017

Mechanism Design with Moral Hazard

Suehyun Kwon

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6621*
Category 11: Industrial Organisation

Mechanism Design with Moral Hazard

Abstract

This paper studies dynamic mechanism design in the presence of moral hazard. Revelation principle extends to models with moral hazard for both full commitment and limited commitment, but I also identify environments in which the principal doesn't benefit from eliciting agents' private information or beliefs. One-shot deviation principle requires the knowledge of agents' private strategies after deviations, and I characterize the necessary and sufficient condition for all IC constraints that requires only the knowledge of agents' equilibrium strategies. I also provide two sufficient conditions for smaller set of IC constraints that require obedience after a single-period deviation to be sufficient for all IC constraints. I illustrate how to apply revelation principle and the smaller set of IC constraints with an application allowing for endogenous state.

Keywords: dynamic mechanism design, adverse selection, moral hazard, revelation principle, one-shot deviation principle, endogenous state.

Suehyun Kwon
Department of Economics
University College London
30 Gordon Street
United Kingdom - London WC1H 0AX
suehyun.kwon@ucl.ac.uk

August 1, 2017
First draft: November 4, 2016. This paper subsumes "Recursive Formulation with Learning" (first draft: November 7, 2013) and "Incentive Compatibility with Endogenous States" (first draft: February 5, 2016). I thank Dilip Abreu, David Ahn, Eduardo Azevedo, Marco Battaglini, Ken Binmore, Simon Board, Patrick Bolton, Antonio Cabrales, Sylvain Chassang, Yeon-Koo Che, Pierre-André Chiappori, Syngjoo Choi, Joyee Deb, Federico Echenique, Camilo García-Jimeno, Daniel Garrett, Johaness Hörner, Emir Kamenica, Navin Kartik, Pat Kehoe, Philipp Kircher, Botond Koszegi, Jihong Lee, David Levine, Lucas Maestri, George Mailath, Miltos Makris, Konrad Mierendorff, Michael Ostrovsky, Sven Rady, Ariel Rubinstein, Yuliy Sannikov, Bruno Strulovici, Balázs Szentes, Juuso Toikka, Juuso Välimäki and participants at UCL research day, EEA-ESEM 2014 Conference, UCL-LSE TheoryWorkshop, University of East Anglia, UCL Brown bag, University of Bristol, Birkbeck, 2016 CESifo Area Conference on Applied Microeconomics, Columbia University, University of Pennsylvania, Yale University, University of Warwick, ESSET 2016, Conference on Market Design and Micro Theory.

1 Introduction

There are many situations where the principal has no private information about the payoff-relevant state but agents may have private information and can take actions unobservable to the principal. If the principal only observes an outcome which depends both on the payoff-relevant state and agents' actions, there is both moral hazard and (potential) adverse selection. In health insurance, the genetic disposition or the current health status can be the payoff-relevant state, and agents can take preventive measures or healthier lifestyle including more exercises and healthy eating.[1] In unemployment insurance, the worker's ability can be the payoff-relevant state, and workers can decide how much effort to put into job search. If the government doesn't know the ability of the unemployed nor effort in job search, it only knows whether the unemployed finds a new job or not, and again, there is both moral hazard and (potential) adverse selection. When agents have private information about the payoff-relevant state, there is adverse selection and moral hazard, and when agents have no private information, there is moral hazard with ex-ante symmetric uncertainty. If agents' actions don't affect the payoff-relevant state, the state is exogenous, but if agents' actions affect the payoff-relevant state, the state is endogenous.

Other examples include endogenous human capital accumulation. If the agent's human capital is the payoff-relevant state and it evolves as a result of the agent's effort, we have endogenous human capital accumulation. If the principal or the government doesn't observe neither the human capital nor the agent's action, then there is moral hazard; there could be adverse selection in case the agent has private information about his own human capital and symmetric uncertainty if the agent doesn't have private information. For another example, consider an agent developing a product for the market. If the agent sees the progress of the product but the principal doesn't, then the agent has private information about the payoff-relevant state (progress of the product). If the development of the product depends on the agent's action, and the principal only sees a noisy signal, for example, the sales data of beta products, then there is also moral hazard and the payoff-relevant state is endogenous. The payoff-relevant state could also be the match quality between the principal and the agent, the cost shock, the productivity shock, the market condition, the firm fundamental

[1] This is different from adverse selection and moral hazard discussed in Einav et al (2013) where moral hazard comes from overconsumption.

or the agent's quality as in Holmström's career concerns.

The case of exogeneous payoff-relevant state with ex-ante symmetric uncertainty has been relatively extensively studied,[2] but if agents have private information or the state is endogenous, there are only three publications I'm aware of.[3] This is in part due to technical reasons; in particular, one must account for the possibility of profitable multi-period deviations when the state is endogenous or agents have private information. Yet in many of the applications mentioned earlier, the payoff-relevant state is endogenous, and agents have private information. I develop tools to handle this class of models in this paper. Specifically, the class of models I consider has four key assumptions: (i) there is a payoff-relevant state, (ii) the principal has no private signal on the payoff-relevant state, (iii) agents may have private signals about the payoff-relevant state (adverse selection or ex-ante symmetric uncertainty) and (iv) agents' actions are unobservable to the principal and other agents (moral hazard).

When agents have private information, the most general class of contracts that the principal can offer to agents is dynamic mechanism design with moral hazard. When agents don't have private information at the beginning of first period, this still can be accomodated by mechanism design approach. In dynamic mechanism design, revelation principle and one-shot deviation principle are two of the main tools. The principal can always elicit agents' private information before choosing an allocation, and when there is more than one period, agents can always deviate in multiple periods.

I first show that revelation principle, both with full commitment and limited commitment, extends to models with moral hazard. Then I define what one-shot deviation principle should mean if there is both adverse selection and moral hazard. Since one-shot deviation principle requires the knowledge of agents' deviation strategies after every private history, it is difficult to apply it to characterizion of the optimal mechanism. I characterize the necessary and sufficient conditon for all IC constraints (dynamic IC) that requires only the knowledge of agents' equilibrium strategies.[4] Next, I define what I call on-path single deviation IC which is stronger

[2] Most of literature on strategic experimentation or dynamic moral hazard with symmetric uncertainty is included in the class of models I study. The quality of the unknown arm is the payoff-relevant state, and at the onset, none of the principal or agents have any private information about the quality of the unknown arm. Agents' actions are their private information, and all players observe an outcome that depends both on the quality of the unknown arm and actions of agents

[3] See Garrett-Pavan (2012), Board-Meyer-ter-Vehn (2013), Halac-Kartik-Liu (2016)

[4] In the single-agent case with full commitment, an equilibrium strategy is the agent's strategy when he has never deviated before. On the equilibrium path or off the equilibrium path are with

than the one-shot deviation IC; on-path single deviation IC only considers strategies of deviating once from the equilibrium strategy then conforming to the principal's expectation thereafter. This might mean the agent deviates from his private strategy infinitely many times and is different from the one-shot deviation in repeated games, but a few examples of it have been referred to as one-shot deviation in the mechanism design or contracting literature. When the principal elicits private information and recommends actions each period, on-path single deviation is the same as obeying recommendations after a single-period deviation. I provide two sufficient conditions for when on-path single deviation IC constraints are sufficient for all IC constraints.

In this class of models, even if agents don't observe the state, there is asymmetric information off the equilibrium path. If an outcome is a noisy signal of the agent's action and the state, the principal updates his belief about the state using the agent's equilibrium strategy, (and the outcome distribution induced by it) and the agent's deviation leads to belief disagreement off the equilibrium path. If the agent has a private signal on the state, the agent has more information than the principal on the equilibrium path. A mechanism designer (or the principal which I will use interchangeably throughout the paper) can elicit agents' private information at the beginning of a period. Revelation principle shows that despite moral hazard, it is always sufficient to elicit agents' private information in a given period. But perhaps what's more surprising is that in the class of models with symmetric uncertainty, the principal can do no better by eliciting the agents' beliefs. Models with symmetric uncertainty are the models where neither the principal nor agents have a private signal on the state, agents' actions are unobservable to the principal and other agents, and the outcome and the allocation are observed publicly. This implies that in strategic experimentation or career concerns, there is no benefit from eliciting agents' beliefs. There is also a class of models in which the principal doesn't benefit from eliciting agents' private information even if agents have private information. The common feature in this class of models is that an agent's equilibrium strategy doesn't depend on the state; in particular, if the state is endogenous and the state transition only depends on the agent's action and not on the state, the principal never benefits from eliciting agent's private information. An immediate corollary is that in the principal-agent version of Board-Meyer-ter-Vehn (2013), there is no benefit from eliciting the firm's private information even though the firm's quality is its private information

respect to whether the agent has deviated.

and consumers don't observe the quality.

After eliciting agents' private information, most mechanism design problems employ the sufficiency of local IC constraints. Intertemporally, this would be the one-shot deviation principle, and intratemporally, this would be the first-order condition IC. One-shot deviation principle holds with respect to private strategies, but it requires the knowledge of agents' deviation strategies after every private history and is difficult to apply.[5] On-path single deviation IC restricts attention to strategies on the equilibrium path where the agent deviates once but conforms to the principal's expectation from the following period on. This IC only requires the knowledge of the agent's equilibrium strategy and is easy to apply. However, this IC is not always sufficient for all IC constraints, and I provide two sufficient conditions for when this is.

The first sufficient condition is when past actions don't affect the outcome distribution nor the state transition and agents observe the payoff-relevant state; then on-path single deviation ICs are sufficient for all IC constraints. Corollary of this condition holds when agents observe the payoff-relevant state and the state and the outcome are Markovian. In particular, this sufficient condition allows for endogenous states, and to the best of my knowledge, this is the first paper that studies endogenous states with both adverse selection and moral hazard in the principal-agent setting.[6] It also follows that if there is only adverse selection or moral hazard in a Markovian environment, on-path single deviation ICs are sufficient. In the pure adverse selection setting, Pavan-Segal-Toikka (2014) refer to this as strongly truthful strategies.

The second sufficient condition requires the agents' beliefs to satisfy the first-order stochastic dominance and each agent's continuation value satisfies increasing differences in the agent's action and the state. This might seem strong, but many existing papers can be verified ex post to satisfy this condition. When the agent doesn't observe the state, existing literature has focused on belief disagreement, informational rent and trade off between efficiency and rent (dynamics over time). Belief disagreement and informational rent hold in the general class of models.

When on-path single deviation ICs are not sufficient, it is necessary to consider

[5]One-shot deviation principle in repeated games literature is with respect to all private strategies. In mechanism design or contracting literature, a few papers refer to examples of strategies for deviating once from the equilibrium strategy and conforming to the principal's expectation from the following period on as one-shot deviation. I define this class of strategies as on-path single deviation.

[6]The only other paper with endogenous states, adverse selection and moral hazard is Board-Meyer-ter-Vehn and they have competitive market.

multi-period deviations of the agent. I derive a necessary condition (dynamic IC) and show that it is sufficient if the principal has limited commitment or continuity at infinity is satisfied. Dynamic IC implies the on-path single deviation IC, and it imposes restrictions on dynamics of any incentive-compatible contract.

Once I develop the tools, I illustrate how to apply revelation principle and the first sufficient condition for on-path single deviation IC constraints to the discrete-time principal-agent version of Board-Meyer-ter-Vehn (2013). Their model has one firm and a continuum of consumers. I model the firm as the agent and consumers effectively being the principal. This application has infinite horizon, endogenous payoff-relevant state, private information and a continuum of actions. I characterize properties of optimal contracts and efficient equilibria. One of the properties (backloading of payments) has been shown in other contexts, and my result highlights the driving force of this property. Other properties haven't been shown in the literature. For example, it follows from theorem 2 that the principal doesn't benefit from eliciting the agent's belief even though the quality of the firm is the agent's private information. I also show that even though the agent is allowed to take any action from a continuum, the agent only takes the lowest action or the highest action on the equilibrium path.

Comparing my application with Board-Meyer-ter-Vehn (2013) shows that the equilibrium definition is crucial for the dynamics of the optimal contract or efficient equilibria. In particular, the reputation dynamics in Board-Meyer-ter-Vehn (2013) is not efficient when the firm and the market of consumers can use history-contingent strategies; their result depends on the Markovian strategies. Any optimal contract allowing for fully history-contingent strategies also doesn't have the reputation dynamics. Furthermore, backloading of payments has to hold for any signal structure including the perfect good news and the perfect bad news cases, which again shows that allowing for history-contingent strategies has a bigger impact than the signal structure the consumers (or the principal in my application) get. In addition, the bang-bang result of the agent's action (the firm's investment) shows that when the model is linear, allowing for binary actions is without loss of generality; even though Board-Meyer-ter-Vehn (2013) allows for a continuum of actions, the IC constraint of the agent and the optimality for the principal require that the agent will only take the lowest action or the highest action on the equilibrium path.

The main policy implication of this application is that linearity of the model, in particular, the shape of the cost function, is of first-order importance that dominates whether the state is endogenous or the agent has private information. But this requires a constant marginal cost. In the motivating examples mentioned earlier, whether the cost of submitting job applications is a linear or convex function of the number of firms matters more for the shape of the optimal contract than whether the unemployed knows his own ability perfectly. Similarly, whether exercising three times a week or five times a week is a linear function or not matters for the shape of the optimal health insurance.

The scope of tools I develop in the first part of the paper is much wider than the application in the second half. Revelation principle, dynamic IC and sufficiency of on-path single deviation IC work with any finite number of agents, and the IC constraints work both in the mechanism design setting and games. In particular, any stochastic game with a payoff-relevant state has to satisfy these ICs. Decomposition of continuation values can also be applied to decision problems and competitive-market settings, and the application illustrates both usages. Since these IC constraints are different from the usual one-shot deviation strategies that have been used for dynamic programming, I provide dynamic programming method to account for the on-path single deviation IC when it's sufficient and for multi-period deviations when the dynamic IC is necessary.

The methodology I develop by itself also highlights economic forces behind this class of models. Essentially, the tools I develop can be applied to any dynamic model with moral hazard where the principal has no private information about the payoff-relevant state and agents may have private information. Detailed discussion of results are in section 6.

The following section discusses related literature, and the rest of the paper is organized as follows. Section 2 presents the model, and revelation principle is shown in section 3. IC constraints and dynamic programming are discussed in section 4, and section 5 presents the application. Section 6 further discusses results, and section 7 concludes. Omitted proofs and a formal derivation of dynamic programming are in the online supplementary material.

1.1 Related Literature

As mentioned earlier, the case when the payoff-relevant state is exogenous and agents have no private information about the state is extensively studied in the literature.[7] If the state is endogenous or the agent has private information, there are very few publications.

Publications with both adverse selection and moral hazard include Board-Meyer-ter-Vehn (2013), Garrett-Pavan (2012) and Halac-Kartik-Liu (2016). I focus on principal-agent setting while Board-Meyer-ter-Vehn (2013) has one firm and a competitive market of consumers; but among three papers, theirs is the only one with endogenous state, and my application is the discrete-time principal-agent version of their model. In Garrett-Pavan (2012), managers can be matched randomly if they are fired, and the paper focuses on the retention policy. Revelation principle provides an alternative proof for their dynamic direct mechanism, and the first sufficient condition for on-path single deviation IC constraints is satisfied in their model, but they have non-linear costs, and their model doesn't overlap with my application. Halac-Kartik-Liu (2016) has both the project quality and the agent quality. Since the principal and the agent start with symmetric uncertainty about the project quality while the agent knows his own quality, my sufficient conditions for on-path single deviation IC doesn't apply to their setting.

