Motivating Cost Reduction in Regulated Industries with Rolling Incentive Schemes

by

Douglas C. Turner* and David E. M. Sappington**

Abstract

We examine whether an incremental rolling incentive scheme (IRIS) can enhance innovation under performance based regulation (PBR). Under an IRIS, the firm is awarded for a full PBR term the incremental profit generated by a cost reduction regardless of when the cost reduction is implemented. An IRIS enhances incentives for innovation toward the end of a PBR plan and also ensures immediate implementation of achieved cost reductions. However, an IRIS can reduce incentives for innovation early in a PBR plan. On balance, an IRIS often reduces innovation when the regulated firm's ability to delay the implementation of achieved cost reductions is limited. More generally, an IRIS can often enhance innovation.

Keywords: incentive regulation; rolling incentive schemes.

JEL Codes: L51, L97, O38

October 2024

* Department of Economics, University of Florida
PO Box 117140, Gainesville, Florida 32611 USA (douglasturner@ufl.edu).

** Department of Economics, University of Florida
PO Box 117140, Gainesville, Florida 32611 USA (sapping@ufl.edu).

1 Introduction.

Performance based regulation (PBR) is now commonplace in regulated industries throughout the world.[1] A central goal of many PBR plans is to motivate the regulated firm to reduce its production costs. This goal is often pursued by implementing what is known as *standard rebasing* (SR). Under SR, the prices a firm charges for its services are not re-set to match realized production costs until the term of the prevailing PBR regime has expired. Consequently, if the firm implements a cost reduction during a PBR regime, the firm is permitted to retain the associated increase in profit for the remainder of the PBR regime. This potential to secure profit during the prevailing PBR regime enhances the firm's incentive to reduce its costs relative to a policy that immediately re-sets the firm's prices to match its realized costs.

Although SR provides the regulated firm with some incentive to reduce its costs, this incentive can be limited. Specifically, as the end of a PBR regime draws near, the regulated firm realizes that any cost reduction it achieves will soon trigger a corresponding reduction in prices. Consequently, the firm anticipates little profit from – and so has limited incentive to pursue and implement – a cost reduction toward the end of a PBR regime under SR.

To address this well-known drawback to SR, some regulators have implemented what the New Zealand Commerce Commission calls an *incremental rolling incentive scheme* (IRIS).[2] Under IRIS,[3] the length of time for which the regulated firm benefits from an achieved cost reduction is the same, regardless of when the cost reduction is implemented. To illustrate how this constant "regulatory lag" is ensured under IRIS, suppose a PBR regime lasts for five years. If the firm implements a cost reduction at the start of year 4 in the five-year regime under SR, the firm only secures the associated increase in profit during years 4 and 5 of the regime. The firm's prices are re-set to match its observed costs – and so any incremental profit engendered by the cost reduction is eliminated – at the start of the next PBR regime. In contrast, under IRIS, the firm secures for a full five years the increase in profit associated with a cost reduction, regardless of the year in which the cost reduction is implemented. In particular, if the firm implements a cost reduction in year 4 of the prevailing five-year PBR regime, the firm's prices are not ratcheted down to reflect the achieved cost reduction

[1] See, for example, Sappingon and Weisman (2010, 2016), Joskow (2014, 2024), and Wilson et al. (2022). The terms "performance based regulation" and "incentive regulation" are often used interchangeably.

[2] See Frontier Economics (2015) and the New Zealand Commerce Commission (2018), for example. The Australian Energy Regulator (2008, 2013) refers to an IRIS as an *efficiency benefit sharing scheme*. Oxera (2021, footnote 16) reports that an early rolling incentive scheme of this type was employed in "price reviews PR04 and PR09 by Ofwat, the water regulator for England and Wales." A rolling incentive scheme was also employed in the fourth phase of PBR in the UK electricity sector (Ofgem, 2009, pp. 37-38, 53-54).

[3] For expositional ease, we will use the phrase "under IRIS" to denote "under an IRIS."

until the third year of the subsequent PBR regime has concluded.[4] This delayed ratcheting ensures that the firm retains the incremental profit generated by the cost reduction for five years,[5] regardless of whether the cost reduction is implemented at the very start or near the end of the prevailing PBR regime.

The purpose of this research is to compare the incentives for cost-reducing innovation and the corresponding levels of expected consumer welfare under SR and IRIS. We find that, as it is designed to do, IRIS enhances incentives for innovation toward the end of a PBR regime. However, IRIS can also diminish incentives for innovation toward the start of a PBR regime. Diminished incentives for "early" innovation can arise under IRIS because the firm recognizes that if it fails to achieve an early innovation, it can still enjoy the full benefit of an innovation achieved later in the PBR regime. The associated reduction in the perceived urgency of achieving an early innovation can diminish early innovation under IRIS (and thereby reduce expected consumer welfare). In this sense, although IRIS can ameliorate the problem of limited innovation that arises late in a PBR regime under SR, IRIS can aggravate the problem of limited innovation that arises early in the PBR regime under SR. Consequently, it is not apparent whether consumer welfare is higher when the firm operates under SR or under IRIS.

We find that SR the discounted present value of expected consumer welfare ($E_d\{W\}$) often is higher under SR than under IRIS when the regulated firm cannot or chooses not to delay the implementation of an achieved cost reduction under SR.[6] This superior performance of SR reflects two drawbacks to IRIS. First, IRIS implements a relatively long lag between when a cost reduction is implemented and when consumer prices are reduced to reflect the cost reduction. Second, IRIS reduces incentives for early innovation for the reason explained above.

In contrast, IRIS often secures a higher $E_d\{W\}$ than does SR when the regulated firm delays the implementation of an achieved cost reduction to the next PBR regime under SR.[7] IRIS eliminates this delay and also enhances innovation late in a PBR regime. These advantages of IRIS often outweigh its disadvantage (i.e., limited incentive for early innovation) when the firm delays the implementation of an achieved cost reduction under SR. In

[4] For expositional ease, we assume that if the firm implements a cost reduction in year t of a PBR regime, the firm does so at the start of year t.

[5] In the present example, the firm secures the incremental profit generated by the cost reduction during years 4 and 5 of the prevailing PBR regime and years 1, 2, and 3 of the subsequent PBR regime, for a total of five years.

[6] The firm will become more inclined to implement an achieved cost reduction immediately under SR as the start of the next PBR becomes more distant and as the firm's relative valuation of future profit declines.

[7] This will often be the case toward the end of a PBR regime.

such settings, IRIS secures a higher $E_d\{W\}$ than does SR when the discount factor is sufficiently large, when innovation is sufficiently onerous for the firm, and when the gains from innovation are sufficiently small. When the discount factor is large, IRIS's advantage in enhancing innovation late in a PBR regime is weighted more heavily when calculating $E_d\{W\}$. When innovation is sufficiently onerous or when the incremental profit from successful innovation is sufficiently small,[8] the firm implements a relatively small success probability late in a PBR regime under IRIS. Anticipating this relatively small probability of late success, the firm becomes more highly motivated to achieve early success, thereby mitigating IRIS's disadvantage in inducing relatively limited early innovation.

In practice, a regulated firm often will be inclined to implement an achieved cost reduction immediately under SR when the innovation is achieved early in the PBR regime. In contrast, the firm will often be inclined to delay to the next PBR the implementation of an innovation achieved late in the current PBR regime. Therefore, the typical PBR regime includes some periods in which IRIS tends to promote a higher level of $E_d\{W\}$ than does SR, and other periods in which IRIS tends to promote a lower level of $E_d\{W\}$ than does SR. We document a range of stylized regulatory settings in which, on balance, $E_d\{W\}$ is higher under SR than under IRIS whenever each PBR regime consists of at least four periods in which the firm can achieve and implement a cost reduction. In practice, PBR regimes typically last for approximately five years (Sappington and Weisman, 2024).

We interpret our findings as suggesting that regulators should proceed with caution when considering whether to replace SR with IRIS. The policy that best serves consumers varies with many elements of the regulatory environment, including the number and nature of innovation opportunities, the firm's ability to strategically delay the implementation of achieved cost reductions, the prevailing discount factor, and the rate of industry demand growth. Replacing SR with IRIS can enhance consumer welfare in some settings, but it can reduce consumer welfare in many other settings.

The existing literature on regulatory lag focuses on the optimal length of a regulatory regime, assuming that SR is employed.[9] This literature observes in part that a longer PBR regime enhances a regulated firm's incentive to reduce its costs, but also delays price reductions to match realized cost reductions. In contrast to this literature, we take as given the length of the PBR regime and examine whether consumer welfare increases when an alternative to SR is employed. The alternative (IRIS) allows the firm to retain the profit

[8] The incremental profit from a reduction in marginal cost declines as the magnitude of the cost reduction or the level of demand for the firm's product declines.

[9] See, for example, Baumol and Klevorick (1970), Pint (1992), Biglaiser and Riordan (2000), Coco and De Vincenti (2004, 2005), Armstrong et al. (1995), and Armstrong and Sappington (2007).

increment engendered by a cost reduction for the (exogenous) length of the PBR regime, regardless of when during the regime the cost reduction is implemented.[10]

Our comparison of the performance of SR and IRIS proceeds as follows. Section 2 analyzes a stylized setting in which the regulated firm can achieve a cost reduction in one of two periods in a single PBR regime. Section 3 provides an analytic characterization of outcomes in this setting. Section 4 presents corresponding numerical characterizations. Section 5 considers alternative regulatory settings, including settings with multiple PBR regimes in which the firm can secure a cost reduction and PBR regimes with more than two periods. Section 6 summarizes our key findings, discusses policy implications, and suggests directions for future research. The Appendix provides the proofs of all formal conclusions in the text.

2 Model Elements

We first consider a simple setting in which three PBR regimes – PBR1, PBR2, and PBR3 – are imposed sequentially. Each PBR regime consists of two periods: PBR1 consists of periods 1 and 2; PBR2 consists of periods 3 and 4; PBR3 consists of periods 5 and 6.[11] All periods have the same length. $\delta \in (0,1)$ denotes the firm's inter-period discount factor.[12]

The regulated firm's marginal cost is $c_0 > 0$ prior to the start of PBR1. The firm has the opportunity to reduce its marginal cost by $\Delta \in (0, c_0)$ during PBR1. The firm can either implement the Δ cost reduction when it is first achieved or delay the implementation to a subsequent period.[13] Unless otherwise noted, the regulator does not become aware of an achieved cost reduction until it is implemented. For simplicity, we assume that if the cost reduction is achieved (or implemented) in period $t \in \{1,2\}$, it is achieved (or implemented) at the start of the period.

Cost reduction is stochastic. ϕ_t is the probability that the firm achieves the Δ cost reduction in period $t \in \{1,2\}$. $K_t(\phi_t)$ is the unmeasured cost the firm incurs in period t to ensure "success probability" ϕ_t in that period, given that the cost reduction has not been achieved in an earlier period. This cost can be viewed as reflecting the effort the firm's

[10] Frontier Economics (2015) and Oxera (2021) explain the design and implementation of IRIS and note its potential benefits relative to SR. We are not aware of any systematic investigation of the performance of IRIS.

[11] Section 5 considers settings with more than three PBRs and settings where each PBR can contain more than two periods.

[12] Recall that δ is the value that the firm derives at the start of one period from a dollar that will be delivered at the start of the next period.

[13] For example, in period t, the firm might discover a more efficient protocol for performing routine network maintenance. The firm might choose to implement the new protocol immediately or postpone its implmentation to a future date.

managers devote to pursuing new and creative ways to restructure the firm's operations in order to reduce production costs. $K_t(\cdot)$ is a strictly increasing, strictly convex function of ϕ_t, with $K_t(0) = 0$ and $K_t'(0) = 0$.[14] At times, it will be convenient to presume that $K_t(\cdot)$ satisfies Assumption K.

Assumption K. $K_t(\phi_t) = \frac{k_t}{\gamma}(\phi_t)^\gamma$ where $k_t > 0$ and $\gamma > 1$.

When Assumption K holds, higher values of γ reflect a lower level of "innovation costs" ($K_t(\cdot)$) and a reduced rate ($K_t'(\cdot)$) at which innovation costs increase with the associated success probability (ϕ_t).[15] For analytic convenience, we assume that $K_t(\phi_t)$ increases with ϕ_t sufficiently rapidly that the firm never finds it optimal to ensure a cost reduction with certainty. Formally, we assume:

$$K_t'(1) > \Delta\left[\,Q_t(c_0) + \delta\, Q_{t+1}(c_0)\,\right] \text{ for } t \in \{1, 2\} \text{ and}$$

$$K_2'(1) > \delta\,\Delta\left[\,Q_3(c_0) + \delta\, Q_4(c_0)\,\right] \tag{1}$$

where $Q_t(p_t) > 0$ is the demand for the firm's product in period $t \in \{1, ..., 6\}$ when the unit price of the firm's product in that period is $p_t > 0$.

Unless otherwise noted, we assume that the demand for the firm's product does not increase too rapidly relative to the firm's discount rate, i.e., that Assumption D holds.[16]

Assumption D. $Q_t(p) > \delta\, Q_{t+1}(p)$ for all $p > 0$, for $t \in \{1, ..., 5\}$. (2)

It will be convenient at times to assume that the demand for the firm's product increases over time at the constant rate g, as specified in Assumption G.

Assumption G. $Q_t(p) = g^{t-1}\, Q_1(p)$ for all $p > 0$ and $t \in \{1, ..., 6\}$, where $g \in (0, \frac{1}{\delta})$.

When Assumption G holds, demand is stationary if $g = 1$, demand increases over time if $g > 1$, and demand declines over time if $g < 1$. The upper bound on g in Assumption G ensures that Assumption D holds when Assumption G holds.[17]

[14] Section 5 considers settings in which the firm's first-period success probability affects its second-period success probability.

[15] This is the case because $\frac{\partial K_t(\phi_t)}{\partial \gamma} = \frac{\partial}{\partial \gamma}\left(\frac{k_t}{\gamma}(\phi_t)^\gamma\right) = \frac{k_t}{\gamma}(\phi_t)^\gamma \ln \phi_t - (\phi_t)^\gamma \gamma^{-2} < 0$ and $\frac{\partial K_t'(\phi_t)}{\partial \gamma} = \frac{\partial}{\partial \gamma}\left(k_t\,(\phi_t)^{\gamma-1}\right) = k_t\,(\phi_t)^{\gamma-1} \ln \phi_t < 0$. (Observe that $\ln \phi_t < 0$ for all $\phi_t \in (0, 1)$.)

[16] As demonstrated below, Assumption D limits the extent to which the firm will delay the implementation of an achieved cost reduction solely to take advantage of rapidly expanding demand for its product.

