DYNAMIC COSTS AND MORAL HAZARD

GUY ARIE - JOB MARKET PAPER

Abstract. The cost of effort often increases in past effort. In sales, for example, the last sales of a quarter are harder to make than the first ones – the pool of easy customers is depleted. In an agency setting with unobservable effort, increasing marginal cost complicates the optimal contract problem. If the agent shirks today, his cost tomorrow will be lower than the principal believes. The optimal contract is characterized as a *dynamic quota*. The main features of the optimal contract are consistent with the popular yet thus far puzzling use of nonlinear incentives for sales agents. Historically, the main obstacle for solving dynamic moral hazard problems with private information was that the one-shot-deviation principle cannot be applied. I develop a duality based representation for dynamic moral hazard problems and use it to obtain a stronger characterization of the optimal contract. In particular, the dynamic dual analysis shows that the *optimal* contract does satisfy a one-shot-deviation condition.

1. Introduction

Increasing marginal costs is a standard component of economic analysis. In organizational settings, the increase in cost often has a dynamic motivation – the worker gets tired or the task gets harder. This paper characterizes the optimal long term contract when effort is unobserved and costs increase with past effort.

To fix ideas, consider incentives for a sales person. Sales performance is measured over a period, typically quarter or year. As the quarter progresses, the agent depletes all the "easy" sales leads and must exert more effort to generate later sales. Sales effort is inherently hard to monitor and pay is often performance based. If the firm would know the agent's true cost, it may want to increase incentives towards the end of the quarter.

Indeed, most sales incentive schemes are nonlinear and provide much stronger incentives for end-of-quarter sales, based on the agent's performance earlier in the quarter.[1]

Date: 11/10/11.

Key words and phrases. Dynamic moral hazard, nonlinear incentives, private information, dynamic mechanism design, duality, linear programming, stochastic programming, dynamic programming.

I am indebted to Bill Rogerson, Mark Satterthwaite and especially Jeroen Swinkels and Michael Whinston for their advice and encouragement throughout this project. I thank Matthias Fahn, Paul Grieco, Jin Li and Rakesh Vohra for helpful suggestions. Financial support from the GM Center at Kellogg is gratefully acknowledged.

[1] 75% of the firms surveyed by Joseph and Kalwani (1998) use a nonlinear compensation scheme. Additional aggregate evidence is provided in Oyer (1998) and echoed in Prendergast (1999). Larkin (2007) cites industry reports that nonlinear commissions and similar schemes are used by "nearly every enterprise software vendor."

A popular nonlinear compensation scheme is a "quota contract" in which the agent gets paid a fixed share of revenue but only for sales above a quarterly threshold. Another common nonlinear scheme is a "convex contract" in which the share an agent gets of the revenue from a sale is a convex and increasing function of the total sales already made in the quarter. For example an agent in the firm studied by Larkin (2007) receives 2% for the first $250K sales but 25% for any sales above $6M.

However, nonlinear contracts distort incentives. Agents may "game the system" – delay some easy sales leads and use them only at the end of the quarter. An agent may also "give up" and stop exerting effort for the quarter if his early performance is not good enough to expect high rewards later in the quarter. These complications have proved difficult to model. As noted by Prendergast (1999) in his excellent survey, "rather remarkably, the theoretical literature has made little progress in understanding the observed (nonlinear) shape of compensation contracts, despite costs associated with nonlinearities."

I study a simple model to capture privately increasing costs and show this can explain the nonlinear features of observed contracts. In the model, a risk neutral agent decides every day whether to exert costly effort or not. The probability of success (a sale) in the day increases with effort. The cost of effort today is a convex function of past *effort*. Effort is unobserved and the principal can commit to a contract at the outset.

The optimal contract can be informally described as a *dynamic quota:* the agent starts in an evaluation stage and eventually moves to a compensation stage. In the compensation stage the agent is paid a fixed piece-rate for each sale and works for an additional fixed time that is independent of any new outcomes. In the evaluation stage the agent is rewarded *only* by changes to the expected fixed piece-rate, the length of the compensation stage, and the conditions for entering the compensation stage. If the agent accumulates enough early successes, his compensation per sale later in the quarter will be high. If the agent did not accumulate enough early successes, the contract leads the agent to stop working. Formally, during the evaluation stage, the contract is a list of contingent compensations. For example: "if you succeed in the next period you move to a four period compensation stage", "if you fail next period and then succeed, you move to a two period compensation stage", "if you fail in the next two periods, you're fired".

To see the incentive problem, suppose that the probability of a sale each period is $\frac{1}{2}$ if the agent exerts effort on a customer and zero otherwise, and that the agent's cost for making the n-th effort is n. If both the principal and the agent consider only current period incentives, a contract paying the agent $2n$ for a sale in day n is incentive compatible and provides the agent zero expected utility – clearly first best. However, if the agent considers future payoffs, this contract is no longer incentive compatible. Shirking in the first period and then working whenever asked obtains the agent an expected utility of 1 *each* period. By shirking today the agent increases his rents from future work. The optimal contract must account for this additional incentive to shirk: the agent's utility difference between success and failure must increase.

To evaluate the effect of privately known and increasing costs, I also compare the optimal contract to the one if the costs each period are known.[2] I find that the private information problem adds to the contract exactly the main features that the empirical sales literature finds difficult to explain: very high volatility of the work decision at the end of the quarter (see Oyer (1998)) and seemingly excessive rewards for successful agents (see Larkin (2007) and Misra and Nair (2009)). Interestingly, the analysis suggests that when costs are the agent's private information, firms should invest more in reducing the later and higher costs. This is consistent with another casual observation about sales

[2]For example, the sales-person may be sent each period to a different location that determines the cost of effort. In such cases, while costs are increasing, they are publicly known.

contracts. Firms often allocate additional resources, such as high quality sports tickets, to attract the more difficult to please customers.

The main difficulty in the existing literature for the analysis of dynamic moral hazard settings with private information is the failure of the one-shot-deviation (OSD) condition. That is, it is wrong to assume that if the agent never profits by deviating from the desired behavior just once, the contract is incentive compatible (IC).[3] The optimal long term contract must in theory consider multiple deviations by the agent, rendering the problem intractable. However, while some IC contracts may violate OSD, if the *optimal contract* does not, it is sufficient to limit attention to single deviations and the problem is again tractable. The main methodological contribution of this paper is a reformulation of the dynamic contract problem that allows checking directly whether the optimal contract satisfies OSD in the presence of private information.

The analysis departs from the existing literature by formulating the original problem as a linear program, deriving its dual and then obtaining a recursive representation – the *dynamic dual*. The dynamic dual analysis allows replacing the standard OSD condition with a local-deviations condition which may be roughly stated as follows. Take the optimal contract and consider making a change that increases the expected profits. This change must violate some incentive constraint. If the constraint that is violated is always a one-shot-deviation constraint, then the optimal contract subject only to OSD must be IC.[4]

The dynamic dual analysis also allows a stronger characterization of the optimal contract.[5] The value of the dynamic dual problem at each history accounts for the forward looking profits *and* the effect of the optimal continuation on the agent's past incentives. As a result, the dual value of a history h is exactly the increase in expected effort at the start of the relationship from the contract starting at the history. In contrast, the primal dynamic value in a history is the expected continuation profits (see e.g. Spear and Srivastava (1987)). Thus, while it is well known that the primal value of a history may be negative – the principal optimally committed to providing the agent very high continuation utility – whenever the dual value of a history is negative, the contract is more profitable by committing to terminate the relationship at the history.

The dual state variables proxy for the *degree of agency frictions* generated by the contract at the history. The model has two agency frictions - one generated by the effect of current utility on past incentives and one by the information asymmetry. Each defines a variable that is updated as a function of the period's control and outcome. For example, the information rent generated by working in a period is higher if the period follows a failure than if it follows a success. A useful implication of this is that the contract's dual value is monotonic in each state variable and that the state variables are substitutes. This is not typically the case in the standard recursive formulation and is critical in proving the OSD result.[6]

[3] See appendix A for an example.

[4] This is very similar to the mechanism design result of local deviations. Intuitively, one can consider each possible private history for an agent as an agent "type". The local deviations condition is that in each public history, the agent "type" that requires the strongest incentives is the "so-far compliant" type.

[5] An appendix analyzing the contract using standard recursive methods assuming the OSD condition is satisfied is available from the author. The results are a subset of these obtained here.

[6] Following Spear and Srivastava (1987), dynamic moral hazard analysis uses the agent's continuation utility as the state. If the agent's continuation utility is exactly his outside option, the contract must typically terminate and the principal obtains his outside option. If the agent's continuation utility is

The dynamic dual representation allows extending moral hazard theory to a new setting – increasing marginal cost. However, dynamic dual analysis may also be useful for more standard dynamic moral hazard models as well. While the development (but not the economic intuitions or analysis) relies heavily on the duality property of linear programming, the basic duality arguments generalize to concave programs as well (see Rockafellar (1997)). Moreover, as with the current model, some existing dynamic moral hazard models may be transformed into a linear program.

Following a short literature review, section 2 lays out the dynamic production model, formulates it as a linear problem and identifies a condition for the sufficiency of local deviations. Section 3 develops the basic dual problem and its dynamic formulation. Section 4 proves the sufficiency of local deviations. Section 5 studies the optimal contract, shows the "dynamic quota" structure and the optimal payment scheme. Section 6 isolates the effect of privately increasing costs on the optimal contract. Section 7 concludes.

1.1. **Relation To Existing Literature.**

Of the models that consider history *independent* technology, the agency setting and result characterization here is closest to Clementi and Hopenhayn (2006) and DeMarzo and Fishman (2007). Both consider a risk neutral agent subject to limited liability. The two papers emphasize that long term contracts increase surplus by allowing the principal to reward outcomes today with future rents instead of current utility. This allows the agent to provide "de-facto" liability - his unpaid effort - in return for future equity in the firm. Once the agent has fully paid for the firm through effort, it is effectively sold to the agent for free - the agent is promised the full fruits of all future labor. This is implemented by asking the agent to work at the first best level and providing all incentives for work by payment in the period.[7] In these models, the contract is always in one of three states, either (i) the agent is given the firm or (ii) the contract is terminated or (iii) the contract is in the process of resulting in one of the two other states. Positive period outcomes move the contract towards the first state (giving the firm) and negative period results move the contract towards the second (termination).

The current analysis shows that to counter the additional incentive to shirk resulting from the agent's private cost information, the optimal contract destroys some surplus after each period, with the destruction much larger after a failure. As a result, the contract can no longer "give the firm" to the agent. Instead, the contract eventually offers the agent an inefficient linear piece-rate.

This paper contributes to recent progress in dynamic agency theory with private history dependent technology. Fernandes and Phelan (2000) consider a agency settings in which today's information or effort affects tomorrow's productivity but limit the history dependence to one last period via a Markov assumption. Nevertheless, the result in Fernandes and Phelan (2000) for moral hazard settings is negative - whenever today's effort affects tomorrows productivity, the one-shot-deviation principle does not apply.

---

very high, the principal must either give away the firm (in limited liability settings, see e.g. Clementi and Hopenhayn (2006)), or provide the agent costly insurance. In all cases, the principal's expected continuation value is highest for some expected agent's continuation utility between the two extremes. In the extension to private information developed by Fernandes and Phelan (2000), it may well be that the state variables are in some cases complements and in others substitutes.

[7] Earlier models with similar intuitions for a risk averse agent with unlimited liability are Rogerson (1985) and Spear and Srivastava (1987). There, however, the risk aversion of the agent introduces other considerations.

Several recent papers follow the modeling approach of Fernandes and Phelan (2000) to analyze dynamic moral hazard settings with payoff relevant private histories. In a recent working paper, DeMarzo and Sannikov (2008) extend the aggregation problem of Holmstrom and Milgrom (1987) to settings in which the agent also obtains private shocks to his productivity that are correlated with past effort. While DeMarzo and Sannikov (2008) is closest to the setting studied here, the additional aggregation problem restricts the results. Appendix C outlines an implementation of the duality method with the aggregation problem in discrete time. Another working paper, Tchistyi (2006), maintains the Markov structure in Fernandes and Phelan (2000) and devises a transformation of the agent continuation payoffs under deviations to obtain sufficiency of local deviations in the presence of unobservable utility shocks to the agent. Williams (2011) focuses on dynamic adverse selection – the agent only reports his income, which privately follows a Brownian motion. Bergemann and Hege (2005) and Bonatti and Hörner (2011) consider the case that surplus is private history dependent but the principal only cares about the first success. This simplifies the agency problem, and restores the one-shot-deviation principle as there is only a single instance in which rewards need to be provided. As a result, more involved questions – alternative contracting frameworks and collaboration between multiple agents can be studied.

The paper's contribution to this literature is twofold. First, it provides new and positive results for an important setting. Second, it introduces an approach that separates the different agency frictions and may be useful to other applications.

Related dual approaches to dynamic problems have been suggested before in other contexts. Benveniste and Scheinkman (1982) consider a macro-economic equilibrium model in continuous time and the classic text by Rockafellar (1997) considers several examples that suggest the methods introduced here should apply to standard concave (non-linear) moral hazard models. Both of these texts predate the dynamic moral hazard analysis. More recently, Vohra (2011) extends the analysis of static adverse selection models by analyzing the dual of the classic adverse selection problem. Among the many new results provided by Vohra (2011), the closest to ours regard the sufficiency of local deviations with respect to the agent's type. Vohra's analysis considers a much richer type and allocation space than ours, but does not consider moral hazard or long term contracts.

2. Model

2.1. **The Setting.**

There is a principal and an agent, both risk neutral. Both have an outside option set to zero. The agent has limited liability – i.e. money can only be transferred to the agent. Time is discrete. In each period the agent either works or not. The agent's work is costly to the agent and unobservable to the principal. The cost of effort in a period is c_n for a commonly known function $c : N \to R_+$ where n denotes the number of actual periods of work. That is, for the first period of work, n is one. For the second period of work, n is two, and so on. If the agent shirks in the first period, n in the second period is still one. The analysis will focus on the case that c_n is an increasing and convex function. However, most parts of the analysis, and specifically the derivation and analysis of the dynamic problem, apply to any general c_n. In particular, the methodology developed can be used to characterize the optimal contract if c_n is fixed or decreasing.

Assumption 1. *c_n is increasing and convex*

A period's production outcome is either success or failure, denoted by $y \in Y = \{s, f\}$. The principal earns a revenue of v from each success and zero from a failure. The probability of the outcome s (resp. f) in a period in which the agent works is $p \in (0,1)$ (resp. $1-p$). If the agent does not work, p is replaced with $p_0 \in [0,p)$. To prevent the principal from making free profits, assume the principal incurs a cost of $v \cdot p_0$ for every period in which the contract is still active.[8]

As costs are increasing, the surplus from working becomes negative after enough effort was exerted. Let N^{FB} denote the maximum number of periods in which consecutive work increases surplus:

$$N^{FB} = \max n : c_n \leq v\left(p - p_0\right) \ .$$

The increase in costs is sufficient to prevent an infinite contract from being optimal. Therefore, the analysis is simplified without loss by assuming that the agent and principal do not discount the future. Appendix B.10 adds a common discount factor and shows the effect amounts to a simple accounting exercise.

By the revelation principle, there is no loss in considering only contracts that specify for each period a work decision and a wage based on the period's history. Before defining the contract, the following result simplifies the exposition and notation. Given the risk-neutrality and no-discounting assumptions, the proof is straightforward and omitted.

Lemma 1. *There is an optimal contract in which:*

(1) *The agent works for at most N^{FB} periods*
(2) *The required work decision is a stopping decision: if the agent is ever asked not to work, the contract terminates.*
(3) *The agent is never paid in a period without work or with outcome f.*

The dynamic dual formulation that will be developed is simplified by the finiteness of the contract but can be easily extended to accommodate infinite horizon models with discounting.

Given lemma 1, the space of (payoff or contract relevant) pubic histories H is the space of previous outcomes:

$$H = \bigcup_{n=0}^{N^{FB}} Y^n \quad .$$

A public history $h \in H$ denotes a sequence of outcomes. Let n_h denote the number of the next period given h (i.e. the number of previous periods plus one). The agent's private information is, for every past period, whether he did actually work. The agent *deviated* in a period if he did not work.

As the cost to the agent of working in a period is a function of the *number* of periods in which the agent actually worked in the past, the only information in the agent's private history that is payoff relevant is the number of past deviations:

Definition 1. The agent's *private* history (h,d) is the public history h and the number of past deviations d.

[8] This assumption only simplifies the exposition and is without loss of generality.

Cost depends on the *private* history. With a slight abuse of notation let $c_{h-d} \equiv c_{n_h-d}$ denote the cost for any history h with past deviations d and $c_h \equiv c_{h-0}$. As the difference in cost between two work periods will play an important role, let $\delta_{h-d} \equiv c_{n_h-d} - c_{n_h-d-1}$ denote the cost difference between the current and previous periods if the agent shirked d times in the past and $\delta_h \equiv \delta_{h-0}$. To simplify the notation later on, set $\delta_1 = c_1$.

The analysis makes extensive use of histories *following* and *preceding* other histories. Let $h = \langle h^1, h^2 \rangle$ denote the history h^1 followed by the history h^2. That is, the sequence of outcomes h^1 happened and then the sequence h^2 happened. For example, if the current history is h, then the next history will be either $\langle h, s \rangle$ or $\langle h, f \rangle$. Say that the history $\langle h^1, h^2 \rangle$ *follows* history h^1 and denote the "follows" relation by $\succeq$. That is

$$\tilde{h} \succeq h \quad \Longleftrightarrow \quad \exists \hat{h} \in H : \tilde{h} = \left\langle h, \hat{h} \right\rangle .$$

The set H also includes the empty set and thus $h \succeq h$.

2.2. **The Contract.**

The contract specifies in each period whether the agent works and the resulting wage. Lemma 1 allows simplifying further and considering the work decision as a stopping rule. This section defines the contract using a simple transformation of the standard decision variables. This transformation will allow a linear formulation of the problem without affecting the interpretation of the resulting contract.[9]

Let $(1 - \alpha_h)$ denote the probability that the contract is terminated in history h, if h is reached. Rather than specifying the contract in terms of the conditional stopping rule α, the contract specifies the *ex-ante* probability that a complying agent will reach history h and will be asked to work in the history. This is denoted q_h, and may be defined recursively from any α:

$$q_\emptyset = \alpha_\emptyset \tag{2.1}$$

$$\begin{aligned} q_{\langle h,s \rangle} &= q_h \cdot p \cdot \alpha_{\langle h,s \rangle} \\ q_{\langle h,f \rangle} &= q_h \cdot (1-p) \cdot \alpha_{\langle h,f \rangle} . \end{aligned} \tag{2.2}$$

To define the payment, let W_h denote the payment to the agent if he is asked to work in history h and succeeds. By lemma 1 the payment to the agent in history h is zero if he is either not asked to work or fails. Thus there is no loss of generality in having the contract specify the ex-ante expected wage for success in the history:

$$w_h = q_h \cdot W_h \ . \tag{2.3}$$

Definition 2. A contract is a pair of functions $\langle q, w \rangle$, with $q : H \to [0,1]$ specifying for each history h the ex-ante probability that the agent will be asked to work in the period and $w : H \to R_+$ specifying the ex-ante expected payment to the agent for a success in history h.

As the agent and principal are risk neutral, there are infinitely many equivalent ways for the optimal contract to pay the agent w dollars. If the probability of success in the next attempt is p, paying a dollar today is equivalent, on equilibrium, to paying $\frac{1}{p}$ more dollars tomorrow only if the agent succeeds tomorrow. To remove this technical duplication of the optimal contracts, the analysis assumes the optimal contract makes

[9]The use of a stopping rule simplifies the linear transformation, but is not necessary.

payments "as early as possible." That is, from all contracts that specify the same work plan, the optimal contract is the one that pays more to the agent as early as possible. This selection rule is justified if the agent is just slightly more impatient or risk averse than the principal and allows the analysis to focus on a unique equilibrium.[10]

A special type of contract will play an important role in the solution:

Definition 3. A contract is linear *to N from history h* if once the contract reaches history h the agent works exactly until period $N \leq N^{FB}$ regardless of new outcomes and is paid $\frac{c_N}{p-p_0}$ for success in all remaining periods.

The wage $\frac{c_N}{p-p_0}$ is optimal in a single period game with cost c_N. Thus if the agent faces a linear contract to N from h, he gets paid in all remaining periods the static optimal for period N. Lemma 11 shows that the optimal contract becomes linear after the first payment is made.

2.3. **The Optimal Contract Problem.**

If the agent complies with the contract (i.e. works when asked to), the ex-ante expected revenue for the principal from a history h is the probability that the history is reached and the contract did not terminate, q_h, multiplied by the expected revenue from work $v \cdot (p - p_0)$. The ex-ante expected payment in a period h for a compliant agent is simply $p \cdot w_h$. Thus, the principal's expected profit from a contract the agent complies with is

$$V(q, w) = \sum_{h \in H} \left[q_h (p - p_0) v - w_h p \right] . \tag{2.4}$$

The agent's strategy specifies, for each private history, the probability that the agent works when asked to. Given lemma 1, there is no loss in considering an agent's *effort plan*, $e \in E$, with a typical element $e_h \in \{0, 1\}$ that specifies the agent's pure action at each *public* history h if the agent is asked to work in the history. As the agent is risk neutral, given lemma 1 there is no loss in assuming the agent does not mix. As the agent can fully reconstruct the private history h, d at any period using the past effort plan and public history h, letting e be only a function of the public history is without loss of generality.

Let $U(h, d; q, w, e)$ denote the agent's expected continuation profit from effort plan e, starting from private history h, d , having been asked to work in the history and facing the contract q, w.[11] Let $U^c(h, d; q, w)$ be the agent's expected continuation utility from complying with the contract in all remaining periods, starting at private history h, d. Note that it is not required that history h, d could be reached if the agent follows the plan e from the start of the relationship. Specifically $U^c(\cdot)$ is well defined even if the agent deviated in the past ($d > 0$).

Let $\emptyset$ denote the starting (null) history. The optimal contract problem is:

---

[10] The early payments condition implies that the payments depend on fewer outcomes. As there is no issue of inter-temporal consumption smoothing this reduces the variation in the agent's compensation and thus increases his expected utility if he is risk averse.

[11] If the agent is not asked to work in the history, his continuation profit is exactly zero.

(2.5)

$$
\begin{aligned}
V^* = \max_{q\geq 0, w\geq 0} \quad & \textstyle\sum_{h\in H}\left[q_h\left(p-p_0\right)v - w_h p\right] & \\
& q_{\emptyset} \leq 1 & \text{(Probability at } h=\emptyset) \\
\forall h \quad & q_{\langle h,s\rangle} \leq q_h p & \text{(Probability after success)} \\
\forall h \quad & q_{\langle h,f\rangle} \leq q_h\left(1-p\right) & \text{(Probability after failure)} \\
\forall h, e \quad & U^c\left(h,0;q,w\right) \geq U\left(h,0;q,w,e\right) & \text{(IC)} \\
\forall h \quad & U^c\left(h,0;q,w\right) \geq 0 & \text{(IR)}
\end{aligned}
$$

The first three constraints follow directly from the definition of q in equation 2.1. The analysis will refer to these as the "Probability" constraints as they reflect the upper bound on the probability of work. As any plan e may detail several deviations, there is no loss in restricting the IC to histories in which the agent never shirked.[12] The IR constraint is presented for completeness. However, as the agent always has a possible effort plan – never work – that provides non-negative expected payoff, it is redundant given IC and thus will be subsequently ignored.