Existing literature on endogenous human capital has focused on models where the government perfectly observes the level of human capital and the investment into it; this would mean that both the payoff-relevant state and the agents' actions are observable to the mechanism designer. (See for example Blundell et al (2016) or Stantcheva (2015))

I also allow all of the private history to affect the outcome distribution or the state transition in a given period. Cumulative actions of the agent can be mapped into the class of models I study; in continuous time, there is Sannikov (2014). Pavan-Segal-Toikka (2014) with unobservable actions can also be mapped into this class.

Related papers that haven't been mentioned so far include Myerson (1982) and Bester-Strausz (2001) for revelation principle. I allow for both full commitment and limited commitment, and proofs follow these two papers relatively closely. On-path single deviation IC is also related to obedience in Myerson (1982).

[7] See for example Bergemann-Hege (1998, 2005), Bhaskar-Mailath (2015), DeMarzo-Sannikov (2015), Hörner-Samuelson (2013), Prat-Jovanovic (2014)

Dynamic IC has never been studied in the literature, and most existing papers satisfy sufficiency of on-path single deviation IC. The closest reference for dynamic IC would be the impulse response function in Pavan-Segal-Toikka (2014). Section 6 discusses the relationship further.

Sufficiency of on-path single deviation IC has been shown in specific cases in the mechanism design or contracting literature, but the first sufficient condition I show has never been shown. The second sufficient condition can be verified ex post to be satisfied in many existing models in strategic experimentation, career concerns or dynamic moral hazard with symmetric uncertainty.

2 Model

I first present a binary model in section 2.1 then the general model in section 2.2. I use the binary model to illustrate the intuition in sections 4.1 and 4.2, but the sufficient conditions for on-path single deviation ICs in section 4.3 require the general model.

2.1 Binary Model

The class of models I study makes four key assumptions: (i) There is a payoff-relevant state, (ii) the principal has no private signal on the payoff-relevant state, (iii) agents may have private signals and (iv) agents' actions are unobservable to the principal and other agents. The binary model in this section captures the simplest case with endogenous state and limited commitment.

Formally, a principal hires an agent over the infinite horizon, $t = 1, 2, \cdots$. 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 $a \in [0, 1]$ which is his private information. An outcome is realized and observed by both parties. The principal makes a payment specified in the contract, and they move to the next period. If the agent rejects, the parties receive their outside options and continue in the following period.

In the binary model in this section, I don't allow the principal to elicit the agent's private information, but since the agent receives no private signal, this is without loss of generality and is shown formally in theorem 1.

There are three relevant variables in each period: they are the underlying state, the agent's action and the outcome. There are two states, ω_1 and ω_2. The states are unobservable to both the principal and the agent, and the parties start with a common prior π^1 in the beginning of period 1. The agent's action is his private information which leads to moral hazard, and it controls the transition probabilities of the state, which makes the state endogenous. When the agent chooses a, the probability of going from ω_2 this period to ω_1 next period is a. The probability of staying in ω_1 is $r(a) \in [0, 1)$, where $r(\cdot)$ is differentiable. Let $P(a)$ be the transition matrix for the states when the agent chooses a; $P_{ij}(a)$ is the probability of going from state ω_i to state ω_j next period. $P(a)$ is given by

$$P(a) := \begin{pmatrix} r(a) & 1 - r(a) \\ a & 1 - a \end{pmatrix}.$$

The cost of effort for the agent is $c(\cdot)$ which is differentiable, weakly convex, increasing and $c(0) = 0$. An outcome is denoted by y, and the set of outcomes is $\mathcal{Y} \subset \mathbb{R}$. 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 the distributions are atomless and satisfy full support, i.e., $f_1(y), f_2(y) > 0$ 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.

If the agent accepts the contract, the state changes at the end of the period after the principal makes a payment. If the principal doesn't offer a contract or the agent rejects it, the state doesn't change. The outside option for each party are normalized to 0. The principal is risk neutral, and the agent's vNM utility function from payment is $u(\cdot)$ which is increasing, concave and $u(0) = 0$.

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 perfect Bayesian equilibria of the game.

Let a_t, y_t, w_t be the agent's action, the outcome and the payment in period t. $\{\bar{w}_t(y)\}_y$ is the contract offered by the principal, and since the principal's history coincides with the public history, I use the shorthand notation and denote the payments as a function of current-period outcomes. The formal definition is in the next paragraph. $d_t = 1(0)$ denotes that the agent accepts (rejects) the contract in period t. The public history, which coincides with the principal's history, at period t after the outcome is realized is

$$h^{t,P} = (\{\bar{w}_0(y)\}_y, d_0, y_0, w_0, \{\bar{w}_1(y)\}_y, d_1, y_1, w_1, \cdots, \{\bar{w}_t(y)\}_y, d_t, y_t) \in \mathcal{H}^{t,P}.$$

If $d_t = 0$, then $y_t = \emptyset = w_t$. If the principal doesn't offer a contract, $\{\bar{w}_t(y)\}_y, d_t, y_t, w_t = \emptyset$. The agent's private history at period t is

$$h^{t,A} = (\{\bar{w}_0(y)\}_y, d_0, a_0, y_0, w_0, \cdots, \{\bar{w}_{t-1}(y)\}_y, d_{t-1}, a_{t-1}, y_{t-1}, w_{t-1}) \in \mathcal{H}^{t,A}.$$

Let $\Sigma^{t,P}$ be the set of all measurable functions $\sigma^{t,P} : \mathcal{H}^{t,P} \to \mathbb{R}$. $\Sigma^{t,P}$ is the set of all contracts $\{\bar{w}_t(y)\}_y$ that the principal can offer in period t, and the principal can also choose not to offer any contract which corresponds to $\emptyset$. The agent's strategy is $\sigma^{t,A} : \mathcal{H}^{t,A} \to \{\emptyset\} \cup [0, 1)$. If the agent chooses $\emptyset$, $d_t = 0$; the agent rejects the contract, and the parties receive their outside options and move to the next period. If the agent chooses $a \in [0, 1)$, then $d_t = 1$ and $a_t = a$; the agent accepts the contract and takes action a. Throughout the paper, hat denotes the agent's private history/strategy the principal believes is the true history/strategy.

I assume the agent plays a pure strategy, and it is also without loss of generality to assume that the principal doesn't randomize over transfers. The principal is allowed to randomize over continuation contracts, and there is a public randomization device.

2.2 General Model

A principal (mechanism designer) hires $N \geq 1$ agents for $t = 1, 2, \cdots, T$ where T can be ∞. The principal's discount factor is $\delta_P \in (0, 1)$, and agents' discount factors are $\delta_A^i \in (0, 1)$, $i = 1, 2, \cdots, N$. Each period, there is a payoff-relevant state $\omega_t \in \Omega_t$, and agent i receives a private signal $s_t^i \in \mathcal{S}_t^i$ then sends a message $m_t^i \in \mathcal{M}_t^i$. The mechanism recommends an action $r_t^i \in \mathcal{A}_t^i$ to agent i, agent i takes an action $a_t^i \in \mathcal{A}_t^i$, an outcome $y_t \in \mathcal{Y}_t$ is realized and there is an allocation $z_t = (x_t^1, w_t^1, \cdots, x_t^N, w_t^N) \in \mathcal{Z}_t = \Pi_{i=1}^N (\mathcal{X}_t^i \times \mathbb{R})$. x_t^i is the allocation for agent i, w_t^i is the transfer from the principal to agent i, and $z_t^i = (x_t^i, w_t^i)$.

Ω_t, $\mathcal{S}_t^i$, $\mathcal{M}_t^i$, $\mathcal{A}_t^i$, $\mathcal{Y}_t$, $\mathcal{X}_t^i$ are metric spaces, and they are allowed to be $\emptyset$. I assume all functions are measurable. Let $g_t : \Omega_t \to \Pi_{i=1}^N \mathcal{S}_t^i$ be the signal structure. The cost of action is $c_t^i : \mathcal{A}_t^i \to \mathbb{R}$. The distribution of outcome and the transition probabilities of the state are allowed to depend on all past states, signals, actions, outcomes and allocations. Let $\mathcal{S}_t = \Pi_{i=1}^N \mathcal{S}_t^i,\ \mathcal{A}_t = \Pi_{i=1}^N \mathcal{A}_t^i$ and denote the outcome distribution and the state transition in period t by $f_t,\ P_t$, respectively:

$$
\begin{aligned}
f_t &: \Pi_{j=1}^{t-1}(\Omega_j \times \mathcal{S}_j \times \mathcal{A}_j \times \mathcal{Y}_j \times \mathcal{Z}_j) \times \Omega_t \times \mathcal{S}_t \times \mathcal{A}_t \to \sigma(\mathcal{Y}_t),\\
P_t &: \Pi_{j=1}^{t}(\Omega_j \times \mathcal{S}_j \times \mathcal{A}_j \times \mathcal{Y}_j \times \mathcal{Z}_j) \to \sigma(\Omega_{t+1})
\end{aligned}
$$

where $\sigma(X)$ denotes the set of all probability measures on X.

The principal values outcome y_t with $v : \mathcal{Y} \to \mathbb{R},\ \mathcal{Y} = \cup_{t=1}^T \mathcal{Y}_t$. I assume the principal is risk neutral with respect to w_t, and agent i values (x_t^i, w_t^i) with vNM utility function $u_t^i : \mathcal{X}_t^i \times \mathbb{R} \to \mathbb{R}$. u_t^i is strictly increasing, weakly concave with respect to the second argument. When the allocation only consists of transfers and the agent is risk neutral with respect to transfer, I assume limited liability to rule out the residual claimant argument,[8] but otherwise, I don't assume limited liability.

I will consider different information structures on the state for agents, but the following is maintained throughout the paper: The principal doesn't observe the state, agents' actions are unobservable to the principal and other agents, and the outcome and the allocation are publicly observed. Messages and recommendations are private. In the beginng of period 1, the principal's prior on the payoff-relevant state, ω_1, is π^P and agent i has prior $\pi^{A,i}$.

With full commitment, the outside options of the principal and agents are $\bar{v}$ and $\bar{u}$, respectively. If the principal doesn't offer a mechanism or any agent doesn't participate, all players receive their outside options. Since the game is over if agents don't participate in the mechanism in the first period, I define histories only for the case when the principal offers the mechanism on the equilibrium path, and all agents participate in the mechanism; the mechanism and participation decisions are omitted from histories. The message space for agent i in period t, $\mathcal{M}_t^i$, is determined when the principal offers the mechanism. The mechanism consists of message spaces $\mathcal{M}_t^i,\ \forall t, i$ and allocation which is specified below. The timing within each period is (i) agents receive private signals (ii) agents send messages (iii) the mechanism recommends an action to each agent (iv) agents take actions (v) an outcome is realized (vi) an allocation is made (vii) the state changes.

[8] When there is both adverse selection and moral hazard, risk neutrality of the agent doesn't always imply that the principal can sell the firm. See Kwon (2017).

The private history of agent i in period t after receiving a signal is

$$h^{t,i,m} = (s_1^i, m_1^i, r_1^i, a_1^i, y_1, z_1, \cdots, s_{t-1}^i, m_{t-1}^i, r_{t-1}^i, a_{t-1}^i, y_{t-1}, z_{t-1}, s_t^i) \in \mathcal{H}^{t,i,m}.$$

The private history of agent i in period t after the mechanism recommends an action is

$$h^{t,i,a} = (s_1^i, m_1^i, r_1^i, a_1^i, y_1, z_1, \cdots, s_{t-1}^i, r_{t-1}^i, a_{t-1}^i, y_{t-1}, z_{t-1}, s_t^i, m_t^i, r_t^i) \in \mathcal{H}^{t,i,a}.$$

The public history in the beginning of period t is $h^t = (y_1, z_1, \cdots, y_{t-1}, z_{t-1}) \in \mathcal{H}^t$. The principal's history after messages are sent in period t is

$$h^{t,P,r} = (m_1, r_1, y_1, z_1, \cdots, m_{t-1}, r_{t-1}, y_{t-1}, z_{t-1}, m_t) \in \mathcal{H}^{t,P,r}$$

where $m_t = (m_t^1, \cdots, m_t^N),\ r_t = (r_t^1, \cdots, r_t^N)$. The principal's history after the outcome is realized in period t is

$$h^{t,P,z} = (m_1, r_1, y_1, z_1, \cdots, m_{t-1}, r_{t-1}, y_{t-1}, z_{t-1}, m_t, r_t, y_t) \in \mathcal{H}^{t,P,z}.$$

The reporting strategy of agent i is $\sigma^{t,i,m} : \mathcal{H}^{t,i,m} \to \sigma(\mathcal{M}_t^i)$, and the action strategy of agent i is $\sigma^{t,i,a} : \mathcal{H}^{t,i,a} \to \mathcal{A}_t^i$. The recommendation by the mechanism is $\sigma^{t,P,r} : \mathcal{H}^{t,P,r} \to \sigma(\times_i \mathcal{A}_t^i)$, and the allocation is $\sigma^{t,P,z} : \mathcal{H}^{t,P,z} \to \mathcal{Z}_t$. I assume agents play pure strategies and the principal doesn't randomize over allocations. The principal is allowed to randomize over continuation contracts, and there is a public randomization device. All strategies are measurable functions. I also define $s_t = (s_t^1, \cdots, s_t^N),\ a_t = (a_t^1, \cdots, a_t^N)$ and

$$\bar{h}^t = (\omega_1, s_1, m_1, r_1, a_1, y_1, z_1, \cdots, \omega_{t-1}, s_{t-1}, m_{t-1}, r_{t-1}, a_{t-1}, y_{t-1}, z_{t-1}, \omega_t, s_t)$$

to be the compilation of private histories in period t. The compilation of private histories is the union of all private histories of agents and the payoff-relevant state. The set of compilation of private histories in period t is denoted by $\bar{\mathcal{H}}^t$. The last notation regarding histories is $\bar{\mathcal{H}}(h^{t,i,\cdot})$ which denotes all compilations of private histories in period t that are consistent with agent i's private history $h^{t,i,\cdot}$. A complitaion of private histories in $\bar{\mathcal{H}}(h^{t,i,\cdot})$ is denoted by $\bar{h}(h^{t,i,\cdot})$.

Let $\mu(h^{t,i,a})$ be the probability distribution over $\bar{\mathcal{H}}(h^{t,i,a})$ conditional on $h^{t,i,a}$ and all other agents playing equilibrium strategies; let $\mu(a_t^{-i}|\bar{h}(h^{t,i,a}))$ be the probability distribution over other agents' actions conditional on $\bar{h}(h^{t,i,a})$ and all other agents playing equilibrium strategies. Throughout the paper, (h^t, h^k) denotes history h^t followed by h^k.

With limited commitment, I assume the principal offers a mechanism at the beginning of each period, and agents decide whether to participate. The mechanism consists of the message space for each agent in the current period and allocation. If agents participate, the rest of the period is the same as with full commitment. If an agent doesn't participate or if the principal doesn't offer a mechanism, the principal and agents receive their per-period outside options $\bar{v}$ and $\bar{u}$, respectively. The state transition in a period where they receive outside options is given by P^0. Analogous results hold also if agents first receive private signals then the principal offers a mechanism.

I focus on undetectable deviations and assume that the principal and agents take their outside options forever once a deviation is detected. Since the mechanism offered by the principal, whether agents participated and allocations are publicly observed, I focus on agents' IC constraints for report and action choices. I focus on perfect Bayesian equilibria of the game.

3 Revelation Principle

This section shows that revelation principle extends to models with moral hazard. Since each agent's choices of report and action in each period is sequential, proofs extend naturally and follow Myerson (1982) and Bester-Strausz (2001) closely. When agents take actions in addition to having private information, the mechanism can recommend actions based on reports. The principal can commit to the mechanism within each period, even with limited commitment, and the mechanism can randomize on behalf of agents by the action profile recommended to agents. The mechanism knows both reports and the recommended action profile, and therefore, the allocation can be chosen as in Myerson or Bester-Strausz.

