Optimal Dynamic Contracting: the First-Order Approach and Beyond

Online Appendix

Abstract

In this appendix we present the sections and proofs omitted in “Optimal Dynamic Contracting: the First-Order Approach and Beyond” by Marco Battaglini and Rohit Lamba.

Marco Battaglini
Department of Economics
Princeton University
Princeton NJ 08544
mbattagl@princeton.edu

Rohit Lamba
Department of Economics
Princeton University
Princeton NJ 08544
rlamba@princeton.EDU
In the following sections the numbers of the figures and equations that are not explicitly stated in this appendix refer to the equations presented in the original paper. To distinguish them from the figures and equations in this appendix, their numbers are preceded by the numbers of the section in which they are presented in the paper. In Section 1 we discuss the relationship between continuous and discrete type models. In the following Sections we presents the proofs omitted from the paper. The proof of Lemma 3 is presented after the proofs of Lemma 4 and Propositions 6 and 7 because it relies on some properties of the solution of the weakly relaxed problem established in these results.

1 From discrete to continuous types

In this section we show that the continuous case can be seen as the limit of the discrete case, so all problems of the FO-approach in the discrete version are inherited by the continuous version and vice versa. To keep the notation simple, we assume two periods and \( u(\theta, q) = \theta q \). Consider a type set \( \Theta = [\theta, \theta] \subset \mathbb{R}^+ \), an associated prior distribution \( \Gamma(\theta) \) at \( t = 1 \) and a conditional distribution \( F(\theta' | \theta) \) at \( t = 2 \) defined on \( \Theta \). We assume \( \Gamma(\theta) \) is differentiable in \( \theta \) with density \( \mu(\theta) \) and \( F(\theta' | \theta) \) is differentiable in both \( \theta \), with derivative \( F_\theta(\theta' | \theta) \), and \( \theta' \), with density \( f(\theta' | \theta) \). By standard methods we can obtain the following envelope formula (3.4):

\[
U'(\theta) = q(\theta) - \int_\theta^{\theta'} q(\theta' | \theta) \cdot F(\theta' | \theta) d\theta'
\]

and then derive the FO-optimal contract:

\[
q(\theta' | \theta) = \theta' + \frac{1 - \Gamma(\theta)}{\mu(\theta)} F_\theta(\theta' | \theta) f(\theta' | \theta) \tag{1}
\]

In the rest of this section, we refer to this as the continuous model. We start with an example.

Example A1. Assume \( F(\theta' | \theta) = (\theta' - \theta)^\gamma / \gamma \) for \( \theta' \in [\theta, \theta] \) where \( \theta - \theta = 1 \). If we assume that the prior on \( \Theta \) is uniform, (1) implies that the FO-optimal contract is \( q(\theta' | \theta) = \theta' + (\theta' - \theta) \ln(\theta' - \theta) / \theta - \theta \). Figure 1 plots the conditional distributions and the associated FO-optimal allocations when \( \gamma = 1 \) and \( \Theta = [5, 6] \) after two histories \( \theta_1 = 5.01 \) and \( \theta_1 = 5.08 \). It is evident that the contract is non-monotonic in the realization at \( t = 2 \), \( \theta_2 \). It is easy to see that this FO-optimal contract violates global constraints at \( t = 2 \) and so it is not incentive compatible.

We now explore the connection between the continuous model and the discrete model studied in the previous sections. The continuous model can be easily derived as the limit of the discrete model of the previous sections as follows. Define \( \Theta^N = \{ \theta_0, \ldots, \theta_N \} \) with \( \theta_0 = \theta, \theta_N = \theta \) and \( \theta_i = \theta_{i+1} + \Delta \theta_N \); and let \( \Gamma^N(\theta_i) = \Gamma(\theta_i) \) and \( F^N(\theta_j | \theta_i) = F(\theta_j | \theta_i) \). Given this, the probability of a type \( j \) at \( t = 1 \) is \( \mu_j^N = \Gamma^N(\theta_j) - \Gamma^N(\theta_{j+1}) \) and the probability of a type \( i \) at \( t = 2 \) after a

1 For example, equation 3.2 is equation 2 in Section 3 of the paper.

2 See Baron and Besanko [1984], Besanko [1985], Laffont and Tirole [1996], Courty and Li [2000], Eso and Szentes [2007], and Pavan, Segal and Toikka [2013].
type $j$ at $t = 1$ is $f^N_j (\theta_j | \theta_i) = F^N_j (\theta_j | \theta_i) - F^N_j (\theta_{j+1} | \theta_i)$.

In the rest of the section, we refer to this as the discrete model.

Consider a sequence of supports $\Omega^N$ for $N \to \infty$ such that $\Delta \theta_N \to 0$ as $N \to \infty$ and $\Omega^N \subset \Omega^{N+1}$, so that along the sequence the finite approximation of $\Omega$ becomes increasingly fine.

Using the formula derived in the paper (3.9), we can write the FO-optimal contract along the sequence as:

$$q_N (\theta_j | \theta_i) = \theta_j - \frac{1 - F^N_j (\theta_j | \theta_i)}{f^N_j (\theta_j | \theta_i)} \frac{F^N_j (\theta_j | \theta_i) - F^N_j (\theta_{j+1} | \theta_i)}{\Delta \theta_N} \Delta \theta_N$$

for any $\theta_j \in \Omega^N$, $\theta_i \in \Omega^N$. Note that $\mu^N_i$ can be written as: $\mu^N_i = \frac{\Gamma (\theta_i) - \Gamma (\theta_{i+1})}{\Delta \theta_N} \cdot \Delta \theta_N$. and $f^N_j (\theta_j | \theta_i) = \frac{F^N_j (\theta_j | \theta_i) - F^N_j (\theta_{j+1} | \theta_i)}{\Delta \theta_N} \Delta \theta_N$. We can therefore rewrite (2) as:

$$q_N (\theta_j | \theta_i) = \theta_j + \left(1 - \Gamma^N (\theta_i)\right) \frac{\left[F^N_j (\theta_j | \theta_i) - F^N_j (\theta_{j+1} | \theta_i)\right]}{\Gamma (\theta_i) - \Gamma (\theta_{i+1})} \frac{\Delta \theta_N}{\Delta \theta_N}$$

This condition immediately implies that

$$\lim_{N \to \infty} q_N (\theta_j | \theta_i) = \theta_j + \frac{1 - \Gamma (\theta_i)}{\mu (\theta_i)} \frac{F^N_j (\theta_j | \theta_i)}{f^N_j (\theta_j | \theta_i)} = q (\theta_j | \theta_i)$$

since $\mu^N_i/\Delta \theta_N = \mu (\theta_i)$ and $f^N_j (\theta_j | \theta_i) / \Delta \theta_N \to f (\theta_j | \theta_i)$ as $N \to \infty$. It follows that the limit of the discrete FO-optimal contracts is equal to the continuous FO-optimal contract.$^5$

3 In both definitions, we are implicitly assuming a dummy “$N + 1$" type with mass 0.

4 For example, consider the sequence $(\theta_0^m, ..., \theta_N^m)$ such that $\theta_0^m = \emptyset$, $\theta_N^m = \overline{\Omega}$, $\theta_i^m - \theta_{i-1}^m = (\overline{\Omega} - \emptyset)/2^m$ and so $N^m = 2^m$.

5 Since $\Omega^N \subset \Omega^{N+1}$, if $\theta_j \in \Omega^N$, $\theta_i \in \Omega^N$, then $\theta_j \in \Omega^M$, $\theta_i \in \Omega^M$ for $M \geq N$, so $\lim_{N \to \infty} q_N^* (\theta | \theta)$ is well defined. To extend the contract for points on the real line that do not appear in the sequence of approximations we can consider, for example, the sequence of linear interpolations of the discrete contract. It is immediate to verify that this is a sequence of equicontinuous curves that converges to (1).
This discussion makes it clear that there is a natural connection between discrete and continuous types dynamic principal-agent models. In the light of this we can revisit the examples we have discussed in the previous sections in their continuous version.