It can be shown that specifying the contract in ex-ante terms using q_h and w_h rather than in conditional terms (the α_h and W_h used in deriving q_h and w_h) allows writing the IC constraint for every possible alternative effort plan e as a linear inequality. Thus, problem 2.5 is a *linear program.* As in any linear program, the difficulty lies in identifying the binding set of constraints. Specifically, the set of IC is too large and the main problem is to identify the relevant subset to consider. This will be the set of Local Deviation Incentive Constraints (LDIC).

To define the local deviations constraints, let $U^D\left(h,d;q,w\right)$ be the agent's expected continuation profit starting at a period with private history h,d if the agent *deviates* in the current period and complies with the contract in all following periods.[13]

Definition 4. The set of Local Deviation Incentive Constraints (LDIC) is

$$\forall h: \quad U^c\left(h,0;q,w\right) \quad \geq \quad U^D\left(h,0;q,w\right) \quad \text{(LDIC)} \tag{2.6}$$

The set LDIC is a clear relaxation of the IC as it limits the agent to a single deviation. The problem would be simpler if the set LDIC was sufficient to imply IC. However, not all contracts that satisfy LDIC are IC.[14] Instead, I will show that the set of *optimal* solutions to the problem considering only LDIC does satisfy all IC and thus there is no loss in considering the relaxed problem. To show this, I first define a larger set of constraints, the Final Deviations Incentive Constraints (FDIC). These constraints require that there is no private history h,d such that the agent prefers deviating in history h,d and complying with the contract in all remaining periods to complying in history h,d and all remaining periods.

[12]As the agent is risk neutral, the IC and IR constraint may be written for only $h=\emptyset$. However, this does not simplify the problem and writing the IC for all histories simplifies the analogy to the alternative problems developed below.
[13]Formally, $U^D\left(h,d;q,w\right)$ is defined only for $d\leq n_h$.
[14]Appendix A provides an example.

Definition 5. The set of Final Deviations Incentive Constraints IC (FDIC) is

$$\forall h,d: \quad U^{c}(h,d;q,w) \quad \geq \quad U^{D}(h,d;q,w) \quad \text{(FDIC)}$$

As a first step, the next lemma shows that FDIC is a stricter set of constraints than IC. Intuitively, any most profitable non-compliant plan must have some period in which it is profitable to make a last deviation and thus violate some FDIC. The proof is standard and relegated to appendix B.1.

Lemma 2. *If a contract is FDIC it is IC*

As FDIC is stricter than IC and LDIC is weaker than IC, the following simple result provides a criterion for the sufficiency of the LDIC.

Corollary 1. *If any optimal contract subject to LDIC satisfies FDIC, then any optimal contract subject to LDIC is optimal subject to IC.*

2.4. **Deriving the FDIC and LDIC.**

This section derives all FDIC as linear inequalities. The LDIC are the subset of FDIC with $d=0$. As the contract terms $\langle q,w\rangle$ are identical for all continuation utilities they are omitted from the parametrization of $U^{D}(\cdot)$ and $U^{C}(\cdot)$.

The FDIC in history h,d requires that the agent's expected continuation utility from following the contract at history h,d is at least his expected continuation utility from making a final deviation in the history. If the agent will never be asked to work in history h ($q_h=0$) then his expected continuation utility is zero for all continuation effort plans and the FDIC trivially holds at h for all d.

For $q_h>0$, working has three effects. First, at the current period, work costs the agent c_{h-d} and increases the expected payment from $p\cdot W_h$ to $p_0\cdot W_h$. Second, working affects the transition probabilities into the next public history. If the agent shirks the probability of moving to the history $\langle h,s\rangle$ is p_0 instead of p , and similarly, for $\langle h,f\rangle$, the probability $(1-p)$ is replaced by $(1-p_0)$. Finally, work today increases the agent's costs in all continuation periods. If the agent would shirk, his cost in all future periods would be lower. The FDIC requires that the total of these effects will be positive. Then the FDIC for private history (h,d) is:

$$\begin{gathered}(p-p_0)\left(W_h+U^{c}(\langle h,s\rangle,d)-U^{c}(\langle h,f\rangle,d)\right)-c_{h-d}\geq \\ p_0\left(U^{c}(\langle h,s\rangle,d+1)-U^{c}(\langle h,s\rangle,d)\right)+(1-p_0)\left(U^{c}(\langle h,f\rangle,d+1)-U^{c}(\langle h,f\rangle,d)\right)\end{gathered} \tag{2.7}$$

As the agent is risk neutral, and $q_h>0$, the FDIC can be evaluated at the outset of the contract by multiplying all elements by q_h:[15]

$$\begin{aligned}&(p-p_0)\,w_h-q_h c_{h-d}\\ &+(p-p_0)\left(\tfrac{q_{\langle h,s\rangle}}{p}U^{c}(\langle h,s\rangle,d)-\tfrac{q_{\langle h,f\rangle}}{1-p}U^{c}(\langle h,f\rangle,d)\right)\\ &+p_0\tfrac{q_{\langle h,s\rangle}}{p}\left(U^{c}(\langle h,s\rangle,d)-U^{c}(\langle h,s\rangle,d+1)\right)\\ &+(1-p_0)\tfrac{q_{\langle h,f\rangle}}{1-p}\left(U^{c}(\langle h,f\rangle,d)-U^{c}(\langle h,f\rangle,d+1)\right)\quad\geq 0\end{aligned} \tag{2.8}$$

The first line is the simple static tradeoff (recall that $w_h=W_hq_h$). The second line is the incentive effect of continuation utilities, evaluated in ex-ante terms and ignoring the change in costs. I will call this the "naive" dynamic incentive. The third and fourth lines

[15]Appendix B.2 provides a detailed derivation.

are the expected losses due to higher future costs from working, adjusted for the correct continuation probabilities. I will call this the "dynamic information rent".

All continuation utility terms, $U^c(\cdot)$, in the FDIC (2.8) are ex-ante continuation utilities from *complying* with the contract starting at some history $\langle h,y\rangle$ ($\langle h,s\rangle$ or $\langle h,f\rangle$) with some private number of deviations d. Thus, they may be directly defined using the model's primitives. For any history $\langle h,y\rangle$, the continuation utility is simply the expected wages less costs:

$$q_{\langle h,y\rangle}\cdot U^c(\langle h,y\rangle,d) = \sum_{\tilde{h}\succeq\langle h,y\rangle} pw_{\tilde{h}} - \sum_{\tilde{h}\succeq\langle h,y\rangle} q_{\tilde{h}}c_{\tilde{h}-d} \quad . \tag{2.9}$$

Note that only the expected costs depend on the private information d, but the expected payment is unaffected by d. This implies that the information rent terms (the last two lines of 2.8), are determined only through the lower costs:

$$\begin{aligned} q_{\langle h,y\rangle}\left(U^c(\langle h,y\rangle,d+1) - U^c(\langle h,y\rangle,d)\right) &= \sum_{\tilde{h}\succeq\langle h,y\rangle} q_{\tilde{h}}\left(c_{\tilde{h}-d} - c_{\tilde{h}-d-1}\right) \\ &= \sum_{\tilde{h}\succeq\langle h,y\rangle} q_{\tilde{h}}\delta_{\tilde{h}-d} \end{aligned} \tag{2.10}$$

Using (2.9) and (2.10) in (2.8) and rearranging to prepare for the dual yields the final form of the FDIC:

$$\begin{aligned} FDIC: \quad & -(p-p_0)\,w_h + q_h c_{h-d} \\ & -\tfrac{p-p_0}{p}\textstyle\sum_{\tilde{h}\succeq\langle h,s\rangle} pw_{\tilde{h}} + \tfrac{p-p_0}{p}\sum_{\tilde{h}\succeq\langle h,s\rangle} q_{\tilde{h}}\cdot c_{\tilde{h}-d} \\ & +\tfrac{p-p_0}{1-p}\textstyle\sum_{\tilde{h}\succeq\langle h,f\rangle} pw_{\tilde{h}} - \tfrac{p-p_0}{1-p}\sum_{\tilde{h}\succeq\langle h,f\rangle} q_{\tilde{h}}c_{\tilde{h}-d} \\ & +\tfrac{p_0}{p}\textstyle\sum_{\tilde{h}\succeq\langle h,s\rangle} q_{\tilde{h}}\cdot\delta_{\tilde{h}-d} \\ & +\tfrac{1-p_0}{1-p}\textstyle\sum_{\tilde{h}\succeq\langle h,f\rangle} q_{\tilde{h}}\cdot\delta_{\tilde{h}-d} \quad \leq 0 \end{aligned} \tag{2.11}$$

The FDIC is linear in the q and w variables. The complete linear program 2.5 subject to the FDIC (2.11) is:

$$\begin{aligned} V^{FD} = \max_{q\geq 0, w\geq 0} \quad & \textstyle\sum_{h\in H}\left[q_h(p-p_0)\,v - w_h p\right] \\ s.t. \quad & \\ & q_{h_0} \leq 1 \\ \forall h \quad & q_{\langle h,s\rangle} - q_h p \leq 0 \\ \forall h \quad & q_{\langle h,f\rangle} - q_h(1-p) \leq 0 \\ \forall h,d \quad & FDIC \quad (2.11) \end{aligned} \tag{2.12}$$

The Local-Deviations Problem is problem 2.12 with the FDIC limited to $d=0$.

Before deriving the dual, the following lemma establishes a useful result. Recall definition (3) of a linear contract – the agent works from history h to period N regardless of new outcomes and is paid $\frac{c_N}{p-p_0}$ for success in all remaining periods.

Define a work plan as *fixed* if it can be part of a linear contract – the agent works exactly until period N, regardless of outcomes.

Lemma 3. *If the work plan is fixed to N from history h, the contract that minimizes the agent's continuation utility is linear to N from history h; that is, the agent is paid $\frac{c_N}{p-p_0}$ for success in all remaining histories. All remaining LDIC bind and all remaining FDIC for $d>0$ are slack.*

Proof. Suppose the work plan is fixed from h to N. As the information rent depends only on the work plan the agent's gain from shirking is given by (see equation 2.10):

$$q_{\langle h,y\rangle}U^c\left(\langle h,y\rangle,d+1\right)-q_{\langle h,y\rangle}U^c\left(\langle h,y\rangle,d+1\right)=$$
$$\textstyle\sum_{n=n_h+1-d}^{N-d}\left(c_{n-d}-c_{n-d-1}\right)=c_{N-d}-c_{n_h-d}\ .$$

A fixed work plan also implies that any difference in continuation utilities between $\langle h,s\rangle$ and $\langle h,f\rangle$ depends only on wages. Because the agent is risk neutral, any incentive compatible future wage difference can be incorporated into the current period wage. Thus there is an optimal contract such that if the work plan is fixed,

$$U^c\left(\langle h,s\rangle;q,w\right)=U^c\left(\langle h,f\rangle;q,w\right)$$

Finally, if the work plan is fixed and the contract is not terminated at the end of the period, $q_{\langle h,s\rangle}=pq_h$ and $q_{\langle h,f\rangle}=(1-p)\,q_{\langle h,y\rangle}$. The FDIC in period h is thus simply

$$\left(p-p_0\right)w_h-q_hc_{h-d}-q_h\left(p_0+\left(1-p_0\right)\right)\left(c_{N-d}-c_{h-d}\right)\quad\geq\quad 0\ ,$$

which simplifies to

$$\frac{w_h}{q_h}\geq\frac{c_{N-d}}{p-p_0}\ .$$

Only the cost in the last period of work matters. By definition of w_h, the payment to the agent for success in history h is exactly $\frac{w_h}{q_h}$. Observe that the constraint for w_h *does not* depend on any future wage. As c_{n-d} decreases in d, all but the LDIC (in which $d=0$) must be slack. The contract that minimizes the agent's continuation utility minimizes the total payments. Thus, all LDIC bind and the payment to the agent for success in history h is

$$W_h=\frac{w_h}{q_h}=\frac{c_N}{p-p_0}\ .$$

□

Remark 1. The wage plan for a fixed work plan is the only result that is qualitatively affected by the no-discounting assumption. Appendix B.10 shows that if the principal and the agent have a common discount factor β, then a utility minimizing wage plan for a fixed work plan is given by:

$$\begin{aligned}W_{N(r)}&=\frac{c_{N(r)}}{p-p_0}\\W_n&=\beta W_{n+1}+(1-\beta)\frac{c_n}{p-p_0}\ .\end{aligned}$$

3. The Dual Dynamic Analysis

Dynamic moral hazard analysis typically proceeds by restating the Optimal Contract Problem 2.5 (or its Local/Finite Deviations alternatives) as a dynamic program. The agent's continuation utility from complying with the contract is the state for period h and the incentive constraint is written in terms of the difference between the continuation utility after success and after failure. In the case of history-dependent costs, the agent's utility if he deviated once in the past is added as a control (see e.g. Fernandes and Phelan (2000)).

Instead, I derive a dynamic program for the *dual* of the LDIC and FDIC optimal contract problems.[16] The dynamic dual program allows a relatively straightforward proof of the sufficiency of local deviations. It is sufficient to show that the agency frictions in any period are largest if the agent never shirked in the past. The dual LDIC program also allows for a stronger characterization of the optimal contract compared to the standard LDIC analysis.

To ease the economic interpretation, the derivation and analysis of the dual is done in steps. This section derives and analyzes the dual of the local deviations problem. Section 4 adds to the dual all possible final deviations and proves that the local deviations problem is sufficient. Therefore, section 4 proves that the analysis of this section describes the optimal contract.

3.1. **The LDIC Dual – Two Period Derivation.**

The derivation of the dual requires intensive notation. To make the reasoning more explicit, the analysis starts with a two period example and then applies an intuitive argument to add more periods. A formal derivation of the multi-period dynamic dual problem is provided in appendix B.4.

If there are only two periods, the possible public histories are $h \in \{\emptyset, s, f\}$. Let μ^h for $h \in \{\emptyset, s, f\}$ be the shadow variable on the probability constraints and λ^h the shadow variable on the LDIC. These are specified next to each of their corresponding constraints in the following two period LDIC primal problem:

$$\max_{(q_h, w_h)\geq 0} \left(q_\emptyset + q_s + q_f\right) v \left(p - p_0\right) - p \left(w_0 + w_f + w_s\right) \tag{3.1}$$

$s.t.$

$$\begin{aligned}
q_\emptyset &\leq 1 && \mu^\emptyset \\
q_s - p q_\emptyset &\leq 0 && \mu^s \\
q_f - (1-p)\, q_\emptyset &\leq 0 && \mu^f \\
-\left(p - p_0\right) w_\emptyset + q_\emptyset c_1 & && \lambda^\emptyset \\
-\tfrac{p-p_0}{p} p w_s + \tfrac{p-p_0}{p} q_s c_2 & && \\
+\tfrac{p-p_0}{1-p} p w_f - \tfrac{p-p_0}{1-p} q_f c_2 & && \\
+\tfrac{p_0}{p} q_s \delta_2 + \tfrac{1-p_0}{1-p} q_f \delta_2 &\leq 0 && \\
-\left(p - p_0\right) w_s + q_s c_2 &\leq 0 && \lambda^s \\
-\left(p - p_0\right) w_f + q_f c_2 &\leq 0 && \lambda^f
\end{aligned}$$

To fix ideas, I first mechanically derive the dual of problem 3.1. As only the left-hand side of the probability constraint for $\emptyset$ is not zero, the objective for the dual is

$$\min \mu^\emptyset$$

[16]See appendix B.3 for a review of the basic duality concepts.

Each primal variable defines a dual constraint. Consider first the wage constraint for $w_\emptyset$. $w_\emptyset$ appears with a coefficient $-p$ in the objective and with a coefficient $-(p-p_0)$ in the LDIC for $h=\emptyset$. The dual constraint is therefore:

$$-(p-p_0)\,\lambda^\emptyset \geq -p\ .$$

The wage w_s appears in the LDICs for histories $\emptyset$ and s, with a coefficient $-(p-p_0)$ in both. The constraint is thus:

$$-(p-p_0)\left(\lambda^s+\lambda^\emptyset\right) \geq -p\ .$$

The wage w_f appears in the LDICs for histories $\emptyset$ and f . The coefficient in the first is $-(p-p_0)$ and on the second $\frac{p-p_0}{1-p}p$. The constraint is thus

$$-(p-p_0)\left(\lambda^f-\lambda^\emptyset\frac{p}{1-p}\right) \geq -p\ .$$

The probability $q_\emptyset$ appears in all the probability constraints and in the LDIC for $\emptyset$. The constraint for $q_\emptyset$ is:

$$\mu^\emptyset - p\mu^s - (1-p)\,\mu^f + \lambda^\emptyset c_1 \quad \geq \quad v\,(p-p_0)\ .$$

The probability q_s appears in its probability constraint, in the two FDIC for s, and twice in the FDIC for the preceding history $\emptyset$ – once for the continuation utility term and once for the shirking gains term. The constraint for q_s is

$$\mu^s + \frac{p-p_0}{p}c_2\lambda^\emptyset + \frac{p_0}{p}\delta_2\lambda^\emptyset + \lambda^s c_2 \geq v\,(p-p_0)\ .$$

The same procedure yields the probability constraint for q_f:

$$\mu^f - \frac{p-p_0}{1-p}c_2\lambda^\emptyset + \frac{1-p_0}{1-p}\delta_2\lambda^\emptyset + \lambda^f c_2 \geq v\,(p-p_0)$$

It will be convenient to rearrange the constraints. The dual to be analyzed is

(3.2) $\min_{(\mu,\lambda)\geq 0}\mu^\emptyset$

s.t.

$$\begin{array}{lcl}
(q_\emptyset): & \mu^\emptyset & \geq p\mu^s + (1-p)\,\mu^f + v\,(p-p_0) - \lambda^\emptyset c_1 \\
(w_\emptyset): & \lambda^\emptyset\,(p-p_0) & \leq p \\
 & & \\
(q_s): & \mu^s & \geq v\,(p-p_0) - \dfrac{p-p_0}{p}c_2\lambda^\emptyset - \dfrac{p_0}{p}\delta_2\lambda^\emptyset - \lambda^s c_2 \\
(w_s): & (p-p_0)\left(\lambda^s+\lambda^\emptyset\right) & \leq p \\
 & & \\
(q_f): & \mu^f & \geq v\,(p-p_0) + \dfrac{p-p_0}{1-p}c_2\lambda^\emptyset - \dfrac{1-p_0}{1-p}\delta_2\lambda^\emptyset - \lambda^f c_2 \\
(w_f): & (p-p_0)\left(\lambda^f-\lambda^\emptyset\frac{p}{1-p}\right) & \leq p
\end{array}$$

3.2. **A First Interpretation of the Dual.**

To interpret the dual problem, consider first the variable μ^h. As μ^h is the shadow price on the primal constraint on q_h, it reflects how much the principal is "willing to pay" to increase q_h by a little more than is feasible. Recall that q_h is the ex-ante committed probability that the agent will be asked to work in history h. Intuitively, the commitment for each period should either increase or decrease the principal's ex-ante expected profit. If the commitment decreases profits, the optimal contract should set $q_h = 0$. The primal constraint on q_h will not bind and $\mu^h = 0$. If the commitment increases profits, increasing q_h marginally will increase the expected profits by exactly the same amount. Thus $\mu^h > 0$ is the increase in expected profits from committing to ask the agent to work in history h. Whenever $\mu^h > 0$, it must be that q_h is set to the highest possible value.

$\mu^\emptyset$ is therefore the ex-ante expected profit from the contract. The min operator in the dual problem "limits" the profit to the feasible value. This is the right hand side of each dual q_h constraint. Consider first the $q_\emptyset$ constraint. The first two elements are the continuation profits $p\mu^s + (1-p)\,\mu^f$. The feasible continuation profit starting today is the feasible profit today plus any feasible profits tomorrow. The next element is the revenue from work $v\,(p - p_0)$. However, work is not free. Recall that λ^h is the shadow price of the IC in history h. For the principal, λ^h is therefore the marginal cost of the agent's effort at history h. If the agent would be asked to work a little less, the IC would be relaxed by c_h. In the $q_\emptyset$ dual constraint, this is the term $-\lambda^\emptyset c_1$. In the other dual q_h constraints, this is the $\left(-\lambda^h c_h\right)$ term at the end of each constraint.

The dual profit in history s, μ^s, is given by the right hand side of the dual q_s constraint. μ^s differs from $\mu^\emptyset$ in several ways. As $'s'$ is a last period, there are no continuation terms. However, there are additional terms in μ^s. Consider first the term $\left(-\frac{p_0}{p}\delta_2\lambda^\emptyset\right)$. By committing to require work in history s, the principal increases the agent's incentives to shirk in the first period in order to obtain lower costs. The term $\frac{p_0}{p}\delta_2$ is exactly the additional gain to the agent. It is multiplied by $\left(-\lambda^\emptyset\right)$, the shadow price on the first period's IC. The second term, $\left(-\frac{p-p_0}{p}c_2\lambda^\emptyset\right)$ reflects the fact that the agent knows he will comply in the future. Thus, at the first period, the agent considers also the additional cost of working in the second period c_2, multiplied by the increased probability of getting to history s from work in the first period: $\frac{p-p_0}{p}$. Considering known results in dynamic moral hazard, it may seem counter intuitive that working after success requires stronger incentives to work in the first period. However, the agent's incentives to work in the first period are relaxed because of the reward for success in the second period, and not because of the work that is required to generate it. The same analysis results in the dual profit μ^f.

It is left to interpret the incentive costs λ^h. At the first period, the principal can relax the IC constraint by $(p - p_0)$ by increasing the agent's wage $w_\emptyset$. The marginal cost of increasing $w_\emptyset$ is the probability of payment (success), p. The dual $w_\emptyset$ constraint requires that the benefit from paying the agent, $\lambda^\emptyset\,(p - p_0)$, never exceeds the cost. In any other history, as the variable w_h already accounts for the probability of arriving at the history, the cost is still p. Paying the agent in the s history, relaxes the λ^s IC by $(p - p_0)$ and also the $\lambda^\emptyset$ IC by the same factor. In contrast, paying the agent in the history f relaxes the λ^f IC but increases the profitability of failing in the first period. Adjusting for the relative probabilities, an increase in w_f costs the principal an increase in $\frac{p}{1-p}\,(p - p_0)$ in first period incentives, and thus the term $\frac{p}{1-p}\,(p - p_0)\cdot\left(-\lambda^\emptyset\right)$ in the dual w_f constraint. If no work is required in history f, no incentives would be provided and $w_f = 0$. This captures the intuition that future work after success increases incentive to work in the first period – not because of the expectation of more work, but in anticipation of higher wages.