When agents have private information, revelation principle extends in a straightforward manner. However, an important implication is when there is symmetric information on the equilibrium path, i.e., agents receive no private signal on the payoff-relevant state. In this case, it is without loss of generality not to ask agents for reports. In light of revelation principle, it almost seems trivial given that the set of private signals for a particular agent in a given period is an empty set. But this implies that in strategic experimentation or career concerns, there is no gain from eliciting agents' beliefs on the payoff-relevant state. In particular, even though agents have no private information on the equilibrium path in this class of models, agents have private information off the equilibrium path. Once an agent starts deviating, he knows his past actions, and in addition, he can update his belief on the payoff-relevant state using the correct probability distributions induced by the actions he took; the principal updates his belief using the probability distributions induced by the agent's equilibrium actions, and the principal has a wrong belief on the payoff-relevant state. Therefore, when the principal doesn't observe agents' actions, which implies that the principal doesn't know whether agents have deviated, it is not immediate that the principal cannot benefit by asking agents to report their beliefs on the payoff-relevant state. My result shows that this intuition is wrong, and the principal cannot benefit by eliciting agents' beliefs.

Any proof that is not in the main text is in the online supplementary material.

Theorem 1 (Revelation Principle)**.** *Regardless of the commitment power, direct mechanisms with* $\mathcal{S}_t^i$ *as the message space for agent* i *in period* t *are without loss of generality. With full commitment, if* $\mathcal{S}_t^i = \emptyset$*, then not eliciting the belief of agent* i *in period* t *is without loss of generality. With limited commitment, if* $\mathcal{S}_t^i = \emptyset$ *for all* i *in a given period, not eliciting beliefs is without loss of generality. When the principal has full commitment, there exists a direct mechanism where truthful reporting is an optimal strategy for each agent, but when the principal has limited commitment, agents might randomize. Agents follow actions recommended by the mechanism after the report.*

The next theorem shows that there is another class of models in which the principal doesn't benefit from eliciting agents' private information, and in this case, it is despite agents having private information on the equilibrium path. This is an example of environments in which the principal doesn't benefit from eliciting private information, and there are other variations as well. This again may not seem particularly interesting as it is, but the discrete-time principal-agent version of Board-Meyer-ter-Vehn (2013) is a special case and further discussed in section 5.

Theorem 2. *Suppose (i) the outcome distribution only depends on the current state, (ii) the state transition P_t is Markovian and (iii) can be decomposed as*

$$P_t = M_t + a_t M_t'$$

where $a_t \in \mathbb{R}$, M_t, M_t' don't depend on a_t and M_t' doesn't depend on the current state ω_t. It is without loss of generality not to elicit agents' private information in period t.

4 IC constraints

Results on IC constraints are in this section. I first show how I decompose agents' continuation values in section 4.1 then derive the necessary and sufficient condition for all IC constraints (dynamic IC) in section 4.2. This condition only requires the knowledge of the agents' equilibrium strategies. Section 4.3 shows sufficient conditions for on-path single deviation IC to be sufficient for all IC constraints, and section 4.4 describes dynamic programming.

I will describe the intuition using the model from section 2.1. Theorem 1 shows that in this case, it is without loss of generality not to elicit the agent's belief, and I focus on the IC constraint for action choices.

4.1 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 the first step in this section. I am going to describe it with the binary model in section 2.1 then present the result for the general model.

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 π^t. The agent's equilibrium strategy is to choose a. 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 a, both the principal and the agent believe

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

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

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

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 π^{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 $a' \neq a$ 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 ω_i, (iii) the agent does what the principal expects him to do from the following period on and (iv) the principal follows his equilibrium strategy. 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 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. I define this class of strategies as on-path single deviation and discuss them further in the following two sections. Furthermore, this strategy includes many of the local IC constraints that have been referred to as one-shot deviation in the contracting 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, on-path single deviation includes these strategies. The formal definition is as follows.

Definition 1. *On-path single deviation is deviating in a single period then following the strategy that the principal believes is the agent's equilibrium strategy given the principal's private history. If the principal elicits the agent's private information and recommends an action each period, on-path single deviation is equivalent to obeying the recommendation every period after a single-period deviation.*

For example, in the binary model from section 2.1, given the principal's history $h^{t,P}$*, there exists a private history of the agent* $\hat{h}^{t,A}$ *that the principal believes is the agent's true private history. On-path single deviation in period* t *is formally a* *strategy* $\sigma^{t+k,A\prime}$ *such that* $\sigma^{t+k,A\prime}(h^{t,A}, d_t, a', y_t, w_t, h^k) = \sigma^{t+k,A}(\hat{h}^{t,A}, d_t, a_t, y_t, w_t, h^k)$ *for all* k, h^k*.*

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 continuation 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, ω_i) because it is conditional on getting outcome y this period and going to state ω_i in the following period. Neither the principal nor the agent observes the state, and even if they happen to be in state ω_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, ω_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 figure 1. There are two outcomes, H or L in period t, and

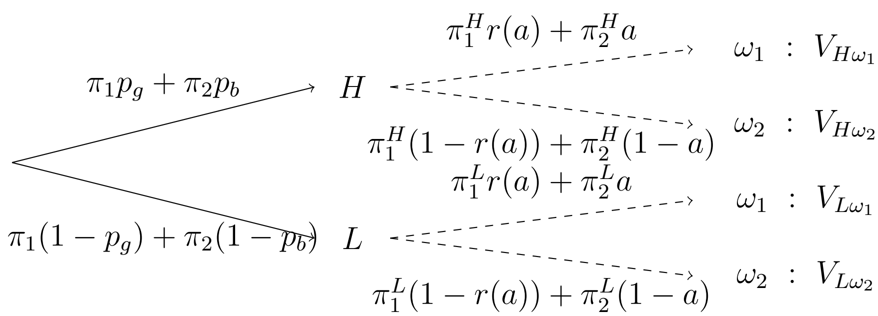

Figure 1: Decomposition with two outcomes

there are two states we could be in period $t+1$, ω_1 and ω_2. π^H, π^L are beliefs on the state after outcome H and L, respectively. We can use figure 1 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 figure 2. What needs a bit more attention is the continuation values

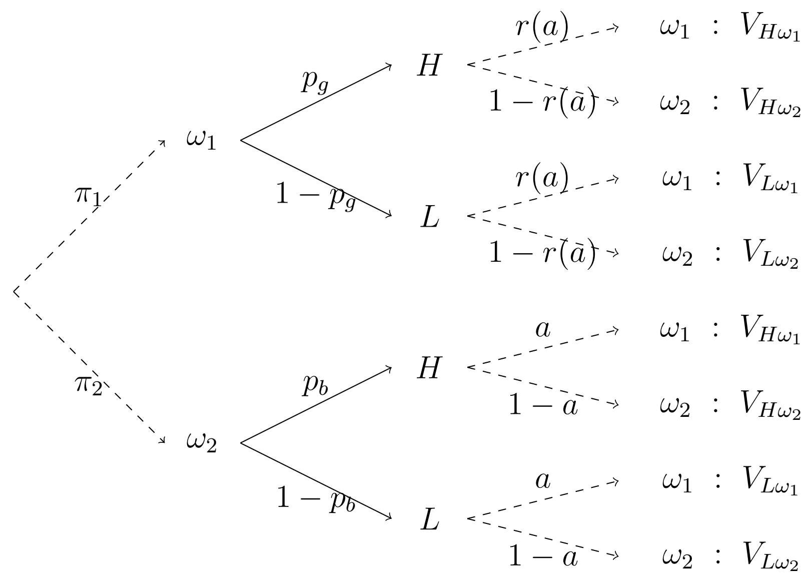

Figure 2: Decomposition with two outcomes (2)

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 $t+1$. Note that the agent's hypothetical continuation value doesn't depend on the current state.[9] 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 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 ω_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:

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

[9]The last point depends on the Markovian assumption of the binary model. The general model allows all of the history to affect the state transition.

$$+\pi_2(p_b(u(w(H))+\delta(aV_{H\omega_1}+(1-a)V_{H\omega_2}))+(1-p_b)(u(w(L))+\delta(aV_{L\omega_1}+(1-a)V_{L\omega_2})))$$

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

This idea generalizes to more than two outcomes.

Proposition 1. *Suppose in period t, the agent accepts the contract and chooses $a \in [0,1)$. 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 ω_i, (iii) the agent's continuation strategy coincides with what the principal expects him to do, i.e., $\sigma^{t+k,A\prime}(h^{t,A}, d_t, a', y_t, w_t, h^k) = \sigma^{t+k,A}(\hat{h}^{t,A}, d_t, a_t, y_t, w_t, h^k)$ for all k, h^k and (iv) the principal follows his equilibrium strategy. Given the agent's belief π in the beginning of period t, the agent's continuation value from period t on is given by*

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

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

$$\begin{aligned} V_1 &= -c(a)+\int u(w(y))+\delta(r(a)V_{y\omega_1}+(1-r(a))V_{y\omega_2})dF_1, \\ V_2 &= -c(a)+\int u(w(y))+\delta(aV_{y\omega_1}+(1-a)V_{y\omega_2})dF_2, \end{aligned}$$

in state ω_1 and ω_2, respectively.

We can also generalize it to the general model in section 2.2. We can represent the agent's continuation value at any point and theorem 3 is one example. Essentially, for each agent i given his private history $h^{t,i,\cdot}$, there is one hypothetical continuation value corresponding to every compilation of private history $\bar{h}^t$ that is consistent with $h^{t,i,\cdot}$. And the multiplicity of representation comes from the timing we take the expectation over all compilations of private histories consistent with $h^{t,i,\cdot}$. If there is only one agent as we saw in the binary model, the agent's expectation is over the payoff-relevant state. If there is more than one agent, each agent also takes an expectation over other agents' private signals and actions. Since it's already notationally heavy, I express the continuation value only in terms of agents' action choices. When agents send messages and the mechanism recommends actions, we need to take expectations over reports and action recommendations as well, but conceptually, we just need to integrate all hypothetical continuation values.

Theorem 3 (General model)**.** *Let $V : \bar{\mathcal{H}}^{t+1,i,a}(h^{t+1,i,a}) \to \mathbb{R}$ be the hypothetical continuation value of agent i if from the following period on, all agents take actions expected by the mechanism designer and the mechanism designer plays the equilibrium strategy. The agent's continuation value from period t on is given by*

$$\int_{\bar{\mathcal{H}}(h^{t,i,a})} \int_{\Pi_{j\neq i}\mathcal{A}_t^j} -c_t^i(\tilde{a}_t^i) + \int_{Y_t} \Bigg(u_t^i(z_t^i(h^{t,P,z})) + \delta_A^i \int_{\Omega_{t+1}\times S_{t+1}} V(\bar{h}(h^{t,i,a}), \tilde{a}_t^i, a_t^{-i}, y_t, z_t, \omega_{t+1}, s_{t+1})$$

$$g_{t+1}(s_{t+1})ds_{t+1}P_t(\omega_{t+1})d\omega_{t+1}\Bigg) f_t(y_t)dy_t d\mu(a_t^{-i}|\bar{h}(h^{t,i,a}))d\mu(h^{t,i,a}).$$

Corollary 1 (Markovian environment with no private signal)**.** *Suppose the outcome distribution and the state transition are Markovian and $\mathcal{S}_t^i = \emptyset$ for all t, i. Let $V : \mathcal{H}^{t+1} \times \Omega_{t+1} \to \mathbb{R}$ be the hypothetical continuation value of agent i conditional on (i) the public history and the state in period $t+1$ and (ii) from period $t+1$ on, all agents take actions expected by the mechanism designer and the mechanism designer plays the equilibrium strategy. If the agent i's prior in the beginning of period t is π, the agent's continuation value from period t on is given by*

$$\int_{\Omega_t} -c_t^i(\tilde{a}_t^i) + \int_{Y_t} \Bigg(u_t^i(z_t^i(h^{t,P,z})) + \delta_A^i \int_{\Omega_{t+1}} V(h^t, y_t, z_t, \omega_{t+1})P_t(\omega_{t+1})d\omega_{t+1}\Bigg) f_t(y_t)dy_t d\pi.$$

Theorem 3 allows for mixed strategies, but I assume agents play pure strategies, and for each compilation of history, $\mu(a_t^{-i}|\bar{h}^{t,i}(h^{t,i}))$ is a degenerate distribution. Corollary 1 shows that the expression simplifies greatly if the environment is Markovian and agents have no private signals. Since we already know from theorem 1 that there is no benefit from eliciting beliefs when agents have no private signal, the mechanism designer doesn't need to ask for reports. The only type of undetectable deviations are action choices, and the expression in corollary 1 is without loss of generality. Since the outcome distribution and the state transition are Markovian, if agent i deviates to $\tilde{a}_t^i$, it affects both f_t and P_t. However, from period $t+1$ onwards, the agent's action in period t doesn't matter for hypothetical continuation values.

The decomposition 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. Consider proposition 1 for the binary model. Because both the agent's prior and the agent's action enter 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

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

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

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

Now we're ready to discuss the IC constraints.

4.2 Dynamic IC

This section presents the necessary and sufficient condition for all IC constraints that requires only the knowledge of agents' equilibrium strategies. I continue describing the result with the binary model in section 2.1.

First consider the on-path single deviation defined in definition 1. The agent hasn't deviated in the first $t-1$ periods. He deviates in period t but conforms to the principal's expectation from period $t+1$ on. This is different from the usual one-step deviation in repeated games; 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 private strategy infinitely many times. But in any case, if the agent does an on-path single deviation, the agent's IC constraint looks like the following:

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

Remember the agent follows a strategy $\sigma^{t+k,A\prime}(h^{t,A}, d_t, a', y_t, w_t, h^k) = \sigma^{t+k,A}(\hat{h}^{t,A}, d_t, a_t, y_t, w_t, h^k)$ for all k, h^k. 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. 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 on the equilibrium path.

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

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

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

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

must hold for $a \neq 0$. 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 ω_1 rather than ω_2 in $t+1$. Therefore, W_1 is the expected benefit of going to ω_1 when the outcome in t hasn't been realized yet and they're in ω_1 in t. W_2 is the analogue for ω_2 in t. When the agent deviates to $a' \neq a$, 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 a, 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 on-path single deviation equalizes the marginal benefit and the marginal cost.

Proposition 2. *Suppose the agent's prior in the beginning of period t is π. Any incentive-compatible contract must satisfy the on-path single deviation IC after every history on the equilibrium path:*

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

Now consider the following deviation. The agent hasn't deviated in the first $t-1$

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 $a' \neq a$, and in $t+1$, the agent deviates to $a'(y) \neq a(y)$ where I used the short-hand notation that after the first deviation in period t, the only new information is the outcome y, and the agent is allowed to condition his second deviation on the realization of y. Consider expanding figure 2 one more period allowing for the continuum of outcomes.