[17] When Assumption G holds, $Q_{t+1}(p) = g\, Q_t(p) \;\Rightarrow\; Q_t(p) = \frac{1}{g}\, Q_{t+1}(p) > [\,\frac{1}{1/\delta}\,]\, Q_{t+1}(p) = \delta\, Q_{t+1}(p)$.

c_t is the firm's marginal cost in period $t \in \{1, ..., 6\}$. For expositional ease, the firm's fixed cost is normalized to 0. Consequently, the firm's total cost of producing Q_t units of output in period t (not counting unmeasured "innovation cost" $K_t(\cdot)$) is $C_t(Q_t) = c_t\, Q_t$.

The firm's price is set equal to c_0 throughout PBR1 (i.e., $p_1 = p_2 = c_0$). The regulated prices during PBR2 and PBR3 vary with the regime under which the firm operates (SR or IRIS) and the firm's observed marginal costs. Under SR, the firm's price remains at c_0 throughout PBR2 and PBR3 (i.e., $p_3 = ... = p_6 = c_0$) if no cost reduction is implemented in PBR1 or PBR2. If the firm first implements the Δ cost reduction in PBR1, the firm's price is reduced to $c_0 - \Delta$ throughout PBR2 and PBR3 (i.e., $p_3 = ... = p_6 = c_0 - \Delta$). If the firm first implements the Δ cost reduction during PBR2, the firm's price is set at c_0 throughout PBR2 (i.e., $p_3 = p_4 = c_0$) and reduced to $c_0 - \Delta$ throughout PBR3 (i.e., $p_5 = p_6 = c_0 - \Delta$).[18] Under IRIS, the firm's price is reduced from c_0 to $c_0 - \Delta$ two periods after the cost reduction is first implemented. Formally, if the cost reduction is first implemented in period $t \in \{1, ..., 5\}$, then $p_1 = ... = p_{t+1} = c_0$ and $p_{t+2} = ... = p_6 = c_0 - \Delta$.

The key difference between IRIS and SR is that under IRIS, the firm retains the increment in profit engendered by a cost reduction for two periods, regardless of when the cost reduction is first implemented. In contrast, under SR: (i) the firm retains this profit increment for two periods if the cost reduction is implemented in the first period of a PBR regime; whereas (ii) the firm retains the profit increment for only one period if it implements the cost reduction in the second period of a PBR regime.

The interaction between the regulator and the firm proceeds as follows. First, the regulator announces whether the firm will operate under SR or IRIS. Next, the firm chooses ϕ_1 at the start of period 1 (in PBR1). The firm then observes whether the Δ cost reduction has been achieved and, if so, decides when to implement the reduction and thereby reduce its marginal cost from c_0 to $c_0 - \Delta$. For expositional ease, we assume that if the firm is indifferent between implementing immediately and delaying the implementation of an achieved cost reduction, the firm implements the cost reduction immediately.

After c_1 is observed at the end of period 1, the firm chooses ϕ_2 at the start of period 2 if the Δ cost reduction was not achieved in period 1. The firm also determines at the start of period 2 whether to implement a cost reduction that is achieved in period 2 or a cost reduction that was achieved, but not implemented, in period 1. c_2 is observed at the end of period 2. Finally, at the start of periods 3 – 6, the firm decides whether to implement a Δ

[18] As will be shown below, Assumption D ensures that the firm never delays to PBR3 the implementation of a cost reduction achieved during PBR1.

cost reduction that was achieved in period 1 or 2, but not yet implemented.[19]

3 Analytic Findings

We now characterize the outcomes that arise in this setting under IRIS and under SR.

A. Implementation Decisions

We begin by examining the firm's decision to implement an achieved cost reduction when it operates under IRIS.

Lemma 1. *When the firm operates under IRIS, it implements immediately any cost reduction it achieves.*

Lemma 1 reflects the fact that when the firm operates under IRIS, it is effectively awarded for two periods the profit increment engendered by a cost reduction, regardless of when the cost reduction is implemented.[20] Therefore, because demand does not increase too rapidly relative to the firm's discount rate (i.e., because Assumption D holds), the firm secures the highest discounted present value (PDV) of profit by implementing the cost reduction (and securing the associated profit increment) when it first becomes feasible to do so.

For corresponding reasons, the firm implements immediately a cost reduction achieved in period 1 when the firm operates under SR. Delaying the implementation to period 3 (or to period 5) would only postpone the two-period increment in profit that the firm can secure immediately by implementing the cost reduction in period 1. Furthermore, delaying the implementation to period 2 (or to period 4 or period 6) would reduce from two to one the number of periods during which the firm secures the profit increment associated with the cost reduction.

The firm's implementation decision is less straightforward when the firm achieves the cost reduction in period 2 under SR. In this case, if the firm implements the reduction immediately, it secures for only one period the profit increment associated with the cost reduction.[21] In contrast, the firm can secure this profit increment for two periods (periods 3 and 4) if it delays the implementation of the cost reduction to period 3.[22] Of course,

[19] The present model considers a single cost reduction and a single "productive" PBR (i.e., a single PBR regime in which it is feasible to achieve a cost reduction) for analytic ease. The two additional "non-productive" PBRs are considered to ensure that consumers always derive some benefit from an achieved cost reduction.

[20] The profit increment that the firm can achieve in period t by first implementing the Δ cost reduction in this period is $\Delta\, Q_t(\cdot)$.

[21] This is the case because the firm's price in this event is set throughout PBR2 and PBR3 to reflect the firm's observed cost at the end of PBR1 (i.e., $p_3 = p_4 = p_5 = p_6 = c_0 - \Delta$).

[22] This is the case because delayed implementation ensures that $p_3 = p_4 = c_0 > c_0 - \Delta$ under SR.

discounting diminishes the PDV of the profit derived from a delayed implementation. Lemma 2 reports that the firm will implement the cost reduction immediately if the value of the associated profit in period 2 ($\Delta\, Q_2(c_0)$) exceeds the PDV of the corresponding profit during periods 3 and 4 ($\delta\, \Delta\, [\, Q_3(c_0) + \delta\, Q_4(c_0)\,]$).

Lemma 2. *Suppose the firm operates under SR. If the firm achieves the cost reduction in period 1, it implements the cost reduction immediately. If the firm achieves the cost reduction in period 2, it implements the cost reduction immediately if*

$$Q_2(c_0) \;\geq\; \delta\, [\, Q_3(c_0) + \delta\, Q_4(c_0)\,] \tag{3}$$

and otherwise implements the cost reduction in period 3.

Corollary. *Suppose Assumption G holds in the setting of Lemma 2. Then if the firm achieves the cost reduction in period 2, it implements the cost reduction immediately if and only if* $\widetilde{\delta} \;\equiv\; \delta\, g \;\leq\; \widehat{\widehat{\delta}} \;=\; \frac{1}{2} \left[\, \sqrt{5} - 1\, \right] \;\approx\; 0.618.$

Lemma 2 and its corollary indicate that when the firm operates under SR, it becomes more likely to delay to period 3 the implementation of a cost reduction achieved in period 2 as the discount factor (δ) increases or the demand growth rate (g) increases, *ceteris paribus.* The latter effect arises because as demand increases, the firm's incremental profit from a specified reduction in marginal cost (Δ) increases.[23]

The ensuing discussion distinguishes between settings in which the firm implements immediately and delays the implementation of a cost reduction achieved in period 2 under SR. In the *NID setting*, the firm implements the cost reduction immediately. In the *ID setting*, the firm delays the implementation to period 3.[24] Lemma 2 implies that the NID setting prevails when inequality (3) holds, whereas the ID setting prevails otherwise.

B. Success Probabilities

We now characterize the "success probabilities" that the firm implements under IRIS and under SR. Formally, success probability ϕ_t^j is the probability that the firm achieves the cost reduction in period $t \in \{1, 2\}$ when it operates under regulatory regime $j \in \{S, I\}$, given that it has not achieved the reduction in an earlier period. Regulatory regime S denotes SR. Regulatory regime I denotes IRIS.

Proposition 1. $0 \,<\, \phi_2^S \,<\, \phi_2^I \,<\, 1.$

[23] This observation reflects the well-known *Arrow Effect* (Arrow, 1962).

[24] "ID" denotes implementation delay. "NID" denotes no implementation delay.

Proposition 1 reflects the fact that when the firm operates under IRIS, it is effectively awarded for two periods (periods 2 and 3) the increment in profit engendered by a cost reduction in period 2. In contrast, the firm only enjoys this increment in profit for a single period (period 2) when it achieves the cost reduction in period 2 in the NID setting under SR. In the ID setting, the firm secures the profit increment for two periods (periods 3 and 4) under SR, but the delay reduces the PDV of the profit increment. Thus, in both the NID setting and the ID setting, a cost reduction achieved in period 2 increases the PDV of the firm's profit to a greater extent under IRIS than under SR. Consequently, the firm implements a higher period 2 success probability under IRIS.

Proposition 2. $0 < \phi_1^I < \phi_1^S < 1$.

Proposition 2 identifies a potential drawback to IRIS. Specifically, the firm secures a lower success probability in period 1 under IRIS than under SR. The firm does so because it effectively perceives a smaller penalty from failing to achieve a cost reduction in period 1 under IRIS than under SR. The reduced penalty arises because, under IRIS, the firm has an additional chance in period 2 to secure a cost reduction that will generate a full two-period profit increment ($\Delta\,[\,Q_2(c_0) + \delta\, Q_3(c_0)\,]$). In contrast, under SR, first-period failure in the NID setting leaves the firm with no chance to secure a full two-period profit increment. Second-period success only secures a single period of incremental profit ($\Delta\, Q_2(c_0)$) for the firm in the NID setting under SR. In the ID setting, second-period success allows the firm to generate two periods of incremental profit (in periods 3 and 4). However, discounting diminishes the PDV of this incremental profit when Assumption D holds.[25] Thus, in both the NID setting and the ID setting, first-period failure leaves the firm with a higher PDV of maximum expected future profit under IRIS than under SR. This diminished penalty for first-period failure under IRIS induces the firm to implement a smaller first-period success probability under IRIS than under SR.

To determine whether, on balance, the firm is more likely to achieve a cost reduction under SR or under IRIS, define the aggregate success probability (Φ) to be the probability that the cost reduction is achieved either in period 1 or in period 2. This probability is the sum of the probability that the cost reduction is achieved in period 1 and the conditional probability that the cost reduction is achieved in period 2, given that it was not achieved in period 1. Formally:

$$\Phi^j \;\equiv\; \phi_1^j + [\,1 - \phi_1^j\,]\,\phi_2^j \;\text{ for }\; j \in \{S, R\}\,. \tag{4}$$

Lemma 3 identifies conditions under which the aggregate success probability is higher

[25] Formally, $\delta\,\Delta\,[\,Q_3(c_0) + \delta\, Q_4(c_0)\,] < \Delta\,[\,Q_2(c_0) + \delta\, Q_3(c_0)\,]$.

under SR than under IRIS and *vice versa.* The lemma refers to the following definitions.

$$G^{NID} \equiv \Delta Q_0 \left\{ [\,2 + 4\widetilde{\delta} + (\widetilde{\delta}\,)^2\,]\,k_2 - g\,[\,(\,1 + \widetilde{\delta}\,)^3 - 1\,]\,\Delta Q_0 \right\} - 2\,k_1\,k_2\,;$$

$$G^{ID} \equiv \Delta Q_0\,[\,1 + \widetilde{\delta}\,] \left\{ [\,2 - \widetilde{\delta} - (\widetilde{\delta}\,)^3\,]\,\widetilde{k}_2 - \widetilde{\delta}\,[\,1 + \widetilde{\delta}\,]\,[\,1 - (\widetilde{\delta}\,)^3\,]\,\Delta Q_0 \right\} - 2\,[\,1 - \widetilde{\delta}\,]\,k_1\,\widetilde{k}_2\,;$$

where $\widetilde{\delta} \equiv g\,\delta$, $\widetilde{k}_2 \equiv k_2\,/\,g$, and $Q_0 \equiv Q(c_0)$.

Lemma 3. *Suppose Assumption G holds and Assumption K with* $\gamma = 2$ *holds. Then:*

$$\Phi^S \gtreqless \Phi^I \;\Leftrightarrow\; G^{NID} \gtreqless 0 \text{ in the NID setting; and}$$

$$\Phi^S \gtreqless \Phi^I \;\Leftrightarrow\; G^{ID} \gtreqless 0 \text{ in the ID setting.} \tag{5}$$

Lemma 3 allows us to identify conditions under which the aggregate success probability is higher when the firm operates under SR than when it operates under IRIS.

Proposition 3. *Suppose Assumption G with* $g = 1$ *and Assumption K with* $\gamma = 2$ *holds. Then* $\Phi^S > \Phi^I$ *in the NID setting if* $\Delta Q_0 > \sqrt{k_1\,k_2}$.

Proposition 4. *Suppose Assumption G with* $g = 1$ *and Assumption K with* $\gamma = 2$ *holds. Then* $\Phi^S > \Phi^I$ *in the ID setting if* $g\left[\,1 + g\,\delta^2\,\right] < 1$ *and* $\Delta Q_0\,[\,1 + g\,\delta\,]$ *is sufficiently close to* $k_1 = k_2$.

Proposition 3 reports that under the specified conditions, the aggregate success probability is higher under SR than under IRIS in the NID setting when the potential net gain from innovation is large in the sense that the potential reduction in total cost (ΔQ_0) is large relative to the firm's innovation costs (k_1 and k_2). This conclusion reflects two primary considerations. First, when ΔQ_0 is large relative to k_1 and k_2, the firm secures a relatively large first-period success probability (ϕ_1) under SR to ensure that it can benefit from the relatively large profit increment associated with a cost reduction for two full periods without any implementation delay. The firm implements a relatively small ϕ_1 under IRIS because it perceives a relatively small penalty for failing to achieve the cost reduction in period 1. The small perceived penalty arises because the firm recognizes that it will implement a relatively large success probability in period 2 in response to the relatively pronounced potential gain from innovation. Second, recall from Proposition 1 that the firm implements a smaller success probability in period 2 under SR than under IRIS (i.e., $\phi_2^S < \phi_2^I$). However, the extent to which a relatively small value of ϕ_2^S diminishes the aggregate success probability under SR (Φ^S) is limited when ϕ_1^S is large. This is the case because $1 - \phi_1^S$ is small when ϕ_1^S is large. Consequently, the probability that a cost reduction is not achieved in period 1 under SR is small.