3.3. Dynamic Formulation.

The interpretation of μ^h lends itself to a dynamic formulation. Technically, for every $\lambda^\emptyset$, the objective for 3.2 is minimized by minimizing the weighted average of μ^s and μ^f. In turn, the dual value μ^s (and similarly μ^f) depends only on the shadow price λ^s (respectively λ^f) and on the effect on past incentives, here through the past shadow price $\lambda^\emptyset$. Isolating μ^s in the constraint for q_s , any optimal μ^s must be the solution to the problem

$$\mu^s\left(\lambda^\emptyset\right) = \max \quad \left[0, \min_{\lambda^s\geq 0} v\left(p-p_0\right) - \tfrac{p-p_0}{p}c_2\lambda^\emptyset - \tfrac{p_0}{p}\delta_2\lambda^\emptyset - \lambda^s c_2\right]$$
$$s.t.$$
$$\left(p-p_0\right)\left(\lambda^s+\lambda^\emptyset\right)\leq p$$

The problem for μ^f only differs in the coefficients for $\lambda^\emptyset$ (the use of λ^f instead of λ^s is semantic):

$$\mu^f\left(\lambda^\emptyset\right) = \max \quad \left[0, \min_{\lambda^f\geq 0} v\left(p-p_0\right) + \tfrac{p-p_0}{1-p}c_2\lambda^\emptyset - \tfrac{1-p_0}{1-p}\delta_2\lambda^\emptyset - \lambda^f c_2\right]$$
$$s.t.$$
$$\left(p-p_0\right)\left(\lambda^f-\tfrac{p}{1-p}\lambda^\emptyset\right)\leq p$$

It would be convenient to solve the same problem for both second period histories, as a function of the past. For that, define $r^s\left(\lambda^\emptyset\right)\equiv\frac{p_0}{p}\lambda^\emptyset$, $r^f\left(\lambda^\emptyset\right)\equiv\frac{1-p_0}{1-p}\lambda^\emptyset$, $i^s\left(\lambda^\emptyset\right)\equiv -\frac{p-p_0}{p}\lambda^\emptyset$, and $i^f\left(\lambda^\emptyset\right)\equiv\frac{p-p_0}{1-p}\lambda^\emptyset$.

The second period problem becomes

(3.3)
$$\mu^h=\mu\left(n=2,i^h,r^h\right)=\max \quad \left[0,\min_{\lambda\geq 0} v\left(p-p_0\right)+ic_n-r\delta_n-\lambda c_n\right]$$
$$s.t. \qquad \left(p-p_0\right)\lambda - i\cdot p\leq p$$

The complete two period dual problem 3.2 can therefore be written recursively as:

$$\mu^\emptyset = \max\left[0,\min_{\lambda\geq 0} v\left(p-p_0\right)-\lambda c_1+p\mu\left(2,i^s,r^s\right)+\left(1-p\right)\mu\left(2,i^f,r^f\right)\right]$$
$$s.t. \qquad \left(p-p_0\right)\lambda\leq p$$
$$i^s=-\tfrac{p-p_0}{p}\lambda \quad ; \quad i^f=\tfrac{p-p_0}{1-p}\lambda$$
$$r^s=\tfrac{p_0}{p}\lambda \quad ; \quad r^f=\tfrac{1-p_0}{1-p}\lambda$$

If the dynamic structure is preserved when adding periods, one can suggest the following formulation:

Definition 6. The LDIC Dynamic Dual is

(3.4)

$$\mu(n,i,r) = \max\left[0, \min_{\lambda\geq 0}\hat{\mu}(n,i,r,\lambda)\right]$$

$$s.t.$$

$$\text{(profit definition)} \quad \hat{\mu}(n,i,r,\lambda) = p\mu(n+1,i^s,r^s) + (1-p)\mu\left(n+1,i^f,r^r\right) + v(p-p_0) - c_n\lambda + c_n l - \delta_n r$$

$$\text{(wage constraint)} \quad \lambda \leq \frac{p}{p-p_0}(1+i)$$

$$\text{(stopping condition)} \quad \hat{\mu}\left(N^{FB}+1,i,r,\lambda\right) = 0$$

$$\text{(law of motion, } i\text{)} \quad i^s = i - \frac{p-p_0}{p}\lambda \quad ; \quad i^f = i + \frac{p-p_0}{1-p}\lambda$$

$$\text{(law of motion, } r\text{)} \quad r^s = r + \frac{p_0}{p}\lambda \quad ; \quad r^f = r + \frac{1-p_0}{1-p}\lambda$$

Problem 3.4 is a dynamic program with three state variables, (n,i,r). n, the period counter is trivial. The law of motion for i and r replicates the effect of the first period shadow price in the second period problem in a general period setting. The wage constraint is the constraint associated with w_h in the primal problem for the history h. The main result of this section is that indeed problem 3.4 is a valid representation of the dual of the LDIC problem. I first state this result and then interpret it.

Lemma 4. *For every h, $\mu^h = \mu\left(n_h, i^h, r^h\right)$. In particular, $\mu(1,0,0)$is the value of the dual of the LDIC problem. $\mu(n,i,r)$ is continuous and convex in (l,r) for every n. $\hat{\mu}(n,i,r,\lambda)$ is continuous and convex in λ for every (n,i,r).*

Proof. A detailed proof is provided in appendix B.4. I sketch the main steps here.

(1) The LDIC problem is problem 2.12 with the FDIC limited to $d=0$ (LDIC).
(2) The dual of the LDIC problem is given by

(3.5) $$\min_{(\mu,\lambda)\geq 0} \mu^{\emptyset}$$

$s.t.\ \forall h:$

$$\lambda^h \leq \tfrac{p}{p-p_0}\left(1+i^h\right) \qquad \text{(wage } (w_h)\text{constraint)}$$

$$\mu^h \geq v(p-p_0) + (1-p)\mu^{\langle h,f\rangle} + p\mu^{\langle h,s\rangle} - c_h\lambda^h + c_h i^h - \delta_h r^h \qquad (q_h\text{constraint})$$

$$i^{\langle h,s\rangle} = i^h - \tfrac{p-p_0}{p}\lambda^h \quad ; \quad i^{\langle h,f\rangle} = i^h + \tfrac{p-p_0}{1-p}\lambda^h \qquad \text{(definition } l\text{)}$$

$$r^{\langle h,s\rangle} = r^h + \tfrac{p_0}{p}\lambda^h \quad ; \quad r^{\langle h,f\rangle} = r^h + \tfrac{1-p_0}{1-p}\lambda^h \qquad \text{(definition } r\text{)}$$

$$i^{\emptyset} = r^{\emptyset} = 0$$

(3) In the dual LDIC of the previous step, whenever $\mu^h > 0$ it must be that the corresponding dual q_h is an equality. If not, reducing μ^h is feasible and allows relaxing the previous history's q_h constraint. Continuously applying the backward reduction will decrease the objective $\mu^{\emptyset}$.
(4) It is now possible to apply the Principle of Optimality to obtain equivalence between problem 3.5 and problem 3.4 starting at $\mu(1,0,0)$.

□

A standard corollary of lemma 4 is:

Corollary 2. *For every state (n,i,r), the set of optimal λ is non-empty, convex and continuous in (l,r).*

3.4. **Interpreting the Dynamic Dual.**

The value of the dynamic dual at each state, $\mu(n,i,r)$, has the same interpretation as was provided for μ^h in the interpretation following the two period problem: the expected increase in the overall profits at the start of the relationship, from the optimal contract starting at history h. The Duality Theorem of Linear Programming implies:

Lemma 5. *$\mu(1,0,0)$ is the principal's expected profit from the optimal LDIC contract.*

Proof. $\mu(1,0,0)$ is the solution to the dual problem of the LDIC. It is immediate that $\mu(1,0,0)$ exists and is bounded. The Duality Theorem (Dantzig (1963)) applies. □

It is important to observe that in a general history $h \neq \emptyset$, $\mu\left(n_h, i^h, r^h\right)$, is *not* the expected continuation profit starting at the history. $\mu(\cdot)$ accounts for the effect of the current history on past incentives, while the continuation profit does not. An illustration of the difference is provided in section 5.

The interpretation of $\lambda(n,i,r)$ is the same as λ^h – the shadow price of the IC in the history that corresponds to the state (n,i,r). The dynamic dual however introduces two state variables, i and r. Each state variable summarizes a different implication of the shadow prices in all histories leading up to h:

Definition 7. The direct *incentive cost* in the public history h is i^h, defined by:

$$i^h = (p-p_0)\left[\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},f\rangle}\frac{\lambda^{\tilde{h}}}{1-p}\right)-\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},s\rangle}\frac{\lambda^{\tilde{h}}}{p}\right)\right]. \tag{3.6}$$

The *information rent* in the public history h is r^h, defined by:

$$r^h = \left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},s\rangle}\frac{p_0}{p}\lambda^{\tilde{h}}\right)+\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},f\rangle}\frac{1-p_0}{1-p}\lambda^{\tilde{h}}\right) \tag{3.7}$$

Consider the effect of giving the agent an extra util in history h, without changing the work plan.[17] For any history $\tilde{h}$ that h , the extra util will affect incentives through the expected continuation utility. The marginal effect on profits is proportional to the shadow cost $\lambda^{\tilde{h}}$. To determine how much will be saved, we must know how much the agent values the continuation utility at $\tilde{h}$. This is given by the second row of the LDIC. Thus, the cost of incentives at $\tilde{h}$ is reduced by $\frac{p-p_0}{p}\lambda^{\tilde{h}}$ if h follows a success in history $\tilde{h}$ and increased by $\frac{p-p_0}{1-p}\lambda^{\tilde{h}}$ if h follows a failure in history $\tilde{h}$.

Definition 8. If i^h is negative, increasing utility in the current period would in total relax the constraints in the past – allow a decrease in past incentives. If i^h is positive,

[17]Recall that the LDIC in a history $\tilde{h}$, given in (2.8), may be written as

$$\begin{aligned}
&-(p-p_0)\,w_{\tilde{h}}+q_{\tilde{h}}c_{\tilde{h}}\\
&+(p-p_0)\left(\frac{q_{\langle\tilde{h},f\rangle}U^c(\langle\tilde{h},f\rangle,0)}{1-p}-\frac{q_{\langle\tilde{h},s\rangle}U^c(\langle\tilde{h},s\rangle,0)}{p}\right)\\
&+\tfrac{p_0}{p}q_{\langle\tilde{h},s\rangle}\left(U^c\left(\left\langle\tilde{h},s\right\rangle,1\right)-U^c\left(\left\langle\tilde{h},s\right\rangle,0\right)\right)\\
&+\tfrac{1-p_0}{1-p}q_{\langle\tilde{h},f\rangle}\left(U^c\left(\left\langle\tilde{h},f\right\rangle,1\right)-U^c\left(\left\langle\tilde{h},f\right\rangle,0\right)\right)\qquad\leq 0
\end{aligned}$$

the utility provided to the agent today required an increase in past incentives. This intertemporal provision of utility and incentives is required because of the agent's limited liability. Without limited liability the contract could simply charge the agent a payment for any future utility difference.

The state variable r^h is the aggregate effect of working this history on the incentives to shirk in past histories in order to decrease costs. In every preceding history, the agent has an incentive, indicated in the third and fourth rows of the LDIC, to shirk and reduce future costs. Requiring the agent to work in history h increases the incentive costs at any history $\tilde{h}$ that preceded h by $\lambda^{\tilde{h}}\delta_h$ multiplied by either $\frac{p_0}{p}$ if $h \succeq \langle h, s\rangle$ or $\frac{1-p_0}{1-p}$ if $h \succeq \langle h, f\rangle$. The increase in incentive costs is therefore the agent's *information rent*.

4. Sufficiency of Local Deviations

This section constructs the dual of the FDIC problem and establishes that it is sufficient to consider only local deviations. This proves that it is sufficient to analyze the LDIC problem, which is done in the next section. As a result, this section is self contained – there is no other use for the FDIC other than proving the sufficiency of local deviations. Intuitively, the construction of the dual for the FDIC problem can be thought of as considering, in each public history, which private history requires the strongest incentives, keeping the state variables fixed. Local deviations are sufficient if the strongest incentives to shirk are when the agent never shirked in the past. Intuitively, if the agent did shirk in the past, his costs this period are lower (c_n is increasing) and the effect of his shirking on future costs is lower (δ_n is increasing). Therefore, if the agent did shirk in the past lower incentives are sufficient. This section formalizes this intuition. I first identify the relation between the dual variables and the sufficiency of local deviations. Then, a "summation by parts" exercise allows qualitative comparison and makes the sufficiency argument transparent.

4.1. **The Dual Variables and Sufficiency of Local Deviations.**

Recall that the dual problem assigns a non-negative variable (the shadow price or multiplier) to each constraint, and a constraint to each of the original variables. Let μ^h be the dual variable associated with the probability constraint in problem 2.12 in which q_h appears with coefficient one. Let λ^h_d be the dual variable associated with the FDIC for h, d.

I first establish the formal relation between the λ variables and the sufficiency of local deviations. For this, let D-FDP denote the dual of problem 2.12. Observe that the dual of problem 2.12 subject only to LDIC is equivalent to D-FDP with the additional constraint

$$\forall h, \forall d > 0 : \ \lambda^h_d = 0 \quad . \tag{4.1}$$

Lemma 6. *If every solution to D-FDP satisfies also the constraint 4.1, then any optimal solution subject to LDIC is an optimal contract in the original problem.*

Proof. If every solution to D-FDP satisfies also the constraint 4.1, then every optimal solution to D-LDP is optimal to D-FDP. By the Duality theorem of Linear Programming it follows that any optimal contract subject to LDIC is feasible subject to FDIC. Thus, the condition of Corollary 1 holds. □

4.2. **From λ to Λ – Private Information and Summation By Parts.**

Recall that the different λ in the same period reflect different possible *private histories*. Equivalently, one may think of each private history as an agent *type*. The analysis tries to determine which agent type requires the strongest incentives – for which type the incentive constraint binds first. It is known from adverse-selection models that integration by parts over the agent's types is useful in answering such questions. The same will apply here. To see why summation by parts is required, consider the two period problem given in section 3.1, but add the two FDIC for the second period histories s and f, with $d = 1$. These are simply

$$\forall y \in \{s, f\} : \qquad -(p - p_0)\, w_y + q_y c_1 \leq 0$$

Letting λ_d denote the multiplier on the FDIC for $d \in \{0, 1\}$, the resulting second period dual problem is

(4.2)

$$\begin{array}{rl} \mu\,(n = 2, i, r) = \max & \left[0, \min_{(\lambda_0, \lambda_1) \geq 0} v\,(p - p_0) + ic_n - \delta_n r - \lambda_0 c_n - \lambda_1 c_{n-1}\right] \\ s.t. & \lambda_0 + \lambda_1 \leq \frac{p}{p - p_0}\,(1 + i) \end{array}$$

As $c_2 > c_1$, it is immediate that $\lambda_1 = 0$. However, there are two drawbacks to this formulation. First, λ_0 and λ_1 have the same sign in the objective. This would complicate determining that $\lambda_0 > 0$ while $\lambda_1 = 0$: if the problem is super-modular in λ_0, it is super-modular in λ_1. Second, the feasible space of $\lambda \times i$ is *not* a sub-lattice.[18] The sub-lattice structure is required to apply monotone comparative static arguments.

These problems are averted by redefining the period control variables and the resulting state variables using partial sums. Let Λ^h be a vector of length n_h such that Λ^h_m for $m = \{0, ..., n_h\}$ is the sum of the last $n_h - m$ elements of λ^h. That is:

$$\Lambda^h_m \equiv \sum_{d=m}^{n} \lambda^h_d \,. \tag{4.3}$$

Note that in the local-deviations problem both λ and Λ have just one element ($d = 0$) and are equal to each other. For the FDIC problem, the following additional constraints verify that $\lambda^h_d \geq 0$:

$$\begin{array}{rcl} \Lambda^h_m - \Lambda^h_{m+1} & \geq & 0 \\ \Lambda^h_m & \geq & 0\,. \end{array} \tag{4.4}$$

The condition (4.1) is equivalent to the condition

$$\forall h,\ \Lambda^h_1 = 0 \quad . \tag{4.5}$$

Corollary 3. *If every solution to D-FDP satisfies also the constraint 4.5, then any optimal contract subject to LDIC is optimal in the original problem.*

---

[18]That is, it may very well be that

$$\lambda_0 + \lambda_1 \leq \frac{p}{p - p_0}\,(1 + i)$$

and

$$\hat{\lambda}_1 + \hat{\lambda}_1 \leq \frac{p}{p - p_0}\left(1 + \hat{i}\right) \quad ,$$

but

$$\max\left[\hat{\lambda}_0, \lambda_0\right] + \max\left[\hat{\lambda}_1, \lambda_1\right] > \frac{p}{p - p_0}\left(1 + \max\left[\hat{i}, i\right]\right) \quad .$$

To see the effect on the recursive formulation, in the two period setting, the second period problem is now

(4.6)

$$\mu(n=2,i,r) = \max \quad \left[0, \min_{(\Lambda_0,\Lambda_1)\geq 0} v(p-p_0) + ic_n - \delta_n r - \Lambda_0 c_n + \Lambda_1 (c_n - c_{n-1})\right]$$
$$s.t. \qquad \Lambda_0 \leq \tfrac{p}{p-p_0}(1+i)$$
$$\Lambda_0 - \Lambda_1 \geq 0$$

The sufficiency of local deviations is now simpler than it was in problem 3.3. Λ_0 decreases the objective while Λ_1 increases it, and setting $\Lambda_1 = 0$ does not impose a limit on Λ_0 . Note also that the space of $\Lambda \times i$ that are feasible is a lattice.

4.3. **The Multiple Periods FDIC Dual.**

To make the difference between the LDIC and FDIC duals explicit, denote the FDIC recursive dual by $F(\cdot)$. Deriving a recursive formulation of the multi-period FDIC dual $F(\cdot)$ for a general history h is mostly a technical extension of previous steps. The first step derives the FDIC dual using λ_d^h and μ^h using the same steps as the LDIC dual was derived. Next, the summation by parts exercise described above is applied. This section states the resulting problem and outlines the intuition. The technical steps are provided in appendix B.4.

To derive the multiple periods Dual of the FDIC problem, observe that the only difference from the LDIC is the introduction of the new λ_d^h variables. For each history h, the original FDIC problem has n_h incentive constraints (one for each d from zero to the maximum number of past shirks, $n_h - 1$).

The variables in the primal are identical for the FDIC and LDIC problems. The dual problem for the FDIC therefore has exactly the same constraints as the LDIC, with the additional variables λ_d^h. For every history h, observe that the FDIC for $d > 0$ is exactly the same as of $d = 0$, with only the subscripts on c_n and δ_n changing to $n - d$. Letting λ be a vector of length n the state variables i, r are vectors of length $n-1$ summarizing the past $\lambda_d^{\tilde{h}}$ using the same law of motion as in the LDIC formulation. That is, the dual q_h constraint is

$$\mu^h \geq p\mu^{\langle h,s\rangle} + (1-p)^{\langle h,f\rangle} + v(p-p_0) - \sum_{d=0}^{n} \lambda_d^h c_{n-d} + \sum_{d=0}^{n-1} \left(i_d^h c_{n-d} - r_d^h \delta_{n-d}\right) .$$

The state variables are[19]

$$i_d^h = (p-p_0)\left[\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},f\rangle} \frac{\lambda_d^{\tilde{h}}}{1-p}\right) - \left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},s\rangle} \frac{\lambda_d^{\tilde{h}}}{p}\right)\right]$$

$$r_d^h = \left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},s\rangle} \frac{p_0}{p}\lambda_d^{\tilde{h}}\right) + \left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},f\rangle} \frac{1-p_0}{1-p}\lambda_d^{\tilde{h}}\right)$$

Now apply the summation by parts exercise from above on λ. That is:

$$\sum_{d=0}^{n} \lambda_d^h c_{n-d} = \Lambda_0^h c_n + \sum_{m=1}^{n+1} \Lambda_m^h (c_{n-m+1} - c_{n-m}) .$$

[19]To save on notation, we set $\lambda_d^{\tilde{h}} = 0$ for all $\tilde{h}, d$ such that $d > n_{\tilde{h}} - 1$.

As summation by parts is preserved under summation, it may be verified that, using Λ as defined above obtains the state variables

$$I_m^h = (p - p_0)\left[\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},f\rangle}\frac{\Lambda_m^{\tilde{h}}}{1-p}\right) - \left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},s\rangle}\frac{\Lambda_m^{\tilde{h}}}{p}\right)\right] \tag{4.7}$$

$$R_m^h = \left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},s\rangle}\frac{p_0}{p}\Lambda_m^{\tilde{h}}\right) + \left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},f\rangle}\frac{1-p_0}{1-p}\Lambda_m^{\tilde{h}}\right). \tag{4.8}$$

The constraint is rewritten as (see appendix for the technical steps):

$$\begin{aligned}\mu^h \quad \geq \quad & p\mu^{\langle h,s\rangle} + (1-p)\,\mu^{\langle h,f\rangle} + v\,(p-p_0) - \Lambda_0^h c_n + \sum_{m=1}^{n+1}\Lambda_m^h\,(c_{n-m+1} - c_{n-m}) \\ & + I_0^h c_n - R_0^h\delta_n - \sum_{m=1}^{n}\left(I_m^h\,(c_{n-m+1} - c_{n-m}) - R_m^h\,(\delta_{n-m+1} - \delta_{n-m})\right).\end{aligned}$$

Therefore, define the FDIC period return, $f\,(n, I, R, \Lambda)$:

(4.9)

$$\begin{aligned}f\,(n, I, R, \Lambda) \quad \equiv \quad & v\,(p-p_0) - c_n\,(\Lambda_0 - I_0) - \delta_n R_0 \\ & + \sum_{m=1}^{n_h}\left[\delta_{n-m}\,(\Lambda_m - I_m) + (\delta_{n-m+1} - \delta_{n-m})\,R_m\right].\end{aligned}$$

And the FDIC dual is

$$\begin{aligned}F\,(n, I, R) \quad = \quad & \max\left[0, \min_{\Lambda\geq 0}\hat{F}\,(n, I, R, \Lambda)\right] \\ s.t. & \\ \hat{F}\,(n, I, R, \Lambda) \quad = \quad & pF\,(n+1, I^s, R^s) \\ & + (1-p)\,F\left(n+1, I^f, R^f\right) \\ & + f\,(n, I, R, \Lambda) \\ s.t. & \\ & \Lambda_0 \leq \frac{p}{p-p_0}\,(1+I_0) \\ & \Lambda_m - \Lambda_{m+1} \geq 0 \\ & I^s = I - \frac{p-p_0}{p}\Lambda \quad ; \quad I^f = I + \frac{p-p_0}{1-p}\Lambda \\ & R^s = R + \frac{p_0}{p}\Lambda \quad ; \quad R^f = R + \frac{1-p_0}{1-p}\Lambda\end{aligned} \tag{4.10}$$

Lemma 7. *For every h, $\mu^h = F\left(n_h, I^h, R^h\right)$. In particular, $F\,(1,0,0)$ and the corresponding optimal Λ's define a solution to the dual of problem (2.12).*

Proof. See appendix B.4. □

4.4. **Sufficiency Of Local Deviations.**

This section proves that in the solution for the problem $F\,(1,0,0)$, $\Lambda_m = 0$ for all $m > 0$. Thus, corollary 3 applies and local deviations are sufficient.