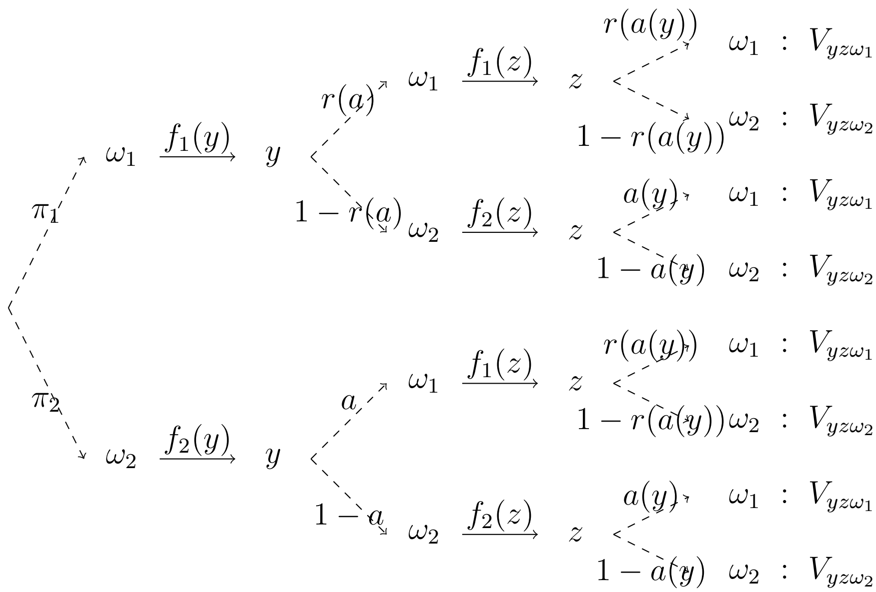

Figure 3: Decomposition for two periods

With a slight abuse of notation, let $V_{yz\omega_i}$ be the agent's hypothetical continuation value if outcomes in period t and $t+1$ are y and z, respectively, and the state in $t+2$ is ω_i. Let $W_{yi} = \int V_{yz\omega_1} - V_{yz\omega_2} dF_i(z)$. If we expand the continuation value of the agent from Proposition 1 for one more period, this is what we get:

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

$$+ (1-a)(-c(a(y)) + \int u(w(z)) + \delta(a(y)V_{yz\omega_1} + (1-a(y))V_{yz\omega_2})dF_2)\Big)\Big)dF_2.$$

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

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

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

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

(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 prof-

itable. However it changes the marginal benefit of the period-$(t+1)$ deviation through $\pi_1 r(a')dF_1 + \pi_2 a' dF_2$ and $\pi_1(1-r(a'))dF_1 + \pi_2(1-a')dF_2$, 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 ω_1 in the following period, conditional on being in ω_1 this period. W_2 is the marginal benefit of increasing the probability of being in ω_1 in the following period if they're in ω_2 this period. W_{y1} is the marginal benefit of increasing the probability of being in ω_1 in $t+2$, conditional on the history up to t and being in ω_1 in $t+1$. When the agent deviates to $a'(y) \neq a(y)$, it changes the probability of being in ω_1 in $t+2$, and $(r(a(y)) - r(a'(y)))W_{y1}$ is the benefit of deviating to $a'(y)$ if they're in ω_1 in $t+1$. Furthermore, if the agent deviated to $a' \neq a$ in period t, $\pi_1 r(a')dF_1 + \pi_2 a' dF_2$ is the probability of having y in period t and being in ω_1 in $t+1$. Therefore, from period-t perspective, the benefit of period-$(t+1)$ deviation conditional on being in ω_1 in $t+1$ is $\delta^2 \int (r(a(y)) - r(a'(y)))W_{y1}(\pi_1 r(a')dF_1 + \pi_2 a' dF_2)$, 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 ω_1 in $t+2$, conditional on the history up to t and being in ω_2 in $t+1$. From period-t perspective, the benefit of period-$(t+1)$ deviation conditional on being in ω_2 in $t+1$ is $\delta^2 \int (a(y) - a'(y))W_{y2}(\pi_1(1-r(a'))dF_1 + \pi_2(1-a')dF_2)$.

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

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

if the agent deviates to $a' \neq a$ 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:

$$
\sum_{n=t}^{t+N-1} \delta^{n-t} \int \Bigg(\delta \Big(\tilde{\pi}_1^n (r(a_n(\hat{h}^{n,A})) - r(a'_n(h^{n,A})))W_1(\hat{h}^{n+1,A})
$$

$$+\tilde{\pi}_2^n(a_n(\hat{h}^{n,A}) - a_n'(h^{n,A}))W_2(\hat{h}^{n+1,A})\Big) - \Big(c(a_n(\hat{h}^{n,A})) - c(a_n'(h^{n,A}))\Big)\Bigg)dG_n(h^{n,A}) \geq 0.$$

where $\tilde{\pi}^n$ is the agent's belief given his private history in the beginning of period n and G_n is the CDF of the agent's private history $h^{n,A}$. $h^{n,A}$ is the agent's private history at the beginning of period n, and $\hat{h}^{n,A}$ is what the principal believes is the agent's private history. a_n is the agent's equilibrium strategy, and a_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

$$\sum_{n=t}^{\infty} \delta^{n-t} \int \Bigg(\delta\Big(\tilde{\pi}_1^n(r(a_n(\hat{h}^{n,A})) - r(a_n'(h^{n,A})))W_1(\hat{h}^{n+1,A})$$
$$+\tilde{\pi}_2^n(a_n(\hat{h}^{n,A}) - a_n'(h^{n,A}))W_2(\hat{h}^{n+1,A})\Big) - \Big(c(a_n(\hat{h}^{n,A})) - c(a_n'(h^{n,A}))\Big)\Bigg)dG_n(h^{n,A}) \geq 0.$$

which I call the dynamic IC. The dynamic IC also turns out to be sufficient for all IC constraints in the binary model from section 2.1. In general, the dynamic IC is sufficient if there is limited commitment or continuity at infinity is satisfied.

Proposition 3. *Suppose the agent's prior in the beginning of period n given his private history is $\tilde{\pi}^n$. A contract is incentive compatible if and only if it satisfies the dynamic IC constraint after every history on the equilibrium path:*

$$\sum_{n=t}^{\infty} \delta^{n-t} \int \Bigg(\delta\Big(\tilde{\pi}_1^n(r(a_n(\hat{h}^{n,A})) - r(a_n'(h^{n,A})))W_1(\hat{h}^{n+1,A})$$
$$+\tilde{\pi}_2^n(a_n(\hat{h}^{n,A}) - a_n'(h^{n,A}))W_2(\hat{h}^{n+1,A})\Big) - \Big(c(a_n(\hat{h}^{n,A})) - c(a_n'(h^{n,A}))\Big)\Bigg)dG_n(h^{n,A}) \geq 0.$$

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

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

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

In the general model, the dynamic IC is given by the following theorem. As in theorem 3, I focus on action choices. Allowing for reports and action recommendations require expanding the integral twice more within each period.

Theorem 4 (General model)**.** *Let $V : \cup_t \bar{\mathcal{H}}^{t+1,i,a}(h^{t+1,i,a}) \to \mathbb{R}$ be the hypothetical continuation value of agent i if from the following period on, all agents take actions expected by the mechanism designer and the mechanism designer plays the equilibrium strategy. The dynamic IC after $h^{t,i,a}$ is given by*

$$\begin{aligned}
&\sum_{n=t}^{\infty} \delta^{n-t} \int_{\bar{\mathcal{H}}(h^{t,i,a}, h^{k,i})} \Bigg(\int_{\Pi_{j \neq i} \mathcal{A}^j_{t+k}} -(c^i_{t+k}(a^i_{t+k}) - c^i_{t+k}(a^{i\prime}_{t+k})) \\
&+ \int_{Y_{t+k}} \Bigg(u^i_{t+k}(z^i_{t+k}(h^{t+k,P})) + \delta^i_A \int_{\Omega_{t+k+1} \times S_{t+k+1}} V(\bar{h}(h^{t+k,i,a}), a_{t+k}, y_{t+k}, z_{t+k}, \omega_{t+k+1}, s_{t+k+1}) \\
&\qquad g_{t+k+1}(s_{t+k+1}) ds_{t+k+1} P_{t+k}(\omega_{t+k+1} | a^i_{t+k}) d\omega_{t+k+1} \Bigg) f_{t+k}(y_{t+k} | a^i_{t+k}) dy_{t+k} \\
&- \int_{Y_{t+k}} \Bigg(u^i_{t+k}(z^i_{t+k}(h^{t+k,P})) + \delta^i_A \int_{\Omega_{t+k+1} \times S_{t+k+1}} V(\bar{h}(h^{t+k,i,a}), (a^{i\prime}_{t+k}, a^{-i}_{t+k}), y_{t+k}, z_{t+k}, \omega_{t+k+1}, s_{t+k+1}) \\
&\qquad g_{t+k+1}(s_{t+k+1}) ds_{t+k+1} P_{t+k}(\omega_{t+k+1} | a^{i\prime}_{t+k}) d\omega_{t+k+1} \Bigg) f_{t+k}(y_{t+k} | a^{i\prime}_{t+k}) dy_{t+k} \\
&d\mu(a^{-i}_{t+k} | \bar{h}(h^{t,i,a}, h^{k,i})) \Bigg) d\mu_(h^{t,i,a}, h^{k,i}) \geq 0
\end{aligned}$$

where a^i_{t+k} is the equilibrium action, and $a^{i\prime}_{t+k}$ is agent i's deviation strategy.

The dynamic IC after each history on the equilibrium path is necessary. It is sufficient if there is limited commitment or continuity at infinity.

Corollary 2 (Markovian environment with no private signal)**.** *Suppose the outcome distribution and the state transition are Markovian and $\mathcal{S}_t^i = \emptyset$ for all t, i. Let $V : \cup_t(\mathcal{H}^{t+1} \times \Omega_{t+1}) \to \mathbb{R}$ be the hypothetical continuation value of agent i conditional on (i) the public history and the state in period $t+1$ and (ii) from period $t+1$ on, all agents take actions expected by the mechanism and the mechanism plays the equilibrium strategy. The dynamic IC in period t is given by*

$$\sum_{n=t}^{\infty} \delta^{n-t} \int \Bigg(\int_{\Omega_n} -(c_n^i(a_n^i) - c_n^i(\tilde{a}_n^i)) + \int_{Y_n} u_n^i(z_n^i(h^{n,P,z}))(f_n(y_n|a_n^i) - f_n(y_n|\tilde{a}_n^i))dy_n$$
$$+ \delta_A^i \Bigg(\int_{Y_n} \int_{\Omega_{n+1}} V(h^n, y_n, z_n, \omega_{n+1}) P_n(\omega_{n+1}|a_n^i) d\omega_{n+1} f_n(y_n|a_n^i) dy_n$$
$$- \int_{Y_n} \int_{\Omega_{n+1}} V(h^n, y_n, z_n, \omega_{n+1}) P_n(\omega_{n+1}|\tilde{a}_n^i) d\omega_{n+1} f_n(y_n|\tilde{a}_n^i) dy_n \Bigg) d\tilde{\pi}_n^i \Bigg) dG_n(h^{n,i,a}) \geq 0$$

where G_n is the distribution of agent i's private history in period n, $\tilde{\pi}_n^i$ is his private belief in period n, a_n^i is the equilibrium action and $\tilde{a}_n^i$ is the deviation in period n.

4.3 Sufficient Conditions for On-path Single Deviation IC

The advantage of decomposition in section 4.1 and the dynamic IC in section 4.2 is that we have an explicit expression for agents' continuation values both on and off the equilibrium path, and the dynamic IC only requires the knowledge of the agents' equilibrium strategies. However, in general one still needs to consider all possible deviations to verify the dynamic IC. I will show in this section when on-path single deviation IC is sufficient for all IC constraints. When it is sufficient, we only need to verify that the agent doesn't want to deviate from the equilibrium strategy in a single period and conform to the principal's expectation thereafter, and we don't even need to worry about the private continuation strategies of agents after a deviation.

The rest of section 4 discusses results with the general model in section 2. I will first state the results then explain the intuition. When the mechanism elicits private information, on-path single deviation is defined as deviating once in (report, action) but reporting truthfully and obeying recommendations by the mechanism from the following period on.

Corollary 3 provides an alternative proof for direct mechanism in Garret-Pavan (2012), but apart from their paper, theorem 5 hasn't been shown in the literature even as special cases. Theorem 6 can be verified ex post to be satisfied in many papers on strategic experimentation, career concerns or dynamic moral hazard with ex-ante symmetric uncertainty, and it identifies the underlying economic intuition across those models.

Theorem 5 (Sufficient Condition 1). *Suppose (i) each agent observes the state every period, ($s_t^i = \omega_t$) (ii) the domain of the outcome distribution f_t is $\Pi_{j=1}^{t-1}(\Omega_j \times \mathcal{Y}_j \times \mathcal{Z}_j) \times \Omega_t \times \mathcal{A}_t$, and (iii) the domain of the state transition P_t is $\Pi_{j=1}^{t}(\Omega_j \times \mathcal{Y}_j \times \mathcal{Z}_j) \times \mathcal{A}_t$. On-path single deviation ICs are sufficient for all IC constraints.*

Corollary 3 (Markovian environment). *Suppose (i) each agent observes the state every period, ($s_t^i = \omega_t$) (ii) the domain of the outcome distribution f_t is $\Omega_t \times \mathcal{A}_t$, and (iii) the domain of the state transition P_t is $\Omega_t \times \mathcal{A}_t \times \mathcal{Y}_t \times \mathcal{Z}_t$. On-path single deviation ICs are sufficient for all IC constraints.*

Theorem 6 (Sufficient Condition 2). *Suppose (i) the outcome distribution and the state transition are Markovian, (allowing for fully persistent states) (ii) $\mathcal{S}_t^i = \emptyset$ for all t, i, (iii) each agent can only deviate downwards, (iv) the posterior of each agent dominates the equilibrium belief in the sense of first-order stochastic dominance after a deviation and (v)*

$$\int_{Y_t} \left(u_t^i(z_t^i(h^{t,P,z})) + \delta_A^i \int_{\Omega_{t+1}} V(h^t, y_t, z_t, \omega_{t+1}) P_t(\omega_{t+1}) d\omega_{t+1} \right) f_t(y_t) dy_t$$

is supermodular in ω_t and $\tilde{a}_t^i$. On-path single deviation ICs are sufficient for all ICs.

Corollary 3 shows that if the outcome distribution and the state transition are Markovian and agents observe the state every period, on-path single deviation ICs are sufficient. In this environment, the optimal continuation strategy of each agent only depends on the current state and the continuation strategy of the mechanism designer and other agents; in particular, past deviations don't matter for the maximum continuation value the agent can generate with any strategy. Therefore, the maximum deviation payoff the agent can get from this period onward coincides with the maximum deviation payoff from on-path single deviations, and on-path single deviation ICs are sufficient for all IC constraints.

Theorem 5 allows non-Markovian environments as long as each agent observes the state every period and past actions don't affect neither the outcome distribution nor the state transition. This still allows past states, outcomes and allocations to affect future outcome distribution and state transition, and the mechanism designer can endogenously affect the environment; it is further discussed in section 6. Theorem 5 also naturally extends to pure moral hazard models where (i) there is no state or (ii) the principal and agents observe the state each period. In pure adverse selection, Pavan-Segal-Toikka (2014) refer to this as strongly truthful strategies in Markov environments.

Generically, if past actions affect the continuation game, through beliefs, the outcome distribution or the state transition, on-path single deviation ICs are not sufficient. In particular, if agents don't know the state, on-path single deviation ICs are generally not sufficient even if the outcome distribution and the state transition is Markovian. This is because of learning. When the agent doesn't perfectly observe the state, the agent learns about the state from an outcome each period, and the agent's deviation has lasting effects through his belief on the state until there is a perfectly informative signal.

The second sufficient condition (theorem 6) shows when we still have sufficiency of on-path single deviation ICs when past actions affect the continuation game. When agents' hypothetical continuation values satisfy increasing differences and an agent's belief satisfies first-order stochastic dominance after a deviation, the agent's continuation value from conforming to the principal's expectation thereafter is minimized on the equilibrium path. If the agent has deviated in the past, his continuation value from conforming to the principal's expectation is weakly greater than what his equilibrium payoff would have been given the public history. Further deviations are taken care of by increasing differences. Theorem 6 is a condition on hypothetical continuation values which are endogenous, but it can be verified ex post after solving for the optimal mechanism subject to on-path single deviation ICs.