**Example 8 and 9 (cont.).** Consider $f_\alpha(\theta'|\theta) = \alpha \cdot e^{-\frac{(\theta' - \theta)^2}{\sigma_\theta(\alpha)}}$ and $f_\alpha(\theta'|\theta) = \frac{\alpha}{1+\sigma_\theta(\alpha)|\theta' - \theta|}$ with all other parameters same as before. Note that $\sigma_\theta(\alpha)$ is chosen so that the probabilities sum to one. The larger is $\alpha$, the higher is the persistence of the types. Figures 2 and 3 show two sample distributions and the associated quantities in period 2 that were plotted for the discretized case in Figures 5.3 and 5.4, respectively. The contract is non-monotonic in two ways: first, for a given history, it is non-monotonic in $\theta_2$. Because of this alone, the FO-optimal contract is not implementable and violates a global constraint. In addition to this, the FO-optimal contract is not monotonic with respect to $\theta_1$; this can be seen from the fact that the contracts with the two different histories cross each other.
2 Proof of Lemma A1

In the proof of Lemma 1 we use the following result:

Lemma A1. In a FO-relaxed problem: $IR_N(h^{t-1})$ can be assumed to hold as equality for all $h^{t-1} \in H^{t-1}$; $IC_{i,i+1}(h^{t-1})$ can be assumed to hold as an equality for all $h^{t-1} \in H^{t-1}$ and $i = 0, 1, ..., N - 1$.

Proof. We proceed in two steps:

Step 1. Suppose that $U(\theta_N|h^{t-1}) = \epsilon > 0$ for some $h^{t-1}$. If $t = 1$, then decreasing $U(\theta_i|h^0)$ by $\epsilon$ for all $i$ does not violate any constraints and increases the monopolist’s profit. If $t > 1$, fix $h^{t-1}$ and decrease $U(\theta_i|h^{t-1})$ by $\epsilon$ for all $\theta_i$. This does not change any of the constraints and keeps the profit of the monopolist the same.

Step 2. Suppose that $IC_{i,i+1}(h^{t-1})$ does not hold as an equality for some $h^{t-1} \in H^{t-1}$ and $i = 0, 1, ..., N - 1$. Then, decrease $U(\theta_k|h^{t-1})$ by $\epsilon$ for each $k \leq i$. If $t = 1$, all the constraints are still satisfied and the monopolist’s profit is strictly higher, giving a contradiction. If $t > 1$, this change does not affect any constraint except $IC_{j-1,j}(h^{t-2})$, where $\theta_j$ is such that $h^{t-1} = (h^{t-2}, \theta_j)$. The right hand side of $IC_{j-1,j}(h^{t-2})$ is reduced by $\delta \sum_{k \leq i}(\alpha_{(j-1)k} - \alpha_{jk})\epsilon = \delta \Delta F(\theta_{i+1} | \theta_j) \epsilon \geq 0$, where the last inequality follows from first order stochastic dominance. Now, repeat the same procedure, decreasing $U(\theta_k|h^{t-1})$ by $\delta \Delta F(\theta_{i+1} | \theta_j) \epsilon$ for each $k \leq j - 1$. We can keep reducing utility vectors backward till the first period, unless $h^{t-1}$ contains $\theta_0$, in which case the backward iteration ends there, to deduce a strictly greater increase in the monopolist’s profit. Thus, the changes do not violate any of the constraints and keep the profit of the monopolist larger than or equal to the change. \]

3 Proof of Lemmata A2-A3

We now prove the lemmata used in the proof of Proposition 2. Recall that $\Delta U(\theta_k| h^{t-1}, \theta_i) = U(\theta_k| h^{t-1}, \theta_i) - U(\theta_k| h^{t-1}, \theta_{i+1})$. We have:

Lemma A2. If $q(\theta|h^{t-1})$ and $\Delta U(\theta_k| h^{t-1})$ are non increasing in, respectively, $i$ and $k$ for any $h^{t-1}$, then (3.5) implies that local upward incentive compatibility constraints are satisfied.

Proof. Condition (3.5) implies that local downward constraints, $IC_{i,i+1}(h^{t-1})$, hold as equality for any $i$ and $h^{t-1}$, that is:

$$U(\theta_i|h^{t-1}) = U(\theta_{i+1}|h^{t-1}) + \Delta q(\theta_{i+1}|h^{t-1}) + \delta \sum_{k=0}^{N} (\alpha_{ik} - \alpha_{(i+1)k}) U(\theta_k|h^{t-1}, \theta_{i+1}).$$
Thus,

\[ U(\theta_{i+1}|h^{t-1}) - U(\theta_i|h^{t-1}) = -\Delta \theta q(\theta_{i+1}|h^{t-1}) - \delta \sum_{k=0}^{N} (\alpha_{i,k} - \alpha_{(i+1),k}) U(\theta_k|h^{t-1}, \theta_{i+1}) \]

\[ = -\Delta \theta q(\theta_i|h^{t-1}) + \delta \sum_{k=0}^{N} (\alpha_{(i+1),k} - \alpha_{i,k}) U(\theta_k|h^{t-1}, \theta_i) \]

\[ + \Delta \theta (q(\theta_i|h^{t-1}) - q(\theta_{i+1}|h^{t-1})) + \delta \sum_{k=0}^{N} (\alpha_{i,k} - \alpha_{(i+1),k}) U(\theta_k|h^{t-1}, \theta_i) \]

\[ \geq -\Delta \theta q(\theta_i|h^{t-1}) + \delta \sum_{k=0}^{N} (\alpha_{(i+1),k} - \alpha_{i,k}) U(\theta_k|h^{t-1}, \theta_i), \]

where the last inequality follows from the fact that \( q(\theta_i|h^{t-1}) \) is non increasing in \( i \) and \( \sum_{k=0}^{N} (\alpha_{i,k} - \alpha_{(i+1),k}) U(\theta_k|h^{t-1}, \theta_i) \geq 0 \). The second observation follows from the fact that \( \Delta U(\theta_k|h^{t-1}, \theta_i) \) is non increasing in \( k \), and that \( \alpha_{i,k} \) first order stochastically dominates \( \alpha_{(i+1),k} \). Thus, \( IC_{i+1,i}(h^{t-1}) \) holds.

**Lemma A3.** If \( q(\theta_i|h^{t-1}) \) and \( \Delta U(\theta_k|h^{t-1}) \) are non increasing in, respectively, \( i \) and \( k \) for any \( h^{t-1} \) and (3.5) holds, then the local incentive compatibility constraints imply the global incentive compatibility constraints.

**Proof.** We show that \( IC_{i,i+2}(h^{t-1}) \) holds. The envelope formula (3.5) is equivalent to assuming that all the local downward incentive compatibility constraints are satisfied as equalities. From \( IC_{i,i+1}(h^{t-1}) \) and \( IC_{i+1,i+2}(h^{t-1}) \) we have:

\[ U(\theta_i|h^{t-1}) - U(\theta_{i+2}|h^{t-1}) \]

\[ = \left[ U(\theta_i|h^{t-1}) - U(\theta_{i+1}|h^{t-1}) \right] + \left[ U(\theta_{i+1}|h^{t-1}) - U(\theta_{i+2}|h^{t-1}) \right] \]

\[ = \Delta \theta q(\theta_{i+1}|h^{t-1}) + \delta \sum_{k=0}^{N} (\alpha_{i,k} - \alpha_{(i+1),k}) U(\theta_k|h^{t-1}, \theta_{i+1}) \]

\[ + \Delta \theta q(\theta_{i+2}|h^{t-1}) + \delta \sum_{k=0}^{N} (\alpha_{(i+1),k} - \alpha_{(i+2),k}) U(\theta_k|h^{t-1}, \theta_{i+2}). \]

It follows that:

\[ U(\theta_i|h^{t-1}) - U(\theta_{i+2}|h^{t-1}) \]

\[ = 2\Delta \theta q(\theta_{i+2}|h^{t-1}) + \delta \sum_{k=0}^{N} (\alpha_{i,k} - \alpha_{(i+2),k}) U(\theta_k|h^{t-1}, \theta_{i+2}) \]

\[ + \Delta \theta (q(\theta_{i+1}|h^{t-1}) - q(\theta_{i+2}|h^{t-1})) + \delta \sum_{k=0}^{N} (\alpha_{i,k} - \alpha_{(i+1),k}) \Delta U(\theta_k|h^{t-1}, \theta_{i+1}) \]

\[ \geq \Delta \theta q(\theta_{i+2}|h^{t-1}) + \delta \sum_{k=0}^{N} (\alpha_{(i+1),k} - \alpha_{(i+2),k}) U(\theta_k|h^{t-1}, \theta_{i+2}), \]
First, we prove a useful lemma that will be invoked in the proof of Lemma 2.