Theorem 1. *Any optimal contract subject to LDIC is an optimal contract.*

Proof. A sketch of the proof is provided here. The complete proof is in appendix B.9.1. By corollary 3 , it is sufficient to show that $\Lambda_m = 0$ for every $m > 0$. Each Λ_m has three effects on $F(\cdot)$ – the period return $f(\cdot)$, the law of motion for I and the law of motion for R. The proof considers each of these separately and shows that all effects choose the lowest possible Λ_m. For this, consider a relaxed problem that allows, for every $m > 0$ to choose separately $\Lambda_m^p, \Lambda_m^I, \Lambda_m^R$ (all non-negative) such that Λ_m^s affects the period return term $f(\cdot)$, Λ_m^I affects the law of motion for I , and Λ_m^R affects the law of motion for R. Moreover, the constraints $\Lambda_m \geq \Lambda_{m+1}$ are ignored. The proof will show that the optimal solution sets $\Lambda_m^p = \Lambda_m^I = \Lambda_m^R = 0$. Thus, in the original problem $F(\cdot)$, it must be that the optimal solution is $\Lambda_m = 0$.

For $f(\cdot)$, any reduction in Λ_m^p for $m > 0$ decreases $f(\cdot)$ and thus $\Lambda_m^p = 0$ must be optimal.

Observing $F(\cdot)$, it is intuitive (and proved in the appendix) that $F(\cdot)$ is increasing in R_m for $m > 0$ and thus $\Lambda_m^R = 0$ is optimal.

Finally, increasing Λ_m^I increases the variance in I between continuations, while not affecting the expected I in next period. As F is convex, this should increase the expected continuation value, and thus be sub-optimal. This last argument requires some subtle technical steps to rule out complementarity effects between Λ_0 and Λ_m that are provided in the appendix. □

5. The Optimal Contract

Having established the sufficiency of local deviations, this section analyzes the optimal contract subject to local deviation incentive compatibility.

5.1. **Basic Properties of the Solution.**

A key advantage of the dual dynamic analysis is that it restores monotonicity of the principal's value in the state variables. Each of the state variables i, r reflects the degree of agency frictions – the past incentive costs and information rent. Intuitively, as frictions increase, profits decrease. Thus, profit should decrease in both i and r. I now show this is the case.

For the information asymmetry state r, this is evident by observing problem 3.4: r only decreases the objective and thus decreases profits. For i, one expects that as the utility in the period is more costly (i increases), the contract is less profitable. However, this is not immediately observable as the coefficient on i is positive in the period return. Nevertheless, every increase in the incentive cost i increases period profits by c_n but allows an increase of $\frac{p}{p-p_0} > 1$ in λ that decreases period profits by more than c_n.

Lemma 8. *Period value , $\mu(n, i, r)$ is strictly decreasing in i and r whenever $\mu > 0$.*

Proof. See above for intuition and appendix B.7 for details. □

The agent works only in a period in which $\hat{\mu}(n, i, r, \lambda) \geq 0$. By the law of motion, both state variables are higher after a failure than after success. The previous lemma implies then that $\mu(n+1, \cdot)$ is lower after failure than after success. Therefore, if $\hat{\mu}(n+1, \cdot) < 0$ after success, the same is also true after failure.

Corollary 4. *If the contract terminates after success, it terminates after failure.*

Monotonicity of μ also implies that whenever feasible, $\lambda > 0$ and therefore the LDIC always binds.

Lemma 9. *Whenever $i > -1$, $\lambda > 0$. Thus, the LDIC binds whenever $i > -1$.*

Proof. If $\lambda = 0$ the continuation contract (and thus the agent's expected continuation utility) from the next period is the same regardless of outcomes. Because the wage constraint does not bind at $\lambda = 0$, by complementary slackness the agent is not getting paid for success ($w_h = 0$). Therefore, the agent's optimal plan must be to shirk in this period, contradicting incentive compatibility. □

Consider now the effect of each friction (i or r) on the other. Suppose the agent's incentive cost friction (i) *decreases*. This means the providing the agent utility today is cheaper. The optimal contract therefore provides higher utility to the agent in return for more work. Therefore, the relative profit impact of any past commitments to destroy surplus, which is the value of r, increases. The next lemma establishes this:

Lemma 10. *i, r are substitutes: $\mu(n, i, r)$ is sub-modular in $(-i, r)$.*

Proof. The proof is a simplified version of the proof of lemma 28, which is provided in the appendix to prove a similar characteristic of the FDIC dual. I provide a sketch here.

In any last period, $\lambda = \frac{p}{p-p_0}(1+i)$ and

$$\mu(n, i, r) = \max\left[0, v(p - p_0) - \frac{p}{p - p_0}c_n - \frac{p_0}{p}c_n \cdot i - \delta_n r\right]$$

The marginal effect of i and r is either fixed and strictly negative or, if $\mu(n, i, r) = 0$, it is zero.

(1) As i (resp. r) increases, the marginal effect of r (resp. i) *increases* from $-\delta_n$ (resp. $-\frac{p_0}{p}c_n$) to zero. Therefore, in any last period, $\mu(n, i, r)$ is sub-modular in $(-i, r)$.
(2) Now suppose that for $\tilde{n} > n$, $\mu(\tilde{n}, i, r)$ is sub-modular in $(-i, r)$. Then both continuation values are sub-modular in $(-i, r)$ and so is the period return. The positive weights sum of sub-modular functions are sub-modular and therefore $\hat{\mu}(n, i, r, \lambda)$ is sub-modular in $(-i, r)$ for any λ. As the feasible set defines a lattice, the sub-modularity is preserved under minimization w.r.t. λ. Thus, $\hat{\mu}(n, i, r, \lambda)$ is sub-modular in $(-i, r)$.
(3) The sub-modularity is preserved through the $\max[0, \hat{\mu}(n, i, r, \lambda)]$ operator because $\hat{\mu}(n, i, r, \lambda)$ is decreasing in (i, r) (see the proof of lemma 30 for a parallel result).

□

The results so far consider only the principal's value. The effect of success on the agent's continuation value is easier to evaluate using the primal (regular) LDIC, which is obtained by setting $d = 0$ in (2.7). As the right hand side of 2.7 is always positive, it follows that in any period in which the agent is not paid for success ($w_h = W_h = 0$), the agent's continuation utility after success must be strictly higher than the agent's utility after failure. Combining this with lemma 8 yields a simple result that, surprisingly, cannot be obtained in the primal analysis:

Corollary 5. *Both the principal and the agent prefer success to a failure in any period that the agent is asked to work.*

5.2. **Linear Contracts and Dynamic Quotas.**

This section shows that if the agent is paid for a success in history h, the contract after payment if a linear piece rate and identifies the implied properties of the optimal contract.[20] Given lemma 11, it is sufficient to show that if the agent is paid for a success in history h, the work plan after payment is fixed.

Suppose that the agent is paid for success in a history h. By complementary slackness, the wage constraint binds. Thus $\lambda^h = \frac{p}{p-p_0}\left(1+i^h\right)$. The law of motion for i ($i^{\langle h,s\rangle} = i^h - \frac{p-p_0}{p}\lambda^h$), now implies that in history $i^{\langle h,s\rangle} = -1$.

The only feasible value for λ in $\langle h,s\rangle$ is zero and thus the state variables i and r remain fixed. Intuitively, if $i^{\langle h,s\rangle} = -1$ all the utility given to the agent from $\langle h,s\rangle$ onwards was used to relax incentive constraints in the past. Therefore, it is impossible to use the next period's utility to relax the current incentive constraint and $\lambda = 0$ in $\langle h,s\rangle$ and all histories that follow it. Critically, i and r remain unchanged *regardless of new outcomes*. Thus, the dual value is the same in the next period regardless of outcomes and the work plan is fixed – independent of new outcomes.

It remains to determine the stopping period. For this, place $\lambda = 0$ and $i = -1$ in $\hat{\mu}(n,i,r,\lambda)$ to determine the dual value of a history after payment:

$$\mu(n,-1,r) = \max\left[0, \hat{\mu}(n,-1,r,0)\right] ,$$

and

$$\hat{\mu}(n,-1,r,0) = \mu(n+1,-1,r) + v(p-p_0) - c_n - r\delta_n \tag{5.1}$$

As both c_n and δ_n are positive and increasing, $\mu(n,-1,r)$ is decreasing in n. Eventually $\hat{\mu}(n,-1,r,0)$ will become negative and production will stop.

The continuation work plan is fixed and, as there is no reason to provide the agent more utility than necessary, lemma 3 implies that the contract is fixed.

Lemma 11. *In the optimal contract, if the agent is paid for success in history h, then starting in history $\langle h,s\rangle$ the contract is fixed to $N(r)$, given by*

$$N(r) = \max n : \quad v(p-p_0) - c_n - r\cdot\delta_n \geq 0 \quad .$$

In any history $\tilde{h} \succeq \langle h,s\rangle$ the agent is paid a fixed rate per success:

$$W_{\tilde{h}} = \frac{c_{N(r)}}{p-p_0} \quad .$$

The value to the principal is the remaining surplus less $r\cdot\left(c_{N(r)} - c_{n-1}\right)$

Proof. See above for the proof that the work plan after payment is fixed to some N. $N(r)$ is defined as the last n in which $\hat{\mu}(n,-1,r,0) \geq 0$. Lemma 3 implies that the contract after a payment that provides the lowest utility to the agent is as stated here. As the agent is paid for success in the current period and is risk neutral, any additional utility can be provided as payment in the current period and the optimal continuation contract is fixed.

[20]An additional property – that agent works in history $\left\langle h,f,\tilde{h}\right\rangle$ only if he works in history $\left\langle h,s,\tilde{h}\right\rangle$ requires additional notation and will be added to the next version. The property also implies that the expected number of remaining work period weakly decreases in i and r for any n.

To obtain the value for the principal, apply backward induction to obtain that

$$\begin{aligned}\mu(n,-1,r) &= \sum_{m=n}^{N(r)} \left(v(p-p_0) - c_n - r\delta_n\right) \\ &= \left(\sum_{m=n}^{N(r)} v(p-p_0) - c_n\right) - r\cdot\left(c_{N(r)} - c_{n-1}\right)\end{aligned}$$

□

The number of work periods after payment is determined by considering the information asymmetry cost $r\cdot\delta_n$ as an additional cost to the current period. The contract is implemented by paying the agent a wage as if production cost is at the highest value for which the agent should still work. This guarantees the agent will only work the desired number of remaining periods. This implies an important separation result. Before payment, incentives are provided *only* through changes to the ensuing work plan. Once a payment was made, incentives are provided only through wages.

Corollary 6. *The period outcome affects the future work plan iff the agent was never paid so far.*

Lemma 11 implies the dynamic quota interpretation. The contract can be divided to two stages. If the agent was not yet paid, he is in the evaluation stage. Outcomes affect the state variables and through them the probability that the wage constraint would soon bind. Once the agent is paid, the contract moves to the compensation stage. The agent is paid in all remaining periods a linear wage as if his cost is the highest cost for which he is still expected to work. This last cost is determined in the evaluation stage.

A possible implementation of the contract is a pre-specified quota with "adjustments". In the sales example, at the start of the quarter, the agent is offered a quota contract. At any period, the agent can come to the firm and "complain" that the quota contract is too aggressive – requires too many successes. Based on the agent's performance, the firm then reduces the threshold to make the quota, but also reduces the linear rate the agent receives when making the quota. As long as the firm can commit in advance to the adjustments (e.g. have a policy in place), a contract arbitrarily close to the optimal contract may be implemented.

Suppose that the agent is paid for success in history h, but not in any history that precedes h. By the law of motion for r, the information rent starting in the next period is strictly larger if the agent fails in h than if he succeeds in h ($r^{\langle h,s\rangle} < r^{\langle h,f\rangle}$). As the information rent only increases between periods, in any history $\tilde{h} \succeq \langle h,f\rangle$, it must be that $r^{\tilde{h}} > r^{\langle h,s\rangle}$. Therefore, if the agent will ever be paid for success in any history that follows a failure in h (i.e. $\tilde{h} \succeq \langle h,f\rangle$), the linear rate would be lower than it is after success in h.

Corollary 7. *If a success in history h implies a linear continuation contract, then any linear rate in any history that follows $\langle h,f\rangle$ is lower than the linear rate that starts in $\langle h,s\rangle$. In particular, if the agent works in any history $\left\langle h,f,\tilde{h}\right\rangle$ then he works in the history $\left\langle h,s,\tilde{h}\right\rangle$.*

A useful implication of lemmas 11 and 3 is that the local IC binds in all periods that follow a payment. This is not surprising – as the agent is paid in all remaining periods, if an IC would not bind, the agent would simply be paid less. Lemma 9 established that the IC binds in all periods up to the first payment period. Combining the two results:

Corollary 8. *The local IC binds in all histories in the optimal contract.*

Finally, it is possible to determine the most efficient possible linear contract. That is, the largest number of periods the agent may be asked to work, over all the possible realizations of outcomes. Define $\overline{N}$ as

$$\overline{N} = \max n : \quad v(p-p_0) - c_n \geq \frac{p_0}{p-p_0} \cdot \delta_n$$

Lemma 12. *In the optimal contract, the agent never works more for more than $\overline{N}$ periods. The agent works for exactly $\overline{N}$ periods if he never fails before the first payment. In particular, if $v(p-p_0) < c_{N^{FB}} + \delta_{N^{FB}} \frac{p_0}{p-p_0}$ then the optimal contract is never ex-post efficient.*

Proof. The linear contract with the highest $N(r)$ is the linear contract with the lowest possible r. As in every period with $\lambda > 0$, $r^s < r^f$, the lowest r for any sequence of λ is obtained if the agent constantly succeeds. Suppose h is the first history for the longest linear contract. Then

$$r^h \geq \frac{p_0}{p} \sum_{h \succeq \tilde{h}} \lambda^{\tilde{h}}$$

The inequality is an equality if the agent never failed in the past.

At the start of the contract $i = 0$. For the contract to move to the linear rate it must be that $i^h = -1$. By the law of motion for i, this implies that the sequence of past λ must satisfy at least

$$\frac{p-p_0}{p} \sum_{h \succeq \tilde{h}} \lambda^{\tilde{h}} \geq 1 \, .$$

The inequality is an equality if the agent never fails. In this case

$$\sum_{h \succeq \tilde{h}} \lambda^{\tilde{h}} = \frac{p}{p-p_0} \, ,$$

and the lowest possible value for r^h is:

$$r^h = \frac{p_0}{p} \frac{p}{p-p_0} = \frac{p_0}{p-p_0} \, .$$

Placing this in $N(r)$ yields the desired result. □

5.3. **Termination Without Payment.**

Once the agent is paid, he expects a linear wage that provides him a net positive utility in any period of work. Suppose now that the agent was not yet paid, and so $i > -1$. In the current history the biggest difference between continuation utilities is if the agent moves to the linear contract after a success and is fired after a failure. For the agent to move to a linear contract, it must be that $\lambda = \frac{p}{p-p_0}(1+i)$. As the information rent increases by $\frac{p_0}{p}\lambda$ following success, it must be that $r^s = r + \frac{p_0}{p-p_0}(1+i)$. Using the definition of $N(r)$ for the linear contract given in lemma 11, the linear contract can require work only if $N(r^s) \geq n+1$:

$$v(p-p_0) - c_{n+1} - \left(r + \frac{p_0}{p-p_0}(1+i)\right)\delta_{n+1} \geq 0 \tag{5.2}$$

Suppose now that i and r are sufficiently large so that 5.2 is violated:

$$\left(\frac{p_0}{p-p_0}i+r\right)\delta_{n+1}\geq v\left(p-p_0\right)-c_{n+1}-\frac{p_0}{p-p_0}\delta_{n+1}\;. \tag{5.3}$$

Then the agent's information rent in any future linear contract is too large and the contract will stop immediately after the first payment to the agent. Any promise of "future utility" is worthless. If he works, the agent must be paid for success in this period and then the contract stops. As the work plan starting in this period is fixed – work only in the current period – lemma 11 can be applied to determine that the wage is $\frac{c_n}{p-p_0}$. The wage constraint binds and so $\lambda=\overline{\lambda}$. To determine whether the agent works in this period, evaluate the μ knowing that the agent cannot work in any future periods:

$$\begin{aligned}\hat{\mu}\left(n,i,r,\overline{\lambda}\right) &= v\left(p-p_0\right)-\frac{p}{p-p_0}\left(1+i\right)c_n+ic_n-r\delta_n\\ &= v\left(p-p_0\right)-\frac{p}{p-p_0}c_n-\frac{p_0}{p-p_0}ic_n-r\delta_n\end{aligned}$$

The agent will work in this period if and only if $\hat{\mu}\left(n,i,r,\overline{\lambda}\right)\geq 0$. Thus, the agent will not work whenever

$$\frac{p_0}{p}ic_n+r\delta_n>v\left(p-p_0\right)-\frac{p}{p-p_0}c_n \tag{5.4}$$

Lemma 13. *For any state (n,i,r):*

(1) *If not (5.4), the agent will work this period.*
(2) *If (5.4) and (5.3), the contract terminates at the start of the period.*
(3) *If not (5.4) but (5.3), the agent will only work this period. The contract will terminate at the start of the next period regardless of this period's outcome. The agent works in the period and is paid the spot wage for success.*

Proof. Suppose (5.3) in state (n,i,r). Setting λ so that the wage constraint binds this period minimizes the period return.

(1) The continuation values are zero:
   (a) $\mu\left(n+1,i^s,r^s\right)\leq 0:$ this is exactly the right hand side of 5.2. By condition 5.3 the inequality 5.2 is violated and so $\mu\left(n+1,i^s,r^s\right)=0$
   (b) $\mu\left(n+1,i^f,r^f\right)\leq 0:$ i,r are both are larger after failure and μ is decreasing in both.
(2) If 5.4 is violated, the period return is still negative and so the agent works in the period. The wage is determined from the first order condition for $\mu\left(n,i,r\right)$:

$$c_n-\omega\left(p-p_0\right)=0$$

(3) If also (5.4) then $\max_{\lambda}\hat{\mu}\left(n,i,r,\lambda\right)<0$ and so the optimal is to not work.
(4) For the contract to terminate in state (n,i,r), it must be that $\hat{\mu}\left(n,i,r,\lambda\right)<0$. By construction $p\mu\left(n+1,i^s,r^s\right)+\left(1-p\right)\mu\left(n+1,i^f,r^f\right)\geq 0$. Therefore if $\hat{\mu}\left(n,i,r,\lambda\right)<0$, it must be that the period return is negative. The period return is minimized by setting $\lambda=\overline{\lambda}$. Therefore, for the contract to terminate it must be that inequality 5.4 holds.

□

How long can the agent fail before being terminated? Every previous failure both increases the incentive cost the current period inflicts on the past periods and the information rent cost. At some point, the cost is too large and the agent will be fired.

Lemma 14. *If*

$$\left(\frac{1}{p-p_0}-1\right) i\cdot\delta_{n+1} \geq v\left(p-p_0\right)-c_{n+1}-\frac{p_0}{p-p_0}\delta_{n+1}$$

The agent will work for at most one more period

Proof. By the law of motion for i and r, if the agent never succeeds before history h,

$$i^h = \frac{p-p_0}{1-p}\sum_{h\succeq\tilde{h}}\lambda^{\tilde{h}} \quad \text{and} \quad r^h = \frac{1-p_0}{1-p}\sum_{h\succeq\tilde{h}}\lambda^{\tilde{h}}$$

Therefore

$$r^h = i^h\cdot\frac{1-p}{p-p_0}$$

□

Placing this in the stopping condition (5.3) and applying the previous lemma obtains the desired result.

6. The Effect of Private Cost Information

To evaluate the effect of private cost information, it is illustrative to compare the optimal contract to the case that costs depend on time rather than effort. That is, the case that the period cost, c_n , is independent of the agent's private history. We state the main results here without proof.

Lemma 15. *If c_n is independent of the agent's private history, the optimal contract problem is given by the solution to $\phi\left(1,0\right)$ with*

$$\phi\left(n,i\right)=\max\left[0,\min_{\lambda\geq0}\hat{\phi}\left(n,i,\lambda\right)\right]$$

and

$$\begin{aligned}\hat{\phi}\left(n,i,\lambda\right) &= p\phi\left(n+1,i-\frac{p-p_0}{p}\lambda\right)+\left(1-p\right)\phi\left(n+1,i+\frac{p-p_0}{1-p}\lambda\right)\\ &\quad +v\left(p-p_0\right)-\lambda c_n+ic_n\\ s.t.\quad &\lambda\leq\frac{p}{p-p_0}\left(1+i\right)\end{aligned}$$

Proof. The construction follows the same steps as the LDIC dual. Sufficiency of local deviations is standard and may also be proved following the same steps as for the complete model. □

Observe that $\phi\left(n,i\right)$ is the same as the LDIC dual, modified so that λ does not affect the continuation r. The result corresponding to lemma 11 now follows. The proof is similar to the proof for the private information case.

Lemma 16. *In the optimal contract without private information, if the agent is paid for success in history h, then starting in history $\langle h,s\rangle$ the work plan is fixed to N^{FB}. In any history $\tilde{h}\succeq\langle h,s\rangle$ the agent is paid the static rate per success:*

$$W_{\tilde{h}}=\frac{c_{n_{\tilde{h}}}}{p-p_0}\quad .$$

The value to the principal is the remaining surplus.

It is illustrative to compare the contracts implied by lemma 11 and lemma 16. First, if costs are publicly known, the agent should work to the first best level whenever he is paid.[21] In contrast, if costs are private information, the agent may be "make his quota" but work much less than first best. That is, if the agent accumulated enough failures, the optimal contract may set $N(r) = n - 1$: the agent gets paid once and is then fired.

A second important difference is in the wage scheme. If costs are known, the optimal wage scheme rewards each success based on the costs for that success. In contrast, if costs are not known, the optimal wage scheme rewards each success based on the costs of the *last* success. Thus, the contract with private costs requires less detailed information about the timing of successes. Once the agent makes his quota, all future results are aggregated.

Finally, making quota in the private information case implies a much higher rent per remaining work period compared to the public costs case: the agent wage per success is higher. This is consistent with the observation that sales contracts often offer extremely excessive rewards for successful agents.

6.1. **Costs and Incentives.**

Suppose the contract is linear from the start – the agent works for N periods regardless of outcomes and then stops. Then by lemma 3, the profits depend only on the agent's cost in the N-th period. Any cost difference between earlier periods and the last only increases the agent's utility. This implies that if the firm could invest in lowering costs, it would invest in lowering only the highest cost. In contrast, if costs were public knowledge (as in the start of this section), by lemma 16, the profits are equally affected by all of the agent's costs. This subsection formalizes this intuitive relation between costs, incentives and private information.

For any contract, let $U^h = U^c(h, 0)$ be the complying agent's expected utility starting from history h. Let $D^h = U^c(h, 1) - U^c(h, 0)$ be the expected additional utility for an agent that will comply from history h onwards but shirked once in the past. Finally, let $P(N, h)$ be the probability that the last period of work will be period N given the current history h and assuming that the agent will comply with the contract. Let $\overline{c}^h$ be the *expected last cost*:

$$\overline{c}_h = \sum_{N=n_h}^{N^{FB}} P(N, h)\, c_N \ .$$

The following equality may be verified:

$$D^h = \overline{c}_h - c_{n_h - 1} \ .$$

[21] Observe that if the agent was paid in the past, in the last period of work (and possibly in a number of periods before) the principal's *continuation profit* may well be negative. As $\frac{p}{p-p_0} > 1$, it is possible that:

$$v(p - p_0) - \frac{p}{p - p_0} c_{N^{FB}} < 0$$

But the dual value is positive

$$v(p - p_0) - c_{N^{FB}} \geq 0$$

In any history h, using 2.7, the LDIC may be written as[22]

$$
\begin{aligned}
(p-p_0)\left(W_h+U^{\langle h,s\rangle}-U^{\langle h,f\rangle}\right) &\geq c_{n_h}+p_0D^{\langle h,s\rangle}+(1-p_0)\,D^{\langle h,f\rangle} \\
&= c_{n_h}+p_0\overline{c}^{\langle h,s\rangle}-p_0c_{n_h}+(1-p_0)\,\overline{c}^{\langle h,f\rangle}-(1-p_0)\,c_{n_h} \\
&= p_0\overline{c}_{\langle h,s\rangle}+(1-p_0)\,\overline{c}_{\langle h,f\rangle}
\end{aligned}
$$

The utility difference between success and failure must compensate the agent as if his cost is the p_0-weighted average of the expected last cost after failure and after success.[23]

Thus, for example, if the agent is guaranteed to work in the second period, his first period cost has no effect on the wage scheme or profits. Clearly, a reduction in the first period costs has no effect on the IC and thus on the firm's profits.