4.4 Dynamic Programming

When the on-path single deviation IC is sufficient, the standard dynamic programming can allow for adverse selection or ex-ante symmetric uncertainty together with moral hazard. However, when the dynamic IC is necessary, i.e., one must account for multi-period deviations, the standard dynamic programming no longer works. In order to characterize the optimal mechanism only with the dynamic IC, one needs to make sure that the dynamic IC is also sufficient. This happens when there is limited commitment or continuity at infinity. I will describe the intuition for dynamic programming when the dynamic IC is necessary and sufficient in this section. The formal derivation is in the online supplementary material, and I will focus on the single agent case in a Markovian environment.

Essentially, one can use transfinite induction to account for multi-period deviations, and readers should feel free to jump to section 5 which discusses the application. Dynamic programming requires technical assumptions, but they can be dispensed with whenever the environment is finite.

The standard dynamic programming starts with the candidate set of payoffs (typically the set of all individually rational payoffs) and apply a contraction mapping until it reaches the fixed point. Existing literature has focused on cases when on-path single deviation ICs are sufficient, and when they are sufficient, the contraction mapping corresponds to the on-path single deviation given the public history up to that point.

With the dynamic IC and multi-period deviations, it is no longer sufficient to apply one contraction mapping until it reaches the fixed point. There is a sequence of operators, which are monotone but not necessarily contraction mappings, and the dynamic programming requires transfinite induction on this sequence of operators. However, since each operator leads to a monotone sequence of set of payoffs, one can still start with the candidate set of payoffs and apply each operator until it reaches the fixed point. The difference from the standard dynamic programming is that once it reaches the fixed point with the N-th operator T_N, it goes on to apply the $(N+1)$-th operator T_{N+1} until it reaches the next fixed point. Figure 4 shows how the transfinite induction works, where $W^{N,0}$ is the initial set for T_N and $W^{N,k} = (T_N)^k(W^{N,0})$.

The state variable for dynamic programming is $(\pi, V(\cdot,\cdot), V^P)$ which takes into account the belief on the state, hypothetical continuation values for the agent and the supremeum of principal's payoffs, all on the equilibrium path. With one agent, the dynamic IC only requires hypothetical continuation values on the equilibrium path, and there is only one dynamic IC after every history on the equilibrium path; it is sufficient to take the variables on the equilibrium path. Multi-period deviations are taken care of within each dynamic IC, and we don't need to worry about belief disagreement off the equilibrium path for the state variable.

In the first stage, the candidate set of payoffs is $W^0 = \{(\pi, V(\cdot,\cdot), V^P) || V(y,\omega)| \leq$

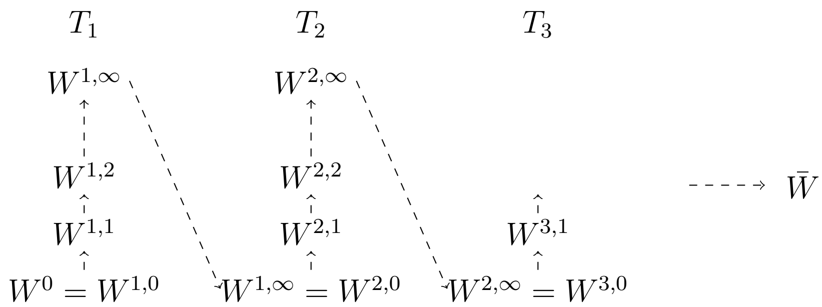

Figure 4: Trans-finite Induction

$\bar{V}\}$ for some $\bar{V}$. All that matters at this point is that there is a uniform upper bound on hypothetical continuation values of the agent, since otherwise, no matter how many times one applies monotone mappings to the initial set, it might never converge to a set with finite measure. With limited commitment or continuity at infinity, the agent's expected utility on the equilibrium path is uniformly bounded. But we still need to show that hypothetical continuation values are bounded; hypothetical continuation values integrate up to a finite expected utility, but in principle, they can still diverge or even be infinite on a set of measure zero. The proof follows from the fact that the hypothetical continuation value is continuous in ω and a, and any continuous function on a compact set is uniformly bounded.

Let T_1 be the operator with the on-path single deviation IC. One can find all points in W that can be generated by W, and this corresponds to the usual mapping in the standard dynamic programming except that it is weakly decreasing and not strictly a contraction mapping. Since $T_1(W) \subseteq W$, $(T_1)^j(W^0) = W^{1,j}$ converges to a set-theoretic limit. We can define T_N to be the mapping with the N-period IC and use the fact that the dynamic IC is equivalent to satisfying the N-period IC for every N. Each T_N is non-increasing, and the limit of a monotone sequence is well-defined.

For each T_N, N-period deviations are taken care of as follows. First, given π^t, we know π^{t+1} when the agent takes action a and outcome y is realized. By induction, we can choose $V_{t+2}(y, \omega)$ for each π^{t+1} such that from period $t+1$ on, $(N-1)$-period deviation ICs are satisfied. Next, choose $w_{t+1}(y)$ such that $V_{t+1}(y, \omega)$ generated by $w_{t+1}(y)$, $V_{t+2}(y, \omega)$ satisfy IR and the promise-keeping constraint. We know the equilibrium beliefs, actions for the next N periods, and we need to verify the N-period deviation IC.

N-period deviation IC is satisfied by backward induction. Let $\tilde{V}_{t+k+1}$ be the agent's maximum deviation payoff from any deviation between period $t+k+1$ and $t+N-1$; the agent conforms to the principal's expectation from period $t+N$ onwards. In period $t+N-1$, given the agent's private belief $\tilde{\pi}^{t+N-1}$ and his hypothetical continuation values $V_{t+N}(y,\omega)$, the agent has the optimal action. In period $t+k$, given the agent's private belief $\tilde{\pi}^{t+k}$ and the maximum deviation payoff from $t+k+1$ on, $\tilde{V}_{t+k+1}$, the agent has the optimal action. We can continue doing the backward induction, and in period t, we have $\tilde{\pi}^t = \pi^t$ and the equilibrium payoff has to be weakly better than the most profitable deviation payoff. The rest of the argument follows from the agent's deviation payoff being a continuous function of his private belief and action; the set of beliefs and the set of actions are compact sets.

Since I take care of all private beliefs the agent might have in period $t+k$ when I verify the N-period deviation IC, there is no need to keep track of the agent's private belief as the state variable.

5 Discrete-Time Principal-Agent Version of Board-Meyer-ter-Vehn (2013)

This section discusses the discrete-time principal-agent version of Board-Meyer-ter-Vehn (2013). It is a special case of the binary model from section 2.1 except allowing for $a \in [0,1]$ and illustrates how revelation principle and the first sufficient condition for on-path single deviation ICs can be applied. I characterize properties of optimal contracts and efficient equilibria.

I will first discretize Board-Meyer-ter-Vehn (2013) then map it into the principal-agent version. In the discrete-time version of Board-Meyer-ter-Vehn, the firm sells one unit of good at price of expected quality each period. The quality (state) is $\theta_t \in \{L, H\}$, where $L = 0$, $H = 1$, and the firm makes investment $a_t \in [0,1]$ at constant marginal cost $c > 0$. Initially, the firm's quality θ_0 is exogenous, and thereafter, $\theta_t = \theta_s$ where $s \leq t$ is the most recent shock. The shock arrives with probability $\lambda > 0$, and when it arrives, the firm's quality changes with $\Pr(\theta_s = H) = a_s$. Consumers don't observe neither the quality θ_t nor investment a_t, and they get a public signal with probability μ_θ each period. The firm knows its own quality and is risk-neutral. They characterize Markov Perfect Equilibria of the game. Since their model is in continuous time, I discretize it in each period in the order of (i) the firm makes investment, (ii) the quality may change, (iii) there may be a public signal, (iv) consumers buy at the given price.

In the principal-agent version, consumers are the principal, and the firm is the agent. Since the firm is risk-neutral, it maps into having a risk-neutral agent, and I assume limited liability to rule out the residual claimant argument. The cost function is $c(a) = ca$, and from $\Pr(\theta_s = H) = a_s$, the state transition is given by

$$P = \begin{pmatrix} \lambda a_t + 1 - \lambda & \lambda(1 - a_t) \\ \lambda a_t & 1 - \lambda a_t \end{pmatrix}$$

where P_{ij} is the probability of transition from quality i to quality j, and the first column and the first row correspond to the high quality. The transition matrix follows from the fact that if the firm is high quality, with probability $1-\lambda$, the quality doesn't change and the firm remains high quality. When the quality changes with probability λ, the firm remains high quality with probability a. If the firm is low quality, it becomes high quality if and only if the shock arrives, which happens with probability λ, and when it does arrive, the probability of becoming a high-quality firm is a.

Since consumers don't observe the firm's quality in Board-Meyer-ter-Vehn, I assume that the principal gets payoff s if there is a public signal and $\emptyset$ if there is no public signal. I normalize values of $s, \emptyset$ so that the expected per-period payoff is 1 if the firm is high quality and 0 if the firm is low quality, i.e., $\mu_H s + (1 - \mu_H)\emptyset = 1$, $\mu_L s + (1 - \mu_L)\emptyset = 0$. The principal can condition his payment on all of the public history up to that point.

The first observation is that the optimal investment doesn't depend on the current quality. This is because of the state transition. If the shock doesn't arrive, the firm's investment is irrelevant to its payoff, and when it arrives, the probability of becoming a high-quality firm is λa for both qualities. In this case, whether the firm knows its own quality or not doesn't matter for strategies, and in particular, this maps into theorem 2 where not eliciting the agent's belief is without loss of generality. The principal doesn't benefit from asking the firm to report its quality.

Second of all, the discrete-time principal-agent version of Board-Meyer-ter-Vehn is a Markovian environment with both adverse selection and moral hazard. From corollary 3, on-path single deviation ICs are sufficient for all IC constraints.

The on-path single deviation IC with high quality is given by

$$\begin{aligned}
& -ca_t + (\lambda a_t + 1 - \lambda)(\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H)) + \lambda(1-a_t)(\mu_L V(s,L) + (1-\mu_L)V(\emptyset,L)) \\
\geq\ & -ca' + (\lambda a' + 1 - \lambda)(\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H)) + \lambda(1-a')(\mu_L V(s,L) + (1-\mu_L)V(\emptyset,L)),
\end{aligned}$$

for all a' and with low quality, it's given by

$$\begin{aligned}
& -ca_t + \lambda a_t(\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H)) + (1-\lambda a_t)(\mu_L V(s,L) + (1-\mu_L)V(\emptyset,L)) \\
\geq\ & -ca' + \lambda a'(\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H)) + (1-\lambda a')(\mu_L V(s,L) + (1-\mu_L)V(\emptyset,L)),\ \forall a'.
\end{aligned}$$

I dropped the public history up to the beginning of the period in the notation and expressed hypothetical continuation values only in terms of current period signal and and the quality. One can see that the on-path single deviation IC is identical for both the high quality and the low quality:

$$\begin{aligned}
& -ca_t + \lambda a_t(\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H)) - \lambda a_t(\mu_L V(s,L) + (1-\mu_L)V(\emptyset,L)) \\
\geq\ & -ca' + \lambda a'(\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H)) - \lambda a'(\mu_L V(s,L) + (1-\mu_L)V(\emptyset,L)),\ \forall a'.
\end{aligned}$$

The third observation is that the on-path single deviation IC is linear in the firm's investment a_t: for all a',

$$(a_t - a')(-c + \lambda((\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H)) - (\mu_L V(s,L) + (1-\mu_L)V(\emptyset,L)))) \geq 0.$$

In particular, let $V_H,\ V_L$ be the firm's payoff conditional on the current quality. We get

$$\begin{aligned}
V_H &= -ca_t + (\lambda a_t + 1 - \lambda)(\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H)) + \lambda(1-a_t)(\mu_L V(s,L) + (1-\mu_L)V(\emptyset,L)), \\
V_L &= -ca_t + \lambda a_t(\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H)) + (1-\lambda a_t)(\mu_L V(s,L) + (1-\mu_L)V(\emptyset,L)).
\end{aligned}$$

It follows that

$$V_H - V_L = (1-\lambda)(\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H) - \mu_L V(s,L) - (1-\mu_L)V(\emptyset,L)),$$

and we get

$$a_t = 1 \text{ if } V_H - V_L > \frac{1-\lambda}{\lambda}c,$$
$$a_t = 0 \text{ if } V_H - V_L < \frac{1-\lambda}{\lambda}c.$$

Unless the coefficient on $a_t - a'$ is 0, i.e., $V_H - V_L = \frac{1-\lambda}{\lambda}c$, linearity implies that the firm chooses the extreme points $a_t = 0$ or $a_t = 1$. But one can easily see that if the coefficient is 0 and the firm chooses an interior $a_t \in (0, 1)$, the principal can offer ϵ more to signal s and let the firm choose $a_t = 1$. The bang-bang result requires linearity of the cost function $c(a)$ and the state transition $r(a)$. Generically, the firm chooses an interior solution if either of $c(a)$, $r(a)$ is non-linear.

Results so far don't depend on details of the model including the signal structure, equilibrium definition, or whether it's a competitive market or principal-agent setting except for the optimality argument in the bang-bang result. The bang-bang result shows that any Markov Perfect Equilibrium has to be characterized by cutoffs in the belief space, and the MPE with perfect learning in Board-Meyer-ter-Vehn are still equilibria in the discrete-time version.

In the principal-agent version, one can also characterize the first best. The first best is $a = 1$ every period for all beliefs if

$$c \leq \frac{\lambda}{1 - \delta(1-\lambda)}.$$

The proof follows from adapting decomposition of continuation values for one-person decision problem, and essentially, since the continuation game doesn't depend on the history nor the current quality, exerting the maximum effort every period is optimal if the cost of effort is not too high.

The last result is backloading; if we don't impose Markov strategies, from risk neutrality and limited liability, an optimal contract (principal-agent) or an efficient equilibrium (competitive market) backloads payments until it can implement the first best ($a = 1$) every period. One can also solve for the minimum rent to implement the first best every period and characterize the evolution of hypothetical continuation values.

Theorem 7. *Consider the discrete-time version of Board-Meyer-ter-Vehn (2013).*

For any signal structure and in any equilibrium (both in competitive market or principal-agent setting) the optimal investment is independent of the firm's quality, and there's no benefit from eliciting the firm's quality. On-path single deviation ICs are sufficient for all ICs. In an optimal contract, the firm always chooses either $0, 1$ *and never chooses an interior* $a \in (0, 1)$*. The first best is* $a = 1$ *every period if*

$$c \leq \frac{\lambda}{1 - \delta(1 - \lambda)}.$$

With full commitment, an optimal contract also backloads payments until the firm can be incentivized to take the first-best action every period with any signal structure.

Compared to Board-Meyer-ter-Vehn (2013), theorem 7 shows that equilibrium definition matters for the dynamics of an optimal contract. Reputation dynamics in Board-Meyer-ter-Vehn is not optimal if consumers (the principal) and the firm can use fully history-contingent strategies. Markov strategy is crucial for their result. The first three points hold for risk-averse firm as well.

It also shows that an optimal contract or an efficient equilibrium shares similarities with models with exogenous states or without adverse selection. Linearity is a stronger driving force than endogeneity of state transition for backloading and dominates the difference between exogenous states and endogenous states.

Theorem 7 also shows that in this class of models, allowing for a continuum of actions leads to the same equilibrium behavior as having binary actions.

6 Discussion

As mentioned earlier, the class of models I can allow for has four key assumptions: (i) there is a payoff-relevant state, (ii) the principal has no private signal on the payoff-relevant state, (iii) agents may have private signals, and (iv) each agent's action is unobservable to the principal and other agents. Within each period, there is a payoff-relevant state, private signal, action, outcome and allocation. I can allow both the outcome distribution and the state transition to be endogenous and depend on everything that has happened up to that point, including cumulative actions.

Main results on IC constraints are that what has been referred to has one-shot deviation in the contracting literature is stronger than the one-shot deviation in repeated games literature, and it is not always sufficient for all IC constraints. I provide two sufficient conditions for when it is sufficient, but when it isn't, the principal must account for multi-period deviations which is captured by the dynamic IC.