Proposition 4 reports that the aggregate success probability can be higher under SR than under IRIS in the ID setting when the demand growth rate (g) is sufficiently small.[26] The incentive to secure a cost reduction declines as demand declines (due to the Arrow Effect). Therefore, IRIS's advantage in motivating a relatively high second-period success probability becomes less pronounced when demand declines over time.[27] In practice, the demand for natural gas is likely to decline in many jurisdictions as governments encourage consumption of electricity (produced with renewable resources) rather than natural gas.[28] Demand has also been declining for many years for some postal services (e.g., first class mail in the U.S.).[29]

C. Consumer Welfare in the NID Setting

We now consider how the regulatory regime under which the firm operates affects consumer welfare. To do so, let $W_t(p) = \int_p^\infty Q_t(\widetilde{p})\, d\widetilde{p}$ denote consumer surplus in period t when the prevailing price is p. Also let p_t^j denote the price that prevails in period $t \in \{1, ..., 6\}$ when the firm operates under regime $j \in \{I, S\}$. Then the discounted present value (PDV) of expected consumer surplus under regime j is $E_d\{W^j\} \equiv E\{\sum_{t=1}^{6} \delta^{t-1} W_t(p_t^j)\}$, where $E\{\cdot\}$ denotes expectation regarding p_t^j.

This subsection compares $E_d\{W^S\}$ and $E_d\{W^I\}$ in the NID setting. The next subsection provides the corresponding comparison in the ID setting.

Proposition 5. *$E_d\{W^S\} > E_d\{W^I\}$ in the NID setting if $\Phi^S > \Phi^I$.*

Proposition 5 reports that the PDV of expected consumer surplus is higher under SR than under IRIS in the NID setting whenever the aggregate success probability (Φ) is higher under SR than under IRIS. This conclusion reflects two considerations. First, the firm never delays the implementation of an achieved cost reduction under SR or under IRIS in the NID setting. Second, a cost reduction in period 2 engenders lower prices for consumers beginning in period 3 under SR, whereas the lower prices are not secured until period 4 under IRIS.[30]

---

[26] Similarly, it can also be shown that $\Phi^S > \Phi^I$ in the NID setting if the conditions in Proposition 4 hold, $k_1 = k_2$, δ is sufficiently large, and g is sufficiently small.

[27] When $Q_0\,[1 + g\,\delta]$ is close to $k_1 = k_2$, the potential net gain from innovation is relatively large, which induces the firm to set ϕ_2 close to 1 under IRIS. The associated limited perceived penalty for first-period failure induces the firm to set a relatively small ϕ_1 under IRIS, which reduces Φ^I, *ceteris paribus.*

[28] See Duma et al. (2024), for example.

[29] United States Postal Service (2024).

[30] A cost reduction in period 1 secures lower prices for consumers beginning in period 3 under both SR and IRIS. (Recall Lemmas 1 and 2.)

This closer temporal link between cost reductions and price reductions, coupled with a higher aggregate probability of a cost reduction, ensures that the PDV of expected consumer surplus is higher under SR than under IRIS in the NID setting.

Because a second-period cost reduction leads to price reductions more rapidly under SR than under IRIS in the NID setting, $E_d\{W^S\}$ can exceed $E_d\{W^I\}$ in the NID setting even if the aggregate probability of a cost reduction is higher under IRIS than under SR. Proposition 6 and its Corollary identify additional conditions under which $E_d\{W^S\}$ exceeds $E_d\{W^I\}$.

Proposition 6. *Suppose Assumptions G and K hold. Then* $E_d\{W^S\} > E_d\{W^I\}$ *in the NID setting if* $\widetilde{\delta}\,[1+\widetilde{\delta}]^{\frac{1}{\gamma-1}} < 1$.[31]

Corollary. *Suppose Assumptions G and K hold. Then* $E_d\{W^S\} > E_d\{W^I\}$ *in the NID setting if* $\gamma \geq 2$.

Proposition 6 indicates that the PDV of expected consumer surplus tends to be higher under SR than under IRIS in the NID setting when the discount factor (δ) or the demand growth rate (g) is relatively small, *ceteris paribus*. IRIS's advantage in inducing a relatively high second-period success probability ($\phi_2^I > \phi_2^S$) becomes less pronounced as demand growth diminishes. The impact of this advantage on expected consumer surplus also is discounted more heavily in the calculation of $E_d\{W^j\}$ as δ declines. The Corollary to Proposition 6 reports that $E_d\{W^S\}$ exceeds $E_d\{W^I\}$ in the NID setting for a wide array of cost structures ($K_t(\cdot)$).

D. Consumer Welfare in the ID Setting

Although the PDV of expected consumer surplus is often higher under SR than under IRIS in the NID setting, IRIS often outperforms SR in this regard in the ID setting. This is the case because in the ID setting, the firm that operates under SR delays to period 3 the implementation of a cost reduction achieved in period 2. This delayed implementation implies that consumers do not benefit from the associated price reduction until period 5 (i.e., after rebasing occurs at the end of PBR2). No such delayed implementation arises under IRIS. (Recall Lemma 1.) Consequently, under IRIS, consumers begin to benefit in period 4 from the cost reduction achieved in period 2. This closer temporal linkage between realized cost reductions and associated price reductions under IRIS, coupled with IRIS's advantage in inducing a higher second-period innovation probability ($\phi_2^I > \phi_2^S$), ensure that the PDV of expected consumer surplus is often higher under IRIS than under SR in the ID setting.

This is the case, for example, when δ is sufficiently large, as Proposition 7 reports. This

[31] Recall that $\widetilde{\delta} \equiv g\,\delta$.

conclusion reflects in part the fact that as δ increases, IRIS's advantage in inducing a higher second-period success probability is weighted more heavily when calculating the PDV of expected consumer surplus.

Proposition 7. *Suppose Assumption K holds, Assumption G with $g = 1$ holds, and $\delta > \widehat{\widehat{\delta}}$. Then $E_d\{W^I\} > E_d\{W^S\}$ when δ is sufficiently large (in the ID setting).*

Proposition 8 reports that the PDV of expected consumer surplus also tends to be higher under IRIS than under SR in the NID setting when first-period innovation costs (k_1) are sufficiently pronounced. This is the case because SR induces a higher probability of first-period success than does IRIS (i.e., $\phi_1^S > \phi_1^I$ from Proposition 2), and this advantage of SR dissipates as first-period innovation costs increase.[32]

Proposition 8. *Suppose Assumption K holds, Assumption G holds, and $\widetilde{\delta} \geq \widehat{\widehat{\delta}}$. Then $E_d\{W^I\} > E_d\{W^S\}$ for sufficiently large k_1 (in the ID setting).*

Proposition 9 identifies additional conditions under which the PDV of expected consumer surplus is higher under IRIS than under SR in the ID setting.

Proposition 9. *Suppose Assumption K with $\gamma = 2$ holds and Assumption G holds. Then $E_d\{W^I\} > E_d\{W^S\}$ when ΔQ_0 is sufficiently small or $k_1 = k_2 \equiv k$ is sufficiently large in the ID setting.*

Proposition 9 reflects the fact that when the potential gain from innovation is relatively limited (because ΔQ_0 is small and/or $k \equiv k_1 = k_2$ is large), the firm chooses a relatively small second-period success probability under IRIS. Consequently, the firm perceives a relatively large penalty from failing to achieve a first-period cost reduction. Therefore, $\phi_1^S - \phi_1^I$ declines, thereby reducing the key potential advantage of SR. ϕ_1^S also tends to be relatively low when the gain from innovation is limited, so $1 - \phi_1^S$ can be relatively high. Consequently, the two main drawbacks to SR – limited incentives for second-period cost reduction and delayed implementation of a cost reduction achieved in period 2 – become more likely to be relevant.

Although the PDV of expected consumer surplus is often higher under IRIS than under SR in the ID setting, this ranking can be reversed, as Proposition 10 reports.

[32] Furthermore, $1 - \phi_1^S$ increases as ϕ_1^S declines in response to an increase in k_1. As $1 - \phi_1^S$ increases, innovation activity in the second period (where $\phi_2^I > \phi_2^S$) becomes more relevant (due to the increased probability that the cost reduction is not achieved in period 1).

Proposition 10. *Suppose Assumption K with $\gamma = 2$ holds, Assumption G holds, $k_2 \leq k_1 g$, and $\widetilde{\delta} > \widehat{\widehat{\delta}}$. Then $E_d\{W^S\} > E_d\{W^I\}$ when $\widetilde{\delta}$ is sufficiently close to $\widehat{\widehat{\delta}}$ and $\frac{\Delta Q_0 g [1 + \widehat{\widehat{\delta}}]}{k_2}$ is sufficiently close to* 1 *(in the ID setting).*

Proposition 10 reflects the fact that ϕ_2^I is close to 1 when $\frac{\Delta Q_0 g [1 + \widehat{\widehat{\delta}}]}{k_2}$ is close to 1 in the identified setting. Therefore, the firm perceives a relatively limited penalty from first-period "failure" under IRIS. In contrast, this penalty is relatively large under SR in the ID setting when δ is relatively small, i.e., close to $\widehat{\widehat{\delta}}$.[33] Consequently, ϕ_1^S is large relative to ϕ_1^I, which promotes a higher PDV of expected consumer surplus under SR than under IRIS in the ID setting.

Proposition 11 identifies additional conditions under which the PDV of expected consumer surplus is higher under SR than under IRIS in the ID setting.

Proposition 11. *Suppose Assumption K with $\gamma = 2$ holds, Assumption G holds, and $\widetilde{\delta} \geq \widehat{\widehat{\delta}}$. Then $E_d\{W^S\} > E_d\{W^I\}$ when k_2 is sufficiently large and $\frac{\Delta Q_0}{k_1} [1 + \widetilde{\delta}]$ is sufficiently close to* 1 *(in the ID setting).*

Proposition 11 reflects the fact that relatively high second-period innovation costs serve to limit IRIS's advantage over SR in securing a relatively high probability of a second-period cost reduction. Furthermore, the condition in Proposition 11 implies that ΔQ_0 is relatively large and/or k_1 is relatively small. The associated relatively large gains from first-period innovation accentuate SR's advantage over IRIS in inducing first-period innovation effort.

4 Numerical Characterizations

To further compare outcomes under SR and IRIS, we consider a *baseline setting* with the following properties. First, Assumptions D, G, and K hold. Second, the demand for the firm's product is linear, so $Q_t(p_t) = g^{t-1} M [a - b p_t]$ for $t \in \{1, .., 6\}$, where $M > 0$, $a > 0$, and $b > 0$ are parameters.[34] Third, the firm's innovation cost function is stationary, so $K_t(\cdot) = \frac{k}{\gamma} (\phi_t)^\gamma$ for $t \in \{1, 2\}$, where $k > 0$ is a parameter.

We consider two variants of this baseline setting. In the *unrestricted baseline setting*, the firm decides when to implement a realized cost reduction, just as it does in the preceding

[33] A relatively small value of δ reduces the firm's expected profit from a second-period cost reduction because the reduction will not be implemented until period 3 under SR in the ID setting.

[34] We continue to assume that three PBR regimes (PBR1, PBR2, and PBR3) are imposed consecutively, and that each PBR regime consists of two periods of equal length. We also continue to assume that the Δ reduction in marginal cost can be achieved in PBR1 (i.e., in period 1 or period 2), but not in PBR2 or PBR3.

analysis. In the *restricted baseline setting*, the firm always implements a cost reduction as soon as it is realized. Thus, by construction, the NID setting always prevails in the restricted baseline setting. In contrast, either the ID setting or the NID setting can prevail in the unrestricted setting, depending upon model parameters.

The restricted baseline setting is intended to reflect environments in which a regulated firm has limited ability to strategically delay the implementation of an achieved cost reduction. This limited ability might stem from the hierarchical structure of the regulated firm and the nature of the regulatory process, for example. In practice, regulators often interact directly with managers of the firms they regulate. Senior executives in a regulated firm are unlikely to direct the firm's managers to delay the introduction of an achieved cost reduction when such a directive is likely to come to light in the course of routine interactions between regulators and the firm's managers.[35,36]

We consider both the restricted and the unrestricted baseline settings to demonstrate that the PDV of expected consumer surplus often is higher under SR than under IRIS when the firm cannot (or chooses not to) delay the implementation of an achieved cost reduction. In contrast, the PDV of expected consumer surplus often is higher under IRIS than under SR when such strategic delay is feasible and enhances the firm's expected profit under SR.

We begin by specifying initial values for the parameters in these baseline settings, and then proceed to consider variation in these parameters. The initial parameter values are specified in Table 1.

Parameter	Parameter Value
M	10
a	10
b	1
g	1

Parameter	Parameter Value
c_0	1
Δ	0.01
k	2
γ	2
δ	0.834

Table 1. Initial Parameter Values in the Baseline Settings

When the initial parameter values in Table 1 prevail, the demand for the firm's product is stationary ($g = 1$) and the maximum amount a consumer is willing to pay for a unit of the firm's product is 10 ($= \frac{a}{b}$). The potential cost reduction ($\Delta = 0.01$) is 1% of the firm's initial

[35] If a regulator were to learn of such strategic actions that harm consumers, she would likely penalize harshly the regulated firm and/or its senior executives.

[36] Scalise and Zech (2013) discuss the types of cost reductions that utilities can achieve in practice, some of which are difficult to conceal once they are discovered.

marginal cost ($c_0 = 1$). The firm's innovation costs are quadratic ($K_t(\cdot) = (\phi_t)^2$ because $k = \gamma = 2$). The inter-period discount factor is 0.834, which corresponds to the 0.93 annual discount factor cited by the Council of Economic Advisors (2017).[37,38]

Table 2 records outcomes in the baseline settings when these initial parameter values prevail. In Table 2, SR-Br (SR-Bu) denotes the restricted (unrestricted) baseline setting when the firm operates under SR. IRIS-B denotes the baseline setting when the firm operates under IRIS.[39] ϕ_t is the probability of a successful cost reduction in period $t \in \{1, 2\}$, given that the cost reduction has not been achieved earlier. Φ is the aggregate success probability ($\phi_1 + [1 - \phi_1]\,\phi_2$). $E_d\{W\}$ is the PDV of expected consumer surplus.[40]

	Setting		
Outcomes	SR-Bu	SR-Br	IRIS-B
ϕ_1	0.628	0.741	0.541
ϕ_2	0.688	0.450	0.825
Φ	0.884	0.858	0.920
$E_d\{W\}$	1.427	1.670	1.555

Table 2. Baseline Setting Outcomes for Initial Parameter Values

Table 2 reports that when the parameter values in Table 1 prevail, $E_d\{W\}$ is highest in the restricted baseline setting when SR prevails (i.e., in the *SR-Br setting*) and lowest in the unrestricted baseline setting when SR prevails (i.e., in the *SR-Bu setting*).[41] The reduced

[37] The Council of Economic Advisors (2017, p. 2) advises consideraton of a 7 percent annual discount rate when calculating the PDV of government projects, noting that this rate is "an estimate of the average before-tax rate of return to private capital in the U.S. economy, sometimes referred to as the social opportunity cost of capital." Office of Management and Budget (2023) provides additional guidance on selecting appropriate discount factors.