To highlight the effect of private costs, consider in contrast the LDIC if costs are public (see the previous section). Then the LDIC is simply

$$(p-p_0)\left(W_h+U^{\langle h,s\rangle}-U^{\langle h,f\rangle}\right)\geq c_{n_h}\ .$$

In this case only current period costs directly affect the period's IC (all future costs affect the IC through their effect on the continuation utilities). In particular, reducing the first period cost will relax the first period IC and increase profits.

Therefore, if costs are private information, the firm should invest more in reducing costs for later periods compared to settings in which costs are public information.

Another implication of the information asymmetry is that the ex-ante variance in the agent's utility and expected work plan is larger when costs are private rather than public information. Recall that the work plan is determined in periods in which the agent is not paid ($W_h=0$). Therefore, the left hand side of the IC is simply the difference in continuation utilities between success and failure, which in turn determines the continuation work plans. In any period h, if the cost is public information this difference is only a function of the current period cost. In contrast, if the cost is private information this difference is a function of the expected last cost $\overline{c}_h$.

7. Conclusion

This paper evaluated the optimal contract for an agent exerting unobservable effort that affects future costs. The optimal contract responded to two agency frictions – the effect of future utility on current incentives (i) and the private information rent (r).

The asymmetric information friction forced the contract to artificially limit the future work after each period, depending on the period's outcome. This guaranteed that the agent never gains by creating a discrepancy between his private information and the public information.

The resulting optimal contract was characterized as a *dynamic quota.* At the start of the contract the agent is not paid for successes. Once the agent is paid, he is paid a fixed linear piece-rate that depends only on his outcomes prior to the first payment.

The optimal contract explains features of real world contracts that puzzled economic observers. The variance in the expected total effort is larger with private cost information

[22] If the agent does not work in history $\langle h,y\rangle$, set $\overline{c}^{\langle h,y\rangle}=c_h$.

[23] Note that a fixed work plan is the special case in which

$$(p-p_0)\,\omega_h=\overline{c}^h$$

than without. Such large variation in ex-post incentives and effort across agents is inefficient and led several authors (see e.g. Oyer (1998); Larkin (2007); Misra and Nair (2009)) to suggest that there is significant room for improvement in either the design of real world incentives or models of the moral hazard setting. The model shows that this variance allows the firm to provide sufficient incentives for effort when it is relatively cheap and to provide high powered incentives when those are required *without fear that agents misrepresent their effort* (delay "easy sales" to the end of the period). The optimal contract must balance between efficiency (having the agents work longer) and profitability. While a high linear commission would guarantee all agents make the efficient level of effort, the firm's profits would all be provided as rents. Consistent with the model, in the firm documented by Larkin (2007) the top end of the reward scale provides the salesperson a 25% commission on *revenues,* a figure very close to the industry's accounting profit margins.

The analysis used duality based arguments to design a dynamic program. The duality based analysis allows applying standard mechanism design techniques to the dynamic private information problem. There is reason to hope that a duality based approach to dynamic moral hazard should prove fruitful in many additional settings as well, especially those with dynamic private information.

The model abstracted from several important aspects of dynamic agency settings. I briefly review the main abstractions and consider their implications.

It was assumed there is no information aggregation problem, as in Holmstrom and Milgrom (1987) and DeMarzo and Sannikov (2008). The aggregation problem occurs if the principal only learns of the aggregate outcome at some final period $\overline{N} \leq N^{FB}$. Critically, the actual order of successes becomes private information. The aggregation problem in the current setting is thus equivalent to constraining the wage to depend only on the *number* of successes. I show in appendix C that the problem may still be written as a linear problem and form the resulting dynamic dual problem. Interestingly, preliminary analysis suggests that the optimal contract in such settings is exactly the convex reward scheme.

The action and outcome space per period was constructed to be the simplest possible – binary effort and binary outcomes. It may be possible to extend the model in these dimensions without breaking any of the assumptions, as in Rogerson (1985). However, the resulting model is very involved notationally. A more promising approach may be to use duality theory of convex programming (see e.g. Rockafellar (1997)) to apply similar methodology using the more convenient first order approach.

It was assumed the principal could commit to a long term contract. Renegotiation with private histories creates issues on many levels that merit further research.

The structure imposed also implied that the optimal contract is finite. However, all the methodology used in creating the dynamic dual problem can be equally applied to infinite horizon models with discounting. Finally, the agent is assumed to be risk neutral, an assumption that clearly simplifies the analysis and was critical to writing the IC as linear inequalities.

The interaction between private history and current technology has several organizational implications for future research. Perhaps the main one is the increased gains from sequential division of labor. In moral hazard settings in which only the *public* history affects current technology, the space of incentive compatible contracts increases when considering the agency relationship as a long term one. It is always possible to offer the agent the optimal short-term contract each period, and the agent will comply. However, when the *private* history affects current technology, as it does here, contracts that are incentive compatible in the short term may not be incentive compatible in the long term – the agent may have a strict incentive to shirk *only* to improve his future utility. Thus, the set of long term incentive compatible contracts no longer includes the set of short term incentive compatible contracts. If private history has a significant effect, it may be profitable to ex-ante limit the duration of the contract, perhaps shifting production to another agent. Such sequential division of labor – an incentive based hierarchy – can allow the principal to extract the agent's private information without paying the high information rent. I explore these ideas in related work.

Finally, conceptually, the principal in this paper attempts to base incentives on histories – pay more when costs *should* be higher. This is closely related to the "ratcheting" approach, that has been traditionally considered problematic (Berliner (1957) introduced the famous "ratchet effect"). Theoretic studies of ratcheting however tended to assume a specific mechanism and study its shortcomings. In the closest related paper, Weitzman (1980) identifies the effort distortion in an intuitive sub-optimal ratcheted contract for a production setting in which the agent privately observes a production shock each period. In contrast, the current paper identifies what can be interpreted as the optimal ratcheting mechanism and shows that it improves on the optimal no-ratcheting – fixed – alternative.

References

Benveniste, L., Scheinkman, J., Jun. 1982. Duality theory for dynamic optimization models of economics: The continuous time case. Journal of Economic Theory 27 (1), 1–19.

Bergemann, D., Hege, U., 2005. The Financing of Innovation: Learning and Stopping. The RAND Journal of Economics 36 (4), 719–752.

Berliner, J. S., 1957. factory and manager in the ussr. Harvard University Press.

Bonatti, A., Hörner, J., Apr. 2011. Collaborating. American Economic Review 101 (2), 632–663.

Clementi, G. L., Hopenhayn, H. a., Feb. 2006. A Theory of Financing Constraints and Firm Dynamics. The Quarterly Journal of Economics 121 (1), 229–265.

Dantzig, G. B., 1963. Linear programming and extensions. Princeton University Press.

DeMarzo, P., Sannikov, Y., 2008. Learning in dynamic incentive contracts. Unpublished Manuscript.

DeMarzo, P. M., Fishman, M. J., Nov. 2007. Optimal Long-Term Financial Contracting. Rev. Financ. Stud. 20 (6), 2079–2128.

Fernandes, A., Phelan, C., Apr. 2000. A Recursive Formulation for Repeated Agency with History Dependence. Journal of Economic Theory 91 (2), 223–247.

Holmstrom, B., Milgrom, P., Mar. 1987. Aggregation and linearity in the provision of intertemporal incentives. Econometrica: Journal of the Econometric Society 55 (2), 303–328.

Joseph, K., Kalwani, M. U., Mar. 1998. The Role of Bonus Pay in Salesforce Compensation Plans. Industrial Marketing Management 27 (2), 147–159.

Larkin, I., 2007. The cost of high-powered incentives: Employee gaming in enterprise software sales. Harvard Business School.

Misra, S., Nair, H., 2009. A structural model of sales-force compensation dynamics: Estimation and field implementation. Quantitative Marketing and Economics, 1–47.

Oyer, P., Feb. 1998. Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality. The Quarterly Journal of Economics 113 (1), 149–185.

Prendergast, C., 1999. The provision of incentives in firms. Journal of economic literature 37 (1), 7–63.

Rockafellar, R. T., 1997. Convex analysis. Princeton Landmarks in Mathematics. Princeton University Press.

Rogerson, W. P., Jan. 1985. Repeated Moral Hazard. Econometrica 53 (1), 69–76.

Spear, S. E., Srivastava, S., Oct. 1987. On Repeated Moral Hazard with Discounting. The Review of Economic Studies 54 (4), 599–617.

Tchistyi, A., 2006. Security design with correlated hidden cash flows: The optimality of performance pricing. Unpublished manuscript, New York University.

Topkis, D. M., 1998. Supermodularity and complementarity. Princeton University Press.

Vohra, R. V., 2005. Advanced Mathematical Economics. Routlege.

Vohra, R. V., 2011. Mechanism Design: A Linear Programming Approach (Econometric Society Monographs). Cambridge University Press.

Weitzman, M. L., 1980. The "Ratchet Principle" and Performance Incentives. The Bell Journal of Economics 11 (1), 302.

Williams, N., 2011. Persistent private information. Econometrica 79 (4), 1233–1275.

Appendix A. Example – LDIC does not imply IC

Suppose $c_n = n$, $p = \frac{1}{2}$ and $p_0 = 0$ and consider the following contract. The agent stops if he succeeds in any of the first two periods. If the agent fails in the first two periods, he is asked to work for eight more periods regardless of new outcomes and is paid 20 for each success. The agent is paid 48 for success in any of the first two periods.

To verify that the contract is LDIC, note that in the last eight periods, the contract is fixed and the agent is paid as if his cost is 10. Lemma 3 establishes the optimality of this wage plan. Letting $U^n(d)$ denote the agent's continuation utility after being asked to work in period n having shirked d times in the past:

$$U^3(d) = \sum_{n=3}^{10} \frac{1}{2} \cdot 20 - \sum_{n=3}^{10} (n-d) = 80 - 52 + 8 \cdot d = 28 + 8d$$

Now consider the agent's problem if he is asked to work in the second period.

(1) If the agent did not shirk in the first period ($d = 0$), he works whenever

$$\begin{aligned} \frac{1}{2} \cdot 48 + \frac{1}{2} \cdot U^3(0) - 2 &\geq U^3(1) \\ 24 + \frac{28}{2} - 2 &\geq 28 + 8 \end{aligned}$$

Which holds as an equality

(2) If the agent did shirk in the first period ($d = 1$), he works whenever

$$\begin{aligned} \frac{1}{2} \cdot 48 + \frac{1}{2} \cdot U^3(1) - 1 &\geq U^3(2) \\ 24 + \frac{28}{2} + \frac{8}{2} - 1 &\geq 28 + 8 + 8 \end{aligned}$$

As the agent works more after failing than after succeeding, the previous shirk is worth more after failing than after succeeding. As a result, the left hand side increased less than the right hand side and the IC is violated.

It is easy to verify that the first period IC just binds:

$$U^2(0) = 24 + \frac{28}{2} - 2 = 36$$

$$U^2(1) = 24 + \frac{28}{2} + \frac{8}{2} - 1 = 41$$

$$\frac{1}{2} \cdot 48 + \frac{1}{2} \cdot U^2(0) - 1 = 24 + 18 - 1 = U^2(1).$$

Appendix B. Detailed Proofs and Derivations

B.1. **Proof that FDIC implies IC.**

Lemma 17. *If a contract is FDIC it is IC*

Proof. Suppose the contract q, w is not IC. Then it violates FDIC:

(1) As the set of possible work plans for the agent is finite and the agent's expected profit is well defined and bounded for each work plan given q, w, there is a set $\hat{E}(q, w)$ of most profitable work plans given q, w.
(2) Suppose $e^c \notin \hat{E}$ and let $\hat{e} \in \hat{E}$, be a most profitable deviating work plan.
(3) Consider the set of histories $\hat{H}$ in which the agent makes a "final deviation" according to $\hat{e}$. That is, $\hat{h} \in \hat{H}$ if $q_{\hat{h}} > 0$, $\hat{e}_{\hat{h}} = 0$ and for every $h \succeq \hat{h}$, $h \neq \hat{h}$, either $q_h = 0$ or $\hat{e}_h = 1$. Let $\hat{d}$ be the number of past deviations at $\hat{h}$ according to $\hat{e}$. Clearly, if the agent profits from making this final deviation, the FDIC for $\hat{h}, \hat{d}$ is violated and the proof is complete.
(4) If the agent does not profit from making this final deviation then the effort plan that complies in this last period provides at least the same expected profit to the agent. Thus, the effort plan with $\hat{e}_{\hat{h}} = 1$ provides at least the same expected profit for the agent. We can now repeat the process of searching for a profitable final deviation after setting $\hat{e}_{\hat{h}} = 1$. As H is a finite set, the process ends either in finding a history in which FDIC is violated or if we change all periods in which $\hat{e}_h = 0$ to $\hat{e}_h = 1$ while weakly increasing the agent's expected profit, implying that $\hat{e}$ was not more profitable than e^c.

□

B.2. **Longer Derivation of the FDIC.**

As the contract is fixed, I omit $\langle q, w \rangle$ from the definition of $U^D(\cdot)$ and $U^C(\cdot)$. The FDIC in history h, d requires that the agent's expected utility from following the contract at history h, d is at least his expected utility from making a final deviation in the history. Write this in the form convenient for taking the dual later:

$$U^D(h, d) - U^c(h, d) \leq 0 . \tag{B.1}$$

If the agent will never be asked to work in history h ($q_h = 0$) then $U^c(\cdot) = U^D(\cdot) = 0$ and (B.1) trivially holds at h for all d. For $q_h > 0$, it will be convenient to write $U^c(h, d)$ and $U^D(h, d)$ in recursive form. The agent's expected utility from complying with the contract when asked to work in history h, d is his expected payment in the period less the cost, plus the expected continuation utility from complying after success and after failure:

(B.2)

$$U^c(h,d) = \frac{1}{q_h}\left[pw_h - q_h c_{h-d} + q_{\langle h,s\rangle}U^c(\langle h,s\rangle, d) + q_{\langle h,f\rangle}U^c(\langle h,f\rangle, d)\right].$$

To construct $U^D(h,d)$, observe that shirking has three effects. First, it saves the agent the cost c_{h-d} in the current period. Second, it replaces the success probability p with p_0. This changes the term $p \cdot w_h$ to $p_0 \cdot w_h$. It also changes the conditional probability of arriving to the public history $\langle h,s\rangle$ from p to p_0 and the probability of moving to the history $\langle h,f\rangle$ from $1-p$ to $1-p_0$. As the term $q_{\langle h,s\rangle}$ (resp. $q_{\langle h,f\rangle}$) assumes the correct probability p (resp. $1-p$), it must be multiplied by $\frac{p_0}{p}$ (resp. $\frac{1-p_0}{1-p}$). Lastly, in the continuation utilities, the total deviations increase by one and so d is replaced with $d+1$:

(B.3)

$$U^D(h,d) = \frac{1}{q_h}\left[p_0 w_h + \frac{p_0}{p}q_{\langle h,s\rangle}U^c(\langle h,s\rangle, d+1) + \frac{1-p_0}{1-p}q_{\langle h,f\rangle}U^c(\langle h,f\rangle, d+1)\right].$$

To facilitate comparison between $U^c(h,d)$ and $U^D(h,d)$, note that for any history, the continuation utility after shirking $d+1$ times in the past equals the continuation utility after shirking d times in the past, plus the utility increase from the last shirk:

$$U^c(\langle h,y\rangle, d+1) = U^c(\langle h,y\rangle, d) + \left(U^c(\langle h,y\rangle, d+1) - U^c(\langle h,y\rangle, d)\right).$$

Thus, we may write

(B.4)

$$\begin{aligned} U^D(h,d) &= \frac{1}{q_h}\left[p_0 w_h + \frac{p_0}{p}q_{\langle h,s\rangle}U^c(\langle h,s\rangle, d) + \frac{1-p_0}{1-p}q_{\langle h,f\rangle}U^c(\langle h,f\rangle, d)\right] \\ &\quad + \frac{1}{q_h}\left[\frac{p_0}{p}q_{\langle h,s\rangle}\left(U^c(\langle h,s\rangle, d+1) - U^c(\langle h,s\rangle, d)\right)\right] \\ &\quad + \frac{1}{q_h}\left[\frac{1-p_0}{1-p}q_{\langle h,f\rangle}\left(U^c(\langle h,f\rangle, d+1) - U^c(\langle h,f\rangle, d)\right)\right] \end{aligned}$$

Placing (B.2) and (B.3) in the FDIC B.1 and multiplying by $q_h > 0$ obtains the linear inequality

$$\begin{gathered} p_0 w_h + \tfrac{p_0}{p}q_{\langle h,s\rangle}U^c(\langle h,s\rangle, d) + \tfrac{1-p_0}{1-p}q_{\langle h,f\rangle}U^c(\langle h,f\rangle, d) \\ + \tfrac{p_0}{p}q_{\langle h,s\rangle}\left(U^c(\langle h,s\rangle, d+1) - U^c(\langle h,s\rangle, d)\right) \\ + \tfrac{1-p_0}{1-p}q_{\langle h,f\rangle}\left(U^c(\langle h,f\rangle, d+1) - U^c(\langle h,f\rangle, d)\right) \\ - \left(pw_h - q_h c_h + q_{\langle h,s\rangle}U^c(\langle h,s\rangle, d) + q_{\langle h,f\rangle}U^c(\langle h,s\rangle, d)\right) \\ \leq 0 \qquad . \end{gathered}$$

As the first and last rows have the same terms, collect terms to obtain:

$$
\begin{aligned}
& -\left(p-p_0\right) w_h + q_h c_{h-d} + \\
& -\tfrac{p-p_0}{p} q_{\langle h,s\rangle} U^c\left(\langle h,s\rangle, d\right) \\
& +\tfrac{p-p_0}{1-p} q_{\langle h,f\rangle} U^c\left(\langle h,f\rangle, d\right) \\
& +\tfrac{p_0}{p} q_{\langle h,s\rangle}\left(U^c\left(\langle h,s\rangle, d+1\right) - U^c\left(\langle h,s\rangle, d\right)\right) \\
& +\tfrac{1-p_0}{1-p} q_{\langle h,f\rangle}\left(U^c\left(\langle h,f\rangle, d+1\right) - U^c\left(\langle h,f\rangle, d\right)\right) \\
& \leq 0 .
\end{aligned} \tag{B.5}
$$

Which is identical to the FDIC (2.8)

B.3. **Duality - Main Theorems.** The classic reference is Dantzig (1963). The results are given in current textbooks on static optimization (see e.g. Vohra (2005)). Any linear problem may be written as

$$\max_{x\geq 0} c\cdot x \qquad s.t. \quad Ax\leq b \quad . \tag{B.6}$$

With c a vector of coefficients and A a matrix that holds in each row the coefficients on a constraint. The dual of the problem is

$$\min_{y\geq 0} y\cdot b \qquad s.t. \quad yA\geq c \tag{B.7}$$

The main results of interest are:

(1) Each primal variable (x) translates to a constraint in the dual problem. Each primal constraint translates to a dual variable (y)
(2) The Duality Theorem: If x^* and y^* are optimal, $y^*\cdot b = c\cdot x^*$ whenever both exist and are finite; and
(3) Complementary Slackness: y_i^* is the Lagrange multiplier in the primal solution for the constraint associated with the i-th row in A . If $y_i^*=0$ then the constraint associated with the i-th row in A does not bind when solving the primal.
(4) The Dual of the Dual is the primal. Therefore, the primal variables x_i^* are the Lagrange multipliers in the dual's solution.

The linearity of the objective implies:

(1) If $y_i^* = 0$ then the solution to problem B.6 is not changed if the constraint associated with the i-th row in A is removed.

While this last result may not have a formal name, it is a combination of the Complementary Slackness result and the Fundamental Theorem of Linear Programming. See e.g. the discussion in Vohra (2005) preceding theorem 4.10 (Complementary Slackness).

B.4. **The Dual LDIC Dynamic Program - Detailed Derivation.**

This section proves that the dual of the optimal contract problem subject to LDIC is given by the LDIC dynamic dual (problem 3.4).

B.4.1. *The Optimal Contract subject to LDIC.*

For convenience, this section explicitly defines the optimal contract problem subject only to LDIC. This problem is given in problem 2.12, with the FDIC replaced by the relaxed LDIC. The dual variable to be associated with each constraint is given next to the constraint.

$$\begin{aligned} V^{LD} = \max_{q\geq 0, w\geq 0} \quad & \textstyle\sum_{h\in H}\left[q_h\left(p-p_0\right)v - w_h p\right] \\ s.t. \quad & \\ & q_\emptyset \leq 1 && \left(\mu^\emptyset\right) \\ \forall h \quad & q_{\langle h,s\rangle} - q_h p \leq 0 && \left(\mu^{\langle h,s\rangle}\right) \\ \forall h \quad & q_{\langle h,f\rangle} - q_h\left(1-p\right) \leq 0 && \left(\mu^{\langle h,f\rangle}\right) \\ \forall h \quad & LDIC && \left(\lambda^h\right) \end{aligned} \tag{B.8}$$

The LDIC for history h is given by setting $d=0$ in 2.11:

(B.9)
$$\begin{aligned} & -(p-p_0)\,w_h + q_h c_h \\ -\tfrac{p-p_0}{p}\textstyle\sum_{\tilde h\succeq\langle h,s\rangle} p w_{\tilde h} &+ \tfrac{p-p_0}{p}\textstyle\sum_{\tilde h\succeq\langle h,s\rangle} q_{\tilde h}\cdot c_{\tilde h} \\ +\tfrac{p-p_0}{1-p}\textstyle\sum_{\tilde h\succeq\langle h,f\rangle} p w_{\tilde h} &- \tfrac{p-p_0}{1-p}\textstyle\sum_{\tilde h\succeq\langle h,f\rangle} q_{\tilde h} c_{\tilde h} \\ & +\tfrac{p_0}{p}\textstyle\sum_{\tilde h\succeq\langle h,s\rangle} q_{\tilde h}\cdot\delta_{\tilde h} \\ & +\tfrac{1-p_0}{1-p}\textstyle\sum_{\tilde h\succeq\langle h,f\rangle} q_{\tilde h}\cdot\delta_{\tilde h} \qquad \le 0 \quad . \end{aligned}$$

B.4.2. *The Dual Problem.*

This section derives the dual of problem B.8 in standard form, using the simplifying variables i^h,r^h, defined in equations 3.6 and 3.7. To facilitate the proof, these are repeated here:

$$i^h = (p-p_0)\left[\left(\sum_{\tilde h: h\succeq\langle\tilde h,f\rangle}\frac{\lambda^{\tilde h}}{1-p}\right)-\left(\sum_{\tilde h: h\succeq\langle\tilde h,s\rangle}\frac{\lambda^{\tilde h}}{p}\right)\right] .$$

and

$$r^h = \left(\sum_{\tilde h: h\succeq\langle\tilde h,s\rangle}\frac{p_0}{p}\lambda^{\tilde h}\right)+\left(\sum_{\tilde h: h\succeq\langle\tilde h,f\rangle}\frac{1-p_0}{1-p}\lambda^{\tilde h}\right) .$$

Recall also that i^h and r^h may be written recursively using $i^\emptyset = r^\emptyset = 0$ and

$$\begin{aligned} i^{\langle h,s\rangle} &= i^h - \tfrac{p-p_0}{p}\lambda^h \quad ; \quad i^{\langle h,f\rangle} = i^h + \tfrac{p-p_0}{1-p}\lambda^h \\ r^{\langle h,s\rangle} &= r^h + \tfrac{p_0}{p}\lambda^h \quad ; \quad r^{\langle h,f\rangle} = r^h + \tfrac{1-p_0}{1-p}\lambda^h \end{aligned}$$

Lemma 18. *The dual to problem B.8 is*

(B.10)
$$\min_{(\mu,\lambda)\ge 0}\mu^\emptyset$$

s.t. $\forall h$:

(B.11)
$$\lambda^h \le \tfrac{p}{p-p_0}\left(1+i^h\right) \qquad \textit{(wage } (w_h)\textit{constraint)}$$

(B.12)
$$\mu^h \ge v(p-p_0) + (1-p)\,\mu^{\langle h,f\rangle} + p\mu^{\langle h,s\rangle} - c_h\lambda^h + c_h i^h - \delta_h r^h \qquad (q_h\,\textit{constraint})$$
$$n_h > N^{FB} \implies \mu^h = 0 \qquad \textit{(stopping condition)}$$

and subject to the definitions of i^h and r^h given above.