4 Proof of Lemma 2

First, we prove a useful lemma that will be invoked in the proof of Lemma 2.

Lemma A8. The optimal solution satisfies: \( q_L \leq \theta_L, q_L(L) \leq \theta_L \) and \( q_M(L) \leq \theta_M \).

Proof. Suppose \( q_L > \theta_L \). Then, decrease \( q_L \) by \( \varepsilon \). Since it only appears on the RHS of incentive constraints and has positive coefficients, this does not violate any of the constraints. Moreover, the change in the monopolist’s profit is proportional to

\[
\left( \theta_L (q_L - \varepsilon) - \frac{1}{2} (q_L - \varepsilon)^2 \right) - \left( \theta_L q_L - \frac{1}{2} q_L^2 \right) = (q_L - \theta_L) \varepsilon - \frac{1}{2} \varepsilon^2.
\]

We can choose \( \varepsilon \) small enough so that the above expression is positive, giving us a contradiction. We can similarly show that \( q_L(L) \leq \theta_L \).

Next, suppose \( q_M(L) > \theta_M \). Note that the second period incentive constraints after history \( L \) give

\[
\Delta \theta q_L(L) \leq u_M(L) - u_L(L) \leq \Delta \theta q_M(L).
\]

Without loss of generality, \( IC_{ML}(L) \) can be assumed to hold as an equality. Suppose \( u_M(L) - u_L(L) > \Delta \theta q_L(L) \). Then, decrease \( u_M(L) \) so that \( IC_{ML}(L) \) holds as an equality. This does not violate any constraints and keeps the profit of the monopolist the same.

If \( IC_{LM}(L) \) holds as an equality, then we must have \( q_M(L) = q_L(L) \leq \theta_L \) and \( q_M(L) \), giving a contradiction. If \( IC_{LM}(L) \) does not hold as an equality, then we can decrease \( q_M(L) \) by \( \varepsilon \) without disturbing any of the constraints. Moreover, the change in the monopolist’s profit is proportional to the following expression:

\[
\left( \theta_M (q_M(L) - \varepsilon) - \frac{1}{2} (q_M(L) - \varepsilon)^2 \right) - \left( \theta_M q_M(L) - \frac{1}{2} q_M(L)^2 \right) = (q_M(L) - \theta_M) \varepsilon - \frac{1}{2} \varepsilon^2.
\]

We can choose \( \varepsilon \) small enough so that the above expression is positive, giving us a contradiction.

Now, we show that \( IR_L \) binds. Suppose not. Decrease \( U_H, U_M, U_L \) by the same small amount. The first period incentive compatibility constraints continue to hold and the second period constraints are unaffected. This increases the profit of the monopolist without disturbing any of the constraints, giving us a contradiction. Thus, \( U_L = 0 \). Next, we show that \( IC_{ML} \) binds. Suppose not. Decrease \( U_M \) by \( \varepsilon \). Then, all the constraints are satisfied and we increase the monopolist’s
profit, giving us a contradiction. Using these two binding constraints we can eliminate \( U_L \) and \( U_M \) from the maximization problem. In particular, \( IC_{HM} \) can now be written as

\[
U_H \geq \Delta \theta (q_M + q_L) + \delta \frac{3\alpha - 1}{2} [u_H(M) - u_M(M)] + (u_M(L) - u_L(L))
\]

Also, \( IC_{HL} \) is given by

\[
U_H \geq 2\Delta \theta q_L + \delta \frac{3\alpha - 1}{2} [u_H(L) - u_L(L)]
\]

First, note that at least one of \( IC_{HM} \) and \( IC_{HL} \) must bind. If not, then we can decrease \( U_H \) and increase the monopolist’s profit. Suppose \( IC_{HM} \) does not bind. Then, \( IC_{HL} \) must bind. Thus, we can eliminate \( U_H \) from the maximization problem. In particular, \( IC_{HM} \) can now be written as

\[
\Delta \theta q_L + \delta \frac{3\alpha - 1}{2} [u_H(L) - u_M(L)] \geq \Delta \theta q_M + \delta \frac{3\alpha - 1}{2} [u_H(M) - u_M(M)] \quad (3)
\]

Second, we claim that if \( IC_{ML} \) and \( IC_{HL} \) bind and \( IC_{HM} \) does not bind, then \( IC_{HM}(L) \) binds. Suppose \( u_H(L) - u_M(L) > \Delta \theta q_M(L) \). Decrease \( u_M(L) \) by \( \varepsilon \) (and so \( U_H \) by \( \delta (\alpha_{HH} - \alpha_{LH}) \varepsilon \) and \( U_M \) by \( \delta (\alpha_{MH} - \alpha_{LM}) \varepsilon \), thereby, increasing the profit of the monopolist without disturbing any of the remaining constraints, giving us a contradiction. Thus, \( IC_{HM}(L) \) must bind.

Using \( IC_{HM}(M) \) and the binding \( IC_{HM}(L) \) we can rewrite (3) to obtain:

\[
\Delta \theta q_L + \delta \frac{3\alpha - 1}{2} \Delta \theta q_M(L) \geq \Delta \theta q_M + \delta \frac{3\alpha - 1}{2} \Delta \theta q_M(M)
\]

Since \( IC_{HM} \) does not bind, it is easy to see that \( q_M = \theta_M \) and \( q_i(M) = \theta_i \) for any \( i \). By Lemma A8, we have \( q_L \leq \theta_L \) (and thus \( q_L < \theta_M \)) and \( q_M(L) \leq \theta_M \). These clearly contradict the above inequality. Thus, we must have that \( IC_{HM} \) binds. \( \blacksquare \)

5 Proof of Lemma 4

For the reminder of the proof, it is useful to state the first-order conditions of the WR-problem. It is easy to see that the \( H \) type always gets the efficient quantity. After history \( H \), moreover, quantities are always efficient, implying: \( q_H = q_H(M) = q_H(L) = \theta_H \) and \( q_H(H) = \theta_H, q_M(H) = \theta_M, q_L(H) = \theta_L \). The remaining first-order conditions are given by:

\[
[q_M] : \quad \mu_M (\theta_M - q_M) - \mu_H \Delta \theta + \lambda \Delta \theta = 0
\]

\[
[q_L] : \quad \mu_L (\theta_L - q_L) - (\mu_H + \mu_M) \Delta \theta - \lambda \Delta \theta = 0
\]

\[
[q_M(M)] : \quad \mu_M \delta \alpha (\theta_M - q_M(M)) - \lambda_{HM}(M) \Delta \theta + \lambda_{LM}(M) \Delta \theta = 0
\]

\[
[q_L(M)] : \quad \mu_M \delta \frac{1 - \alpha}{2} (\theta_M - q_M(M)) - \lambda_{ML}(M) \Delta \theta = 0
\]

\[
[q_M(L)] : \quad \mu_L \delta \frac{1 - \alpha}{2} (\theta_M - q_M(L)) - \lambda_{HM}(L) \Delta \theta + \lambda_{LM}(L) \Delta \theta = 0
\]
We conclude that for Case 1 (Region A1).

Next, note that we must have

\[ q_H(L) = \theta_L - \frac{\mu_H + \mu_M}{\mu_L} \Delta \theta \]

Also, it is easy to show that \( \lambda_{LM}(M) = \lambda_{LM}(L) = 0 \).