[38] There are two periods in each PBR in our model. As noted above, PBR regimes typically last for approximately five years in practice (Sappington and Weisman, 2024). When 0.93 is the relevant annual discount factor over a five-year time interval, the corresponding discount factor (δ) over a two-period time interval of identical duration is determined by $\delta^2 = (0.93)^5$, so $\delta = (0.93)^{5/2} \approx 0.834$.

[39] When the firm operates under IRIS, outcomes are the same in the restricted baseline setting and the unrestricted baseline setting because the firm always prefers to implement a cost reduction as soon as it is achieved. (Recall Lemma 1.)

[40] For expositional ease, $E_d\{W\}$ is reported in Table 2 as the difference between the PDV of expected consumer surplus in the identified settings and the PDV of expected consumer surplus that arises when innovation is not feasible, so $c_t = c_0$ and $\phi_t = 0$ for all $t \in \{1, ..., 6\}$.

[41] It can be shown that when the parameter values in Table 1 prevail, the PDV of the firm's expected profit is highest in the IRIS-B setting (0.861) and lowest in the SR-Br setting (0.718). This profit measure is 0.789 in the SR-Bu setting.

urgency for innovation that arises under IRIS (recall Proposition 2), coupled with the two-period delay in passing achieved cost reductions on to consumers in the form of lower prices, lead to a lower PDV of expected consumer surplus under IRIS than under SR when the firm cannot strategically delay the implementation of an achieved cost reduction. This is the case even though the inability to delay an achieved cost reduction substantially reduces the firm's second-period cost-reducing effort under SR.[42,43]

The potential for strategic implementation delay under SR induces the firm to increase its second-period cost-reducing effort in the SR-Bu setting. However, this effort remains below the effort induced under IRIS. Consequently, $E_d\{W\}$ is lower in the SR-Bu setting than in the IRIS-B setting when the parameter values in Table 1 prevail.[44]

Figures 1 – 3 further compare $E_d\{W\}$ in these three baseline settings as selected parameter values change, holding all other parameter values at the levels specified in Table 1.[45] Figure 1 reports that the difference between the PDV of expected consumer surplus in the SR-Br setting and in the IRIS-B setting increases as Δ increases, *ceteris paribus*.[46] The firm's incentive to achieve a cost reduction increases as the magnitude of the cost reduction increases. The corresponding induced increase in ϕ_2 induces the firm to reduce ϕ_1 under IRIS, which causes $E_d\{W\}$ to increase less rapidly as Δ increases in the IRIS-B setting than in the SR-Br setting.

Figure 2 demonstrates that $E_d\{W\}$ can be higher in both the SR-Bu setting and the SR-Br setting than in the IRIS-B setting when g is sufficiently small. When demand declines over time, IRIS's advantage in inducing a relatively high second-period success probability (ϕ_2) translates into a smaller increase in $E_d\{W\}$. In contrast, SR's advantage in inducing a relatively high first-period success probability (ϕ_1) secures a relatively large increase in $E_d\{W\}$ when demand declines over time.

Figure 3 reports that the difference between the PDV of expected consumer surplus in

[42] As Table 2 reports, when the parameter values in Table 1 prevail, $\phi_2 = 0.450$ in the SR-Br setting whereas $\phi_2 = 0.688$ in the SR-Bu setting and $\phi_2 = 0.825$ in the IRIS-B setting.

[43] Anticipating the relatively low second-period success probability it will implement in the SR-Br setting, the firm implements a relatively high first-period success probability ($\phi_1 = 0.741$).

[44] The PDV of expected consumer surplus is lower (higher) in the SR-Bu (SR-Br) setting than in the IRIS-B setting even though the aggregate success probability (Φ) is higher in the SR-Bu setting than in the SR-Br setting.

[45] As in Table 2, $E_d\{W\}$ in Figures 1 – 3 reflects the difference between the PDV of expected consumer surplus in the identified setting and the PDV of expected consumer surplus that arises when innovation is not feasible.

[46] This difference also increases as M or a increases. Like an increase in Δ, an increase in M or a increases the increment in profit generated by a reduction in marginal cost.

the IRIS-B setting and in the SR-Bu setting increases as the discount factor (δ) increases. This conclusion arises in part because as δ increases, IRIS's advantage in inducing a higher second-period success probability is weighted more heavily when calculating the PDV of expected consumer surplus. (Recall Proposition 7.)

5 Extensions

The analysis to this point has abstracted from any inter-period effects of cost-reducing activity, restricted the number of periods in a PBR regime to two, considered only a single productive PBR, and permitted only a single cost reduction. The ensuing analysis relaxes these restrictions.

A. Innovation Persistence

The foregoing analysis assumes that the firm's unsuccessful cost-reducing activity in period 1 has no effect on the likelihood of achieving a second-period cost reduction. This simplifying assumption is not the source of the key qualitative conclusions drawn above, as Propositions 12 and 13 (below) demonstrate. The propositions characterize outcomes in the *setting with innovation persistence.* This setting parallels the setting analyzed in section 3, with one exception. If the firm does not achieve the Δ unit cost reduction in period 1, the fraction $\alpha \geq 0$ of the first-period innovation probability persists in period 2. Formally, when the firm implements first-period success probability ϕ_1, no cost reduction is achieved in period 1, and the firm implements incremental success probability ϕ_2 in period 2, the total probability of success in period 2 is $\widetilde{\phi}_2 = \phi_2 + \alpha\, \phi_1$. For analytic and expositional ease, all success probabilities (ϕ_1, ϕ_2, and $\widetilde{\phi}_2$) are assumed to be interior (i.e., strictly positive and strictly less than 1) in this setting.[47]

Proposition 12. *$\phi_1^I < \phi_1^S$ and $\phi_2^I > \phi_2^S$ in the setting with innovation persistence.*

Proposition 12 implies that the key conclusions established in Propositions 1 and 2 persist in the setting with innovation persistence. Specifically, the first-period success probability is higher under SR than under IRIS, whereas the second-period success probability is higher under IRIS than under SR. The rationale for these conclusions is precisely the rationale developed in section 3.

Proposition 13. *Suppose Assumptions G and K hold. Then when the NID setting prevails, $E_d\{W^S\} > E_d\{W^I\}$ if $\widetilde{\delta}\,[\,1 + \widetilde{\delta}\,]^{\frac{1}{\gamma - 1}} < 1$ in the setting with innovation persistence.*

[47] This will be the case, for example, if Assumption K holds, k_t is sufficiently small, and $K_t'(1)$ is sufficiently large for $t \in \{1, 2\}$.

Corollary. Suppose Assumptions G and K hold. Then when the NID setting prevails, $E_d\{W^S\} > E_d\{W^I\}$ if $\gamma \geq 2$ in the setting with innovation persistence.

Proposition 13 and its corollary parallel Proposition 6 and its corollary. When the firm always chooses to implement immediately a cost reduction achieved under SR (as it will when $\delta\, g < \widetilde{\delta}$), the PDV of expected consumer surplus often is higher under SR than under IRIS in the setting with innovation persistence. The lower value of $E_d\{W\}$ under IRIS arises in part because the firm continues to choose a lower first-period success probability under IRIS than under SR in the presence of innovation persistence.

B. T $\geq$ 2 Periods in Each PBR Regime

For analytic ease, the foregoing analysis has taken T, the number of periods in a PBR regime, to be 2. Figure 4 compares the PDV of expected consumer surplus ($E_d\{W\}$) in the SR-Bu setting and the IRIS-B setting as T varies when all parameter values other than δ are as specified in Table 1. In Figure 4, δ adjusts as T varies to hold constant the inter-PBR discount factor. Specifically, the annual discount factor is taken to be $\delta_a = 0.93$ and a PBR is assumed to last for five years. Therefore, the inter-PBR discount factor is $(\delta_a)^5 \approx 0.6957$. When there are T periods in a PBR (i.e., T opportunities for the firm to achieve the single Δ cost reduction), the corresponding inter-period discount rate, δ_p, is determined by $(\delta_p)^T = (\delta_a)^5$. Thus, $\delta_p = (\delta_a)^{5/T} \approx (0.93)^{5/T}$.

Figure 4 reports that $E_d\{W^S\} - E_d\{W^I\}$ increases as T increases in these baseline settings. In this sense, the simplifying assumption that each PBR consists of only two periods may promote an understatement of SR's relative performance in generating consumer surplus if, in practice, regulated firms typically have more than two opportunities to achieve and implement a cost reduction during a PBR regime.[48] SR's relative performance in enhancing the PDV of expected consumer surplus improves as the number of opportunities to achieve and implement the Δ unit cost reduction in PBR1 increases for two reasons.

First, the NID setting becomes more likely to prevail in the early stages of PBR1 as T increases. To see why, suppose the firm achieves the cost reduction in period 2 ($< T$) of PBR1. If the firm implements the achieved cost reduction immediately, it can benefit from the associated profit increment for a relatively long time period during PBR1 when T is large. Specifically, the firm can benefit from the profit increment for $T-1$ periods, which constitutes the fraction $\frac{T-1}{T}$ of the length of PBR1. In contrast, the firm must wait a relatively long time (i.e., the fraction $\frac{T-1}{T}$ of the length of PBR1) to realize the profit increment if it delays the

[48] As noted above, the length of the typical PBR regime is approximately five years (Sappington and Weisman, 2024).

implementation of the cost reduction to the start of PBR2. The PDV of the delayed profit increment is relatively small when the associated delay is relatively long. Therefore, as T increases, the firm that operates under SR becomes more likely to implement immediately a cost reduction achieved early in PBR1. Consequently, the firm becomes less likely to engage in the strategic implementation delay that can reduce $E_d\{W\}$ under SR.

Second, as T increases, the firm is afforded more opportunities to achieve the Δ cost reduction during PBR1. The increased number of such opportunities effectively reduces the urgency the firm perceives to achieve the cost reduction early in PBR1. The reduced urgency is particularly prevalent under IRIS, where the firm retains for T periods the profit increment associated with the cost reduction regardless of when the reduction is achieved. Therefore, an increase in T tends to reduce the firm's cost-reducing effort during the early portion of PBR1 under IRIS.

Figure 4 reports that $E_d\{W^S\}$ exceeds $E_d\{W^I\}$ in the identified setting whenever there are at least four periods in each PBR regime. As noted above, PBR regimes typically last for approximately five years. Figures A1 – A4 in the Appendix show that $E_d\{W^S\}$ continues to exceed $E_d\{W^I\}$ when $T = 5$ as the parameters Δ, M, k, and γ vary substantially from the levels specified in Table 1.[49] Figures A5 and A6 reveal that $E_d\{W^I\}$ exceeds $E_d\{W^S\}$ when the annual discount rate (δ) exceeds 0.95 or the annual demand growth rate (g) exceeds 1.02. When $\delta\, g$ is sufficiently close to 1, the firm delays to PBR2 the implementation of a cost reduction achieved after period 1 in PBR1. Such delayed implementation, coupled with the relatively limited incentive for innovation that arises under SR late in a PBR regime, reduces $E_d\{W^S\}$ below $E_d\{W^I\}$.

C. Multiple Productive PBRs

For analytic ease, the analysis to this point has considered only a single productive PBR.[50] Figure 5 illustrates how the comparison between $E_d\{W^S\}$ and $E_d\{W^I\}$ changes as the number (N) of productive PBRs changes. The figure compares $E_d\{W\}$ in the SR-Bu setting and in the IRIS-B setting when the parameter values in Table 1 prevail.[51] The dashed line toward the bottom of Figure 5 depicts $E_d\{W^S\} - E_d\{W^I\}$ in these settings as N changes. The line indicates that $E_d\{W^S\} - E_d\{W^I\}$ varies little as N changes in this

[49] k and γ are the parameters in the firm's stationary cost function $K_t(\phi_t) = \frac{k}{\gamma}\,(\phi_t)^{\gamma}$ for $t = 1, ..., 5$.

[50] Recall that a productive PBR is one in which it is feasible to achieve a cost reduction.

[51] In all cases, two unproductive PBRs (in which a cost reduction can be implemented, but not achieved) follow the N productive PBRs. As in Figures 1 – 3, $E_d\{W\}$ in Figure 5 reflects the difference between the PDV of expected consumer surplus in the identified setting and the PDV of expected consumer surplus that arises when innovation is not feasible.

setting.[52] Figure 5 thereby suggests that the qualitative conclusions drawn above are not an artifact of the simplifying focus on a single productive PBR.

D. Multiple Cost Reductions

The preceding analysis assumed that the firm could achieve at most a single cost reduction. We now consider a variation on the baseline setting in which two cost reductions are possible. With one exception, the setting parallels the baseline setting with the parameter values specified in Table 1. The exception is that there are now two productive PBR regimes and two possible cost reductions. Specifically, the firm can achieve an initial cost reduction $\Delta_1 = 0.01$ in PBR1 or PBR2 (i.e., in any one of periods 1, 2, 3, or 4). If it achieves the first cost reduction (Δ_1) in period $t \in \{1, 2, 3\}$, the firm can attempt to achieve a second cost reduction, $\Delta_2 = 0.01$, in period $t' \in \{t + 1, ..., 4\}$. In addition to the two productive PBR regimes (PBR1 and PBR2), the present setting continues to include two additional PBR regimes (PBR3 and PBR4) in which the firm cannot achieve (but may implement) a cost reduction.

Table 3 records outcomes in this environment for the three baseline settings considered in section 4 (SR-Bu, SR-Br, and IRIS-B). The table employs the following notation. ϕ_{1t} is the probability that the firm achieves cost reduction Δ_1 in period $t \in \{1, ..., 4\}$, given that it has not achieved the reduction in an earlier period. ϕ_{2t} is the probability that the firm achieves cost reduction Δ_2 in period $t \in \{2, 3, 4\}$, given that it has achieved the Δ_1 cost reduction in an earlier period but has not achieved the Δ_2 cost reduction in an earlier period.[53] Φ is the aggregate probability that the firm achieves both cost reductions. $E_d\{W\}$ continues to denote the PDV of expected consumer welfare (in excess of the PDV of expected consumer welfare that arises when innovation is not feasible).

[52] $E_d\{W^S\} - E_d\{W^I\}$ also varies little as N changes when parameter values diverge from their levels in Table 1.