Most existing papers on this class of models satisfy the second sufficient condition, and Garrett-Pavan (2012) is the only one that satisfies the first sufficient condition. The sufficiency of on-path single deviation IC has been proved for individual cases in these papers, and by providing sufficient conditions for bigger class of models, I identify the underlying economic force for the sufficiency of on-path single deviation IC.

The first sufficient condition essentially requires that the optimal continuation strategy doesn't depend on agents' private information. This happens when each agent observes the state perfectly, which breaks the intertemporal linkage through agents' private beliefs on the payoff-relevant state, and the outcome distribution and the state transition only depend on the current period action profile. This allows the outcome distribution and the state transition to depend on anything that is publicly observable, i.e., all past outcomes and allocations. From each agent's perspective, the environment is effectively Markovian because their past actions don't matter for the continuation game, but the principal can endogenously control the environment through allocation. The second sufficient condition requires increasing differences of each agent's hypothetical continuation value and first-order stochastic dominance of each agent's belief on the payoff-relevant state.

Proofs for both theorems 5 and 6 are not mathematically involved. Once we decompose each agent's continuation value as a linear combination of hypothetical continuation values, proofs follow immediately. However, the decomposition of continuation value is one of the main innovations of this paper that led to the characterization of all results. In particular, if one represents the agent's continuation value as a function of his private belief, then any change (from deviation) to the agent's private belief requires all his private beliefs in future periods to be updated differently. This continues until there is a perfectly informative signal on the state which resets all beliefs of the principal and agents to common prior. If one were to follow this route, particularly if there is no perfectly informative signal, one needs to keep track of the change in beliefs through Bayesian updating which is nonlinear and depends on the signal and outcome distributions in each period. Normal distribution makes it easier to track changes in beliefs, which I believe is part of the reason why most papers on this class of models either have binary environment or linear-Gaussian environment.

In my decomposition, I express the agent's continuation value as a linear combination of hypothetical continuation values, and this breaks the linkage through the agent's private belief. In a Markovian environment, a hypothetical continuation value is conditional on the payoff-relevant state in the following period and continuation strategies of the principal and agents from the following period onwards. If the environment is not Markovian, it is conditional on the compilation of private histories. However, once each agent's continuation value is decomposed into hypothetical continuation values, then the agent's private belief is the weight on hypothetical continuation values in the current period, and current period only. The agent's private belief doesn't affect hypothetical continuation values, and there is no need to keep track of how beliefs get updated.

However, in the usual way of defining continuation value as a function of private belief, the belief matters for the continuation value in two ways. The belief matters for the probability of each history, but it also affects the agent's continuation strategy. The argument in the previous paragraph shows how hypothetical continuation value circumvents the linkage through the probability of events, but it also takes care of the second source as well. When I define hypothetical continuation values, it is conditional on the continuation strategies of the principal and agents. Once we fix the continuation strategies, then beliefs don't matter for strategies, and any change in private belief of the agent only matters as weights on the linear combination.

This definition of hypothetical continuation value requires one to fix the continuation strategies of the principal and agents. However, what has been referred to as one-shot deviation in the contracting literature does precisely this; it assumes that the agent will follow the continuation strategy the principal is expecting given the principal's history. In addition, the biggest challenge in this class of models is the private strategy. Agents might receive private information each period, but even in the ex-ante symmetric uncertainty case where there is no private signal, there is asymmetric information off the equilibrium path because the payoff-relevant state is unobservable to the principal and agents. The lack of papers on this class of models in a more general setting is linked to the lack of papers on private monitoring repeated games with a fixed discount factor bounded away from one. Most papers on private monitoring repeated games either study folk theorem or belief-free equilibrium.

In order to account for private strategies with the one-shot deviation principle as in repeated games, one needs to know the agent's continuation strategy after a deviation. Without characterizing all possible private strategies after deviations, the one-shot deviation principle doesn't help very much. Hypothetical continuation values require fixing continuation strategies of the principal and agents, but if we fix the continuation strategies to be the equilibrium strategies, then one can express each agent's continuation value only using the equilibrium strategies, and one must characterize equilibrium strategies anyway.

Results on IC constraints apply to stochastic games with a payoff-relevant state as well. Suppose there is a payoff-relevant state unobservable to all players, and players may receive private signals. Players may communicate before taking actions, and after players take actions, which are unobservable to other players, an outcome is realized and observed by all players. Each player's payoff can depend on all past payoff-relevant states and private histories of all players. Results on revelation principle from section 3 require a mediator. But IC constraints apply immediately to this class of models.

When on-path single deviation ICs are not sufficient, one needs to take care of profitable multi-period deviations, and this requires the dynamic IC. The dynamic IC takes an infinite sum, but still each expression only requires equilibrium strategies of the principal and agents. Essentially, each expression in the infinite sum captures the benefit of one more deviation given private history up to that point; after the last deviation, the agent follows the strategy the principal expects him to follow given the principal's private history. When on-path single deviation ICs are sufficient, the net benefit of additional deviations are weakly negative. After the initial deviation, any further deviation gives a weakly lower benefit than the net benefit on the equilibrium path conditional on the public history up to that point. Therefore, even though the dynamic IC still has to be satisfied in principle, it is sufficient to verify the on-path single deviation ICs, and the on-path single deviation ICs imply the dynamic IC which in turn implies all IC constraints. On-path single deviation ICs are not sufficient if there can be profitable multi-period deviations. Even if all on-path single deviation ICs are satisfied, they don't necessarily imply that the dynamic ICs are satisfied, and even if each agent never benefits from a single-period deviation, there can be profitable multi-period deviations.

In this class of models, once an agent deviates, it affects the outcome distribution, and therefore the allocation this period. It can also affect the state transition this period, and even if the state is exogenous, the agent knows his own deviation and updates his private belief accordingly, while the principal and all other agents update their posterior beliefs using the wrong outcome distribution induced by the agent's equilibrium strategy. I allow the outcome distribution and the state transition in subsequent periods to depend on this period's deviation, but the intuition can be seen in a Markovian environment. In the dynamic IC, each expression captures the benefit of one additional deviation, but this benefit depends on the agent's private belief given his own private strategy. When the on-path single deviation IC given the public history is satisfied, the net benefit given the agent's private belief on the equilibrium path is satisfied. The agent has a different private belief after a deviation, and therefore, the net benefit from the additional deviation is different and can be positive. However, since each expression is also integrated with respect to the probability of each private history, the integrand of the dynamic IC ends up being the weighted sum of hypothetical continuation values with respect to the probability of private histories. At the end of the day, the net benefit from one more deviation comes from different outcome distribution and state transition through different probabilities of private histories.

I study full implication of the dynamic IC in another project, but the closest reference to the dynamic IC in the literature is the impulse response function in Pavan-Segal-Toikka (2014). Compared to their impulse response function, the dynamic IC of my model is not strictly an orthogonalization, because the axes I project the continuation value on are not orthogonal to each other. However, one can rewrite the dynamic IC as a series of projections, where each projection corresponds to one additional deviation. This requires rewriting the infinite sum in the form of $x_1 + (x_2 + (x_3 + \cdots$ and requires separating out the probability of each private history by individual periods.

7 Conclusion

I study mechanism design with moral hazard in this paper. I show that revelation principle extends to models with moral hazard, both with full commitment and limited commitment, and not eliciting agents' beliefs is without loss of generality if there is symmetric information on the equilibrium path or the agent's optimal strategy doesn't depend on the payoff-relevant state. One-shot deviation principle as in the repeated games literature holds in this environment, but it requires the knowledge of agents' private strategies after every history; I show a necessary and sufficient condition (dynamic IC) that requires only the knowledge of agents' equilibrium strategies. The dynamic IC however still has to be satisfied against every possible deviation and has to take into account multi-period deviations. I then show two sufficient conditions for on-path single deviation ICs to be sufficient for all ICs. I also provide dynamic programming for the dynamic IC.

The class of models I study allows for any degree of commitment power, risk aversion, persistence and any objective function by the principal. Decomposition of continuation value, the tool I used to derive the IC constraints and develop dynamic programming, can also be adapted to decision problems and competitive-market settings. I allow for non-Markovian environments, endogenous states, and agents may or may not have private information on the payoff-relevant state.

The scope of the tools I develop in this paper is much wider than the application in section 5; the application (the principal-agent version of Board-Meyer-ter-Vehn) is an example of how the revelation principle and the sufficiency of on-path single deviation ICs can be applied. It also shows how decomposition of continuation values can be adapted to decision problems, and I show the bang-bang result of the agent's action and the backloading of payments. The bang-bang result also shows why in linear models, having binary actions is without loss and allowing for an interval only leads to the extreme actions on the equilibrium path. Together with backloading, it shows that linearity is a stronger driving force than endogeneity of state transition.

References

[1] Bergemann, Dirk, and Ulrich Hege. "Venture capital financing, moral hazard, and learning." *Journal of Banking & Finance* 22.6 (1998): 703-735.

[2] Bergemann, Dirk, and Ulrich Hege. "The financing of innovation: Learning and stopping." *RAND Journal of Economics* (2005): 719-752.

[3] Bester, Helmut, and Roland Strausz. "Contracting with imperfect commitment and the revelation principle: the single agent case." *Econometrica* 69.4 (2001): 1077-1098.

[4] Bhaskar, V. "The ratchet effect re-examined: A learning perspective." (2014) Working paper.

[5] Bhaskar, V. and George Mailath. "The curse of long horizons." (2017) Working paper.

[6] Blundell, Richard and Monica Costa-Dias, Costas Meghir and Jonathan Shaw. (2016) "Female labour supply, human capital and welfare reform." *forthcoming Econometrica.*

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

[8] DeMarzo, Peter M. and Yuliy Sannikov. "Learning, termination, and payout policy in dynamic incentive contracts." (2016) Working paper.

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

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

[11] Garrett, Daniel F., and Alessandro Pavan. "Managerial turnover in a changing world." *Journal of Political Economy* 120.5 (2012): 879-925.

[12] Halac, Marina, Navin Kartik, and Qingmin Liu. "Optimal contracts for experimentation." *forthcoming Review of Economic Studies*

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

[14] He, Zhiguo, Bin Wei, Jianfeng Yu, and Feng Gao. "Optimal long-term contracting with learning." (2016) *forthcoming Review of Financial Studies.*

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

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

[17] Kapička, Marek. "Pareto efficient taxation with learning by doing." (2015) Working paper.

[18] Kapička, Marek and Julian Neira. "Optimal taxation with risky human capital." (2016) Working paper.

[19] Kwon, Suehyun. "Adverse selection and moral hazard with multidimensional types." (2017) Working paper.

[20] Makris, Miltiadis, and Alessandro Pavan. "Incentives for endogenous types: taxation under learning-by-doing." (2017) Working paper.

[21] Pavan, Alessandro, Ilya Segal, and Juuso Toikka. "Dynamic mechanism design: A Myersonian approach." *Econometrica* 82.2 (2014): 601-653.

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

[23] Sannikov, Yuliy. "Moral hazard and long-run incentives." (2014) Working paper.

[24] Stantcheva, Stefanie. "Optimal taxation and human capital policies over the life cycle." (2016) *forthcoming Journal of Political Economy.*

Supplementary Material

A Dynamic Programming with Dynamic IC

This section presents a formal derivation of dynamic programming in section 4.4. When the dynamic IC is necessary, one needs to do transfinite induction to find the set of payoffs that can be obtained by any incentive-compatible mechanism. I consider a Markovian environment with endogenous state and a single agent. If the agent perfectly observes the state each period, then we know from theorem 5 that on-path single deviation IC constraints are sufficient for all IC constraints. I consider the ex-ante symmetric information case in this section, i.e., both the principal and the agent start with a common prior in the beginning of the game, and there is no private signal on the payoff-relevant state.

A.1 Setup

Existence of optimal mechanism requires some technical assumptions as described below. However, they are automatically satisfied whenever the environment is finite.
$\omega \in \Omega$: set of states (non-empty compact Borel subset of a Polish (complete, separable, metric) space)
$p \in A$: set of actions (non-empty compact Borel subset of a Polish space)
$y \in Y$: set of outcomes (non-empty compact Borel subset of a Polish space)
$P_{\omega\omega'}(p)$: probability of going from state ω to ω' when the agent chooses p. (problility measure on ω' given ω, p) Assume P is jointly continuous in ω, ω', p.
$c(p) \in \mathbb{R}$: cost of action p. Assume it's continuous in p.
$f_\omega(y)$: pdf of outcome y in state ω. (measurable, non-negative and $\int_Y f_\omega(y)dy = 1$ for all $\omega \in \Omega$) Assume $f_\omega(y)$ is a continuous function of ω, y and $\int_Y y f_\omega(y)dy < M$ for some M, all $\omega \in \Omega$. Also assume full support, atomless, positive density everywhere.
$\pi(\omega)$: belief on state ω.

Assume f, P are such that if we start with a uniformly bounded π, then resulting beliefs in all subsequent periods are also uniformly bounded. Or more precisely, sup norm is well-defined for resulting beliefs. If there are a finitely many states, it is automatically satisfied.

A.2 Quick Summary of Dynamic IC

After outcome y is observed, the belief is updated from $\pi(\cdot)$ to

$$\pi^0(\omega) = \frac{\pi(\omega) f_\omega(y)}{\int_\Omega \pi(\omega') f_{\omega'}(y) d\omega'},$$

and in the following period, the belief is

$$\tilde{\pi}(\hat{\omega}) = \int_\Omega \frac{\pi(\omega) f_\omega(y) P_{\omega\hat{\omega}}(p)}{\int_\Omega \pi(\omega') f_{\omega'}(y) d\omega'} d\omega.$$

For π^0 to be well defined, I need $f_\omega(y)$ to be measurable in ω. For $\tilde{\pi}$ to be well defined, I need P given $\hat{\omega}, p$ to be measurable in ω.

Since P is jointly continuous in ω, p and Ω is compact, for given $\hat{\omega}$ and $\delta > 0$, there exists $\epsilon(\hat{\omega}, \delta) > 0$ such that $|P_{\omega\hat{\omega}}(p') - P_{\omega\hat{\omega}}(p)| < \delta$ for all $|p' - p| < \epsilon(\hat{\omega}, \delta)$. Since P is also continuous in $\hat{\omega}$, we can find $\epsilon(\hat{\omega}, \delta)$ continuous in $\hat{\omega}$, and together with the compactness of Ω, we get $\epsilon(\delta) > 0$ such that $|\tilde{\pi_{p'}}(\hat{\omega}) - \tilde{\pi}_p(\hat{\omega})| < \delta$ for all $\hat{\omega}$, $|p' - p| < \epsilon(\delta)$, and $\tilde{\pi}$ is a continuous function of p. $\tilde{\pi}$ is a continuous function of π^0. (P is jointly continuous on a compact set) When Ω, Y are compact and $f_\omega(y)$ is continuous in ω, y, π^0 is a continuous function of π. Up to here, I used pointwise convergence and sup norm.

The hypothetical continuation value of the agent is

$$\int_\Omega \int_Y -c(p) + u(w(y)) + \delta \int_\Omega V(y, \hat{\omega}) P_{\omega\hat{\omega}}(p) d\hat{\omega} f_\omega(y) dy \pi(\omega) d\omega.$$

I need to make sure $u(w(y))$ is measurable in y and $V(y, \omega)$ is measurable in y, ω. But it seems reasonable to assume that the principal can offer $w(y)$ as a measurable function of y, and by our definition of $V(y, \omega)$, it should be measurable in both y, ω. The limited liability from section 2.2 is sufficient if $w(y) \geq -\underline{M}$ for some $\underline{M}$ sufficiently large.