We therefore have \( \lambda_{ML}(L) = \lambda_{LM}(M) \). We have two possible cases:

**Case 1 (Region A1).** \( \lambda_{ML}(M) = \lambda_{LM}(M) = 0 \). In this case:

\[ q_M(L) = \theta_M - \frac{\mu_H}{\mu_M} 3\alpha - 0 \Delta \theta \quad \text{and} \quad q_L(M) = \theta_L - \frac{\mu_H + \mu_M}{\mu_L} 3\alpha - 1 \Delta \theta. \]

For this to be a solution, we must have \( \theta_M - \frac{\mu_H}{\mu_M} 3\alpha - 1 \Delta \theta \geq \theta_L \), so \( \alpha \leq \alpha_0(\mu_M) \) where

\[ \alpha_0(\mu_M) = \frac{\mu_H}{3\mu - 2\mu_M}. \]

We conclude that for \( \alpha \leq \alpha_0(\mu_M) \) the solution is given by

\[ q_H = \theta_H, \quad q_H(j) = \theta_H, \quad q_H(j) = \theta_j \quad \text{for all} \quad j = H, M, L \quad \text{in addition to (4)-(6)}. \]

**Case 2 (Region A2).** \( \lambda_{ML}(M) = \lambda_{LM}(M) > 0 \). Then, \( q_M(M) \) and \( q_M(L) \) are both equal to a constant \( q \). From the first order condition with respect to \( q_M(M) \) and \( q_L(M) \) we have:

\[ q_M(M) = q_L(M) = \frac{2\alpha}{1 + \alpha} \theta_M - \frac{\mu_H}{\mu_M} 3\alpha - 1 \Delta \theta. \]

We conclude that for \( \alpha > \alpha_0(\mu_M) \) the solution is given by

\[ q_H = \theta_H, \quad q_H(j) = \theta_H, \quad q_H(j) = \theta_j \quad \text{for all} \quad j = H, M, L, \quad \text{(4)-(5)} \quad \text{and} \quad \text{(7)}. \]
To characterize the necessary and sufficient condition for $\lambda = 0$, we need to verify that given the solution defined above, $IC_{HL}$ is satisfied. Plugging in the values of Case 1, we obtain:

$$\theta_M - \frac{\mu_H}{\mu_M} \Delta \theta + \frac{3\alpha - 1}{2} \left( \theta_M - \frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta \right) \geq \theta_L - \frac{\mu_H + \mu_M}{\mu_L} \Delta \theta + \frac{3\alpha - 1}{2} \theta_M, \quad (8)$$

that is,

$$\mu_M \geq \frac{\mu_L (1 - \mu_L) \left(1 + \frac{\frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta}{1 + \frac{\frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta}{1 + \frac{\frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta}}\right)}{1 + \mu_L \left(1 + \frac{\frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta}{1 + \frac{\frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta}}\right)} = \mu_1^* (\alpha) \quad (9)$$

Plugging in the values of Case 2, we obtain:

$$\theta_M - \frac{\mu_H}{\mu_M} \Delta \theta + \frac{3\alpha - 1}{2} \left( \frac{2\alpha}{1 + \alpha} \theta_M + \frac{1 - \alpha}{1 + \alpha} \theta_L - \frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{1 + \alpha} \Delta \theta \right) \geq \theta_L - \frac{\mu_H + \mu_M}{\mu_L} \Delta \theta + \frac{3\alpha - 1}{2} \theta_M, \quad (10)$$

that is,

$$\mu_M \geq \frac{\mu_L (1 - \mu_L) \left(1 + \frac{\frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta}{1 + \frac{\frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta}{1 + \frac{\frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta}}\right)}{1 + \mu_L \left(1 - \delta \frac{\frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta}{1 + \frac{\frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta}}\right)} = \mu_2^* (\alpha) \quad (11)$$

Let us define $\mu^* (\alpha) = \min \{\mu_1^* (\alpha), \mu_2^* (\alpha)\}$. We have the following result.

**Lemma A9.** If $\alpha, \mu_M$ is such that $\mu_M \geq \mu^* (\alpha)$ and $\alpha \leq \alpha_0 (\mu_M)$ then the optimal contract is as described in Case 1 presented above. If $\mu \geq \mu^* (\alpha)$ and $\alpha > \alpha_0 (\mu_M)$ then the optimal contract is as described in Case 2 presented above.

**Proof.** We first prove that when $\alpha \leq \alpha_0 (\mu_M)$, then $\mu_M \geq \mu^* (\alpha)$ implies $\mu_M \geq \mu_1^* (\alpha)$. To this end, we prove the counterpositive: when $\alpha \leq \alpha_0 (\mu_M)$, $\mu_M < \mu_1^* (\alpha)$ implies $\mu_M < \mu^* (\alpha)$. Note that: 1. the left hand side of (8) and (10) are the same; 2. the right hand side of (8) is not larger than the right hand side of (10) if and only if $\frac{\mu_M}{\mu_H} \leq \frac{2\alpha}{3\alpha - 1}$, that is if $\alpha \leq \alpha_0 (\mu_M)$. It follows that if $\mu_M < \mu_1^* (\alpha)$, then neither (8) nor (10) hold, implying $\mu_M < \mu_2^* (\alpha)$ as well: we therefore conclude that $\mu_M < \mu^* (\alpha)$. Given this, the conditions $\mu_M \geq \mu^* (\alpha)$ and $\alpha \leq \alpha_0 (\mu_M)$ imply the conditions $\mu_M \geq \mu_1^* (\alpha)$ and $\alpha \leq \alpha_0 (\mu_M)$, so by the discussion presented above, the allocation described in Case 1 is an optimal solution of the WR-problem. By a similar argument, we can prove that when $\alpha > \alpha_0 (\mu_M)$, then $\mu_M \geq \mu^* (\alpha)$ implies $\mu_M \geq \mu_2^* (\alpha)$. This implies that when we have $\mu_M \geq \mu^* (\alpha)$ and $\alpha > \alpha_0 (\mu_M)$, then the allocation described in Case 2 is an optimal solution of the WR-problem. ■

Finally note that Case 1 and Case 2 described above are the only possible allocations consistent with $\lambda = 0$. So, if $\mu_M < \mu^* (\alpha)$, the Largrange multiplier of $IC_{HL}$ must be binding.

### 6 Proof of Proposition 6

The result follows from Lemma A9. ■
7 Proof of Proposition 7

We first prove a useful lemma.

**Lemma A10.** The optimal solution satisfies: 
$q_L \leq \theta_L - \frac{\mu_H + \mu_M}{\mu_L} \Delta \theta$, 
$q_L(L) \leq \theta_L - \frac{\mu_H + \mu_M}{\mu_L} \frac{3\alpha - 1}{2\alpha} \Delta \theta$

and 
$q_L(M) \leq \theta_L$.

**Proof.** We proceed in 3 steps.

**Step 1.** Suppose $q_L > \theta_L - \frac{\mu_H + \mu_M}{\mu_L} 2 \Delta \theta$. Now, decrease $q_L$ by $\epsilon$. All the constraints are still satisfied. The change in the monopolist’s profit is given by

$$
\Delta \theta \mu_L \left( -\theta_L \epsilon - \frac{1}{2} \left( (q_L - \epsilon)^2 - (q_L)^2 \right) \right) + (\mu_H + \mu_M) \Delta \theta \epsilon 
$$

which is greater than zero for small enough $\epsilon$, giving us a contradiction.

**Step 2.** Suppose $q_L(L) > \theta_L - \frac{\mu_H + \mu_M}{\mu_L} \frac{3\alpha - 1}{2\alpha} \Delta \theta$. Now, decrease $q_L(L)$ by $\epsilon$ and $\omega_{ML}(L)$ by $\Delta \theta \epsilon$. All the constraints are still satisfied. The change in the monopolist’s profit is given by

$$
\Delta \theta \mu_L \left( -\theta_L \epsilon - \frac{1}{2} \left( (q_L(L) - \epsilon)^2 - (q_L(L))^2 \right) \right) + (\mu_H + \mu_M) \Delta \theta \epsilon 
$$

which is greater than zero for small enough $\epsilon$, giving us a contradiction.

**Step 3.** Suppose $q_L(M) > \theta_L$. Now, decrease $q_L(M)$ by $\epsilon$ and $\omega_{ML}(M)$ by $\Delta \theta \epsilon$. All the constraints are still satisfied. The change in the monopolist’s profit is given by

$$
\Delta \theta \mu_M \delta \left( -\theta_L \epsilon - \frac{1}{2} \left( (q_L(M) - \epsilon)^2 - (q_L(M))^2 \right) \right) = \mu_M \delta \frac{1 - \alpha}{2} \left( (q_L(M) - \theta_L) \epsilon - \frac{1}{2} \epsilon^2 \right),
$$

which is greater than zero for small enough $\epsilon$, giving us a contradiction. ■

Keep in mind that $\lambda > 0 \Rightarrow \lambda_{HM}(L) > 0$. It follows from the first order condition with respect to $\omega_{HM}(L)$. Next, in order to characterize the quantities after history $M$, we prove a useful lemma.