Proof. The only primal constraint with a non zero right hand side is $q_\emptyset \le 1$. This establishes the objective. The stopping condition is a technical simplification. The original problem limits attention to histories not longer than N^{FB} and so there are no primal constraints for any longer histories. By setting these dual variables to zero, we do not have to specify a separate q_h constraint to histories of length N^{FB}.

The wage and q_h constraints are established in the next two steps.

(1) The following steps establish the dual constraint associated with w_h (constraint B.11):
   (a) Rewrite constraint B.11 as:

(B.13)
$$\forall h: \quad -(p-p_0)\,\lambda^h + p i^h \ge -p$$

   (b) For any h, the right hand side of B.13 is the coefficient on w_h in the objective of B.8, $-p$, as required.

(c) To construct the left hand side, observe that w_h appears with a coefficient $-(p-p_0)$ in each LDIC (B.9) for h. The variable w_h also appears with a coefficient $\frac{p-p_0}{p}p = (p-p_0)$ in all LDIC for $\tilde{h}$ such that $h \succeq \left\langle \tilde{h}, s \right\rangle$ and with a coefficient $\frac{p-p_0}{1-p}p$ in all LDIC for $\tilde{h}$ such that $h \succeq \left\langle \tilde{h}, f \right\rangle$. Summing up, the constraint is:

$$\forall h \quad -(p-p_0)\lambda^h - (p-p_0)\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},s\rangle}\lambda^{\tilde{h}}\right) + (p-p_0)\frac{p}{1-p}\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},f\rangle}\lambda^{\tilde{h}}\right) \quad \geq \quad -p\,.$$

Collecting terms:

$$\forall h \quad -(p-p_0)\lambda^h + p\cdot\left[(p-p_0)\left(\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},f\rangle}\frac{\lambda^{\tilde{h}}}{1-p}\right) - \left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},s\rangle}\frac{\lambda^{\tilde{h}}}{p}\right)\right)\right] \geq -p\,.$$

Observe that the bracketed term is exactly i^h. Placing i^h in the inequality obtains B.13.

(2) The following steps establish the dual constraint associated with q_h (constraint B.12):

(a) Rewrite constraint B.12 as :

$$\mu^h - (1-p)\mu^{\langle h,f\rangle} - p\mu^{\langle h,s\rangle} + c_h\lambda^h - c_h i^h + \delta_h r^h \quad \geq v(p-p_0)\,. \tag{B.14}$$

(b) The variable q_h appears in the objective of B.8 with coefficient $v(p-p_0)$, obtaining the right hand side.

(c) In the primal constraints, q_h appears in the three probability constraints and in all the LDIC for histories that h follows. The probability constraints generate the first terms in B.14:

$$\mu^h - (1-p)\mu^{\langle h,f\rangle} - p\mu^{\langle h,s\rangle} \tag{B.15}$$

(d) In the LDIC for history h, q_h appears with a coefficient c_h . This generates the term

$$c_h\lambda^h \tag{B.16}$$

(e) The variable q_h also appears twice in each of the LDIC for $\tilde{h}$ such that $h \succeq \tilde{h}$, once as part of the continuation utility term (in either the second or third row of B.9) and once as part of the future gains from shirking term (in either the fourth or fifth row of B.9). The continuation utility term will determine the coefficient on l. The shirking gains term will determine the coefficient on r.

- In the continuation utility term in the LDIC for all histories that h follows, the coefficient for $\lambda^{\tilde{h}}$ is the cost at history h multiplied by the same coefficient as the wage: $\frac{p-p_0}{p}\cdot c_h$ if $h \succeq \left\langle \tilde{h}, s \right\rangle$ and $-\frac{p-p_0}{1-p}c_h$ if $h \succeq \left\langle \tilde{h}, f \right\rangle$. Summarizing and using the definition of i^h:

$$c_h(p-p_0)\left[\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},s\rangle}\frac{\lambda^{\tilde{h}}}{p}\right) - \left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},f\rangle}\frac{\lambda^{\tilde{h}}}{1-p}\right)\right] = c_h\left(-i^h\right)\,. \tag{B.17}$$

- In the shirking gains term in the LDIC for all histories that h follows, the coefficient for $\lambda^{\tilde{h}}$ is $\frac{p_0}{p}\delta_{h-d}$ if $h \succeq \left\langle \tilde{h}, s \right\rangle$ and $\frac{1-p_0}{1-p}\delta_{h-d}$ if $h \succeq \left\langle \tilde{h}, f \right\rangle$. Again the coefficients depend only on whether h follows a success or failure in $\tilde{h}$. Summarizing and using the definition of r^h:

$$\delta_h \cdot \left(\frac{p_0}{p} \left(\sum_{\tilde{h}: h \succeq \left\langle \tilde{h}, s \right\rangle} \lambda_{\tilde{h},d} \right) + \frac{1-p_0}{1-p} \left(\sum_{\tilde{h}: h \succeq \left\langle \tilde{h}, f \right\rangle} \lambda_{\tilde{h},d} \right) \right) = \delta_h r^h \ . \tag{B.18}$$

Combining B.15,B.16,B.17 and B.18 obtains the left hand side of B.14. □

The following result will be used in what follows

Lemma 19. *In the optimal solution to the dual problem B.10*

$$\mu^h = \max \left[0, v\left(p - p_0\right) + \left(1 - p\right) \mu^{\langle h,f \rangle} + p \mu^{\langle h,s \rangle} - c_h \lambda^h + c_h i^h - \delta_h r^h \right]$$

Proof. By construction, $\mu^h \geq 0$. The lemma states that if $\mu^h > 0$ then

$$\mu^h = v\left(p - p_0\right) + \left(1 - p\right) \mu^{\langle h,f \rangle} + p \mu^{\langle h,s \rangle} - c_h \lambda^h + c_h i^h - \delta_h r^h$$

Suppose the statement is false.

(1) If $\mu^{\emptyset}$ violates the condition, decrease $\mu^{\emptyset}$. This is feasible and decreases the objective. Therefore, $\mu^{\emptyset}$ was not optimal.
(2) If any other history violates the condition, there must be an h that violates the condition and that in all histories that precede h the condition holds. Decrease μ^h by ε. As the constraint in the previous period binds, this allows decreasing the previous history's μ by either $\varepsilon \cdot p$ or $\varepsilon\left(1 - p\right)$. Continuing backwards, this will decrease $\mu^{\emptyset}$. This is a feasible decreases of the objective.

□

B.4.3. *Dynamic Representation of the Dual.*

Define the sub-problem

$$D\left(\hat{h}, l, r\right) = \qquad \min_{(\hat{\mu},\hat{\lambda}) \geq 0} \mu^{\hat{h}}$$

s.t.

$$\hat{i}^{\hat{h}} = i \quad , \quad \hat{r}^{\hat{h}} = r$$

$\forall h \succeq \hat{h}$:

(1) $\hat{\lambda}^h \leq \frac{p}{p-p_0}\left(1 + \hat{i}^h\right)$ (wage (w_h)constraint)

(2) $\hat{\mu}^h \geq v\left(p - p_0\right) + \left(1 - p\right) \hat{\mu}^{\langle h,f \rangle} + p \hat{\mu}^{\langle h,s \rangle} - c_h \hat{\lambda}^h + c_h \hat{i}^h - \delta_h \hat{r}^h$ ($q_h$constraint)

(3) $\hat{i}^{\langle h,s \rangle} = \hat{i}^h - \frac{p-p_0}{p}\hat{\lambda}^h \quad ; \quad \hat{i}^{\langle h,f \rangle} = \hat{i}^h + \frac{p-p_0}{1-p}\hat{\lambda}^h$ (law of motion for i)

(4) $\hat{r}^{\langle h,s \rangle} = \hat{r}^h + \frac{p_0}{p}\hat{\lambda}^h \quad ; \quad \hat{r}^{\langle h,f \rangle} = \hat{r}^h + \frac{1-p_0}{1-p}\hat{\lambda}^h$ (law of motion for r)

$\forall h, n_h > N^{FB} : \ \hat{\mu}^h = 0$ (stopping condition)

Let (μ, λ, i, r) be an optimal solution for the dual problem B.10. The next lemma states that for any history $\hat{h}$, $\mu^{\hat{h}}$ must be the optimal solution to $D\left(\hat{h}, i^{\hat{h}}, r^{\hat{h}}\right)$. This is a replication of the Principle of Optimality.

To save on notation, define σ as a triplet : $\sigma = (h, i, r)$ and let $(\mu_\sigma, \lambda_\sigma, i_\sigma, r_\sigma)$ be the optimal solution for $D(\sigma)$.[24] Let $\sigma(h) = \left(h, i^h, r^h\right)$ where $\left(i^h, r^h\right)$ are given by the optimal solution to the dual problem B.10.

Lemma 20. *For every h , setting $\mu^h = \mu^h_{\sigma(h)}$, $\lambda^h = \lambda^h_{\sigma(h)}$ obtains an optimal solution to the dual problem B.10.*

Proof. Let $\sigma = \sigma(h)$ and consider the solution to problem $D(\sigma)$.

(1) A feasible solution to problem $D(\sigma)$ is to set, for every history $\hat{h} \succeq h$: $\mu^{\hat{h}}_\sigma = \mu^{\hat{h}}$ and $\lambda^{\hat{h}}_\sigma = \lambda^{\hat{h}}$. Therefore, $\mu^h_\sigma \leq \mu^h$.

(2) If $\mu^h_\sigma = \mu^h$ the proof is done. The remainder of the proof shows that if $\mu^h_\sigma < \mu^h$ then μ^h could not be optimal.

(3) For every history $\hat{h} \succeq h$, replace $\left(\mu^{\hat{h}}, \lambda^{\hat{h}}, i^{\hat{h}}, r^{\hat{h}}\right)$ in the solution to the full dual B.10 with those specified by the solution to $D(\sigma)$ – $\left(\mu^{\hat{h}}_\sigma, \lambda^{\hat{h}}_\sigma, i^{\hat{h}}_\sigma, r^{\hat{h}}_\sigma\right)$. I first show the replacement is feasible:

   (a) For any history $\hat{h} \succeq h$, all the constraints in the dual B.10 are satisfied after the replacement by construction.

   (b) For the history just before h, suppose $h = \left\langle \hat{h}, f \right\rangle$. The reduction in μ^h is a reduction in $\mu^{\langle \hat{h}, f\rangle}$, which relaxes the q-constraint in $\hat{h}$. The same applies if $h = \left\langle \hat{h}, s \right\rangle$.

   (c) For any other history $\hat{h}$, neither $i^{\hat{h}}, r^{\hat{h}}, \mu^{\langle \hat{h}, s\rangle}$ or $\mu^{\langle \hat{h}, f\rangle}$ are affected by the change. Therefore, the original solution is still feasible.

(4) We have obtained a feasible solution with the q-constraint slack at the history just before h. It is clear that the new feasible solution did not increase the objective $\mu^\emptyset$. Moreover, if $\mu^{\hat{h}} > 0$ for all histories between the first history and h , the change allows decreasing $\mu^\emptyset$, as in the proof of lemma 19: reduce μ in the history just before h (in which the q constraint is slack, now reduce μ in the history just before, and so on until $\mu^\emptyset$.

□

The dynamic dual problem is:

(B.19)

$$\begin{aligned}
\mu(n, i, r) \quad &= \max \left[0, \quad \min_{\lambda \geq 0} \quad \begin{matrix} p\mu\left(n+1, i^s, r^s\right) + (1-p)\,\mu\left(n+1, i^f, r^f\right) \\ +v\left(p - p_0\right) - c_n\lambda + c_n i - \delta_n r \end{matrix} \right] \\
&\quad s.t. \\
&\qquad \lambda \leq \frac{p}{p-p_0}(1+i) \\
&\qquad i^s = i - \frac{p-p_0}{p}\lambda \quad ; \quad i^f = i + \frac{p-p_0}{1-p}\lambda \\
&\qquad r^s = r + \frac{p_0}{p}\lambda \quad ; \quad r^f = r + \frac{1-p_0}{1-p}\lambda \\
\mu\left(N^{FB}+1, i, r\right) \quad &= 0
\end{aligned}$$

Let $\lambda^*(n, i, r)$ denote the set of maximizers of problem B.19 for a state (n, i, r)

[24] Note that this is a vector specifying a value for each history. For example, $\mu^{\tilde{h}}_\sigma$ specifies the value obtained for $\mu^{\tilde{h}}$ in the problem $D(\sigma)$.

Lemma. *For every h let μ^h, λ^h be the solution to the dual problem B.10. Then $\mu^h = \mu\left(n_h, i^h, r^h\right)$ and $\lambda^h \in \lambda^*\left(n_h, i^h, r^h\right)$. In particular, $\mu(1,0,0) = \mu^{\emptyset}$.*

Proof. It may be observed that the solution to problem $\mu\left(n_h, i^h, r^h\right)$ must be the solution to problem $D(\sigma(h))$. Therefore, the previous lemma provides the result. The proof verifies that $\mu\left(n_h, i^h, r^h\right) = D(\sigma(h))$.

(1) The recursive definitions of i and r in problem B.19 are as in the definition of problem $D(\sigma(h))$
(2) The constraint $\lambda \leq \frac{p}{p-p_0}(1+i)$ in problem B.19 is identical to the wage constraint in problem $D(\sigma(h))$. Thus, λ^h is feasible in problem $D(\sigma(h))$ iff λ is feasible in problem B.19 in state (n, i, r).
(3) If $n_h = N^{FB} + 1$, $\mu\left(n_h, i^h, r^h\right) = 0 = D(\sigma(h))$
(4) Suppose that for every $\tilde{h}$ such that $n_{\tilde{h}} \geq n+1$, $\mu\left(n_{\tilde{h}}, i^{\tilde{h}}, r^{\tilde{h}}\right) = D\left(\sigma\left(\tilde{h}\right)\right)$ and consider h such that $n_h = n$.
   (a) By step 2, in the solution for $\mu\left(n_h, i^h, r^h\right)$, it is always feasible to set $\lambda = \lambda^h_{\sigma(h)}$. By the induction assumption, the objective is now $D(\sigma(h))$. Therefore, $\mu\left(n_h, i^h, r^h\right) \leq D(\sigma(h))$
   (b) By step 2, in the solution for $D(\sigma(h))$ it is always feasible to set $\lambda = \lambda^h_{\sigma(h)}$. By the induction assumption, the right hand side of the q-constraint for history h in $D(\sigma(h))$ is now exactly $\mu\left(n_h, i^h, r^h\right)$. As the objective of $D(\sigma(h))$ is to minimize μ^h, the constraint will bind. Therefore $\mu\left(n_h, i^h, r^h\right)$ is an upper bound on $D(\sigma(h))$: $D(\sigma(h)) \leq \mu\left(n_h, i^h, r^h\right)$
   (c) Combining the last two steps obtains $\mu\left(n_h, i^h, r^h\right) = D(\sigma(h))$ for every h.

□

B.5. **Derivation of the FDIC Dual.** This section derives the recursive dual of the FDIC problem. Formally, it proves that $V^{FD} = F(1,0,0)$ where V^{FD} is given in problem 2.12 and $F(n,I,R)$ is given in problem 4.10.

B.5.1. *The Optimal Contract subject to FDIC.*

The primal problem 2.12 is reproduced here for convenience, with the dual variable given next to the associated primal constraint:

$$\begin{array}{rlcl} V^{FD} = & \max\limits_{q\geq 0, w\geq 0} & \sum_{h\in H}\left[q_h\left(p-p_0\right)v - w_h p\right] & \\ & s.t. & & \\ & & q_{h_0}\leq 1 & \mu^{\emptyset} \\ & \forall h & q_{\langle h,s\rangle} - q_h p \leq 0 & \mu^{\langle h,s\rangle} \\ & \forall h & q_{\langle h,f\rangle} - q_h\left(1-p\right) \leq 0 & \mu^{\langle h,f\rangle} \\ & \forall h,d & FDIC \quad (2.11) & \lambda_d^h \end{array}$$

The FDIC is (2.11) is:

$$\begin{array}{llr} FDIC \quad (2.11): & -\left(p-p_0\right)w_h + q_h c_{h-d} & \\ & -\frac{p-p_0}{p}\sum_{\tilde{h}\succeq\langle h,s\rangle} p w_{\tilde{h}} + \frac{p-p_0}{p}\sum_{\tilde{h}\succeq\langle h,s\rangle} q_{\tilde{h}}\cdot c_{\tilde{h}-d} & \\ & +\frac{p-p_0}{1-p}\sum_{\tilde{h}\succeq\langle h,f\rangle} p w_{\tilde{h}} - \frac{p-p_0}{1-p}\sum_{\tilde{h}\succeq\langle h,f\rangle} q_{\tilde{h}} c_{\tilde{h}-d} & \\ & +\frac{p_0}{p}\sum_{\tilde{h}\succeq\langle h,s\rangle} q_{\tilde{h}}\cdot\delta_{\tilde{h}-d} & \\ & +\frac{1-p_0}{1-p}\sum_{\tilde{h}\succeq\langle h,f\rangle} q_{\tilde{h}}\cdot\delta_{\tilde{h}-d} & \leq 0 \quad . \end{array}$$

B.5.2. *The Dual FDIC in standard form.*

This section derives the dual of problem 2.12 in standard form, using the simplifying variables I^h, R^h and Λ^h.

Λ, I, R are vectors for each history h. The control variables, Λ are the partial sums of the dual variables for the FDIC, λ_d^h. For a detailed exposition see section 4.2. Λ^h is defined in equation 4.3:

$$(4.3): \qquad \Lambda_m^h \equiv \sum_{d=m}^{n}\lambda_d^h$$

The state variables I, R are implicitly defined as

$$I^{\emptyset} = R^{\emptyset} = 0$$

$$I^{\langle h,s\rangle} = I^h - \frac{p-p_0}{p}\Lambda^h \qquad\qquad I^{\langle h,f\rangle} = I^h + \frac{p-p_0}{1-p}\Lambda^h$$

$$R^{\langle h,s\rangle} = R^h + \frac{p_0}{p}\Lambda^h \qquad\qquad R^{\langle h,f\rangle} = R^h + \frac{1-p_0}{1-p}\Lambda^h$$

The proof is clearer if I, R are given in explicit form:

$$I_m^h = \left(p-p_0\right)\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},f\rangle}\frac{\Lambda_m^{\tilde{h}}}{1-p} - \sum_{\tilde{h}:h\succeq\langle\tilde{h},s\rangle}\frac{\Lambda_m^{\tilde{h}}}{p}\right) \tag{B.20}$$

$$R_m^h = p_0\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},s\rangle}\frac{\Lambda_m^{\tilde{h}}}{p}\right) + \left(1-p_0\right)\left(\sum_{\tilde{h}:h\succeq\langle\tilde{h},f\rangle}\frac{\Lambda_m^{\tilde{h}}}{1-p}\right) \tag{B.21}$$

Lemma 21. *The dual problem for V^{FD} is*

$$\begin{array}{rl} \text{(B.22)} & \min_{\Lambda^h \geq 0} \quad \mu^{\emptyset} \\ \textit{s.t., for every} \quad h \in H & \\ \textit{The state variables} & B.20, B.21 \\ \lambda_d^h \geq 0: & \Lambda_m^h - \Lambda_{m+1}^h \geq 0 \\ \textit{The } q_h \textit{ dual constraint} & B.25 \\ \textit{The } w_h \textit{ dual constraint} & B.23 \\ \textit{stopping condition:} & n_h > N^{FB} \implies \mu^h = 0 \end{array}$$

Proof. The only difference from the LDIC dual is in the w_h and q_h constraints, which are derived in the next two subsections □

B.5.3. *Derivation of the w_h constraint.*

Lemma. *The constraint associated with w_h in the FDIC dual is*

$$\forall h \quad \Lambda_0^h \leq \frac{p}{p-p_0}\left(1+I_0^h\right) \tag{B.23}$$

Proof. The right hand side of each constraint is the coefficient on w_h in the objective of V^{FD}: $-p$.