[53] The firm's cost of achieving success probability ϕ_{jt} in period t, given that the Δ_j cost reduction is feasible in period t, is $K_t(\phi_{jt}) = (\phi_{jt})^2$ for all $j \in \{1, 2\}$ and $t \in \{1, ..., 4\}$. The firm can implement an achieved cost reduction immediately or delay the implementation to a subsequent period.

	Setting		
Outcomes	SR-Bu	SR-Br	IRIS-B
ϕ_{11}	0.711	0.724	0.598
ϕ_{12}	0.404	0.395	0.664
ϕ_{13}	0.826	0.825	0.826
ϕ_{14}	0.688	0.450	0.825
ϕ_{22}	0.359	0.150	0.466
ϕ_{23}	0.628	0.742	0.542
ϕ_{24}	0.689	0.451	0.826
Φ	0.880	0.806	0.930
$E_d\{W\}$	3.043	3.491	3.411

Table 3. Outcomes when Two Cost Reductions are Possible

Table 3 reports that in this setting with two possible cost reductions, $E_d\{W\}$ continues to be highest in the SR-Br setting, where the firm cannot delay the implementation of an achieved cost reduction.[54] SR generates the highest PDV of expected consumer welfare in this setting even though the inability to delay the implementation of an achieved cost reduction substantially reduces the success probability the firm implements at the end of each productive PBR (i.e., in periods 2 and 4).[55]

6 Conclusions

We have examined the ability of an incremental rolling incentive scheme (IRIS) to enhance cost-reducing innovation and consumer welfare in regulated industries. Under IRIS, the regulated firm is permitted to retain the incremental profit engendered by a cost reduction for the full length of a PBR regime, regardless of when the cost reduction is implemented. The promise of a fixed period of incremental profit induces the firm to implement achieved cost reductions immediately and also enhances the firm's incentive to achieve a cost reduction late in a PBR regime. However, this promise can also reduce innovation early in the PBR regime (by reducing the perceived urgency of such innovation). On balance, IRIS can either increase or reduce innovation and consumer welfare relative to the levels that arise under standard rebasing (SR), wherein the firm's prices are recalibrated to match observed costs at the start of each PBR regime.

[54] The same qualitative conclusion generally persists in the presence of "diminishing returns to innovation," i.e., when the second potential cost reduction (Δ_2) is smaller than the first potential cost reduction (Δ_1).

[55] Observe from the penultimate row in Table 3 that the relatively limited incentive for innovation in the SR-Br setting leads to a relatively small aggregate probability of achieving both cost reductions (Φ).

We found that IRIS generally reduces the discounted present value of consumer welfare ($E_d\{W\}$) below the level achieved under SR in settings where the regulated firm either chooses not to, or can be compelled not to, delay the implementation of achieved cost reductions under SR. In contrast, IRIS can enhance $E_d\{W\}$ when the firm has both the ability and the incentive to delay the implementation of a cost reduction achieved late in a PBR regime under SR. IRIS is relatively likely to increase $E_d\{W\}$ above the level achieved under SR when the discount factor (δ) is large, when cost-reducing innovation is onerous for the firm, and when the increased profit admitted by a cost reduction is relatively small.

Our findings imply that the appropriate choice between IRIS and SR depends in part on the ability of the regulated firm to delay the implementation of achieved cost reductions. In settings where the regulator can deter such strategic delay, $E_d\{W\}$ often is higher under SR than under IRIS. Consequently, regulators, may sometimes be better able to increase $E_d\{W\}$ by instituting policies that limit opportunities for implementation delay than by instituting IRIS. Such policies might entail requiring the regulated firm to report regularly on its efforts to reduce costs, for example. Regular progress reports may help to limit the firm's ability to conceal achieved cost reductions until the start of a new PBR regime.

Additional research is required to fully assess the relative merits of IRIS and SR. To illustrate, it would be useful to analyze settings in which the magnitude of a potential cost reduction is endogenous and/or settings in which the firm can pursue multiple cost reductions simultaneously. The basic forces at play in our model seem likely to persist in these model extensions. However, the details remain to be explored.[56]

IRIS effectively allows the regulated firm to retain for a fixed period of time the incremental profit admitted by a cost reduction, regardless of when the reduction is implemented. In contrast, under SR, this period of enhanced profitability is the remainder of the prevailing PBR regime, which declines steadily as the PBR regime progresses. Future research might examine whether some alternative pattern of enhanced profitability could increase consumer welfare above both the level achieved under SR and the level achieved under IRIS.[57]

Future research might also consider whether stochastic renewal of PBR regimes under SR can enhance consumer welfare above the level achieved under deterministic renewal. Uncertainty about whether PBR will be renewed can limit the regulated firm's incentive to strategically delay the implementation of an achieved cost reduction. As we have shown,

[56] Future research might also analyze settings in which the firm can deliver unobserved effort to reduce fixed costs of operation or enhance service quality.

[57] For example, the period of enhanced profitability might be shorter the later in the PBR regime a cost reduction is implemented. However, a non-trivial period of enhanced profitability might be promised even when a cost reduction is achieved at the very end of a PBR regime.

consumer welfare often is higher under SR when the firm does not engage in strategic delay of achieved cost reductions.

Appendix

Part A of this Appendix sketches the proofs of the formal conclusions in the text.[58] Part B compares the discounted present value of expected consumer welfare under SR and under IRIS when each PBR contains five periods.

A. Proofs of Formal Conclusions in the Text.

Proof of Lemma 1. Under IRIS, if the firm first implements the achieved cost reduction in period $\widehat{t} \in \{1, ..., 5\}$, then $p_t = c_0$ for $t = 1, ..., \widehat{t} + 1$ and $p_t = c_0 - \Delta$ for $t = \widehat{t} + 2, ..., 6$.[59] Suppose the firm achieves the Δ cost reduction in period $t \in \{1, 2\}$. If the firm implements the cost reduction immediately, the discounted present value (PDV) of its profit is $\Delta \left[Q_t(c_0) + \delta \, Q_{t+1}(c_0) \right]$. If the firm delays the implementation to period $t + l$, the PDV of its profit is $\delta^l \, \Delta \left[Q_{t+l}(c_0) + \delta \, Q_{t+l+1}(c_0) \right]$. Therefore, the firm will implement the achieved cost reduction immediately if:

$$\begin{aligned} & \Delta \left[Q_t(c_0) + \delta \, Q_{t+1}(c_0) \right] \;\geq\; \delta^l \, \Delta \left[Q_{t+l}(c_0) + \delta \, Q_{t+l+1}(c_0) \right] \\ \Leftrightarrow \quad & Q_t(c_0) + \delta \, Q_{t+1}(c_0) \;\geq\; \delta^l \left[Q_{t+l}(c_0) + \delta \, Q_{t+l+1}(c_0) \right]. \end{aligned} \tag{6}$$

The inequality in (6) holds because Assumption D implies:

$$Q_t(c_0) \;>\; \delta \, Q_{t+1}(c_0) \;\geq\; ... \;\geq\; \delta^l \, Q_{t+l}(c_0) \;\; \text{for all} \;\; l \in \{ 1, ..., 6 - t \}, \;\; \text{and}$$

$$\delta \, Q_{t+1}(c_0) \;>\; \delta^2 \, Q_{t+2}(c_0) \;\geq\; ... \;\geq\; \delta^{l+1} \, Q_{t+l+1}(c_0) \;\; \text{for all} \;\; l \in \{ 1, ..., 6 - t - 1 \}. \;\; \blacksquare$$

Proof of Lemma 2. The proof consists of three Conclusions (A, B, and C). Each Conclusion pertains to the setting where the firm operates under SR.

Conclusion A. The firm always implements immediately a cost reduction achieved in period 1.

Proof. If the firm implements the cost reduction achieved in period 1 immediately, the PDV of its profit is $\pi_1 \equiv \Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right]$ because $p_1 = p_2 = c_0$ and $p_3 = p_4 = p_5 = p_6 = c_0 - \Delta$.

If the firm first implements in period 2 the cost reduction achieved in period 1, the PDV of its profit is $\pi_2 \equiv \delta \, \Delta \, Q_2(c_0)$ because $p_1 = p_2 = c_0$ and $p_3 = p_4 = p_5 = p_6 = c_0 - \Delta$. It is apparent that $\pi_2 = \pi_1 - \Delta \, Q_1(c_0) < \pi_1$.

If the firm first implements in period 3 the cost reduction achieved in period 1, the PDV of its profit is $\pi_3 \equiv \delta^2 \, \Delta \left[Q_3(c_0) + \delta \, Q_4(c_0) \right]$ because $p_1 = p_2 = p_3 = p_4 = c_0$ and $p_5 = p_6 = c_0 - \Delta$. Assumption D implies:

$$\pi_3 \;=\; \delta \, \Delta \left[\delta \, Q_3(c_0) + \delta^2 \, Q_4(c_0) \right] \;<\; \delta \, \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right]$$

[58] Turner and Sappington (2024) provides detailed proofs.

[59] The firm will not delay the implementation of an achieved cost reduction to period 6. The discounted present value (PDV) of the firm's profit from such a delay is $\delta^5 \, \Delta \, Q_6(c_0)$. The PDV of the firm's profit from implementing the cost reduction in period 5 is $\delta^4 \, \Delta \left[Q_5(c_0) + \delta \, Q_6(c_0) \right] > \delta^5 \, \Delta \, Q_6(c_0)$.

$$= \Delta \left[\delta \, Q_2(c_0) + \delta^2 \, Q_3(c_0) \right] \; < \; \Delta \, [\, Q_1(c_0) + \delta \, Q_2(c_0) \,] \; = \; \pi_1 \, .$$

If the firm first implements in period 4 the cost reduction achieved in period 1, the PDV of its profit is $\pi_4 \equiv \delta^3 \, \Delta \, Q_4(c_0)$ because $p_1 = p_2 = p_3 = p_4 = c_0$ and $p_5 = p_6 = c_0 - \Delta$. It is apparent that $\pi_4 = \pi_3 - \delta^2 \, \Delta \, Q_3(c_0) < \pi_3 \, (< \pi_1)$.

If the firm implements in period 5 the cost reduction achieved in period 1, the PDV of its profit is $\pi_5 \equiv \delta^4 \, \Delta \, [\, Q_5(c_0) + \delta \, Q_6(c_0) \,]$ because $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = c_0$. Assumption D implies:

$$\begin{aligned} \pi_5 \; &= \; \delta^3 \, \Delta \left[\delta \, Q_5(c_0) + \delta^2 \, Q_6(c_0) \right] \; < \; \delta^3 \, \Delta \, [\, Q_4(c_0) + \delta \, Q_5(c_0) \,] \\ &= \; \delta^2 \, \Delta \left[\delta \, Q_4(c_0) + \delta^2 \, Q_5(c_0) \right] \; < \; \delta^2 \, \Delta \, [\, Q_3(c_0) + \delta \, Q_4(c_0) \,] \; = \; \pi_3 \; (< \pi_1) \, . \end{aligned}$$

If the firm implements in period 6 the Δ cost reduction achieved in period 1, the PDV of its profit is $\pi_6 \equiv \delta^5 \, \Delta \, Q_6(c_0)$ because $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = c_0$. It is apparent that $\pi_6 = \pi_5 - \delta^4 \, \Delta \, Q_5(c_0) < \pi_5 \, (< \pi_1)$. □

Conclusion B. The firm never delays beyond period 3 the implementation of a cost reduction achieved in period 2.

Proof. If the firm implements in period 3 the cost reduction it achieves in period 2, the PDV of its profit is $\pi_L \equiv \delta \, \Delta \, [\, Q_3(c_0) + \delta \, Q_4(c_0) \,]$ because $p_1 = p_2 = p_3 = p_4 = c_0$ and $p_5 = p_6 = c_0 - \Delta$. We will show that the maximum PDV of profit the firm can secure by delaying the implementation of the achieved cost reduction beyond period 3 is always less π_L.

If the firm delays to period 4 the implementation of the cost reduction achieved in period 2, the PDV of its profit is $\delta^2 \, \Delta \, Q_4(c_0)$ because $p_1 = p_2 = p_3 = p_4 = c_0$ and $p_5 = p_6 = c_0 - \Delta$. It is apparent that $\delta^2 \, \Delta \, Q_4(c_0) < \delta \, \Delta \, [\, Q_3(c_0) + \delta \, Q_4(c_0) \,] = \pi_L$.

If the firm delays to period 5 the implementation of the cost reduction achieved in period 2, the PDV of its profit is $\delta^3 \, \Delta \, [\, Q_5(c_0) + \delta \, Q_6(c_0)]$ because $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = c_0$. Assumption D implies:

$$\begin{aligned} \delta^3 \, \Delta \, [\, Q_5(c_0) + \delta \, Q_6(c_0) \,] \; &= \; \delta^2 \, \Delta \left[\delta \, Q_5(c_0) + \delta^2 \, Q_6(c_0) \right] < \; \delta^2 \, \Delta \, [\, Q_4(c_0) + \delta \, Q_5(c_0) \,] \\ &= \; \delta \, \Delta \left[\delta \, Q_4(c_0) + \delta^2 \, Q_5(c_0) \right] \; < \; \delta \, \Delta \, [\, Q_3(c_0) + \delta \, Q_4(c_0) \,] \; = \; \pi_L \, . \end{aligned} \tag{7}$$

If the firm delays to period 6 the implementation of the Δ cost reduction achieved in period 2, the PDV of its profit is $\delta^4 \, \Delta \, Q_6(c_0)$ because $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = c_0$. It is apparent that:

$$\delta^4 \, \Delta \, Q_6(c_0) \; < \; \delta^3 \, \Delta \, [\, Q_5(c_0) + \delta \, Q_6(c_0) \,] \; < \; \delta \, \Delta \, [\, Q_3(c_0) + \delta \, Q_4(c_0) \,] \; = \; \pi_L \, . \tag{8}$$

The last inequality in (8) reflects (7). □

Conclusion C. If the firm achieves the cost reduction in period 2, it implements the cost reduction immediately if (3) holds, and otherwise implements the cost reduction in period 3.