The on-path single deviation IC is equivalent to

$$\begin{array}{rc} & \int_\Omega \int_Y -c(p) + u(w(y)) + \delta \int_\Omega V(y, \hat{\omega}) P_{\omega\hat{\omega}}(p) d\hat{\omega} f_\omega(y) dy \pi(\omega) d\omega \\ \geq & \int_\Omega \int_Y -c(p') + u(w(y)) + \delta \int_\Omega V(y, \hat{\omega}) P_{\omega\hat{\omega}}(p') d\hat{\omega} f_\omega(y) dy \pi(\omega) d\omega \\ \Leftrightarrow & \int_\Omega \int_Y c(p') - c(p) + \delta \int_\Omega V(y, \hat{\omega})(P_{\omega\hat{\omega}}(p) - P_{\omega\hat{\omega}}(p')) d\hat{\omega} f_\omega(y) dy \pi(\omega) d\omega \geq 0 \end{array}$$

If we assume $P_{\omega\hat{\omega}}(p)$ is differentiable in p, we can take the left and right limits (after dividing by $p' - p$) and get the equality constraint.

The dynamic IC is given by

$$\sum_{n=t}^{\infty} \delta^{n-t} \int \int_{\Omega} \int_{Y} c(p'_n(\tilde{h^n})) - c(p_n(\hat{h^n})) + \delta \int_{\Omega} V_{y\hat{\omega}}(P_{\omega\hat{\omega}}(p_n(\hat{h^n})) - P_{\omega\hat{\omega}}(p'_n(\tilde{h^n})))d\hat{\omega} f_{\omega}(y)dy\tilde{\pi}^n(\omega)d\omega dG \geq 0$$

where G is the cdf of reaching each history given the agent's true private history.

A.3 Dynamic programming

Hypothetical continuation values $V(y, \omega)$ can be thought of as a bounded function $V : Y \times \Omega \to \mathbb{R}$. It is bounded because the principal has no commitment power and the expected outcome in each state is uniformly bounded. If the principal has full-commitment power, I need to show that the expected utility is bounded. (this is necessary both for the agent's optimal action to be well-defined and also for continuity at infinity)

Lemma 1. *Hypothetical continuation values on the equilibrium path are uniformly bounded.*

Proof. We know that the equilibrium payoffs are uniformly bounded, but a priori, we cannot rule out having unbounded hypothetical continuation values on a set of measure zero; I restrict attention to equivalent classes that coincide on a set of measure one. I'll show that $V(y_{t-1}, \omega_t)$ is a continuous function of ω_t and p_t. Given $V(y_{t-1}, \omega_t)$, we have $w(y_t), V_{t+1}(y_t, \omega_{t+1})$ such that $V(y_{t-1}, \omega_t) = \int_Y -c(p_t) + u(w(y_t)) + \delta \int_{\Omega} V(y_t, \omega_{t+1}) P_{\omega_t \omega_{t+1}}(p_t) d\omega_{t+1} f_{\omega_t}(y_t) dy_t$. I'll first show that $\int_{\Omega} V(y_t, \omega_{t+1}) P_{\omega_t \omega_{t+1}}(p_t) d\omega_{t+1}$ is continuous in ω_t, p_t. Define $g(\omega, p) = \int_{\Omega} V(y, \omega') P_{\omega\omega'}(p) d\omega'$ and consider a sequence $\omega^n \to \omega$. Since P is jointly continuous, $V(y, \omega') P_{\omega^n \omega'}(p) \to V(y, \omega') P_{\omega\omega'}(p)$ almost everywhere. Together with the compactness of Ω, the joint continuity of P implies that for given ω, there exists a neighborhood $\mathcal{N}_{\omega}$ such that $P_{\omega^n \omega'}(p) \leq \kappa P_{\omega\omega'}(p)$ for some $\kappa > 0$ and all ω', p and $\omega^n \in \mathcal{N}_{\omega}$. We already know that $V(y, \omega') P_{\omega\omega'}(p)$ is integrable, and limited liability implies that hypothetical continuation values are bounded from below. Then there exists a neighborhood $\mathcal{N}'_{\omega} \subseteq \mathcal{N}_{\omega}$ such that $|V(y, \omega')|$ is also integrable on $\mathcal{N}'_{\omega}$. By the dominated convergence theorem, $g(\omega^n, p) \to g(\omega, p)$. The proof for continuity in p_t is similar. The proof of continuity of $V(y_{t-1}, \omega_t)$ in ω_t, p_t is similar, and we use the fact that $-c(p_t) + u(w(y_t)) + \delta \int_{\Omega} V(y_t, \omega_{t+1}) P_{\omega_t \omega_{t+1}}(p_t) d\omega_{t+1} f_{\omega_t}(y_t)$ is bounded from below. Therefore, the hypothetical continuation value is a continuous function on a compact set and is bounded. □

I'll use the fact that the dynamic IC is equivalent to satisfying any N-period IC for all N.

Theorem 8. *The dynamic programming is well-defined, i.e., there exists a sequence of set operations such that the set-theoretic limit is the largest self-generating set. The agent's optimal actions for any N-period IC in the largest self-generating set is well-defined. The supremum of the principal's payoff is well-defined.*

Proof. I will set up the dynamic programming problem and show that at each iteration, the most profitable deviation for the agent is well-defined. The sequence of sets we get after each iteration is non-increasing and has a well-defined limit in the set-theoretic sense. The largest self-generating set is non-empty because the agent choosing the cheapest action and the principal making no payment is an equilibrium. With the compact action set and the continuous cost function, the cheapest action exists. With no commitment, the relevant constraints for the dynamic programming are (i) the principal offers the equilibrium contract (ii) the agent accepts/rejects according to the equilibrium strategy (iii) the agent's dynamic IC (iv) the principal makes the payment (v) the promise-keeping constraint. With within-period commitment, the relevant constraints for the dynamic programming are (i) the principal offers the equilibrium contract (ii) the agent accepts/rejects according to the equilibrium strategy (iii) the agent's dynamic IC (iv) the promise-keeping constraint. Whether the principal has no commitment power or within-period commitment power matters for the minmax NE. I'm not going to specify the minmax NE here, but if either the principal or the agent prefers his outside option over the minmax NE, then they'll take their outside options. This pins down the lower bound on payoffs for the IR constraints. For the rest of the proof, I assume within-period commitment power and ignore the principal's incentives to make payments he promised; with no commitment power, this will put an upper bound on the payment the principal can make (the continuation payoff minus the minmax NE or the outside option). Computationally, I can just impose the outside options and see the minmax NE from the largest self-generating set. If the minmax NE is better than the outside option for both the principal and the agent, then I need to use the minmax NE instead of the outside options. I could also just assume that the minmax NE we get by imposing IR with outside options is worse than taking their outside options.

I will construct a sequence of operations so that the limit is the largest self-generating set we want. I don't think I need the agent's optimal action to be unique, but I still need the agent to play a pure strategy. The state space for the dynamic programming is $(\pi, V(\cdot,\cdot), V^P)$ where V^P is the principal's expected payoff. The argument I'm going to use for N-period ICs works as long as there is continuity at infinity. If there is continuity at infinity, there must be a profitable N-period deviation for N sufficiently large, and we can do backward induction.

I need to specify the sequence of operations: I start with the on-path single deviation IC, and for each N, I iterate the operation for the N-period deviation IC until I reach the limit. Once I have the limit for N-period deviations, I continue with $(N+1)$-period deviations. And I take the limit as $N \to \infty$. Let's start with $W^0$where $V(\cdot,\cdot)$ are just assumed to be bounded by the uniform bound on the hypothetical continuation values. The iteration for the N-period deviation IC is T_N, and the limit of T_N starting with W^{N-1} is W^N. Also define $T_N(W^{N,i-1}) = W^{N,i}$, $W^{N,0} = W^{N-1}$. T_1 is just the standard largest self-generating set with the on-path single deviation IC constraint. Among the constraints, (i) and (ii) just mean that the payoffs are weakly greater than the outside options (or minmax NE). (iv) can be taken care of as follows: Suppose we have V_{t+2}. When we choose $w_{t+1}(y)$ for each y, we can pin down V_{t+1} that is consistent with the promise-keeping constraint. First find the largest self-generating set subject to (ii), (iv) and the on-path single deviation IC (without worrying about the principal's payoff). Once we have W^1, we know V^P for each pair of $(\pi, V(\cdot,\cdot))$ and can keep only those that satisfy (i). If there are multiple V^Ps corresponding to $(\pi, V(\cdot,\cdot))$ then choose the supremum of V^P (following the principle of optimality). At this point, we haven't shown that whether the supremum can be obtained as the maximum. But we also know that once we have V^P we can generate any $\hat{V^P} < V^P$ as long as it's weakly greater than the principal's outside option. For T_1, we can show that the agent's optimal action is well-defined because the agent's expected utility is bounded and is a continuous function of his action. Generally speaking, to show that the most profitable N-period deviation is well-defined, I need to show a version of selection theorem, and I need Ω to be a Borel subset of a Polish space, A to be a compact metric space and the agent's expected utility from an N-period deviation to be bounded and upper semi-continuous.

T_N for $N \geq 2$ are defined as follows: Fix π^t and an action p, and we can find the beliefs π^{t+1} that are consistent with π^t, p. Choose $V_{t+2}(y, \omega)$ from $W^{N,i}$ and $w_{t+1}(y)$ for each π^{t+1} such that V_{t+1}, V^P_{t+1} given by the promise-keeping constraint also satisfies (i), (ii) and the on-path single deviation IC constraint for period t is satisfied at p. By construction, there are no profitable $(N-1)$-period deviations starting with $V_{t+2}(y, \omega)$, and we can find optimal actions for the next $N-1$ periods and $V_{t+N}(y, \omega)$ from period $t+N$ on. We should have the payments after each history from period $t+1$ to period $t+N-1$. We also know the equilibrium belief after each history. Fix $\pi^{\tilde{t+N-1}}$ for each history and we can find the most profitable deviation for the agent in $t+N-1$ and therefore assign the agent's maximum deviation payoff from period $t+N-1$ on as a function of $\pi^{\tilde{t+N-1}}$ given the continuation game from $t+N-1$ on the equilibrium path. Fix $\pi^{\tilde{t+N-2}}$ then conditional on the agent's action p'_{t+N-2} we know the agent's beliefs $\pi^{\tilde{t+N-1}}$and his maximum deviation payoff. We can find the most profitable deviation for the agent in period $t+N-2$ and assign the agent's deviation payoff from $t+N-2$ on as a function of $\pi^{\tilde{t+N-2}}$ and the continuation game from $t+N-2$ on the equilibrium path. We can repeat this until we reach $\pi^{\tilde{t}}$. To show that the most profitable deviation is well-defined, suppose we are in period $t+n$ with $\pi^{\tilde{t+n}}, p_{t+n}, w_{t+n}(\cdot), \tilde{V}_{t+n+1}(\cdot,\cdot)$ where $\tilde{V}_{t+n+1}$ is the agent's maximum deviation payoff from period $t+n+1$ on. (in period $t+N-1$, these will just be the hypothetical continuation values from period $t+N$ on) In period $t+n$, the agent's expected utility from period $t+n$ on is a continuous function of $\pi^{\tilde{t+n}}, p_{t+n}$ and we know that it is finite for $\pi^{\tilde{t+n}} = \pi^{t+n}$ and the equilibrium action p_{t+n}. The proof follows the proof of lemma 1 closely, and we know from limited liability that $\tilde{V}_{t+n+1}$ is bounded from below. Since there is no profitable $(N-1)$-period deviation, if the agent starts with π^{t+n} and chooses p_{t+n}, his maximum deviation payoffs from the next period on coincides with his equilibrium payoffs; it follows that the agent's expected utility from period $t+n$ on is bounded. Since the set of priors and the set of actions is compact, we know that the product of the two is compact (Tychonoff's Theorem). Therefore, the agent's deviation payoff from period $t+n$ on is a continuous function on a compact set and is bounded. Therefore, for given $\pi^{\tilde{t+n}}$, there exists the maximum deviation payoff for the agent, and the agent has the most profitable deviation. But in period t, the principal and the agent share the same prior $\pi^{\tilde{t}} = \pi^t$. Keep $V_{t+1}(y, \omega)$ for π^t if and only if it is incentive compatible with respect to the maximum deviation payoff. (I use backward induction to find the most profitable N-period deviation for the agent, but I also use backward induction to show that the maximum deviation payoff for the agent is a continuous function of his belief and his action in the given period. This needs a proof because the agent's maximum deviation payoff from the next period on depends on his belief in the next period)

Since each operation $T_N : W^{N,i-1} \to W^{N,i}$ satisfies $W^{N,i} = T(W^{N,i-1}) \subseteq W^{N,i-1}$, we have a monotone sequence, and the limit is well-defined in the set-theoretic sense. By construction, it is the largest self-generating set satisfying all four conditions. It also follows from the previous paragraph that the agent's optimal actions for any N-period IC is well-defined. The supremum of the principal's payoff for any given π is well-defined. □

B Proofs

Proof of Theorem 1. First consider the case when the principal has full commitment. If the agent i's strategy $\sigma^{t,i,m}$ involves randomizing over messages, the principal can ask the agent to report his private information and randomize on behalf of the agent. We just need to take care of the possibility that the agent conditioned his action on his report and his private history in the original contract.

Formally, given a contract $(\times_i \mathcal{M}_t^i, \sigma^{t,P,r}, \sigma^{t,P,z})_t$, let histories and strategies denoted as in section 2. Given a compilation of private histories $\bar{h}^t$, $h^{t,i,m}(\bar{h}^t), h^{t,i,a}(\bar{h}^t), h^{t,P,r}(\bar{h}^t), h^{t,P,z}(\bar{h}^t)$ denote histories consistent with $\bar{h}^t$; for given $\bar{h}^t$, there is a unique private history at each point consistent with $\bar{h}^t$, and these are well-defined functions.

The principal can offer an alternative contract with $\tilde{\mathcal{M}}_t^i = \mathcal{S}_t^i$ each period. Starting from period 1, define $\bar{\mathcal{H}}^t(s_t)$ to be the compilation of private histories that contain s_t, the messages, recommended actions until period $t-1$ and $r_k^i = a_k^i$ for all $k < t, i$. It is a subset of $\bar{\mathcal{H}}^t$ and there is a natural probability distribution over the payoff-relevant state $\omega_k, k \leq t$ on the equilibrium path. Let μ be the measure on $\bar{\mathcal{H}}^t(s_t)$ induced by this probability distribution. With probability $\int_{\bar{\mathcal{H}}^t(s_t)} \times_i \sigma^{t,i,m}(m_t^i | h^{t,i,m}(\bar{h}^t)) \times \sigma^{t,P,r}(r_t | h^{t,P,r}(\bar{h}^t)) d\mu$, the mechanism recommends r_t^i along with the hypothetical message m_t^i to agent i. If each agent has been reporting truthfully and obeying the recommendation by the mechanism, then the randomization probabilities this period are the same, and the expected utility of each agent doesn't change; truthful reporting is incentive-compatible for each agent this period as well.

With limited commitment, the contract is essentially the combination of Bester-Strausz (2001) and the above argument. When the principal doesn't commit to the contract, we need to satisfy the principal's incentives for recommendation and allocation. The principal's incentives for allocation is the same as in the original contract, and for recommendation, the same argument as in Bester-Strausz (2001) ensures that the principal's belief after messages are sent is the same across all messages that are sent with positive probabilities.