**Lemma A11.** $\lambda > 0 \Rightarrow \lambda_{HM}(M) > 0$.

**Proof.** Assume by contradiction that $\lambda_{HM}(M) = 0$. Then, we must have $\lambda_{ML}(M) = \lambda_{LM}(M) = 0$. Assuming them strictly positive gives us $q_M(M) = q_L(M)$. Also, from the first order condition for $q_M(M)$, we obtain $q_M(M) > \theta_M$, implying $q_L(M) > \theta_M > \theta_L$, a contradiction to Lemma A9. Thus, $\lambda = \mu_H$ and $q_M = q_M(M) = \theta_M$.

Next, we note that if $\lambda > 0$, then $q_M(M) < \theta_M$. To see this point, consider the first-order condition with respect to $q_M(L)$. Since, $\lambda_{LM}(L) > 0$, if $\lambda_{LM}(L) = 0$ then it follows immediately that $q_M(L) < \theta_M$. If $\lambda_{LM}(L) > 0$, then $q_M(L) = q_L(L) < \theta_L < \theta_M$, where the first inequality follows from Lemma A10.
Using these facts, we can now write:

\[
\Delta \theta q + \delta \frac{3\alpha - 1}{2} \omega_M(M) = \Delta \theta \cdot \theta_M + \delta \frac{3\alpha - 1}{2} \omega_M(M) \geq \Delta \theta \cdot \theta_M + \delta \frac{3\alpha - 1}{2} \Delta \theta q_M(M) \tag{12}
\]

\[
= \Delta \theta \cdot \theta_M + \delta \frac{3\alpha - 1}{2} \Delta \theta q_M \geq \Delta \theta q_L + \delta \frac{3\alpha - 1}{2} \omega_M(L).
\]

The strict inequality proven in (12) contradicts \( \lambda > 0 \). Thus, we must have \( \lambda_M > 0 \) as requested. This completes the proof of Lemma A11.

We divide the proof of Proposition 4 into two steps. First, in Section 7.1, we assume that \( IC_{LM}(L) \) is not binding and we characterize the parameter region in which this assumption is correct. This will allow us to define the regions B1 and B2 described in Proposition 7. Then, in Section 7.2, we characterize the optimal contract when \( IC_{LM}(L) \) is binding, region B3.

### 7.1 Characterization of Regions B1 and B2

Let us assume \( \lambda_M(L) = 0 \). Since \( \mu_M < \mu^*(\alpha) \), we have \( \lambda > 0 \). From the first order conditions, we obtain:

\[
q_M = \theta_M - \frac{\mu_H - \lambda}{\mu_M} \Delta \theta, \quad q_L = \theta_L - \frac{\mu_H + \mu_M + \lambda}{\mu_L} \Delta \theta \tag{13}
\]

\[
q_M(L) = \theta_M - \frac{\lambda}{\mu_L} \frac{3\alpha - 1}{2(1 - \alpha)} \Delta \theta, \quad q_L(L) = \theta_L - \frac{\mu_H + \mu_M}{\mu_L} \frac{3\alpha - 1}{2\alpha} \Delta \theta \tag{14}
\]

Since \( \lambda > 0 \), we have \( \lambda_{HM}(M) > 0 \) and \( \lambda_{HM}(L) > 0 \). Thus,

\[
q_M + \delta \frac{3\alpha - 1}{2} q_M(M) = q_L + \delta \frac{3\alpha - 1}{2} q_M(L) \tag{15}
\]

There are two relevant cases. We use \( \lambda_1 \) to denote \( \lambda \) from Case 1 and \( \lambda_2 \) from Case 2.

**Case 1 (Region B1).** \( \lambda_{ML}(M) = \lambda_{LM}(M) = 0 \). Then, from the first-order conditions:

\[
q_M(M) = \theta_M - \frac{\mu_H - \lambda_1}{\mu_M} \frac{3\alpha - 1}{2} \Delta \theta \quad \text{and} \quad q_M(L) = \theta_L
\]

Substituting, the values from (13)-(14) and (16) in equation (15) we obtain:

\[
\frac{1 + \lambda_1}{\mu_L} + \delta \frac{3\alpha - 1}{2} \frac{\lambda_1}{\mu_L} \frac{3\alpha - 1}{2(1 - \alpha)} = \frac{\mu_H - \lambda_1}{\mu_M} + \delta \frac{3\alpha - 1}{2} \frac{\mu_H - \lambda_1}{\mu_M} \frac{3\alpha - 1}{2\alpha} \tag{17}
\]

which gives:

\[
\lambda_1 = \lambda_1(\alpha) = \frac{\mu_H}{\mu_M} \left( 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{2\alpha} \right) - \frac{1}{\mu_L} \frac{1}{\mu_M} \left( 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{2\alpha} \right) \tag{18}
\]
Clearly, for this case to be valid, we must justify the assumption that \( \lambda_{ML}(M) = \lambda_{LM}(M) = 0 \). A necessary and sufficient condition for this is \( q_M(M) \geq q_L(M) \). Given (16), this condition can be rewritten as: \( \frac{\mu_H - \lambda_1 3\alpha - 1}{2\alpha} \leq 1 \), where \( \lambda_1 \) is given by (18). This condition is implied by:

\[
\mu_M \geq \frac{1 + (1 - \mu_L) b_0(\alpha) - \mu_L c_0(\alpha) a_0(\alpha)}{b_0(\alpha) (1 + c_0(\alpha))} = \mu_0(\alpha),
\]

where

\[
a_0(\alpha) = 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{2\alpha}, \quad b_0(\alpha) = 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{1 - \alpha}, \quad \text{and} \quad c_0(\alpha) = \frac{2\alpha}{3\alpha - 1}
\]

It follows that (under the assumption that \( \lambda_{LM}(L) = 0 \)) the solution is given by (13)-(14), (16) and (18) when \( \mu_M \geq \mu_0(\alpha) \).

**Case 2 (Region B2).** For \( \mu_M < \mu_0(\alpha) \) we must have \( \lambda_{ML}(M) = \lambda_{LM}(M) > 0 \). In this case, we must have:

\[
q_M(M) = q_L(L) = \frac{2\alpha}{1 + \alpha} \theta_M + \frac{1 - \alpha}{1 + \alpha} \theta_L - \frac{\mu_H - \lambda_2 3\alpha - 1}{\mu_M} \Delta \theta
\]

Substituting \( q_M(M) \) and \( q_M(L) \) equation (15) we obtain:

\[
\frac{1 + \lambda_2}{\mu_L} + \delta \frac{3\alpha - 1}{2} \left( \frac{\lambda_2}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} - \frac{1 - \alpha}{1 + \alpha} \right) = \frac{\mu_H - \lambda_2}{\mu_M} + \delta \frac{3\alpha - 1}{2} \frac{\mu_H - \lambda_2 3\alpha - 1}{\mu_M} \frac{1 + \alpha}{1 + \alpha}
\]

which gives

\[
\lambda_2 = \frac{\frac{\mu_H}{\mu_M} \left( 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{2\alpha} \right) - \left( \frac{1}{\mu_L} - \delta \frac{3\alpha - 1}{2} \frac{1 - \alpha}{1 + \alpha} \right)}{\frac{1}{\mu_M} \left( 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{1 - \alpha} \right) + \frac{1}{\mu_L} \left( 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{1 - \alpha} \right)}
\]

It follows that (under the assumption that \( \lambda_{LM}(L) = 0 \)) the solution is given by (13)-(14), (19) and (21) when \( \mu_M < \mu_0(\alpha) \).