Proof. If the firm implements the achieved cost reduction in period 2, the PDV of its profit is $\Delta\, Q_2(c_0)$ because $p_1 = p_2 = c_0$ and $p_3 = p_4 = p_5 = p_6 = c_0 - \Delta$. If the firm delays the implementation the achieved cost reduction in period 2 to period 3, the PDV of its profit is $\delta\,\Delta\,[\,Q_3(c_0) + \delta\, Q_4(c_0)\,]$ because $p_1 = p_2 = p_3 = p_4 = c_0$ and $p_5 = p_6 = c_0 - \Delta$. Therefore, Conclusion B implies that the firm will implement the cost reduction immediately if the inequality in (3) holds, and otherwise delay the implementation to period 3. □ ■

Proof of the Corollary to Lemma 2. Define $Q_0 \equiv Q_1(c_0)$. Then when Assumption G holds, the inequality in (3) holds if and only if:

$$\delta\, g^2\, Q_0 + \delta^2\, g^3\, Q_0 \;\leq\; g\, Q_0 \;\Leftrightarrow\; \widetilde{\delta}^2 + \widetilde{\delta} - 1 \;\leq\; 0 \;\Leftrightarrow\; \widetilde{\delta} \;\leq\; \frac{1}{2}\left[\sqrt{5} - 1\right] \;\approx\; 0.618\,. \quad ■$$

Proof of Proposition 1. First consider the firm's problem in period 2 of the NID setting after no cost reduction is achieved in period 1. Under SR in this setting, the firm retains the full benefit of a cost reduction that is achieved in period 2 only for that period. Therefore, the firm's problem is:

$$\underset{\phi_2}{\text{Maximize}} \;\; \phi_2\, \Delta\, Q_2(c_0) - K_2(\phi_2)$$

$$\Rightarrow \;\; K_2'(\phi_2^S) \;=\; \Delta\, Q_2(c_0) \;\; \text{at an interior optimum.} \qquad (9)$$

Under IRIS in this setting and in the ID setting, the firm retains the full benefit of a cost reduction achieved in period 2 during both period 2 and period 3. Therefore, the firm's problem in period 2 is:

$$\underset{\phi_2}{\text{Maximize}} \;\; \phi_2\, \Delta\,[\,Q_2(c_0) + \delta\, Q_3(c_0)\,] - K_2(\phi_2)$$

$$\Rightarrow \;\; K_2'(\phi_2^I) \;=\; \Delta\,[\,Q_2(c_0) + \delta\, Q_3(c_0)\,] \;\; \text{at an interior optimum.} \qquad (10)$$

First suppose that $\phi_2^I = 0$. Then (10) implies that $\Delta\,[\,Q_2(c_0) + \delta\, Q_3(c_0)\,] \leq K_2'(0)$, which violates the maintained assumption that $K_2'(0) = 0$. Therefore, $\phi_2^I > 0$.

Next suppose that $\phi_2^I = 1$. Then (10) implies that $K_2'(1) \leq \Delta\,[\,Q_2(c_0) + \delta\, Q_3(c_0)\,]$, which violates the maintained assumption that $K_2'(1) > \Delta\,[\,Q_2(c_0) + \delta\, Q_3(c_0)\,]$. Therefore, $\phi_2^I < 1$.

The proof (by contradiction) that $\phi_2^S \in (0,1)$ is analogous. Because $\phi_2^S \in (0,1)$ and $\phi_2^I \in (0,1)$, (9) and (10) imply that $K_2'(\phi_2^I) > K_2'(\phi_2^S) \;\Rightarrow\; \phi_2^I > \phi_2^S$. The conclusion here reflects the convexity of $K_2(\cdot)$.

Now consider the firm's problem in period 2 of the ID setting after no cost reduction is achieved in period 1. Under SR in this setting, the firm delays to period 3 the implementation of a cost reduction achieved in period 2. Therefore, the firm's problem is:

$$\underset{\phi_2}{\text{Maximize}} \;\; \phi_2\, \delta\, \Delta\,[\,Q_3(c_0) + \delta\, Q_4(c_0)\,] - K_2(\phi_2)$$

$$\Rightarrow \;\; K_2'(\phi_2^S) \;=\; \delta\, \Delta\,[\,Q_3(c_0) + \delta\, Q_4(c_0)\,] \;\; \text{at an interior optimum.} \qquad (11)$$

The proof that $0 < \phi_2^S < \phi_2^I < 1$ in this setting is analogous. ■

Proof of Proposition 2. (9) implies that the firm's problem in period 1 under SR in the NID setting is:

$$\underset{\phi_1}{\text{Maximize}} \quad \phi_1 \, \Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right] + \left[1 - \phi_1 \right] \, \delta \left[\phi_2^S \, \Delta \, Q_2(c_0) - K_2(\phi_2^S) \right] - K_1(\phi_1) \, . \quad (12)$$

(12) implies that at an interior solution to this problem:

$$K_1'(\phi_1^S) \; = \; \Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right] - \delta \left[\phi_2^S \, \Delta \, Q_2(c_0) - K_2(\phi_2^S) \right] . \quad (13)$$

(10) implies that the firm's problem in period 1 under IRIS in both the NID setting and the ID setting is:

$$\underset{\phi_1}{\text{Maximize}} \quad \phi_1 \, \Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right]$$
$$+ \left[1 - \phi_1 \right] \, \delta \left\{ \phi_2^I \, \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] - K_2(\phi_2^I) \right\} - K_1(\phi_1) \, . \quad (14)$$

(14) implies that at an interior solution to this problem:

$$K_1'(\phi_1^I) \; = \; \Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right] - \delta \left\{ \phi_2^I \, \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] - K_2(\phi_2^I) \right\} . \quad (15)$$

Observe that:

$$\phi_2^I \, \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] - K_2(\phi_2^I) \; = \; \max_{\phi_2} \left\{ \phi_2 \, \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] - K_2(\phi_2) \right\}$$
$$> \; \phi_2^S \, \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] - K_2(\phi_2^S) \; > \; \phi_2^S \, \Delta \, Q_2(c_0) - K_2(\phi_2^S) \, . \quad (16)$$

The first inequality in (16) holds because $\phi_2^S \neq \phi_2^I$, from Proposition 1.

(14) implies that $\phi_1^I \; > \; 0$ in the NID setting if:

$$\Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right] - \delta \left[\phi_2^I \, \Delta \left(Q_2(c_0) + \delta \, Q_3(c_0) \right) - K_2(\phi_2^I) \right] \; > \; K_1'(0) \, . \quad (17)$$

Because $K_1'(0) \, = \, 0$ by assumption, the inequality in (17) holds if:

$$\Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right] \; > \; \delta \left[\phi_2^I \, \Delta \left(Q_2(c_0) + \delta \, Q_3(c_0) \right) - K_2(\phi_2^I) \right] .$$

Assumption D implies that this inequality holds.

(14) implies that $\phi_1^I \; < \; 1$ in the NID setting if:

$$\Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right] - \delta \left\{ \phi_2^I \, \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] - K_2(\phi_2^I) \right\} \; < \; K_1'(1) \, . \quad (18)$$

$\Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right] \; < \; K_1'(1)$, by assumption. Furthermore, $\phi_2^I \, \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] - K_2(\phi_2^I) \, \geq \, 0$ because $\phi_2^I \; = \; \arg\max_{\phi} \left\{ \phi \, \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] - K_2(\phi) \right\}$ and $K_2(0) \, = \, 0$. Therefore, the inequality in (18) holds.

(12) implies that $\phi_1^S > 0$ in the NID setting if:

$$\Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right] - \delta \left[\phi_2^S \, \Delta \, Q_2(c_0) - K_2(\phi_2^S) \right] \; > \; K_1'(0) \, . \quad (19)$$

$K_1'(0) \, = \, 0$ by assumption. Therefore, (19) holds if:

$$\Delta\left[\,Q_1(c_0) + \delta\, Q_2(c_0)\,\right] - \delta\left[\,\phi_2^S\,\Delta\, Q_2(c_0) - K_2(\phi_2^S)\,\right] \;>\; 0\,.$$

This inequality holds because:

$$Q_1(c_0) + \delta\, Q_2(c_0) \;>\; \delta\, Q_2(c_0) \quad\Rightarrow\quad Q_1(c_0) + \delta\, Q_2(c_0) \;>\; \phi_2^S\,\delta\, Q_2(c_0)$$

$$\Rightarrow\quad \Delta\left[\,Q_1(c_0) + \delta\, Q_2(c_0)\,\right] - \delta\left[\,\phi_2^S\,\Delta\, Q_2(c_0) - K_2(\phi_2^S)\,\right] \;>\; 0\,.$$

(12) implies that $\phi_1^S < 1$ in the NID setting if:

$$\Delta\left[\,Q_1(c_0) + \delta\, Q_2(c_0)\,\right] - \delta\left[\,\phi_2^S\,\Delta\, Q_2(c_0) - K_2(\phi_2^S)\,\right] \;<\; K_1'(1)\,. \tag{20}$$

$\Delta\left[\,Q_1(c_0) + \delta\, Q_2(c_0)\,\right] \;<\; K_1'(1)$, by assumption. Furthermore, $\phi_2^S\,\Delta\, Q_2(c_0) - K_2(\phi_2^S) \geq 0$ because $\phi_2^S = \arg\max_\phi \{\,\phi\,\Delta\, Q_2(c_0) - K_2(\phi)\,\}$ and $K_2(0) = 0$. Therefore, the inequality in (20) holds.

To prove that $\phi_1^I < \phi_1^S$ in the NID setting, observe that:

$$\begin{aligned}
\phi_1^I \;&=\; \arg\max_\phi \{\,\phi\,\Delta\left[\,Q_1(c_0) + \delta\, Q_2(c_0)\,\right] \\
&\qquad\qquad + \left[\,1-\phi\,\right]\,\delta\left[\,\phi_2^I\,\Delta\left(\,Q_2(c_0) + \delta\, Q_3(c_0)\,\right) - K_2(\phi_2^I)\,\right] - K_1(\phi)\,\} \\
&<\; \arg\max_\phi \{\,\phi\,\Delta\left[\,Q_1(c_0) + \delta\, Q_2(c_0)\,\right] \\
&\qquad\qquad + \left[\,1-\phi\,\right]\,\delta\left[\,\phi_2^S\,\Delta\, Q_2(c_0) - K_2(\phi_2^S)\,\right] - K_1(\phi)\,\} \;=\; \phi_1^S\,.
\end{aligned} \tag{21}$$

The equalities in (21) reflect (13) and (15) since $\phi_1^S \in (0,1)$ and $\phi_1^I \in (0,1)$. The inequality in (21) reflects (16) and the fact that the firm's profit-maximizing choice of ϕ_1 increases as the firm's expected profit following first-period failure to achieve a cost reduction declines, holding constant the firm's expected profit following first period success in securing a cost reduction.[60]

The proof that $0 < \phi_1^I < \phi_1^S < 1$ in the NID setting is analogous. ■

<u>Proof of Lemma 3</u>. Define $Q_t \equiv Q_t(c_0)$ for $t = 1, ..., 6$. (9) and (10) imply that under the maintained assumptions in the NID setting:

$$\phi_2^S \;=\; \frac{\Delta}{k_2}\, Q_2 \;=\; \frac{\Delta\, g}{k_2}\, Q_0 \quad\text{and}$$

$$\phi_2^I \;=\; \frac{\Delta}{k_2}\left[\,Q_2 + \delta\, Q_3\,\right] \;=\; \frac{\Delta}{k_2}\left[\,g\, Q_0 + \delta\, g^2\, Q_0\,\right] \;=\; \frac{\Delta\, g\left[\,1 + \delta\, g\,\right]}{k_2}\, Q_0\,. \tag{22}$$

(13) implies that in the NID setting:

$$\phi_1^S \;=\; \frac{1}{k_1}\left\{\,\Delta\left[\,1 + \delta\, g\,\right] Q_0 - \delta\left[\,\phi_2^S\,\Delta\, g\, Q_0 - K_2(\phi_2^S)\,\right]\,\right\}. \tag{23}$$

[60] Formally, if $\phi_1 \in (0,1) = \arg\max_\phi \{\,\phi\, A + \left[\,1-\phi\,\right] B - K_1(\phi)\,\}$, then $A - B = K_1'(\phi_1) \Rightarrow \frac{d\phi_1}{dB} = -\frac{1}{K_1''(\phi_1)} < 0$.

(15) implies:

$$\phi_1^I \;=\; \frac{1}{k_1}\left\{\,\Delta\,[\,1+\delta\,g\,]\,Q_0 - \delta\left[\,\phi_2^I\,\Delta\,g\,(1+\delta\,g)\,Q_0 - K_2(\phi_2^I)\,\right]\,\right\}. \tag{24}$$

(22) implies:

$$K_2(\phi_2^S) \;=\; \frac{k_2}{2}\left[\frac{\Delta\,g}{k_2}\;Q_0\right]^2 \;=\; \frac{\Delta^2\,g^2}{2\,k_2}\,[\,Q_0\,]^2\;;$$

$$K_2(\phi_2^I) \;=\; \frac{k_2}{2}\left[\frac{\Delta\,g\,(1+\delta\,g)}{k_2}\;Q_0\right]^2 \;=\; \frac{\Delta^2\,g^2}{2\,k_2}\,[\,1+\delta\,g\,]^2\,[\,Q_0\,]^2. \tag{25}$$

(22) and (25) imply:

$$\phi_2^S\,\Delta\,Q_2(c_0) - K_2(\phi_2^S) \;=\; \frac{\Delta\,g}{k_2}\;Q_0\;\Delta\,g\,Q_0 - \frac{\Delta^2\,g^2}{2\,k_2}\,[\,Q_0\,]^2 \;=\; \frac{\Delta^2\,g^2}{2\,k_2}\,[\,Q_0\,]^2. \tag{26}$$

(22) and (25) also imply:

$$\phi_2^I\,\Delta\,[\,Q_2+\delta\,Q_3\,] - K_2(\phi_2^I) \;=\; \frac{\Delta^2\,g^2}{2\,k_2}\,[\,1+\delta\,g\,]^2\,[\,Q_0\,]^2. \tag{27}$$

(23) and (26) imply:

$$\phi_1^S \;=\; \frac{\Delta}{k_1}\,[\,1+\delta\,g\,]\,Q_0 - \frac{\delta}{k_1}\left[\frac{\Delta^2\,g^2}{2\,k_2}\right][\,Q_0\,]^2 \;=\; \frac{\Delta\,Q_0}{k_1}\left[1+\delta\,g - \frac{\delta\,\Delta\,g^2}{2\,k_2}\;Q_0\right]. \tag{28}$$

(22) and (28) imply:

$$\phi_1^S\,\phi_2^S \;=\; \frac{\Delta^2\,g}{k_1\,k_2}\left[1+\delta\,g - \frac{\delta\,\Delta\,g^2}{2\,k_2}\;Q_0\right][\,Q_0\,]^2. \tag{29}$$

(24) and (27) imply:

$$\phi_1^I \;=\; \frac{\Delta\,Q_0}{k_1}\left[1+\delta\,g - \frac{\delta\,\Delta\,g^2}{2\,k_2}\,(\,1+\delta\,g\,)^2\;Q_0\right]. \tag{30}$$

(22) and (30) imply:

$$\phi_1^I\,\phi_2^I \;=\; \frac{\Delta^2\,g\,[\,1+\delta\,g\,]}{k_1\,k_2}\left[1+\delta\,g - \frac{\delta\,\Delta\,g^2}{2\,k_2}\,(\,1+\delta\,g\,)^2\;Q_0\right][\,Q_0\,]^2\,. \tag{31}$$