To construct the left hand side, observe that w_h appears with a coefficient $-(p-p_0)$ in all the FDIC (2.11) for h. The variable w_h also appears with a coefficient $\frac{p-p_0}{p}p = (p-p_0)$ in all FDIC for $\tilde{h}$ such that $h \succeq \left\langle \tilde{h}, s \right\rangle$ and with a coefficient $\frac{p-p_0}{1-p}p$ in all FDIC for $\tilde{h}$ such that $h \succeq \left\langle \tilde{h}, f \right\rangle$. Summing up:

(B.24)

$$\forall h \qquad \textstyle\sum_{d=0}^{n_h}\left[-(p-p_0)\lambda_d^h - (p-p_0)\left(\sum_{\tilde{h}:h\succeq\left\langle \tilde{h},s\right\rangle}\lambda_d^{\tilde{h}}\right) + (p-p_0)\frac{p}{1-p}\left(\sum_{\tilde{h}:h\succeq\left\langle \tilde{h},f\right\rangle}\lambda_d^{\tilde{h}}\right)\right] \geq -p\,.$$

Open the square brackets and factor out $-p$ from the last two terms (as in the LDIC derivation) to get

$$-(p-p_0)\sum_{d=0}^{n_h}\lambda_d^h + p(p-p_0)\left[\left(\sum_{\tilde{h}:h\succeq\left\langle \tilde{h},f\right\rangle}\sum_{d=0}^{n_h}\frac{\lambda_d^{\tilde{h}}}{1-p}\right) - \left(\sum_{\tilde{h}:h\succeq\left\langle \tilde{h},s\right\rangle}\sum_{d=0}^{n_h}\frac{\lambda_d^{\tilde{h}}}{p}\right)\right] \geq -p\,.$$

Using the definition of Λ (4.3) for the summation over λ_d^h obtains

$$-(p-p_0)\Lambda_0^h + p(p-p_0)\left[\left(\sum_{\tilde{h}:h\succeq\left\langle \tilde{h},f\right\rangle}\frac{\Lambda_0^{\tilde{h}}}{1-p}\right) - \left(\sum_{\tilde{h}:h\succeq\left\langle \tilde{h},s\right\rangle}\frac{\Lambda_0^{\tilde{h}}}{p}\right)\right] \geq -p\,.$$

The definition of I (B.20) can now be applied as was i in the LDIC dual to obtain B.23 □

B.5.4. *Derivation of the q_h constraint.*

Lemma 22. *The constraint associated with q_h in the FDIC dual is*

(B.25)

$$\mu^h - (1-p)\,\mu^{\langle h,f\rangle} - p\mu^{\langle h,s\rangle} + c_h\Lambda_0^h - c_h I_0^h + \delta_h R_0^h$$
$$-\sum_{m=1}^{n_h}\left[\delta_{h-m}\left(\Lambda_m^h - I_m^h\right) + \left(\delta_{h-m+1} - \delta_{h-m}\right) R_m^h\right] \quad \geq \quad v\,(p-p_0)$$

Proof. The variable q_h appears in the objective with coefficient $v\,(p-p_0)$. This obtains the right hand side. The left hand side follows the same construction as for the LDIC, with the additional step of summation by parts:

(1) q_h appears in the three probability constraints as in the LDIC problem. These generate the first three terms in B.25:

$$\mu^h - (1-p)\,\mu^{\langle h,f\rangle} - p\mu^{\langle h,s\rangle} \tag{B.26}$$

(2) In all the FDIC for history h (i.e., for each d), q_h appears with a coefficient c_{h-d} . This generates the term

$$\sum_{d=0}^{n_h} c_{h-d}\lambda_d^h\,.$$

To obtain the term in Λ, apply summation by parts:[25]

$$\begin{aligned}\textstyle\sum_{d=0}^{n_h} c_{h-d}\lambda_d^h &= \\ c_h\Lambda_0^h - \textstyle\sum_{m=1}^{n_h}\left(c_{h-m+1} - c_{h-m}\right)\Lambda_m^h &= \\ c_h\Lambda_0^h - \textstyle\sum_{m=1}^{n_h}\delta_{h-m+1}\Lambda_m^h &.\end{aligned} \tag{B.27}$$

(3) The variable q_h also appears twice in each of the FDIC for $\tilde{h}, d$ such that $h \succeq \tilde{h}$, once as part of the continuation utility term (in either the second or third row of 2.11) and once as part of the future gains from shirking term (in either the fourth or fifth row of 2.11). The continuation utility term will determine the coefficients on I. The shirking gains term will determine the coefficients on R.

(a) In the continuation utility term in the FDIC for all histories that h follows, the coefficient for $\lambda_d^{\tilde{h}}$ is the current cost multiplied by $\frac{p-p_0}{p}$ if $h \succeq \left\langle \tilde{h}, s\right\rangle$ and by$-\frac{p-p_0}{1-p}$ if $h \succeq \left\langle \tilde{h}, f\right\rangle$. Summarizing these terms obtains the sum:

$$\textstyle\sum_{d=0}^{n_h} \frac{c_{h-d}}{p}\left(p-p_0\right)\left[\left(\sum_{\tilde{h}:h\succeq\left\langle\tilde{h},s\right\rangle}\lambda_d^{\tilde{h}}\right) - \frac{p}{1-p}\left(\sum_{\tilde{h}:h\succeq\left\langle\tilde{h},f\right\rangle}\lambda_d^{\tilde{h}}\right)\right] \quad . \tag{B.28}$$

By summation by parts:

$$\begin{aligned}\sum_{d=0}^{n_h} c_{h-d}\cdot\lambda_d^{\tilde{h}} &= c_h\cdot\sum_{d=0}^{n_h}\lambda_d^{\tilde{h}} - \sum_{d=1}^{n_h}\left(c_{h-d+1} - c_{h-d}\right)\cdot\left(\sum_{m=d}^{n_h}\lambda_d^{\tilde{h}}\right) \\ &= c_h\Lambda_0^{\tilde{h}} - \sum_{d=1}^{n_h}\delta_{h+1-d}\Lambda_d^{\tilde{h}}\end{aligned}$$

---

[25] This may be verified directly. Intuitively, this is the same as letting λ be the derivative of Λ and δ the derivative of c. Then

$$\int_0^{n_h} c_t\lambda_t^h dt = \Lambda^h c_{n_h} - \int_0^{n_h}\delta_t\Lambda_t dt$$

Thus, rewrite B.28:

$$(p-p_0)\cdot c_h\left[\left(\sum\nolimits_{\tilde{h}:h\succeq\left\langle\tilde{h},s\right\rangle}\frac{\Lambda_0^{\tilde{h}}}{p}\right)-\left(\sum\nolimits_{\tilde{h}:h\succeq\left\langle\tilde{h},f\right\rangle}\frac{\Lambda_0^{\tilde{h}}}{1-p}\right)\right]$$
$$-(p-p_0)\sum\nolimits_{m=1}^{n_h}\delta_{h+1-m}\left[\left(\sum\nolimits_{\tilde{h}:h\succeq\left\langle\tilde{h},s\right\rangle}\frac{\Lambda_m^{\tilde{h}}}{p}\right)-\left(\sum\nolimits_{\tilde{h}:h\succeq\left\langle\tilde{h},f\right\rangle}\frac{\Lambda_m^{\tilde{h}}}{1-p}\right)\right].$$

As should be expected, the terms in the square brackets are the incentive costs I_m^h. Using the definition of I_m^h in (B.20) to replace the internal summations, the continuation utility term becomes

$$-c_hI_0^h+\sum_{m=1}^{n_h}\delta_{h-m+1}I_m^h \tag{B.29}$$

(b) In the information rents terms in each FDIC constraint (the last two rows in 2.11), the coefficient for $\lambda_d^{\tilde{h}}$ is $\frac{p_0}{p}\delta_{h-d}$ if $h\succeq\left\langle\tilde{h},s\right\rangle$ and $\frac{1-p_0}{1-p}\delta_{h-d}$ if $h\succeq\left\langle\tilde{h},f\right\rangle$. Again the coefficients depend only on whether h follows a success or failure in $\tilde{h}$. Adding up all the relevant terms obtains:

$$\sum_{d=0}^{n_h}\delta_{h-d}\left(\frac{p_0}{p}\left(\sum_{\tilde{h}:h\succeq\left\langle\tilde{h},s\right\rangle}\lambda_d^{\tilde{h}}\right)+\frac{1-p_0}{1-p}\left(\sum_{\tilde{h}:h\succeq\left\langle\tilde{h},f\right\rangle}\lambda_d^{\tilde{h}}\right)\right). \tag{B.30}$$

Summation by parts implies:

$$\begin{aligned}\sum_{d=0}^{n_h}\delta_{h-d}\cdot\lambda_d^{\tilde{h}} &= \delta_h\sum_{d=0}^{n_h}\lambda_d^{\tilde{h}}-\sum_{d=1}^{n_h}\left(\delta_{h-d+1}-\delta_{h-d}\right)\cdot\left(\sum_{m=d}^{n_h}\lambda_d^{\tilde{h}}\right).\\ &= \delta_h\Lambda_0^{\tilde{h}}-\sum_{d=1}^{n_h}\left(\delta_{h-d+1}-\delta_{h-d}\right)\cdot\Lambda_d^h\end{aligned}$$

Thus, rewrite B.30 as:

$$\delta_h\left(\frac{p_0}{p}\left(\sum\nolimits_{\tilde{h}:h\succeq\left\langle\tilde{h},s\right\rangle}\Lambda_0^{\tilde{h}}\right)+\frac{1-p_0}{1-p}\left(\sum\nolimits_{\tilde{h}:h\succeq\left\langle\tilde{h},f\right\rangle}\Lambda_0^{\tilde{h}}\right)\right)$$
$$+\sum\nolimits_{d=1}^{n_h}\left(\delta_{h-d+1}-\delta_{h-d}\right)\left(\frac{p_0}{p}\left(\sum\nolimits_{\tilde{h}:h\succeq\left\langle\tilde{h},s\right\rangle}\Lambda_d^{\tilde{h}}\right)+\frac{1-p_0}{1-p}\left(\sum\nolimits_{\tilde{h}:h\succeq\left\langle\tilde{h},f\right\rangle}\Lambda_d^{\tilde{h}}\right)\right)$$

Applying the definition of R in B.21 obtains the "information rent" term:

$$\delta_hR_0^h-\sum_{m=1}^{n_h}\left(\delta_{h-m+1}-\delta_{h-m}\right)R_m^h \tag{B.31}$$

Combine B.15 , B.27, B.29 and B.31 to obtain the dual constraint for q_h B.25. □

B.5.5. *Recursive Formulation for the FDIC.*

For convenience, the recursive formulation from the text is reproduced here:

$$\begin{aligned}
F(n,I,R) &= \max\left[0, \min_{\Lambda\geq 0}\hat{F}(n,I,R,\Lambda)\right] \\
s.t. & \\
\hat{F}(n,I,R,\Lambda) &= pF(n+1,I^s,R^s) \\
&\quad + (1-p)F\left(n+1,I^f,R^f\right) \\
&\quad + f(n,I,R,\Lambda) \\
s.t. & \\
&\Lambda_0 \leq \frac{p}{p-p_0}(1+I_0) \\
&\Lambda_m - \Lambda_{m+1} \geq 0 \\
&I^s = I - \frac{p-p_0}{p}\Lambda \quad ; \quad I^f = I + \frac{p-p_0}{1-p}\Lambda \\
&R^s = R + \frac{p_0}{p}\Lambda \quad ; \quad R^f = R + \frac{1-p_0}{1-p}\Lambda
\end{aligned}$$

With

$$\begin{aligned}
f(n,I,R,\Lambda) &\equiv v(p-p_0) - c_n(\Lambda_0 - I_0) - \delta_n R_0 \\
&\quad + \sum_{m=1}^{n_h}\left[\delta_{n-m}(\Lambda_m - I_m) + (\delta_{n-m+1} - \delta_{n-m})R_m\right] \quad .
\end{aligned}$$

Lemma 23. *For every h, $\mu^h = F\left(n_h, I^h, R^h\right)$. In particular,$\mu^h = F(1,0,0)$.*

Proof. The proof is a step by step repetition of the proof in section B.4.3, with the variables renamed. I note here the main steps:

(1) The definition of $I^h$and R^h (equations B.20 and B.21) are consistent with the law of motion for the state variables given in the definition of $F(n,I,R)$
(2) The q constraint in the dual B.22 binds whenever $\mu^h > 0$.
(3) The application of the Principle of Optimality is as in the proof for the LDIC.

□

B.6. **Proof for Lemma 4.**

Lemma 24. *$\mu(n,i,r)$ is continuous and convex in (i,r). $\hat{\mu}(n,i,r,\lambda)$ is continuous and convex in λ for every i,r. The optimal λ is continuous in i and r.*

Proof. Separately for each claim.

(1) For $\mu(n,i,r)$:
   (a) In any last period, $\mu(n,i,r)$ is linear and thus continuous and convex.
   (b) Assume that $\mu(n+1,i,r)$ is continuous and convex. As the positive sum of three continuous and convex functions is continuous and convex, for every λ, $V(n,i,r,\lambda)$ is convex in (i,r). As the feasible set is convex and the objective is to minimize a convex function, $\mu(n,i,r)$ is continuous and convex.

(2) For $\hat{\mu}(n,i,r,\lambda)$, the period return is linear in λ and so it is sufficient to show that the continuation is convex in λ. I show this for the continuation after success – $\mu\left(n+1, i-\frac{p-p_0}{p}\lambda, r+\frac{p_0}{p}\lambda\right)$. The same proof applies for the continuation after failure. As the sum of convex functions is convex, this completes the proof. Let λ^1 and λ^2 be feasible solutions. Then by convexity of $\mu(n+1,i,r)$ for any $\alpha \in (0,1)$:

$$\alpha\mu\left(n+1, i-\tfrac{p-p_0}{p}\lambda^1, r+\tfrac{p_0}{p}\lambda^1\right) + (1-\alpha)\mu\left(n+1, i-\tfrac{p-p_0}{p}\lambda^2, r+\tfrac{p_0}{p}\lambda^2\right) \leq$$
$$\mu\left(n+1, \alpha\left(i-\tfrac{p-p_0}{p}\lambda^1\right) + (1-\alpha)\left(i-\tfrac{p-p_0}{p}\lambda^2\right), \alpha\left(r+\tfrac{p_0}{p}\lambda^1\right) + (1-\alpha)\left(r+\tfrac{p_0}{p}\lambda^2\right)\right) =$$
$$\mu\left(n+1, i-\tfrac{p-p_0}{p}\left(\alpha\lambda^1+(1-\alpha)\lambda^2\right), r+\tfrac{p_0}{p}\left(\alpha\lambda^1+(1-\alpha)\lambda^2\right)\right)$$

(3) Given the previous result, the optimal λ is either unique or an interval. Continuity is a standard result (see e.g. Stokey-Lucas).

□

B.7. **Proof for Lemma 8: $\mu(n,i,r)$ decreases in i,r.**

Proof. I first show the result for r and then for i.

- For r:
  - In any last period if $\mu(n,i,r)>0$, $\mu_r = -\delta_n < 0$.
  - Suppose $\mu(n+1,i,r)$ weakly decreases in r . Let λ^* be optimal at the state (n,i,r). As the constraint is not affected by r, λ^* is feasible for $(n,i,r+\varepsilon)$. Therefore

$$\begin{aligned}\mu(n,i,r+\varepsilon) \quad &\leq \quad p\mu\left(n+1, i-\frac{p-p_0}{p}\lambda^*, r+\varepsilon+\frac{p_0}{p}\lambda^*\right)\\ &\quad +(1-p)\mu\left(n+1, i-\frac{p-p_0}{1-p}\lambda^*, r+\varepsilon+\frac{1-p_0}{1-p}\lambda^*\right)\\ &\quad +v(p-p_0)-\lambda^* c_n + ic_n - r\delta_n - \varepsilon\delta_n\\ &< \quad \mu(n,i,r)\end{aligned}$$

    Therefore, for whenever $\mu(n,i,r)>0$, μ strictly decreases in r .
- For i:
  - In any last period, if $\mu(n,i,r)>0$, $\mu_i = -\frac{p_0}{p-p_0}c_n < 0$
  - Suppose $\mu(n+1,i,r)$ weakly decreases in i . Let λ^* be optimal at the state (n,i,r). Then $\lambda^* + \frac{p}{p-p_0}\varepsilon$ is feasible for $(n,i+\varepsilon,r)$. Therefore

$$\begin{aligned}
\mu(n,i+\varepsilon,r) \leq \quad & p\cdot\mu\left(n+1,i+\varepsilon-\tfrac{p-p_0}{p}\lambda^*-\varepsilon,r+\tfrac{p_0}{p}\lambda^*+\varepsilon\tfrac{p_0}{p-p_0}\right)\\
& +(1-p)\cdot\mu\left(n+1,i+\varepsilon+\tfrac{p-p_0}{1-p}\lambda^*+\tfrac{p}{1-p}\varepsilon,r+\tfrac{1-p_0}{1-p}\lambda^*+\varepsilon\tfrac{1-p_0}{1-p}\cdot\tfrac{p}{p-p_0}\right)\\
& v(p-p_0)-c_n\left(\lambda^*+\varepsilon\tfrac{p}{p-p_0}-i-\varepsilon\right)-r\delta_n\\
\leq \quad & p\cdot\mu\left(n+1,i-\tfrac{p-p_0}{p}\lambda^*,r+\tfrac{p_0}{p}\lambda^*\right)\\
& +(1-p)\cdot\mu\left(n+1,i+\tfrac{p-p_0}{1-p}\lambda^*,r+\tfrac{1-p_0}{1-p}\lambda^*\right)\\
& v(p-p_0)-c_n\left(\lambda-l+\tfrac{p_0}{p-p_0}\varepsilon\right)+r\delta_n\\
= \quad & \mu(n,l,r)-c_n\varepsilon\tfrac{p_0}{p-p_0}
\end{aligned}$$

The second inequality removes the added ε terms on r^y and i^y as they are positive and so increase the function by the result on r and the induction assumption. □

B.8. **Contract After Payment.**

Lemma. *11 In the optimal contract, if the agent is paid for success in history h, then starting in history $\langle h,s\rangle$ the contract is fixed to $N(r)$. All remaining LDIC bind and the agent is paid $\frac{c_{N(r)}}{p-p_0}$ for each additional success.*

$$N(r)=\max n: \quad v(p-p_0)-c_n-r\cdot\delta_n\geq 0$$

Proof. The discussion in the text show that after a payment, $\lambda=0$ and the period profit is $\pi(n,r)$ as in equation (5.1). As both c_n and δ_n are increasing, eventually period profit will become negative and production will stop. It therefore remains to show that the principal does not gain by randomizing in any last period. As $\pi(n,r)$ is independent of any new outcomes, this is always true. To see this, write the dual for the last two periods in explicit form:

$$\begin{aligned}
\min_{\mu_h,\mu_{\langle h,s\rangle},\mu_{\langle h,f\rangle}} \quad & \mu_h\\
s.t. \quad & \\
\mu_h-p\mu_{\langle h,s\rangle}-\mu_{\langle h,f\rangle} & \geq\pi(n,r)\\
\mu_{\langle h,s\rangle} & \geq\pi(n+1,r)\\
\mu_{\langle h,f\rangle} & \geq\pi(n+1,r)
\end{aligned}$$

With $\pi(n,r)\geq 0$ and $\pi(n,r+1)<0$. The primal is simply

$$\begin{aligned}
\max_{q\geq 0} \quad & q_h\pi(n,r)+\left(q_{\langle h,s\rangle}+q_{\langle h,f\rangle}\right)\pi(n,r+1)\\
s.t. \quad & \\
& q_h\leq 1\\
& q_{\langle h,s\rangle}\leq pq_h\\
& q_{\langle h,f\rangle}\leq(1-p)\,q_h
\end{aligned}$$

It is immediate that if $\pi(n,r+1)<0$ the optimal solution sets $q_{\langle h,y\rangle}=0$ and as $\pi(n,r)\geq 0$, an optimal solution is to set $q_h=1$. □

B.9. **Sufficiency of Local Deviations.** The proof will use the following lemmas:

Lemma 25. *$F(n,I,R)$ is convex in (I,R). $\hat{F}(n,I,R,\Lambda)$ is convex in Λ for every I,R.*

Proof. The proof is identical to the proof of lemma 4 □

Lemma 26. *$F(n,I,R)$ decreases in R_0 and increases in R_m for any $m>0$*

Proof. By δ_n increasing, the same proof (backward induction) as in the LDIC case applies, with the natural extension to R_m. See section B.7. □

Theorem. *Any optimal contract subject to LDIC is an optimal contract.*

Proof. By corollary 3 , it is sufficient to show that in the solution of the dual FDIC (i.e. problem 4.10 starting from $F(1,0,0)$), $\Lambda_m=0$ for every $m>0$ at every state.

The objective in each state is to minimize $F(\cdot)$. Each Λ_m has three effects on $F(\cdot)$ – the period return $f(\cdot)$, the law of motion for I and the law of motion for R. The proof considers each of these separately and shows that for any $m>0$, $F(\cdot)$ increases through each of these effects and thus the optimal Λ_m is the lowest possible: $\Lambda_m=0$.

For this, consider a relaxed problem that allows, for every $m>0$ to choose separately $\Lambda_m^p,\Lambda_m^I,\Lambda_m^R$ (all non-negative) such that Λ_m^p affects the period return term $f(\cdot)$, Λ_m^I affects the law of motion for I , and Λ_m^R affects the law of motion for R. Moreover, the constraints $\Lambda_m\geq\Lambda_{m+1}$ are ignored. The proof will show that the optimal solution sets $\Lambda_m^p=\Lambda_m^I=\Lambda_m^R=0$. Thus, in the original problem $F(\cdot)$, it must be that the optimal solution is $\Lambda_m=0$.

We first establish that for any $m>0$, it is optimal to set $\Lambda_m^p=\Lambda_m^R=0$. Recall that all the Λ variables are chosen to minimize $F(\cdot)$. Then:

- Λ_m^p only appears in $f(\cdot)$ with a positive coefficient as for all n, $\delta_n\geq 0$. Therefore, the only effect of a reduction in Λ_m^p for $m>0$ is a decrease in $f(\cdot)$. As setting $\Lambda_m^p=0$ is feasible it must be optimal.
- Λ_m^R only appears in the law of motion for R_m. Moreover, both R_m^s and R_m^f increase with Λ_m^R. By lemma 26 above, $F(n+1,\cdot)$ is increasing in R_m for $m>0$. Thus, $\Lambda_m^R=0$ is feasible and optimal.

It remains to consider Λ_m^I. As the purpose of the analysis is to show that $\Lambda_m^I=0$ for all $m>0$, let $\overline{\Lambda}$ be the vector $\left(\Lambda_1^I,\Lambda_2^I,...,\Lambda_{n+1}^I\right)$ and $\overline{I}$ be the vector $(I_1,...,I_n)$ so that $\overline{\Lambda}_m=\Lambda_m^I$ and similarly $\overline{I}_m=I_m$.

By the law of motion for R, if $\Lambda_m^R=0$ at all states then in all states along the optimal solution $R_m=0$. Define the implied problem $G\left(n,I_0,R_0,\overline{I}\right)$ derived by removing from $F(\cdot)$ all elements that are known to be zero (R_m,Λ_m^p and Λ_m^R for $m>0$) and using the definition of $\overline{I}$ and $\overline{\Lambda}$ as above. As we will use monotone comparative static results, the existing results will be more familiar for maximizing $-F(\cdot)$ instead of minimizing $F(\cdot)$. Thus, the problem is:

$$G\left(n,I_0,R_0,\overline{I}\right) \quad = \quad \min\left[0,\max_{\Lambda_0\geq 0,\overline{\Lambda}\geq 0}\hat{G}\left(n,I_0,R_0,\overline{I},\Lambda_0,\overline{\Lambda}\right)\right]$$

With

$$\hat{G}\left(n, I_0, R_0, \overline{I}, \Lambda_0, \overline{\Lambda}\right) = pG\left(n+1, I_0^s, R_0^s, \overline{I}^s\right) + (1-p)\,G\left(n+1, I_0^f, R_0^f, \overline{I}^f\right) - v\,(p-p_0) + c_n\,(\Lambda_0 - I_0) + \delta_n R_0 + \sum_{m=1}^{n_h} \delta_{n-m}\overline{I}_m$$

s.t.

$$\Lambda_0 \leq \frac{p}{p-p_0}\,(1+I_0)$$

$$I_0^s = I_0 - \frac{p-p_0}{p}\Lambda_0 \quad ; \quad I_0^f = I_0 + \frac{p-p_0}{1-p}\Lambda_0$$

$$\overline{I}^s = I - \frac{p-p_0}{p}\overline{\Lambda} \quad ; \quad \overline{I}^f = \overline{I} + \frac{p-p_0}{1-p}\overline{\Lambda}$$

$$R_0^s = R_0 + \frac{p_0}{p}\Lambda_0 \quad ; \quad R_0^f = R_0 + \frac{1-p_0}{1-p}\Lambda_0$$

It remains to show that in the solution to $G\,(1,0,0,0)$, $\overline{\Lambda} = 0$ is optimal for all feasible states. Observe that in any last period N^{FB}, the optimal solution sets $\Lambda_0 = \frac{p}{p-p_0}\,(1+I_0)$. Therefore

$$G\left(N^{FB}, L_0, R_0, \overline{L}\right) = \min\left[0, -v\,(p-p_0) + \tfrac{p_0}{p-p_0}c_{N^{FB}} + \tfrac{p_0}{p-p_0}c_{N^{FB}}L_0 + \textstyle\sum_{m=1}^{N^{FB}-1}\delta_{N^{FB}-m+1}\overline{L}_m + \delta_n R_0\right] \tag{B.32}$$

The following preliminary results will be used in the proof.