We now complete the analysis of this section by characterizing the conditions under which we can ignore the \( IC_{LM}(L) \) constraint and so \( \lambda_{LM}(L) = 0 \). It is easy to see that \( IC_{LM}(L) \) is satisfied if and only if \( q_M(L) \geq q_L(L) \). We have \( q_M(L) \geq q_L(L) \) if and only if:

\[
\lambda_1 \leq \left( \frac{1}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} \right)^{-1} \left( 1 + \frac{1 - \mu_L}{\mu_L} \frac{3\alpha - 1}{2\alpha} \right)
\]

Thus, for Case 1 we have,

\[
\frac{\frac{\mu_H}{\mu_M} \left( 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{2\alpha} \right) - \frac{1}{\mu_L}}{\frac{1}{\mu_M} \left( 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{2\alpha} \right) + \frac{1}{\mu_L} \left( 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{1 - \alpha} \right)} \leq \left( \frac{1}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} \right)^{-1} \left( 1 + \frac{1 - \mu_L}{\mu_L} \frac{3\alpha - 1}{2\alpha} \right)
\]

Define

\[
a_1(\alpha, \mu_L) = 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{2\alpha}, \quad b_1(\alpha, \mu_L) = 1 + \delta \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{1 - \alpha}, \quad \text{and}
\]

12
\[ c_1(\alpha, \mu_L) = \left( \frac{1}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} \right)^{-1} \left( 1 + \frac{1 - \mu_L}{\mu_L} \frac{3\alpha - 1}{2\alpha} \right) \]

We can then write the previous inequality as:

\[ \mu_M \geq \frac{\mu_L a_1(\alpha, \mu_L)[1 - \mu_L - c_1(\alpha, \mu_L)]}{1 + a_1(\alpha, \mu_L) \mu_L + b_1(\alpha, \mu_L) c_1(\alpha, \mu_L)} = \mu_1^{**}(\alpha) \]

Next, for Case 2, we have:

\[ \frac{\mu_H}{\mu_M} \left( 1 + \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{1 + \alpha} \right) - \left( \frac{1}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} \right)^{-1} \left( 1 + \frac{1 - \mu_L}{\mu_L} \frac{3\alpha - 1}{2\alpha} \right) \leq \frac{L}{c} \left( 1 + \frac{1 - \mu_L}{\mu_L} \frac{3\alpha - 1}{2\alpha} \right) \]

Define

\[ a_2(\alpha, \mu_L) = 1 + \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{1 + \alpha}, \quad b_2(\alpha, \mu_L) = \frac{1}{\mu_L} - \frac{3\alpha - 1}{2} \frac{1 - \alpha}{1 + \alpha} \]

\[ c_2(\alpha, \mu_L) = \left( \frac{1}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} \right)^{-1} \left( 1 + \frac{1 - \mu_L}{\mu_L} \frac{3\alpha - 1}{2\alpha} \right), \quad d_2(\alpha, \mu_L) = 1 + \frac{3\alpha - 1}{2} \frac{3\alpha - 1}{1 - \alpha} \]

Rearranging, we obtain:

\[ \mu_M \geq \frac{\mu_L a_2(\alpha, \mu_L)[1 - \mu_L - c_2(\alpha, \mu_L)]}{1 + a_2(\alpha, \mu_L) \mu_L + b_2(\alpha, \mu_L) c_2(\alpha, \mu_L)} = \mu_2^{**}(\alpha) \]

Let us define \( \mu^{**} (\alpha) = \min \{ \mu^* (\alpha), \mu_1^{**} (\alpha), \mu_2^{**} (\alpha) \} \). We have:

**Lemma A12.** If \( \mu_M \in [\mu^{**} (\alpha), \mu^* (\alpha)] \) and \( \mu_M \geq \mu_0(\alpha) \), then the solution of the WR-problem is given by the solution in Case 1 presented above. If \( \mu_M \in [\mu^{**} (\alpha), \mu^* (\alpha)] \) and \( \mu_M < \mu_0(\alpha) \), then the solution of the WR-problem is given by the solution in Case 2 presented above.

**Proof.** We first show that if \( \mu_M \in [\mu^{**} (\alpha), \mu^* (\alpha)] \) and \( \mu_M \geq \mu_0(\alpha) \), then \( \mu_M \in [\mu_1^{**} (\alpha), \mu^* (\alpha)] \) and \( \mu_M \geq \mu_0(\alpha) \). This implies that the solution is given by Case 1. Assume \( \mu_M < \mu_1^{**} (\alpha) \). In this case, (22) does not hold with \( \lambda_1 \). This implies that (22) does not hold with \( \lambda_2 \) as well if \( \lambda_2 \geq \lambda_1 \). Subtracting equation (20) from equation (17), we get

\[ (\lambda_1 - \lambda_2) \left[ \frac{1}{\mu_L} + \frac{1}{\mu_M} + \frac{3\alpha - 1}{2} \left( \frac{3\alpha - 1}{\mu_L 1 - \alpha} + \frac{3\alpha - 1}{\mu_M 1 + \alpha} \right) \right] = \delta \left[ \frac{3\alpha - 1}{2} \frac{1 - \alpha}{1 + \alpha} \left( \frac{\mu_H - \lambda_1}{\mu_M} \frac{3\alpha - 1}{2\alpha} - 1 \right) \right] \]

So, we have that \( \lambda_2 \geq \lambda_1 \) if:

\[ \frac{\mu_H - \lambda_1}{\mu_M} \frac{3\alpha - 1}{2\alpha} - 1 \leq 0, \]

which is implied by \( \mu_M \geq \mu_0(\alpha) \). It follows that if \( \mu_M < \mu_1^{**} (\alpha) \), then \( \mu_M < \mu^{**} (\alpha) \), a contradiction. We conclude that it must be \( \mu_M \geq \mu_1^{**} (\alpha) \).

We now show that if \( \mu_M \in [\mu^{**} (\alpha), \mu^* (\alpha)] \) and \( \mu_M < \mu_0(\alpha) \), then \( \mu_M \in [\mu_2^{**} (\alpha), \mu^* (\alpha)] \) and \( \mu_M < \mu_0(\alpha) \). This implies that the solution is given by Case 2. Assume \( \mu_M < \mu_2^{**} (\alpha) \). In this case, (22) does not hold with \( \lambda_2 \). This implies that (22) does not hold with \( \lambda_1 \) as well if \( \lambda_1 \geq \lambda_2 \). From (23) we have that this always true if \( \mu_M < \mu_0(\alpha) \). It follows that if \( \mu_M < \mu_2^{**} (\alpha) \), then \( \mu_M < \mu^{**} (\alpha) \), a contradiction. We conclude that it must be \( \mu_M \geq \mu_2^{**} (\alpha) \).  

\[ \square \]
7.2 Characterization of Region B3

Finally, we characterize the contract when $\mu_M < \mu^{**}(\alpha)$ and so both $\lambda > 0$ and $\lambda_{LM}(L) > 0$. This is region B3. In this case:

$$q_M = \theta_M - \frac{\mu_H - \lambda}{\mu_M} \Delta \theta \quad \text{and} \quad q_L = \theta_L - \frac{\mu_H + \mu_M + \lambda}{\mu_L} \Delta \theta$$

(24)

We also have that $\lambda_{LM}(L) > 0$ implies $q_M(L) = q_L(L)$, so:

$$q_M(L) = q_L(L) = \frac{1 - \alpha}{1 + \alpha} \theta_M + \frac{2\alpha}{1 + \alpha} \theta_L - \frac{\mu_H + \mu_M + \lambda}{\mu_L} \frac{3\alpha - 1}{2\alpha} \Delta \theta$$

(25)

From Lemma A9, we have $q_L(L) \leq \theta_L - \frac{\mu_H + \mu_M}{\mu_L} \frac{3\alpha - 1}{2\alpha} \Delta \theta$. Also, when $\lambda_{LM}(L) > 0$, the above inequality is strict. Thus, substituting the optimal value of $q_L(L)$, we obtain:

$$1 - \frac{\lambda}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} + \frac{\mu_H + \mu_M}{\mu_L} \frac{3\alpha - 1}{2\alpha} \Delta \theta < 0$$

(26)

Note that as $\lambda_{LM}(L)$ converges to zero, (26) is the exact violation of $\mu_M \geq \mu^{**}(\alpha)$, that is, inequality (22).

To characterize the quantities after history $M$, we now show that $\lambda_{ML}(M) = \lambda_{LM}(L) > 0$.