When the principal has full commitment and agent i has no private signal in period t, the agent's message space in the new mechanism is an empty set, and it is without loss of generality not to ask for reports. When the principal has limited commitment and all agents have no private signal in period t, it is without loss of generality not to ask for reports. $\square$

Proof of Theorem 2. Since the outcome distribution only depends on the state and the state transition is Markovian, the agent's action this period only affects the current-period state transition once we decompose the agent's continuation value. From the assumption on the state transition, the optimal action from both the principal's perspective and the agent's perspective don't depend on the current state, and the principal doesn't benefit from eliciting the agent's belief. $\square$

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, ω_i) such that the agent's continuation value from period $t+1$ on conditional on being in ω_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 $a \in [0,1)$. The agent's action affects the probability of (y, ω_i), and the agent's expected payoff is

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

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

$$V_1 = -c(a) + \int u(w(y)) + \delta(r(a)V_{y\omega_1} + (1-r(a))V_{y\omega_2})dF_1,$$

$$V_2 = -c(a) + \int u(w(y)) + \delta(aV_{y\omega_1} + (1-a)V_{y\omega_2})dF_2.$$

$\square$

Proof of Theorem 3. A compilation of private histories contains all private information up to that point. If all agents and the principal follow the equilibrium strategy from the following period on, there is a unique hypothetical continuation value for each agent, and the rest is just taking the expectation over the compilation of private histories consistent with agent i's private history up to that point. $\square$

Proof of Corollary 1. Theorem 1 shows that it is without loss of generality not to ask for messages. The proof follows from theorem 3 after taking care of degenerate distributions. $\square$

Proof of Proposition 2. Suppose the agent's prior at the beginning of period t is π. The agent's expected payoff from choosing a is

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

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

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

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

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

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

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 $a' \to a$, and we have

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

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

$$\delta(\pi_1 r'(0) W_1 + \pi_2 W_2) \leq c'(0).$$

□

Proof of Proposition 3. Suppose the agent's prior in the beginning of period t is π and the agent deviates N periods. After N periods, the agent conforms to the principal's expectation. Let $a(h^n)$ denote the agent's equilibrium action and $a'(h^n)$ 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 $t+N-1$, $X(\tilde{a}_t, \cdots, \tilde{a}_{t+N-1})$. The agent's IC constraint is

$$X(a_t, \cdots, a_{t+N-1}) - X(a'_t, \cdots, a'_{t+N-1}) \geq 0,$$

which is equivalent to

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

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 one more deviation is

$$X(a'_t,\cdots,a'_{n-1},a_n,\cdots,a_{t+N-1})-X(a'_t,\cdots,a'_n,a_{n+1},\cdots,a_{t+N-1}).$$

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

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

where $V_i(\hat{h}^n)$ is the hypothetical continuation value of the agent conditional on history $\hat{h}^n$ and being in state ω_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

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

The $N-$period IC constraint is

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

This dynamic IC constraint has to be satisfied for any N. Furthermore, given the $N-$period IC, we can always let $a'_{t+N-1}=a_{t+N-1}$, 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:

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

I'll next show the sufficiency of the dynamic IC. 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 σ' from period t on. Suppose σ' provides ϵ more to the agent than his equilibrium strategy σ. By continuity at infinity, there exists N sufficiently large and another deviation strategy σ'' such that the agent's payoff from σ'' is at least $\frac{\epsilon}{2}$ higher than his equilibrium payoff and

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

σ'' coincides with σ' for the first N periods and coincides with the principal's expectation from period $t+N$ on; $\hat{h}$ is the private history of the agent the principal believes is the agent's true private history. It 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 Theorem 4. Compared to proposition 3, we just need to take expectations over all compilations of private histories that are consistent with agent i's private history. In the binary model from section 2.1, the principal has limited commitment which provides continuity at infinity, and the dynamic IC is necessary and sufficient in proposition 3. Otherwise, the dynamic IC is necessary but it is sufficient only if there is limited commitment or continuity at infinity. □

Proof of Corollary 2. The proof follows from theorem 4 after accounting for degenerate distributions. □

Proof of Theorem 5. Theorem 5 assumes (i) each agent observes the state every period, ($s_t^i=\omega_t$) (ii) the domain of the outcome distribution f_t is $\Pi_{j=1}^{t}\Omega_j\times\mathcal{A}_t\times\Pi_{j=1}^{t-1}\mathcal{Y}_j\times\Pi_{j=1}^{t-1}\mathcal{Z}_j$, and (iii) the domain of the state transition P_t is $\Pi_{j=1}^{t}\Omega_j\times\mathcal{A}_t\times\Pi_{j=1}^{t}\mathcal{Y}_j\times\Pi_{j=1}^{t}\mathcal{Z}_j$. Conditions (ii) and (iii) imply that past actions of the agent doesn't matter for the outcome distribution nor the state transition. No matter which action the agent takes in period t, after observing the state in period $t+1$, the agent's continuation value only depends on the continuation strategies of the principal and the agent. In particular, given the public history up to the beginning of period $t+1$, the equilibrium strategy of the agent corresponding to the private history the principal believes is the agent's true private history is an optimal continuation strategy. Therefore, on-path single deviation ICs are sufficient for all IC constraints. $\square$

Proof of Corollary 3. Corollary 3 is a special case of theorem 5 where f_t, P_t are further restricted to be Markovian. $\square$

Proof of Theorem 6. Suppose (i) the outcome distribution and the state transition are Markovian, (allowing for fully persistent states) (ii) $\mathcal{S}_t^i = \emptyset$ for all t, i, (iii) each agent can only deviate downwards, (iv) the posterior of each agent dominates the equilibrium belief in the sense of first-order stochastic dominance after a deviation and (v)

$$\int_{Y_t} \left(u_t^i(z_t^i(h^{t,P,z})) + \delta_A^i \int_{\Omega_{t+1}} V_{t+1}(h^t, \tilde{a}_t^i, a_t^{-i}, y_t, z_t, \omega_{t+1}) P_t(\omega_{t+1}) d\omega_{t+1} \right) f_t(y_t) dy_t$$

is supermodular in ω_t and $\tilde{a}_t^i$.

When agents have no private information on the state, theorem 1 shows that it is without loss of generality not to ask for messages. The dynamic IC on the equilibrium path from corollary 2 is

$$\begin{aligned}
&\sum_{n=t}^{\infty} \delta^{n-t} \int \Bigg(\int_{\Omega_n} -(c_n^i(a_n^i) - c_n^i(\tilde{a}_n^i)) \\
&+ \Bigg(\int_{Y_n} u_t^i(z_n^i(h^{n,P,z})) + \delta_A^i \int_{\Omega_{n+1}} V_{n+1}(h^n, a_n, y_n, z_n, \omega_{n+1}) P_n(\omega_{n+1}|a_n^i) d\omega_{n+1} f_n(y_n|a_n^i) dy_n \\
&- \int_{Y_n} u_t^i(z_n^i(h^{n,P,z})) + \delta_A^i \int_{\Omega_{n+1}} V_{n+1}(h^n, \tilde{a}_n^i, a_n^{-i}, y_n, z_n, \omega_{n+1}) P_n(\omega_{n+1}|\tilde{a}_n^i) d\omega_{n+1} f_n(y_n|\tilde{a}_n^i) dy_n \Bigg) d\tilde{\pi}_n^i \Bigg) \\
&dG_n(h^{n,i,a}) \geq 0
\end{aligned}$$

Conditional on history $h^{n,i,a}$, if agent i has already deviated in the past, his belief $\tilde{\pi}_n^i$ dominates the equilibrium belief π_n^i in the sense of first-order stochastic dominance.

Since

$$\int_{Y_n} u_t^i(z_n^i(h^{n,P,z})) + \delta_A^i \int_{\Omega_{n+1}} V_{n+1}(h^n, a_n, y_n, z_n, \omega_{n+1}) P_n(\omega_{n+1}|a_n^i) d\omega_{n+1} f_n(y_n|a_n^i) dy_n$$

is supermodular in ω_t and a_t^i and the agent can only deviate downwards, if on-path single deviation IC holds on the equilibrium path, we have

$$\begin{aligned}
&\arg\max_{\tilde{a}_n^i} \Bigg(\int_{\Omega_n} -(c_n^i(a_n^i) - c_n^i(\tilde{a}_n^i)) \\
&+ \Bigg(\int_{Y_n} u^i(z_n^i(h^{n,P,z})) + \delta_A^i \int_{\Omega_{n+1}} V_{n+1}(h^n, a_n, y_n, z_n, \omega_{n+1}) P_n(\omega_{n+1}|a_n^i) d\omega_{n+1} f_n(y_n|a_n^i) dy_n \\
&- \int_{Y_n} u^i(z_n^i(h^{n,P,z})) + \delta_A^i \int_{\Omega_{n+1}} V_{n+1}(h^n, \tilde{a}_n^i, a_n^{-i}, y_n, z_n, \omega_{n+1}) P_n(\omega_{n+1}|\tilde{a}_n^i) d\omega_{n+1} f_n(y_n|\tilde{a}_n^i) dy_n \Bigg) d\tilde{\pi}_n^i \Bigg) \\
=\ & a_n^i.
\end{aligned}$$

On-path single deviation ICs are sufficient for all ICs. □

Proof of Theorem 7. The first three points in the theorem is already proved in the main text. I replicate them here for completeness.

The first observation is that the optimal investment doesn't depend on the current quality. This is because of the state transition. If the shock doesn't arrive, the firm's investment is irrelevant to its payoff, and when it arrives, the probability of becoming a high-quality firm is λa for both qualities. In this case, whether the firm knows its own quality or not doesn't matter for strategies, and in particular, this maps into the symmetric uncertainty case in theorem 1 where not eliciting the agent's belief is without loss of generality. The principal doesn't benefit from asking the firm to report its quality.

Second of all, the discrete-time principal-agent version of Board-Meyer-ter-Vehn is a Markovian environment with both adverse selection and moral hazard. From corollary 3, on-path single deviation ICs are sufficient for all IC constraints.

The on-path single deviation IC with high quality is given by

$$\begin{aligned}
&-ca_t + (\lambda a_t + 1 - \lambda)(\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H)) + \lambda(1-a_t)(\mu_L V(s,L) + (1-\mu_L)V(\emptyset,L)) \\
\geq\ &-ca' + (\lambda a' + 1 - \lambda)(\mu_H V(s,H) + (1-\mu_H)V(\emptyset,H)) + \lambda(1-a')(\mu_L V(s,L) + (1-\mu_L)V(\emptyset,L)),
\end{aligned}$$

for all a' and with low quality, it's given by

$$\begin{aligned}&-ca_t+\lambda a_t(\mu_H V(s,H)+(1-\mu_H)V(\emptyset,H))+(1-\lambda a_t)(\mu_L V(s,L)+(1-\mu_L)V(\emptyset,L))\\ \geq\ &-ca'+\lambda a'(\mu_H V(s,H)+(1-\mu_H)V(\emptyset,H))+(1-\lambda a')(\mu_L V(s,L)+(1-\mu_L)V(\emptyset,L)),\ \forall a'.\end{aligned}$$

I dropped the public history up to the beginning of the period in the notation and expressed hypothetical continuation values only in terms of current period signal and and the quality. One can see that the on-path single deviation IC is identical for both the high quality and the low quality:

$$\begin{aligned}&-ca_t+\lambda a_t(\mu_H V(s,H)+(1-\mu_H)V(\emptyset,H))-\lambda a_t(\mu_L V(s,L)+(1-\mu_L)V(\emptyset,L))\\ \geq\ &-ca'+\lambda a'(\mu_H V(s,H)+(1-\mu_H)V(\emptyset,H))-\lambda a'(\mu_L V(s,L)+(1-\mu_L)V(\emptyset,L)),\ \forall a'.\end{aligned}$$

The third observation is that the on-path single deviation IC is linear in the firm's investment a_t: for all a',

$$(a_t-a')(-c+\lambda((\mu_H V(s,H)+(1-\mu_H)V(\emptyset,H))-(\mu_L V(s,L)+(1-\mu_L)V(\emptyset,L))))\geq 0.$$

In particular, let $V_H,\ V_L$ be the firm's payoff conditional on the current quality. We get

$$\begin{aligned}V_H&=-ca_t+(\lambda a_t+1-\lambda)(\mu_H V(s,H)+(1-\mu_H)V(\emptyset,H))+\lambda(1-a_t)(\mu_L V(s,L)+(1-\mu_L)V(\emptyset,L)),\\ V_L&=-ca_t+\lambda a_t(\mu_H V(s,H)+(1-\mu_H)V(\emptyset,H))+(1-\lambda a_t)(\mu_L V(s,L)+(1-\mu_L)V(\emptyset,L)).\end{aligned}$$

It follows that

$$V_H-V_L=(1-\lambda)(\mu_H V(s,H)+(1-\mu_H)V(\emptyset,H)-\mu_L V(s,L)-(1-\mu_L)V(\emptyset,L)),$$

and we get

$$\begin{aligned}a_t&=1\text{ if }V_H-V_L>\frac{1-\lambda}{\lambda}c,\\ a_t&=0\text{ if }V_H-V_L<\frac{1-\lambda}{\lambda}c.\end{aligned}$$

Unless the coefficient on a_t-a' is 0, i.e., $V_H-V_L=\frac{1-\lambda}{\lambda}c$, linearity implies that the firm chooses the extreme points $a_t=0$ or $a_t=1$. But one can easily see that if the coefficient is 0 and the firm chooses an interior $a_t \in (0,1)$, the principal can offer ϵ more to signal s and let the firm choose $a_t = 1$. The bang-bang result requires linearity of the cost function $c(a)$ and the state transition $r(a)$. Generically, the firm chooses an interior solution if either of $c(a)$, $r(a)$ is non-linear.

Results so far don't depend on details of the model including the signal structure, equilibrium definition, or whether it's a competitive market or principal-agent setting. The bang-bang result shows that any Markov Perfect Equilibrium has to be characterized by cutoffs in the belief space, and the MPE with perfect learning in Board-Meyer-ter-Vehn are still equilibria in the discrete-time version.

In the principal-agent version, one can also characterize the first best as follows. Let X_H, X_L be the total surplus from the current period on if the current state is the high quality or the low quality, respectively. We have

$$\begin{aligned}
X_H &= -ca + (1-\lambda+\lambda a)(1+\delta(\mu_H X_H(s) + (1-\mu_H)X_H(\emptyset))) \\
&\quad + \lambda(1-a)\delta(\mu_L X_L(s) + (1-\mu_L)X_L(\emptyset)), \\
X_L &= -ca + \lambda a(1+\delta(\mu_H X_H(s) + (1-\mu_H)X_H(\emptyset))) \\
&\quad + (1-\lambda a)\delta(\mu_L X_L(s) + (1-\mu_L)X_L(\emptyset)), \\
\Rightarrow \quad X_H - X_L &= (1-\lambda)(1+\delta(\mu_H X_H(s) + (1-\mu_H)X_H(\emptyset)) - \delta(\mu_L X_L(s) + (1-\mu_L)X_L(\emptyset)))
\end{aligned}$$

where $X_H(s), X_H(\emptyset), X_L(s), X_L(\emptyset)$ are total surplus from the next period on conditional on the state in the next period and the signal in the current period. If $a = 1$ every period for all beliefs, we have $X_H = X_H(s) = X_H(\emptyset),\ X_L = X_L(s) = X_L(\emptyset)$. Plugging them into above equalities, we know that the first best is $a = 1$ every period for all beliefs if

$$c \leq \frac{\lambda}{1-\delta(1-\lambda)}.$$

Since both the principal and the firm are risk-neutral, the principal can always delay payments and provide better incentives. We just need to make sure that the promise-keeping constraint is satisfied, and as long as the contract doesn't terminate, after sufficient number of periods, the continuation value of the agent from the delayed payment becomes bigger than the continuation value to implement the first-best action every period. But even when the contract dictates that the agent should be fired, the principal can provide zero incentives from there on until the continuation value from the delayed payment becomes sufficiently big at which point the agent can be incentivized to take the first-best action every period. Therefore, in an optimal contract, the payments are backloaded until the agent can be incentivized to take the first-best action every period. This doesn't depend on the signal structure. $\square$