Lemma 27. *$G\left(n, I_0, R_0, \overline{I}\right)$ and $\hat{G}\left(n, I_0, R_0, \overline{I}, \Lambda_0, \overline{\Lambda}\right)$ are increasing continuous and concave in I_0, R_0 and $\overline{I}_m$ for any m. $\hat{G}\left(n, I_0, R_0, \overline{I}, \Lambda_0, \overline{\Lambda}\right)$ is concave in $\Lambda_0, \overline{\Lambda}$.*

Proof. As $F\,(n,\cdot)$ and $\hat{F}\,(n,\cdot)$ are convex and continuous in all their arguments for every n, $G\,(n,\cdot)$ and $\hat{G}\,(n,\cdot)$ are concave and continuous in all their arguments.

To prove the remaining claims it is sufficient to prove that in any last period $G\left(N^{FB}, I_0, R_0, \overline{I}\right)$ is increasing inI_0, R_0,$\overline{I}$ and that backward induction implies $\hat{G}\left(n, I_0, R_0, \overline{I}, \Lambda_0, \overline{\Lambda}\right)$ is increasing in $I_0, R_0, \overline{I}$. The last period result is observable in equation B.32. The backward induction step:

(1) For R_0 (and $\overline{I}_m$): For any n, suppose that $G\,(n+1,\cdot)$ increases in R_0 (and $\overline{I}_m$). Then starting with a higher R_0 (and $\overline{I}_m$) does not affect any constraint but for any $\left(\Lambda_0, \overline{\Lambda}\right)$ this increases both the period return and the continuation values $(R^s, R^f, \overline{I}^s, \overline{I}^f)$. Thus, the proof is complete for R_0 and $\overline{I}$.

(2) For I_0 : For any n, suppose that $G\,(n+1,\cdot)$ increases in I_0 and R_0. Let $\Lambda_0^*, \overline{\Lambda}^*$ be optimal for I_0. For any $\varepsilon > 0$ increase in I_0 , it is feasible to increase Λ_0^* by $\varepsilon\frac{p}{p-p_0} > \varepsilon$. The period return increases by $\varepsilon\frac{p_0}{p-p_0}c_n$. Thus, it remains to show that all the state variables weakly increase. The continuation states $\overline{I}^s$,$\overline{I}^f$ are unaffected as $\overline{\Lambda}^*$ did not change. The continuation R^s, R^f increase as they both increase with Λ_0. I_0^f increases with I_0 and with Λ_0. Finally, the effect on I_0^s is exactly zero for any ε:

$$\begin{aligned} I_0^s &= I_0 + \varepsilon - \Lambda_0^*\frac{p-p_0}{p} - \varepsilon\frac{p}{p-p_0}\frac{p-p_0}{p} \\ &= I_0 - \Lambda_0^*\frac{p-p_0}{p} \end{aligned}$$

□

Given the concavity result in the previous lemma, it is sufficient to evaluate the first order effect on $\hat{G}(n,\cdot)$ of any possible marginal increase in $\overline{\Lambda}$ starting from $\overline{\Lambda}=0$ to determine whether $\overline{\Lambda}=0$ is the optimal solution for $\hat{G}(n,\cdot)$.

For this, define $G_{\overline{\Lambda}^+}(n,\cdot)$ as the positive gradient of $G\left(n,I_0,R_0,\overline{I}\right)$ along any direction $\overline{\Lambda}$. That is

$$G_{\overline{\Lambda}^+}\left(n,I_0,R_0,\overline{I}\right)=\lim_{\alpha\to 0,\alpha>0}\frac{G\left(n+1,I_0,R_0,\overline{I}+\alpha\overline{\Lambda}\right)-G\left(n+1,I_0,R_0,\overline{I}\right)}{\alpha}\ .$$

Define $G_{\overline{\Lambda}^-}(n,\cdot)$ as the negative gradient along direction $\overline{\Lambda}$ (the standard definition given $G_{\overline{\Lambda}^+}(n,\cdot)$).

For any Λ_0, let

$$\hat{G}_{\overline{\Lambda}^+}\left(n,I_0,R_0,\overline{I},\Lambda_0,0\right)=\lim_{\alpha\to 0,\alpha>0}\frac{\hat{G}\left(n+1,I_0,R_0,\overline{I},\Lambda_0,\alpha\overline{\Lambda}\right)-\hat{G}\left(n+1,I_0,R_0,\overline{I},\Lambda_0,0\right)}{\alpha}\ .$$

A sufficient condition for $\overline{\Lambda}=0$ to be optimal is that, for any $\Lambda_0\geq 0$

$$\hat{G}_{\overline{\Lambda}^+}\left(n,I_0,R_0,\overline{I},\Lambda_0,0\right)\leq 0\ .$$

Given the definitions of $G_{\overline{\Lambda}^+}(n,\cdot)$ and $G_{\overline{\Lambda}^-}(n,\cdot)$:

$$\hat{G}_{\overline{\Lambda}^+}\left(n,I_0,R_0,\overline{I},\Lambda_0,0\right)=(p-p_0)\left(G_{\overline{\Lambda}^+}\left(n+1,I_0^f,R_0^f,\overline{I}\right)-G_{\overline{\Lambda}^-}\left(n+1,I_0^s,R_0^s,\overline{I}\right)\right)\ .$$

As $p>p_0$, it remains to show that for any n and $\left(\Lambda_0,\overline{\Lambda}\right)$ non-negative:

$$G_{\overline{\Lambda}^+}\left(n,I_0^f,R_0^f,\overline{I}\right)\leq G_{\overline{\Lambda}^-}\left(n,I_0^s,R_0^s,\overline{I}\right) \tag{B.33}$$

As $G(n,\cdot)$ is concave in $\left(I_0,R_0,\overline{I}\right)$, $G_{\overline{\Lambda}^+}(n,\cdot)$ and $G_{\overline{\Lambda}^-}(n,\cdot)$ exist everywhere and for any I_0,R_0:

$$G_{\overline{\Lambda}^+}\left(n,I_0,R_0,\overline{I}\right)\leq G_{\overline{\Lambda}^-}\left(n,I_0,R_0,\overline{I}\right). \tag{B.34}$$

Thus, a sufficient condition for B.33 is

$$G_{\overline{\Lambda}^+}\left(n,I_0^f,R_0^f,\overline{I}\right)\leq G_{\overline{\Lambda}^+}\left(n,I_0^s,R_0^s,\overline{I}\right) \tag{B.35}$$

at $\Lambda_0=0$ $i^f=i^s$ and $R^s=R^f$ and thus condition B.35 trivially holds.

For any $\Lambda_0>0$, recall that $I_0^f>I_0^s$ and $R_0^f>R_0^s$. Thus it is sufficient to show that as I_0 and R_0 increase $G_{\overline{\Lambda}^+}\left(n,I_0,R_0,\overline{I}\right)$ decrease. This is equivalent to the requirement that $G\left(n,I_0,R_0,\overline{I}\right)$ is supermodular in $\left(-I_0,\overline{I}_m\right)$ and in $\left(-R_0,\overline{I}_m\right)$ for every m. The next result therefore concludes the proof.

Lemma 28. *$G\left(n,I_0,R_0,\overline{I}\right)$ is supermodular in $\left(-I_0,\overline{I}_m\right)$ and in $\left(-R_0,\overline{I}_m\right)$ for every m.*

Proof. By backward induction.

- In any last period N^{FB}, the optimal solution sets $\Lambda_0=\frac{p}{p-p_0}(1+I_0)$. Therefore

$$G\left(N^{FB},L_0,R_0,\overline{L}\right)= \tag{B.36}$$
$$\min\left[0,-v(p-p_0)+\tfrac{p_0}{p-p_0}c_{N^{FB}}+\tfrac{p_0}{p-p_0}c_{N^{FB}}L_0+\textstyle\sum_{m=1}^{N^{FB}-1}\delta_{N^{FB}-m+1}\overline{L}_m+\delta_nR_0\right]$$

  (1) If $G\left(N^{FB},\cdot\right)<0$ then $\frac{\partial G\left(N^{FB},\cdot\right)}{\partial\overline{I}_m}=\delta_{N^{FB}-m+1}>0.$

(2) As I_0 or R_0 increases, eventually $G\left(N^{FB},\cdot\right)=0$ and $\frac{\partial G\left(N^{FB},\cdot\right)}{\partial I_m}$ decreases to $\frac{\partial G\left(N^{FB},\cdot\right)}{\partial I_m}=0$.

(3) Therefore, for every I_m, $G\left(N^{FB},I_0,R_0,\overline{I}\right)$ is supermodular in $\left(-I_0,\overline{I}_m\right)$ and in $\left(-R_0,\overline{I}_m\right)$ for every m.

Suppose that $G\left(n+1,I_0,R_0,\overline{I}\right)$ is supermodular in $\left(-I_0,\overline{I}_m\right)$ and in $\left(-R_0,\overline{I}_m\right)$ for every m.

For any $\Lambda_0,\overline{\Lambda}$, $G\left(n+1,I_0^s,R_0^s,\overline{I}^s\right)$ and $G\left(n+1,I_0^f,R_0^f,\overline{I}^f\right)$ are also supermodular as required.

All remaining parts of the objective of $\hat{G}\left(n,I_0,R_0,\overline{I},\Lambda_0,\overline{\Lambda}\right)$ are linear in all variables.

Therefore, for any $\left(\Lambda_0,\overline{\Lambda}\right)$, the objective of $\hat{G}\left(n,I_0,R_0,\overline{I},\Lambda_0,\overline{\Lambda}\right)$ is supermodular as required.

As the feasible set is a sub-lattice, supermodularity is preserved under maximization (see section B.9.1 below for details). Therefore, $\max_{\left(\Lambda_0,\overline{\Lambda}\right)}\hat{G}\left(n,I_0,R_0,\overline{I},\Lambda_0,\overline{\Lambda}\right)$ is supermodular in $\left(-I_0,\overline{I}_m\right)$ and in $\left(-R_0,\overline{I}_m\right)$ for every m.

It remains to show that $G\left(n,I_0,R_0,\overline{I}\right)=\min\left[0,\max_{\Lambda_0,\overline{\Lambda}}\hat{G}\left(n,I_0,R_0,\overline{I},\Lambda_0,\overline{\Lambda}\right)\right]$ is supermodular in $\left(-I_0,\overline{I}_m\right)$ and in $\left(-R_0,\overline{I}_m\right)$ for every m. Lemma 30 shows that, given the supermodularity result for $\hat{G}(\cdot)$, it is sufficient to show that $\max_{\Lambda_0,\overline{\Lambda}}\hat{G}\left(n,I_0,R_0,\overline{I},\Lambda_0,\overline{\Lambda}\right)$ is increasing in I_0,R_0 and $\overline{I}_m$ for any m. This was proved in lemma 27. □

□

B.9.1. *Monotone Comparative Static Results.*

The result used in the proof is an application of Theorem 2.7.6 in Topkis (1998). I repeat the relevant construction here.

- For any $y\in Y\subset R^n$ and $X\subset R^m$, let $X(y)\subset X$ denote a subset of X for each y
- For any $x,x^{'}\in R^n$, let $x\wedge x^{'}$ denote the meet (pairwise minimum) of x and $x^{'}$and $\vee$ denote the join (pairwise maximum).
- The space $X(y)\times Y$ is a sub-lattice iff for any $x\in X(y)$ and $x^{'}\in X\left(y^{'}\right)$, $x\wedge x^{'}\in X\left(y\wedge y^{'}\right)$ and $x\vee x^{'}\in X\left(y\vee y^{'}\right)$
- For $f(x,y):X\times Y\to R$ suppose that $X(y)\times Y$ is a sub-lattice and let $h(y)=\max_{x\in X(y)}f(x,y)$. Then if $f(x,y)$ is supermodular in y_i,y_j , so is $h(y)$.

Note that the commonly used result is simpler as it assumes the feasible set does not depend on y.

Lemma 29. *The set of Λ and I such that Λ is feasible for I in $\hat{F}(\cdot)$ or $\hat{G}(\cdot)$ defined above is a sub-lattice.*

Proof. The statement is equivalent to the following: For every two pairs I,Λ and $I^{'},\Lambda^{'}$ such that in problem $\hat{F}(n,\cdot)$, Λ is feasible for I and $\Lambda^{'}$ is feasible for $I^{'}$, it must be that $\Lambda\wedge\Lambda^{'}$ is feasible for $I\wedge I^{'}$ and $\Lambda\vee\Lambda^{'}$ is feasible for $I\vee I^{'}$.

This may be verified directly, and is also worked out in part (d) of example 2.6.2 in Topkis (1998) as all constraints are of the form $x_i-x_j\leq b$. □

Lemma 30. *Suppose $z(x):R^n\to R$ is increasing and continuous in x_i,x_j and supermodular in $(-x_i,x_j)$. Then $\min[0,z(x)]$ is supermodular in $(-x_i,x_j)$*

Proof. By standard monotone comparative static results, it is sufficient to consider the two variable function $z(x_i, x_j)$. For every $x_i^{'} \geq x_i$ and $x_j^{'} \geq x_j$, the lemma's assumption is that

$$z\left(x_i, x_j^{'}\right) - z(x_i, x_j) \geq z\left(x_i^{'}, x_j^{'}\right) - z\left(x_i^{'}, x_j\right)$$

We need to show that the inequality is preserved under application of the min operator for each element:

$$\min\left[0, z\left(x_i, x_j^{'}\right)\right] - \min\left[0, z(x_i, x_j)\right] \geq \min\left[0, z\left(x_i^{'}, x_j^{'}\right)\right] - \min\left[0, z\left(x_i^{'}, x_j\right)\right] \tag{B.37}$$

Consider each possible case separately:

(1) If $z(x_i, x_j) \geq 0$ then by $z()$ increasing, all min operators bind and both sides of B.37 are zero.

(2) If $z\left(x_i^{'}, x_j^{'}\right) \leq 0$ then by $z()$ increasing all min operators are redundant and the inequality holds by assumption.

(3) If $z\left(x_i^{'}, x_j^{'}\right)$ is the only non-negative number, by $z()$ increasing there is some $\tilde{x}_j \in \left(x_j, x_j^{'}\right)$ such that $z\left(x_i^{'}, \tilde{x}_j\right) = 0 = \min\left[0, z\left(x_i^{'}, x_j^{'}\right)\right]$ and so

$$z\left(x_i^{'}, \tilde{x}_j\right) - z\left(x_i^{'}, x_j\right) = \min\left[0, z\left(x_i^{'}, x_j^{'}\right)\right] - z\left(x_i^{'}, x_j\right) .$$

By $x_j^{'} \geq \tilde{x}_j$,

$$z\left(x_i^{'}, x_j^{'}\right) - z\left(x_i^{'}, x_j\right) \geq z\left(x_i^{'}, \tilde{x}_j\right) - z\left(x_i^{'}, x_j\right)$$

Therefore, by assumption

$$z\left(x_i, x_j^{'}\right) - z(x_i, x_j) \geq z\left(\tilde{x}_i, x_j^{'}\right) - z\left(x_i^{'}, x_j\right) = \min\left[0, z\left(x_i^{'}, x_j^{'}\right)\right] - z\left(x_i^{'}, x_j\right) .$$

(4) If only $z\left(x_i^{'}, x_j^{'}\right)$ and $z\left(x_i, x_j^{'}\right)$ are positive then B.37 is the same as

$$z\left(x_i^{'}, x_j\right) \geq z(x_i, x_j)$$

which holds as $z(x_i, x_j)$ is increasing in x_i.

(5) If only $z\left(x_i^{'}, x_j^{'}\right)$ and $z\left(x_i^{'}, x_j\right)$ are positive the B.37 is the same as

$$z\left(x_i, x_j^{'}\right) \geq z(x_i, x_j)$$

which holds as $z(x_i, x_j)$ is increasing in x_j.

(6) If $z\left(x_i^{'}, x_j^{'}\right)$, $z\left(x_i^{'}, x_j\right)$ and $z\left(x_i, x_j^{'}\right)$ are positive than B.37 is the same as $z(x_i, x_j) \leq 0$ which holds by assumption.

□

B.10. **Discounting.**

Suppose the principal and agent discount the future by a factor $\beta \in (0, 1)$. Discounting requires some technical adjustments – accounting for the right value of future utility and shirking gains, and allows the principal to punish the agent by delaying work. I first consider the required technical adjustments. Suppose the contract is still limited to stopping contracts – the principal is not allowed to "pause" work. Then we can adjust the definition of q_h from (2.1) to

$$\begin{aligned} q_{\langle h,s\rangle} &= \beta \cdot q_h \cdot p \cdot \alpha_{\langle h,s\rangle} \quad ; \\ q_{\langle h,f\rangle} &= \beta \cdot q_h \cdot (1-p) \cdot \alpha_{\langle h,f\rangle}. \end{aligned} \tag{B.38}$$

The discount factor has two effects on the remaining analysis.

First, in the primal probability constraints, q_h is multiplied by β. The technical implication for the dual is simple: in the objective for $V(n)$, the continuation values $V(n+1,\cdot)$ are multiplied by the discount factor. As $V(n,i,r)$ is concave, it is clear that the optimal contract never delays work (asks the agent not to work for a period and then work in the next period). As both the agent and principal are risk neutral and have the same discount factor, this is equivalent to increasing the probability of termination in a period. As the latter is never optimal without discounting, so is the former.

Corollary 9. *If the principal and the agent have a common discount factor $\beta \in (0,1)$, the optimal contract is still a finite stopping contract.*

Second, the derivation of the wage contract for a fixed plan in lemma 3 used the relevant part of the primal problem assuming $q_h = 1$. This must be updated to account for the discounting as $q_h = \beta^{n_h}$. The resulting problem is

$$\max_{w_n} \sum_{n=n_0}^{N(r)} -\beta^{n-n_0} p w_n$$

s.t.

$$(p-p_0)\, w_{N(r)} \geq c_{N(r)}$$

$$\forall n_1 \in \{n_0, ..., N(r)-1\} \quad (p-p_0)\, n \geq c_n + \sum_{n=n_1+1}^{N(r)} \beta^{n-n_1} (c_n - c_{n-1})$$

It is immediate that all constraints still bind. Standard calculations obtain the wage for a fixed contract with discount factor β.

Lemma 31. *If the principal and agent have a common discount factor $\beta \in (0,1)$, then the optimal contract after payment is fixed through $N(r)$, where $N(r)$ is determined as in lemma 11 and the wage is given by:*

$$\begin{aligned} W_{N(r)} &= \frac{c_{N(r)}}{p-p_0} \\ W_n &= \beta W_{n+1} + (1-\beta)\frac{c_n}{p-p_0} \end{aligned}$$

Appendix C. Aggregation Model

Suppose the principal only observes the aggregate outcome after N^{FB} periods. This section outlines the implications for the model. It is clear that the wage scheme can only depend on the number of outcomes. I state without proof that this constrains the contract enough to imply that LDIC are sufficient for IC.

Let ω_h denote the extra wage for a success in a period h. Then if the agent is supposed to work also at $\langle h,f\rangle$, $\omega_h = \omega_{\langle h,f\rangle}$. If the agent will not be asked to work at $\langle h,f\rangle$, $\omega_{\langle h,f\rangle} = 0$ and requiring $\omega_h \geq \omega_{\langle h,f\rangle}$ does not constrain the contract. As the optimal contract minimizes wages, I can relax the constraint $\omega_h = \omega_{\langle h,f\rangle}$ to[26]

$$\omega_h \geq \omega_{\langle h,f\rangle} \,.$$

The model uses $w_h = \omega_h q_h$. We assume that $q_{\langle h,f\rangle} \in \{(1-p)\,q_h, 0\}$ and so whenever $q_{\langle h,f\rangle} \neq 0$, the constraint may be written as

$$\frac{w_h}{q_h} \geq \frac{w_{\langle h,f\rangle}}{q_{\langle h,f\rangle}} = \frac{w_{\langle h,f\rangle}}{q_h\,(1-p)}$$

[26] It is possible to verify later that indeed in the optimal solution this will bind when $q_{\langle h,f\rangle} = 1$.

Multiplying by q_h obtains the following additional set of constraints to the linear problem:

$$\forall h \quad w_h (1-p) \geq w_{\langle h,f \rangle}$$

Call this the aggregation constraint. To obtain the constraint in the most similar form to the wage constraints, multiply it by $(p-p_0)$ and collect sides:

$$\forall h \quad (p-p_0) \left(w_{\langle h,f \rangle} - w_h (1-p) \right) \leq 0 \; .$$

Let ϕ^h be the multiplier on the aggregation constraint in which w_h appears with coefficient $(p-p_0)(1-p)$.

For any history h, let ϕ^- denote the value of ϕ in the previous period. Recall that the dual wage constraint is multiplied by (-1). The aggregation constraint adds to the wage constraint of every period that follows a failure the term :

$$(p-p_0) \left(-\phi^- + (1-p) \phi^h \right) \; ,$$

and to every period that follows a success the term

$$(p-p_0)(1-p)\phi^h \; .$$

It can be shown that in this setting any LDIC contract is IC – the payment for success cannot depend on the period in which the effort was exerted and so the agent has no informational gain from shirking. Then amend the LDIC problem 3.4 to account for the added constraint by adding the state variable f (for "fixed" wage):

$$\begin{aligned}
\mu(n,f,i,r) \quad &= \quad \max \left[0, \min_{\lambda,\phi \geq 0} \hat{\mu}(n,f,i,r,\lambda) \right] \\
s.t. \quad & \\
\hat{\mu}(n,f,i,r,\lambda) = \quad & p\mu \left(n+1, 0, i - \frac{p-p_0}{p}\lambda, r + \theta \frac{p_0}{p}\lambda \right) \\
& + (1-p)\mu \left(n+1, \phi, i + \frac{p-p_0}{1-p}\lambda, r + \theta \frac{1-p_0}{1-p}\lambda \right) \\
& v(p-p_0) - \lambda c_n + i \cdot c_n - r\delta_n
\end{aligned}$$

$$\lambda + (1-p)\phi \leq \frac{p}{p-p_0}(1+i) + f$$

Paying a higher wage today increases the feasible wage tomorrow. As a result, the only effect of ϕ is an increase in the f after failure. As f only relaxes the wage constraint, it is immediate that $\mu_f \leq 0$. Thus ϕ is set to the highest possible value and the wage constraint binds in all periods – every success earns the agent a direct increase in pay.