**Lemma A13.** $\lambda, \lambda_{LM}(L) > 0 \Rightarrow \lambda_{ML}(M) = \lambda_{LM}(L) > 0$.

**Proof.** Suppose $\lambda_{ML}(M) = \lambda_{LM}(M) = 0$. Then,

$$q_M(M) = \theta_M - \frac{\mu_H - \lambda}{\mu_M} \frac{3\alpha - 1}{2\alpha} \Delta \theta \quad \text{and} \quad q_L(M) = \theta_L.$$

From $\theta_M - \frac{2\alpha}{3\alpha - 1} \frac{\mu_H - \lambda}{\mu_M} \Delta \geq \theta_L$, we have:

$$\frac{2\alpha}{3\alpha - 1} \frac{\mu_H - \lambda}{\mu_M} \geq 0.$$  

(27)

Since $\lambda, \lambda_{LM}(L) > 0$, using $q_M(M) \geq q_L(M) = \theta_L > q_L(L) = q_M(L)$, we get $q_L > q_M$. This implies

$$\left(1 - \frac{\mu_H - \lambda}{\mu_M} + \frac{\mu_H + \mu_M + \lambda}{\mu_L}\right) < 0.$$

Using equation (27), we get

$$\frac{\mu_H + \mu_M + \lambda}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} < 1.$$  

(28)

Now, inequality (26) can be written as

$$1 < \frac{\mu_H + \mu_M + \lambda}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} - \frac{\mu_H + \mu_M}{\mu_L}(3\alpha - 1) \left(\frac{1}{1 - \alpha} + \frac{1}{2\alpha}\right)$$

$$= \frac{\mu_H + \mu_M + \lambda}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} - \frac{\mu_H + \mu_M}{\mu_L} \frac{3\alpha - 1 + 1 + \alpha}{2\alpha}$$

14
which contradicts condition (28). ■

It follows that
\[ q_M(M) = q_L(M) = \frac{2\alpha}{1 + \alpha} \theta_M + \frac{1 - \alpha}{1 + \alpha} \theta_L - \frac{\mu_H - \lambda}{\mu_M} \left( 3\alpha - 1 \right) \Delta \theta. \]

Finally, substituting the optimal values in IC\(_{HL}\) as equality, we obtain:
\[ \left( 1 - \frac{\mu_H - \lambda}{\mu_M} \right) + \frac{\mu_H + \mu_M + \lambda}{\mu_L} = 0 \] (29)

that implies \( q_M = q_L \). Note that equation (29) gives the value of \( \lambda \), which uniquely defines the solution at the optimum. In particular note that type \( M \) and \( L \) are treated as one, that is,
\[ q_M = q_L \] and \( q_M(M) = q_M(L) = q_L(L) \) (30)

We conclude that the solution of the \( WR\)-problem in region B3 \((\mu_M < \mu^{**}(\alpha))\) is given by: (24),(25), (30) and (29). ■

Table 1 summarizes the solution of the \( WR\)-problem describing the optimal allocation for each possible case.

8 Proof of Lemma 3

We prove the lemma as follows. Let \( U = U(h^t) \) be the vector of expected utilities, mapping an history \( h^t \) to the corresponding agent’s expected utility. First, we construct a vector of utilities \( U \) using the solution of the \( WR\)-problem, \( (\omega, q) \). We then show that the solution \( (U, q) \) satisfies all the constraints of the seller’s profit maximization problem and it maximizes profits. We proceed in two steps:

Step 1. We set \( u_L(M), u_L(L), u_L(H) \) all equal to zero. We also define:
\[ u_M(M) = \omega_{ML}(M), u_M(L) = \omega_{ML}(L), u_M(H) = \Delta \theta q_L(H) \]
\[ u_H(M) = \omega_{ML}(M) + \omega_{HM}(M), u_H(L) = \omega_{ML}(L) + \omega_{HM}(L), u_H(H) = \Delta \theta (q_L(H) + q_M(H)) \]

Since \( IR_L, IC_{ML} \) and \( IC_{HM} \) hold as an equality, we must have:
\[ U_L = 0, \]
\[ U_M = \Delta \theta q_L + \delta \frac{3\alpha - 1}{2} \omega_{ML}(L), \text{ and} \]
\[ U_H = U_M + \Delta \theta q_M + \delta \frac{3\alpha - 1}{2} \omega_{HM}(M). \]

Step 2. We now show that \( (U, q) \) satisfies all the constraints of the profit maximizing problem. By construction it is immediate that \( (U, q) \) satisfies all the constraints in the \( WR\)-problem.
remains to be shown that it also satisfies the other constraints,

\[ IR_H, IR_M, IC_{MH}, IC_{LM}, IC_{LH}, \]

\[ IC_{HM}(H), IC_{ML}(H), IR_L(H), IR_L(M), IR_L(L) \]

\[ IC_{MH}(H), IC_{LM}(H), IC_{LH}(H), IC_{HL}(H), IC_{MH}(M), \]

\[ IC_{LH}(M), IC_{HL}(M), IC_{MH}(L), IC_{LH}(L), IC_{HL}(L). \]

First, we show that \( IR_M \) is satisfied. From \( IC_{ML} \) we have

\[
U_M = U_L + \Delta \theta q_L + \delta \frac{3\alpha - 1}{2} [u_M(L) - u_L(L)]
\]

\[
= \Delta \theta q_L + \delta \frac{3\alpha - 1}{2} [u_M(L) - u_L(L)] \quad [\text{Using } IR_L]
\]

\[
\geq \Delta \theta q_L + \delta \frac{3\alpha - 1}{2} \Delta q_L(L) > 0 \quad [\text{Using } IC_{ML}(L)]
\]

Similarly, we can show that \( IR_H \) is satisfied. To prove the remaining constraints we need the following properties of the solution of the WR-problem.

**Lemma A14.** For all parameter configurations, in the solution to the WR-problem we have: 1. \( q_i(H) = \theta \), for \( i = M, L, H \), \( q_M(M) < \theta_M, q_L(M) \leq \theta_L \), and \( q_L(L) \geq q_L(L) \). 2. \( \omega_{HM}(M) = \Delta \theta q_M(M) \) and, without loss of generality, \( \omega_{ML}(M) = \Delta \theta q_L(M) \), \( \omega_L(L) = \Delta \theta q_M(L) \); 3. quantities at \( t = 2 \) are nondecreasing in type after any history; 4. \( q_H \geq q_M \geq q_L \).

**Proof.** Point 1 follow from the solution characterized in Propositions 6 and 7 (for convenience the quantities are reported in Table 1). The first part of Point 2 \( (IC_{HM}(M) \) always binds) follows from Lemma 4 (when \( \lambda = 0 \) and Lemma A11 (when \( \lambda > 0 \)). The second part follows from the fact that \( IC_{ML}(M) \) can be assumed to hold as an equality. Suppose \( \omega_{ML}(M) > \Delta \theta q_L(M) \). Then can decrease \( \omega_{ML}(M) \) so that this holds as an equality. No constraint is violated and the profit of the monopolist is unaffected. Similarly, we show that \( IC_{HM}(L) \) can be assumed to hold as an equality, implying \( \omega_{HM}(L) = \Delta \theta q_M(L) \). Point 3 follows from incentive compatibility constraints for the second (terminal) period. We now turn to Point 4. From the fact that in the solution to the WR-problem, \( q_i = \theta \) and the fact that (as shown in Propositions 6 and 7) \( q_i \leq \theta_i \) for \( i = H, M, L \), we have \( q_H \geq q_i \) \( i = M, L \). We, therefore, only need to prove that \( q_M \geq q_L \). We will show this result case by case for all regions A1, A2, B1, B2 and B3. In cases A1 and A2, from (4) we have \( q_M \geq q_L \) if and only if

\[
1 - \frac{\mu_H}{\mu_M} + \frac{\mu_H + \mu_M}{\mu_L} \geq 0,
\]

that is, \( \frac{1}{\mu_L} \geq \frac{\mu_H}{\mu_M} \). In regions A1 and A2 we have \( \mu_M \geq \mu^*(\alpha) \), as defined in Lemma 4. This condition can be written as

\[
\frac{1}{\mu_L} \geq \frac{\mu_H}{\mu_M} + \delta \frac{3\alpha - 1}{2} \frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{2\alpha} \quad \text{and} \quad \frac{1}{\mu_L} \geq \frac{\mu_H}{\mu_M} + \delta \frac{3\alpha - 1}{2} \left( \frac{1 - \alpha}{1 + \alpha} + \frac{\mu_H}{\mu_M} \frac{3\alpha - 1}{1 + \alpha} \right).
\]