(4), (22), (28), and (29) imply:

$$\begin{aligned}\Phi^S \;=\; \phi_1^S+\phi_2^S-\phi_1^S\,\phi_2^S \;=\; &\frac{\Delta\,Q_0}{k_1}\left[1+\delta\,g - \frac{\delta\,\Delta\,g^2}{2\,k_2}\;Q_0\right] + \frac{\Delta\,g}{k_2}\;Q_0\\ &- \frac{\Delta^2\,g}{k_1\,k_2}\left[1+\delta\,g - \frac{\delta\,\Delta\,g^2}{2\,k_2}\;Q_0\right][\,Q_0\,]^2\,.\end{aligned} \tag{32}$$

(4), (22), (30), and (31) imply:

$$\Phi^I \; = \; \phi_1^I + \phi_2^I - \phi_1^I \, \phi_2^I \; = \; \frac{\Delta \, Q_0}{k_1} \left[1 + \delta \, g - \frac{\delta \, \Delta \, g^2}{2 \, k_2} \, (1 + \delta \, g)^2 \; Q_0 \right] + \frac{\Delta \, g \, [\, 1 + \delta \, g \,]}{k_2} \, Q_0$$

$$- \; \frac{\Delta^2 \, g \, [\, 1 + \delta \, g \,]}{k_1 \, k_2} \left[1 + \delta \, g - \frac{\delta \, \Delta \, g^2}{2 \, k_2} \, (1 + \delta \, g)^2 \; Q_0 \right] [\, Q_0 \,]^2 \; . \quad (33)$$

(32) and (33) imply that because $\delta > 0$:

$$\begin{aligned} \Phi^S - \Phi^I \; &= \; \frac{\Delta \, Q_0}{k_1} \left[\frac{\delta \, \Delta \, g^2}{2 \, k_2} \right] Q_0 \left[\, (1 + \delta \, g)^2 - 1 \, \right] - \frac{\Delta \, g}{k_2} \, Q_0 \, [\, 1 + \delta \, g - 1 \,] \\ &\quad + \frac{\Delta^2 \, g \, [\, 1 + \delta \, g \,]}{k_1 \, k_2} \left[1 + \delta \, g - \frac{\delta \, \Delta \, g^2}{2 \, k_2} \, (1 + \delta \, g)^2 \; Q_0 \right] [\, Q_0 \,]^2 \\ &\quad - \frac{\Delta^2 \, g}{k_1 \, k_2} \left[1 + \delta \, g - \frac{\delta \, \Delta \, g^2}{2 \, k_2} \, Q_0 \right] [\, Q_0 \,]^2 \\ &\stackrel{s}{=} \; \frac{\delta^2 \, \Delta \, g^2 \, Q_0}{2 \, k_1} \, [\, 2 + \delta \, g \,] - \delta \, g \\ &\quad + \frac{\Delta \, Q_0}{k_1} \left\{ \delta \, g \, [\, 1 + \delta \, g \,] - \frac{\delta \, \Delta \, g^2}{2 \, k_2} \left[\, (1 + \delta \, g)^3 - 1 \, \right] Q_0 \right\} \\ &\stackrel{s}{=} \; \Delta \, Q_0 \, \left\{ \left[\, 2 + 4 \, \delta \, g + (\delta \, g)^2 \, \right] k_2 - g \left[\, (1 + \delta \, g)^3 - 1 \, \right] \Delta \, Q_0 \right\} - 2 \, k_1 \, k_2 \, . \end{aligned} \quad (34)$$

(11) implies that under the maintained conditions in the ID setting:

$$\phi_2^S \; = \; \frac{\Delta \, \delta \, [\, Q_3(c_0) + \delta \, Q_4(c_0) \,]}{k_2} \; = \; \frac{\Delta \, \delta \, Q_0 \, [\, g^2 + \delta \, g^3 \,]}{k_2} \; = \; \frac{\Delta \, Q_0 \, \widetilde{\delta} \, [\, 1 + \widetilde{\delta} \,]}{\widetilde{k}_2} \, . \quad (35)$$

When Assumption G holds and the ID setting prevails, the PDV of the firm's profit in period 2 when it achieves the Δ cost reduction in that period is:

$$\pi_2^S \; = \; \Delta \, \delta \, [\, Q_3(c_0) + \delta \, Q_4(c_0) \,] \; = \; \Delta \, \delta \left[\, g^2 \, Q_0 + \delta \, g^3 \, Q_0 \, \right] = \; \Delta \, Q_0 \, g \; \widetilde{\delta} \, [\, 1 + \widetilde{\delta} \,]. \quad (36)$$

(35) and (36) imply:

$$\phi_2^S \, \pi_2^S - K(\phi_2^S) \; = \; \frac{g \left[\, \Delta \, Q_0 \, \widetilde{\delta} \, (\, 1 + \widetilde{\delta} \,) \, \right]^2}{2 \, \widetilde{k}_2} \, . \quad (37)$$

(37) implies that under the specified conditions:

$$\begin{aligned} \phi_1^S \; &= \; \frac{1}{k_1} \, \left\{ \, \Delta \, [\, Q_1(c_0) + \delta \, Q_2(c_0) \,] - \delta \left[\, \phi_2^S \, \pi_2^S - K(\phi_2^S) \, \right] \right\} \\ &= \; \frac{\Delta \, Q_0 \, [\, 1 + \widetilde{\delta} \,]}{k_1} \left[1 - \frac{(\widetilde{\delta})^3}{2 \, \widetilde{k}_2} \, \Delta \, Q_0 \, (\, 1 + \widetilde{\delta} \,) \right] . \end{aligned} \quad (38)$$

(10) implies that under the specified conditions:

$$\phi_2^I \;=\; \frac{\Delta\,[\,Q_2(c_0)+\delta\,Q_3(c_0)\,]}{k_2} \;=\; \frac{\Delta\,Q_0\,[\,g+\delta\,g^2\,]}{k_2} \;=\; \frac{\Delta\,Q_0\,[\,1+\widetilde{\delta}\,]}{\widetilde{k}_2}\,. \tag{39}$$

(39) implies:

$$\phi_2^I\,\Delta\,Q_0\,g\,[\,1+\delta\,g\,]-K(\phi_2^I) \;=\; \frac{g}{2\,\widetilde{k}_2}\left[\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\right]^2. \tag{40}$$

(15) and (40) imply:

$$\begin{aligned}\phi_1^I \;&=\; \frac{1}{k_1}\left\{\,\Delta\,[\,Q_1(c_0)+\delta\,Q_2(c_0)\,]-\delta\left[\,\phi_2^I\,\Delta\,g\,Q_0\,(1+g\,\delta)-K(\phi_2^I)\,\right]\right\}\\ &=\; \frac{\Delta\,Q_0\,[\,1+\widetilde{\delta}\,]}{k_1}\left[\,1-\frac{\widetilde{\delta}}{2\,\widetilde{k}_2}\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\right].\end{aligned} \tag{41}$$

(35) and (38) imply:

$$\phi_1^S\,\phi_2^S \;=\; \frac{\widetilde{\delta}\,[\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\,]^2}{k_1\,\widetilde{k}_2}\left[\,1-\frac{(\widetilde{\delta}\,)^3}{2\,\widetilde{k}_2}\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\right]. \tag{42}$$

(39) and (41) imply:

$$\phi_1^I\,\phi_2^I \;=\; \frac{[\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\,]^2}{k_1\,\widetilde{k}_2}\left[\,1-\frac{\widetilde{\delta}}{2\,\widetilde{k}_2}\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\right]. \tag{43}$$

(4), (35), (38), and (42) imply:

$$\begin{aligned}\Phi^S \;=\;& \frac{\Delta\,Q_0\,[\,1+\widetilde{\delta}\,]}{k_1}\left[\,1-\frac{(\widetilde{\delta}\,)^3}{2\,\widetilde{k}_2}\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\right]+\frac{\Delta\,Q_0\,\widetilde{\delta}\,[\,1+\widetilde{\delta}\,]}{\widetilde{k}_2}\\ &-\;\frac{\widetilde{\delta}\,[\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\,]^2}{k_1\,\widetilde{k}_2}\left[\,1-\frac{(\widetilde{\delta}\,)^3}{2\,\widetilde{k}_2}\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\right].\end{aligned} \tag{44}$$

(4), (39), (41), and (43) imply:

$$\begin{aligned}\Phi^I \;=\;& \frac{\Delta\,Q_0\,[\,1+\widetilde{\delta}\,]}{k_1}\left[\,1-\frac{\widetilde{\delta}}{2\,\widetilde{k}_2}\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\right]+\frac{\Delta\,Q_0\,[\,1+\widetilde{\delta}\,]}{\widetilde{k}_2}\\ &-\;\frac{[\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\,]^2}{k_1\,\widetilde{k}_2}\left[\,1-\frac{\widetilde{\delta}}{2\,\widetilde{k}_2}\,\Delta\,Q_0\,(\,1+\widetilde{\delta}\,)\right].\end{aligned} \tag{45}$$

(44) and (45) imply:

$$\Phi^S - \Phi^I = \frac{\widetilde{\delta}}{2\,k_1\,\widetilde{k}_2} \left[\Delta\,Q_0\,(\,1+\widetilde{\delta}\,) \right]^2 [\,1-(\,\widetilde{\delta}\,)^2\,] - \frac{\Delta\,Q_0\,[\,1+\widetilde{\delta}\,]}{\widetilde{k}_2}\,[\,1-\widetilde{\delta}\,]$$

$$+ \frac{\left[\Delta\,Q_0\,(\,1+\widetilde{\delta}\,) \right]^2}{k_1\,\widetilde{k}_2}\,[\,1-\widetilde{\delta}\,] - \frac{\left[\Delta\,Q_0\,(\,1+\widetilde{\delta}\,) \right]^3}{k_1\,\widetilde{k}_2}\,\frac{\widetilde{\delta}}{2\,\widetilde{k}_2}\,[\,1-(\,\widetilde{\delta}\,)^3\,]$$

$$\stackrel{s}{=} \Delta\,Q_0\,[\,1+\widetilde{\delta}\,] \left\{ [\,2 - \widetilde{\delta} - (\,\widetilde{\delta}\,)^3\,]\,\widetilde{k}_2 - \widetilde{\delta}\,[\,1+\widetilde{\delta}\,]\,[\,1-(\,\widetilde{\delta}\,)^3\,]\,\Delta\,Q_0 \right\}$$

$$-\,2\,[\,1-\widetilde{\delta}\,]\,k_1\,\widetilde{k}_2\,. \quad \blacksquare \tag{46}$$

Proof of Proposition 3. When $g = 1$, the first term in G^{NID} (as defined below (4)) is :

$$\Delta\,Q_0\,\left[\,\left(2+4\,\delta+\delta^2\right)k_2 - \left(3+3\,\delta+\delta^2\right)\delta\,\Delta\,Q_0\,\right]$$

$$> \Delta\,Q_0\,\left[\,\left(2+4\,\delta+\delta^2\right)Q\,(1+\delta)\,\Delta - \left(3+3\,\delta+\delta^2\right)\delta\,\Delta\,Q_0\,\right] \tag{47}$$

$$= \Delta^2\,(Q_0)^2\left[\,2+3\,\delta+2\,\delta^2\,\right]. \tag{48}$$

The inequality in (47) reflects the maintained assumption that for $t \in \{2,3\}$, $K_2'(1) = k_2 > \Delta\,[\,Q_t(c_0) + \delta\,Q_{t+1}(c_0)\,] = \Delta\,[\,1+\delta\,]\,Q_0 \Rightarrow Q_0 < \frac{k_2}{[\,1+\delta\,]\,\Delta}$. (48) and Lemma 3 imply that $\Phi^S > \Phi^I$ if:

$$\Delta^2\,(Q_0)^2\left[\,2+3\,\delta+2\,\delta^2\,\right] > 2\,k_1\,k_2 \;\Leftrightarrow\; \Delta\,Q_0 > \sqrt{\frac{2\,k_1\,k_2}{2+3\,\delta+2\,\delta^2}}\,. \tag{49}$$

The inequality in (49) holds if $\Delta\,Q_0 > \sqrt{k_1\,k_2}$ because $2+3\,\delta+2\,\delta^2 > 2$.

Let Q denote Q_0 in the ensuing analysis. Then Lemma 3 implies that when $g = 1$, $\Phi^S > \Phi^I$ in the ID setting if:

$$\Lambda \equiv [\,1+\delta\,]\,\Delta\,Q\left\{\left[\,2-\delta-\delta^3\,\right]k_2 - \Delta\;\delta\;[\,1+\delta\,]\left[\,1-\delta^3\,\right]Q\right\} - 2\,[\,1-\delta\,]\,k_1\,k_2 > 0\,. \tag{50}$$

The maintained assumption that $K_2'(1) > \max\{\,\Delta\,[\,Q_2(c_0)+\delta\,Q_3(c_0)\,]\,,\,\Delta\,[\,Q_3(c_0)+\delta\,Q_4(c_0)\,]\,\}$ implies that $Q \leq \frac{k_2}{[\,1+\delta\,]\,\Delta}$ in the present setting, which, in turn, implies:

$$\Lambda \geq [\,1+\delta\,]\,\Delta\,Q\left\{\left[\,2-\delta-\delta^3\,\right][\,1+\delta\,]\,\Delta\,Q - \Delta\;\delta\;[\,1+\delta\,]\left[\,1-\delta^3\,\right]Q\right\} - 2\,[\,1-\delta\,]\,k_1\,k_2$$

$$= [\,1+\delta\,]^2\,[\,1-\delta\,]\left[\,2-\delta^3\,\right][\,\Delta\,Q\,]^2 - 2\,[\,1-\delta\,]\,k_1\,k_2$$

$$> 0 \;\text{ if }\; \Delta\,Q > \sqrt{\frac{2\,k_1\,k_2}{[\,1+\delta\,]^2\left[\,2-\delta^3\,\right]}}\,. \tag{51}$$

The conclusion in the Proposition follows from (50) and (51) if $[\,1+\delta\,]^2\left[\,2-\delta^3\,\right] \geq 2$ for all $\delta \in (0,1)$, which is readily proved. $\blacksquare$

Proof of Proposition 4. Lemma 3 implies that $\Phi^S > \Phi^I$ in the ID setting when $k_1 = k_2 = \Delta Q_0 [1 + g\delta]$ if:

$$[1 + \delta g] \Delta Q_0 \left\{ \left[2 - \delta g - (\delta g)^3\right] \frac{k_2}{g} - \Delta\, g \delta [1 + g\delta] \left[1 - (g\delta)^3\right] Q_0 \right\}$$

$$> \; 2 [1 - g\delta] k_1 \frac{k_2}{g} \tag{52}$$

$$\Leftrightarrow \; k_1 \left\{ \left[2 - \delta g - (\delta g)^3\right] \frac{k_1}{g} - k_1 g \delta \left[1 - (g\delta)^3\right] \right\} > 2 \frac{1}{g} [1 - g\delta] [k_1]^2 \tag{53}$$

$$\Leftrightarrow \; 1 - \delta^2 g^2 - g + g^4 \delta^3 > 0. \tag{54}$$

(53) reflects the assumption that $k_1 = k_2 = \Delta Q_0 [1 + g\delta]$. The inequality in (54) holds because $1 - \delta^2 g^2 - g = 1 - g\left[1 + g\delta^2\right] > 0$, by assumption.