clearly implying \( \frac{1}{\mu_L} \geq \frac{\mu_H}{\mu_M} \). For case B3, we show in Proposition 7 that \( q_M = q_L \). We now show that in regions B1 and B2 we have \( q_M \geq q_L \) as well. In these region we have \( \mu \in [\mu^*(\alpha), \mu^*(\alpha)] \).
We have $q_M \geq q_L$ if and only if $1 - \frac{\mu_H - \lambda}{\mu_M} + \frac{\mu_H + \mu_M + \lambda}{\mu_L} \geq 0$. First order conditions in Lemma 4 clearly show that $\lambda > 0$ implies $\lambda_{HM}(L) > 0$, thus, $\omega_{HM}(L) = \Delta q_M(L)$. Therefore, we have in regions B1 and B2,

$$q_M + \frac{3\alpha - 1}{2} q_M(M) = q_L + \frac{3\alpha - 1}{2} q_M(L).$$

When $\mu_M \geq \mu_0 (\alpha)$, substituting optimal values (summarized in Table 1) we have

$$1 - \frac{\mu_H - \lambda_1}{\mu_M} + \frac{\mu_H + \mu_M + \lambda_1}{\mu_L} + \frac{3\alpha - 1}{2} \left[ \frac{\lambda_1}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} - \frac{\mu_H - \lambda_1}{\mu_M} \frac{3\alpha - 1}{2\alpha} \right] = 0.$$

That can be re written as:

$$\left( 1 - \frac{\mu_H - \lambda_1}{\mu_M} + \frac{\mu_H + \mu_M + \lambda_1}{\mu_L} \right) \left( 1 + \delta \frac{(3\alpha - 1)^2}{4\alpha} \right) = \delta \frac{(3\alpha - 1)^2}{4\alpha} \left[ 1 + \frac{\mu_H + \mu_M}{\mu_L} - \frac{\lambda_1}{\mu_L} \frac{3\alpha - 1}{2\alpha} \right].$$

We know from (22) that right hand side of the above equation is non-negative. Thus, $1 - \frac{\mu_H - \lambda_1}{\mu_M} + \frac{\mu_H + \mu_M + \lambda_1}{\mu_L} \geq 0$.

When $\mu_M < \mu_0 (\alpha)$, substituting optimal values again (see Table 1) we have

$$1 - \frac{\mu_H - \lambda_2}{\mu_M} + \frac{\mu_H + \mu_M + \lambda_2}{\mu_L} + \frac{3\alpha - 1}{2} \left[ \frac{\lambda_2}{\mu_L} \frac{3\alpha - 1}{1 - \alpha} - \frac{\mu_H - \lambda_2}{\mu_M} \frac{3\alpha - 1}{1 + \alpha} \right] = 0.$$

That can be rewritten as:

$$\left( 1 - \frac{\mu_H - \lambda_2}{\mu_M} + \frac{\mu_H + \mu_M + \lambda_2}{\mu_L} \right) \left( 1 + \delta \frac{(3\alpha - 1)^2}{2(1 + \alpha)} \right) = \delta \frac{(3\alpha - 1)^2}{2(1 + \alpha)} \left[ 1 + \frac{\mu_H + \mu_M}{\mu_L} - \frac{\lambda_2}{\mu_L} \frac{3\alpha - 1}{2\alpha} \right].$$

We know that (22) is always verified in the relevant range. Using this condition we can see that right hand side of the above equation is non-negative. Thus, we have $1 - \frac{\mu_H - \lambda_2}{\mu_M} + \frac{\mu_H + \mu_M + \lambda_2}{\mu_L} \geq 0$. 

Consider the first period constraints. To show that $IC_{LM}$ holds it is sufficient to prove:

$$0 = U_L \geq \theta_L q_M + \delta \left[ \alpha u_L(L) + \frac{1 - \alpha}{2} u_M(M) + \frac{1 - \alpha}{2} u_H(M) \right] \quad (32)$$

$$= U_M - \Delta \theta q_M - \delta \frac{3\alpha - 1}{2} u_L(M)$$

$$= U_M - \Delta \theta q_M - \delta \frac{3\alpha - 1}{2} q_L(M)$$

Since $U_M = \Delta \theta q_M + \delta \frac{3\alpha - 1}{2} q_L(L)$, (32) can be written as:

$$q_M + \frac{3\alpha - 1}{2} q_M(M) \geq q_L + \frac{3\alpha - 1}{2} q_L(L)$$

The fact that this inequality is satisfied follows from Point 1 and 4 in Lemma A14. (In the following, when we mention a point, we refer to the points of Lemma A14.)

Next, we show that $IC_{MH}$ holds. From $IC_{HM}$ we have:

$$U_H = U_M + \Delta \theta q_M + \delta \frac{3\alpha - 1}{2} [u_H(M) - u_M(M)]$$

17
Thus,

\[ U_M = U_H - \Delta \theta q_M - \delta \frac{3\alpha - 1}{2} [u_H(M) - u_M(M)] \]
\[ = U_H - \Delta \theta q_H - \delta \frac{3\alpha - 1}{2} [u_H(H) - u_M(H)] \]
\[ + \Delta \theta (q_H - q_M) + \delta \frac{3\alpha - 1}{2} [(u_H(H) - u_M(H)) - (u_H(M) - u_M(M))] \]
\[ > U_H - \Delta \theta q_H - \delta \frac{3\alpha - 1}{2} [u_H(H) - u_M(H)]. \]

The last inequality follows from the observation that:

\[ u_H(H) - u_M(H) \geq \Delta \theta q_M(H) = \Delta \theta q_M(M) = u_H(M) - u_M(M), \tag{33} \]

where the first inequality follows from the definition of \( u_i(H) \), the first equality and the second inequality follow from Point 1. From (33) and the fact that \( q_H > q_M \) (Point 4), it follows that \( IC_{MH} \) holds. We now turn to \( IC_{LH} \). Using \( IC_{LM} \) first and then \( IC_{MH} \), we have:

\[ U_L \geq U_M - \Delta \theta q_M - \delta \frac{3\alpha - 1}{2} [u_M(M) - u_L(M)] \]
\[ \geq U_H - \Delta \theta q_H - \delta \frac{3\alpha - 1}{2} [u_H(H) - u_M(H)] - \Delta \theta q_M - \delta \frac{3\alpha - 1}{2} [u_M(M) - u_L(M)] \]
\[ = U_H - 2\Delta \theta q_H - \delta \frac{3\alpha - 1}{2} [u_H(H) - u_M(H)] \]
\[ + \Delta \theta (q_H - q_M) + \delta \frac{3\alpha - 1}{2} [(u_M(H) - u_L(H)) - (u_M(M) - u_L(M))] \]
\[ > U_H - 2\Delta \theta q_H - \delta \frac{3\alpha - 1}{2} [u_H(H) - u_L(H)]. \]

The last inequality follows from the observation that:

\[ u_M(H) - u_L(H) \geq \Delta \theta q_L(H) = \Delta \theta q_L(M) = u_M(M) - u_L(M), \tag{34} \]

where the first inequality follows from the definition of \( u_i(H) \), the first equality and the second inequality follow from Point 1. From (34) and \( q_H > q_M \) (Point 4), it follows that \( IC_{LH} \) holds.

Consider now the second period constraints. The constraints \( IR_L(M) \), \( IR_L(L) \), \( IR_L(H) \), \( IC_{ML}(H) \), and \( IC_{HM}(H) \) follow immediately by the definition of the utilities at \( t = 2 \). The proof that \( \langle U, q \rangle \) solves the seller’s problem is therefore completed if we prove that it satisfies the constraints in the last two lines of (31). This result follows from the fact that the local downward incentive constraints are satisfied in period 2 and quantities are weakly monotonic after any history (Point 3). Finally, to see that the contract is optimal, we note that it maximizes expected profits in the less restricted \( WR \)-problem, so it must be optimal in the seller’s problem. Note moreover that since the original problem is concave in \( q \) this is in fact the unique solution (in quantities).