The inequality in (52) will continue to hold when $k_1 = k_2$ is increased marginally to ensure that $\Delta Q_0 [1 + g\delta] < \min\{k_1, k_2\}$. ■

Proof of Proposition 5. Under the specified conditions, when the firm operates under SR in the NID setting: (i) $p_1 = p_2 = c_0$; (ii) $p_3 = p_4 = p_5 = p_6 = c_0$ if the firm never achieves a cost reduction; and (iii) $p_3 = p_4 = p_5 = p_6 = c_0 - \Delta$ if the firm ever achieves a cost reduction. Therefore, the PDV of expected consumer surplus under SR in this setting is:

$$\begin{aligned} E_d\{W^S\} \;=\; & W_1(c_0) + \delta W_2(c_0) + \delta^2 W_3(c_0) + \delta^3 W_4(c_0) + \delta^4 W_5(c_0) + \delta^5 W_6(c_0) \\ & + \Phi^S \delta^2 [W_3(c_0 - \Delta) - W_3(c_0)] + \Phi^S \delta^3 [W_4(c_0 - \Delta) - W_4(c_0)] \\ & + \Phi^S \delta^4 [W_5(c_0 - \Delta) - W_5(c_0)] + \Phi^S \delta^5 [W_6(c_0 - \Delta) - W_6(c_0)]. \end{aligned} \tag{55}$$

Under the specified conditions, when the firm operates under IRIS: (i) $p_1 = p_2 = c_0$; (ii) $p_5 = p_6 = c_0$ if the firm never achieves success; (iii) $p_5 = p_6 = c_0 - \Delta$ if the firm ever achieves success; (iv) $p_3 = c_0 - \Delta$ if the firm achieves success in period 1; (v) $p_3 = c_0$ if the firm does not achieve success in period 1; (vi) $p_4 = c_0 - \Delta$ if the firm achieves success (in period 1 or period 2); and (vii) $p_4 = c_0$ if the firm does not achieve success. Therefore, the PDV of expected consumer surplus under IRIS in this setting is:

$$\begin{aligned} E_d\{W^I\} \;=\; & W_1(c_0) + \delta W_2(c_0) + \delta^2 W_3(c_0) + \delta^3 W_4(c_0) + \delta^4 W_5(c_0) + \delta^5 W_6(c_0) \\ & + \phi_1^I \delta^2 [W_3(c_0 - \Delta) - W_3(c_0)] + \delta^3 \Phi^I [W_4(c_0 - \Delta) - W_4(c_0)] \\ & + \delta^4 \Phi^I [W_5(c_0 - \Delta) - W_5(c_0)] + \delta^5 \Phi^I [W_6(c_0 - \Delta) - W_6(c_0)]. \end{aligned} \tag{56}$$

(55) and (56) imply:

$$\begin{aligned} E_d\{W^S\} - E_d\{W^I\} \;=\; & \left[\Phi^S - \phi_1^I\right] \delta^2 [W_3(c_0 - \Delta) - W_3(c_0)] \\ & + \left[\Phi^S - \Phi^I\right] \{\delta^3 [W_4(c_0 - \Delta) - W_4(c_0)] + \delta^4 [W_5(c_0 - \Delta) - W_5(c_0)] \end{aligned}$$

$$+ \ \delta^5 \, [\, W_6(c_0 - \Delta) - W_6(c_0) \,] \, \}. \tag{57}$$

If $\Phi^S > \Phi^I$, then $\Phi^S > \phi_1^I$. Consequently, (57) implies that $E_d\{W^S\} > E_d\{W^I\}$ when $\Phi^S > \Phi^I$ (because $\delta > 0$, by assumption). ■

Proof of Proposition 6. (10) and (11) imply that under the specified conditions:

$$k_2 \left(\phi_2^S\right)^{\gamma-1} \ = \ \Delta \ g\, Q_0 \ \Rightarrow \ \phi_2^S \ = \ \left[\frac{\Delta\, g\, Q_0}{k_2} \right]^{\frac{1}{\gamma-1}} \ = \ \left[\frac{\Delta\, Q_0}{\widetilde{k}_2} \right]^{\frac{1}{\gamma-1}} \quad \text{and}$$

$$k_2 \left(\phi_2^I\right)^{\gamma-1} \ = \ \Delta \left[\, g\, Q_0 + g^2\, \delta\, Q_0 \,\right]$$

$$\Rightarrow \ \phi_2^I \ = \ \left[\frac{\Delta\, Q_0\, g \left(1 + \widetilde{\delta}\right)}{k_2} \right]^{\frac{1}{\gamma-1}} \ = \ [\, 1 + \widetilde{\delta}\,]^{\frac{1}{\gamma-1}} \left[\frac{\Delta\, Q_0}{\widetilde{k}_2} \right]^{\frac{1}{\gamma-1}} \ = \ \phi_2^S\, [\, 1 + \widetilde{\delta}\,]^{\frac{1}{\gamma-1}}. \tag{58}$$

When Assumption G holds, $W_t(p) \ = \ g\, W_t(p)$. Therefore, (57) implies:

$$E_d\{W^S\} \ > \ E_d\{W^I\} \ \text{ if } \ \Phi^S - \phi_1^I + \left[\, \Phi^S - \Phi^I \,\right] [\, \widetilde{\delta} + \widetilde{\delta}^2 + \widetilde{\delta}^3 \,] \ > \ 0\,. \tag{59}$$

First suppose that $\Phi^S \geq \Phi^I$. Proposition 2 implies that $\Phi^S > \phi_1^I$. Therefore, (59) implies that $E_d\{W^S\} > E_d\{W^I\}$ when $\Phi^S \geq \Phi^I$.

Now suppose that $\Phi^S < \Phi^I$. (59) holds in this case if:

$$\Phi^S - \phi_1^I + \left[\, \Phi^S - \Phi^I \,\right] \, [\, \widetilde{\delta} + \widetilde{\delta}^2 + \widetilde{\delta}^3 + \widetilde{\delta}^4 + \ldots \,] \ > \ 0$$

$$\Leftrightarrow \ \ \Phi^S - \phi_1^I + \left[\, \Phi^S - \Phi^I \,\right] \, \frac{\widetilde{\delta}}{1 - \widetilde{\delta}} \ > \ 0$$

$$\Leftrightarrow \ \ \left[1 - \widetilde{\delta} \right] \, [\, \phi_1^S + \phi_2^S \left(1 - \phi_1^S\right) - \phi_1^I]$$

$$+ \ \widetilde{\delta}\, [\, \phi_1^S + \phi_2^S \left(1 - \phi_1^S\right) - \left(\phi_1^I + \phi_2^I \, [\, 1 - \phi_1^I \,]\right) \,] \ > \ 0$$

$$\Leftrightarrow \ \ [\, 1 - \phi_1^I \,] \left[1 - \widetilde{\delta}\, \phi_2^I \right] - [\, 1 - \phi_1^S \,] \, [\, 1 - \phi_2^S \,] \ > \ 0\,. \tag{60}$$

Proposition 2 implies:

$$1 - \phi_1^I \ > \ 1 - \phi_1^S\,. \tag{61}$$

(58) implies that when $\widetilde{\delta}\, [\, 1 + \widetilde{\delta}\,]^{\frac{1}{\gamma-1}} < 1$:

$$\phi_2^S - \widetilde{\delta}\, \phi_2^I \ = \ \phi_2^S - \widetilde{\delta}\, \phi_2^S\, [\, 1 + \widetilde{\delta}\,]^{\frac{1}{\gamma-1}} \ = \ \phi_2^S \left[1 - \widetilde{\delta} \left(1 + \widetilde{\delta}\right)^{\frac{1}{\gamma-1}} \right] \ > \ 0$$

$$\Rightarrow \ \ \phi_2^S \ > \ \widetilde{\delta}\, \phi_2^I \ \ \Rightarrow \ \ 1 - \widetilde{\delta}\, \phi_2^I \ > \ 1 - \phi_2^S\,. \tag{62}$$

(61) and (62) imply that the inequality in (60) holds. ■

Proof of the Corollary to Proposition 6.

The Corollary follows directly from Proposition 6 because $\widetilde{\delta}\,[\,1+\widetilde{\delta}\,]^{\frac{1}{\gamma-1}} < 1$ under the specified conditions. This is the case because:

$$\widetilde{\delta}\,[\,1+\widetilde{\delta}\,]^{\frac{1}{\gamma-1}} \;\leq\; \widetilde{\delta}\,[\,1+\widetilde{\delta}\,] \;<\; 1\,.$$

The first inequality here holds because $\gamma \geq 2$, by assumption. The last inequality here holds because $\widetilde{\delta} = g\,\delta < \widehat{\widehat{\delta}}$ in the NID setting and because $\widehat{\widehat{\delta}}\,[\,1+\widehat{\widehat{\delta}}\,] = 1$, by definition. ■

Proof of Proposition 7. Lemma 2 implies that when the firm operates under SR in the ID setting: (i) $p_1 = p_2 = c_0$; (ii) $p_5 = p_6 = c_0$ if the firm never achieves success;[61] (iii) $p_5 = p_6 = c_0 - \Delta$ if the firm ever achieves success; (iv) $p_3 = p_4 = c_0 - \Delta$ if the firm achieves success in period 1; and (v) $p_3 = p_4 = c_0$ if the firm does not achieve success in period 1. Therefore, expected consumer surplus under SR in this setting is:

$$\begin{aligned}
E_d\{W^S\} \;=\;& W_1(c_0) + \delta\,W_2(c_0) + \phi_1^S\left[\,\delta^2\,W_3(c_0-\Delta) + \delta^3\,W_4(c_0-\Delta)\,\right] \\
&+ \left[\,1-\phi_1^S\,\right]\left[\,\delta^2\,W_3(c_0) + \delta^3\,W_4(c_0)\,\right] + \left[\,1-\Phi^S\right]\left[\,\delta^4\,W_5(c_0) + \delta^5\,W_6(c_0)\,\right] \\
&+ \Phi^S\left[\,\delta^4\,W_5(c_0-\Delta) + \delta^5\,W_6(c_0-\Delta)\,\right] \\
=\;& W_1(c_0) + \delta\,W_2(c_0) + \delta^2\,W_3(c_0) + \delta^3\,W_4(c_0) + \delta^4\,W_5(c_0) + \delta^5\,W_6(c_0) \\
&+ \delta^2\,\phi_1^S\,[\,W_3(c_0-\Delta) - W_3(c_0)\,] + \delta^3\,\phi_1^S\,[\,W_4(c_0-\Delta) - W_4(c_0)\,] \\
&+ \delta^4\,\Phi^S\,[\,W_5(c_0-\Delta) - W_5(c_0)\,] + \delta^5\,\Phi^S\,[\,W_6(c_0-\Delta) - W_6(c_0)\,]\,.
\end{aligned} \tag{63}$$

(56) and (63) imply that in the ID setting $E_d\{W^S\} - E_d\{W^I\}$ is:

$$\begin{aligned}
&\delta^2\left[\,\phi_1^S - \phi_1^I\,\right][\,W_3(c_0-\Delta) - W_3(c_0)\,] + \delta^3\left[\,\phi_1^S - \Phi^I\,\right][\,W_4(c_0-\Delta) - W_4(c_0)\,] \\
&\quad + \delta^4\left[\,\Phi^S - \Phi^I\,\right][\,W_5(c_0-\Delta) - W_5(c_0)\,] + \delta^5\left[\,\Phi^S - \Phi^I\,\right][\,W_6(c_0-\Delta) - W_6(c_0)\,] \\
&\stackrel{s}{=} \left[\,\phi_1^S - \phi_1^I\,\right][\,W_3(c_0-\Delta) - W_3(c_0)\,] + \delta\left[\,\phi_1^S - \Phi^I\,\right][\,W_4(c_0-\Delta) - W_4(c_0)\,] \\
&\quad + \delta^2\left[\,\Phi^S - \Phi^I\,\right][\,W_5(c_0-\Delta) - W_5(c_0)\,] + \delta^3\left[\,\Phi^S - \Phi^I\,\right][\,W_6(c_0-\Delta) - W_6(c_0)\,]\,.
\end{aligned} \tag{64}$$

Define $Q_0 \equiv Q(c_0)$. (10) implies that under the specified conditions:

$$k_2\left(\phi_2^I\right)^{\gamma-1} \;=\; \Delta\left[\,g\,Q_0 + g^2\,\delta\,Q_0\,\right] \;\;\Rightarrow\;\; \phi_2^I \;=\; \left[\frac{\Delta\,Q_0\,g\,(1+g\,\delta)}{k_2}\right]^{\frac{1}{\gamma-1}}. \tag{65}$$

(11) implies that under the maintained conditions:

$$\phi_2^S \;=\; \left[\frac{\Delta\,Q_0\,\delta\,g^2\,(1+g\,\delta)}{k_2}\right]^{\frac{1}{\gamma-1}} \;=\; \phi_2^I\,(\delta\,g)^{\frac{1}{\gamma-1}}\,. \tag{66}$$

[61] The firm achieves "success" when it achieves the Δ cost reduction.

Define $\phi_2^{Lim} \equiv \left(\frac{2\,\Delta\,Q_0}{k_2} \right)^{\frac{1}{\gamma-1}}$. (65) and (66) imply that under the specified conditions:

$$\phi_2^I \;\rightarrow\; \phi_2^{Lim} \;\text{ and }\; \phi_2^S \;\rightarrow\; \phi_2^{Lim} \text{ as } \delta \rightarrow 1 \quad \Rightarrow \quad \lim_{\delta \rightarrow 1} \left(\phi_2^I - \phi_2^S \right) \;=\; 0\,. \tag{67}$$

Define $\phi_1^{Lim} \equiv \left(\frac{2\,\Delta\,Q_0 - \left[\,2\,\phi_2^{Lim}\,\Delta\,Q_0 - K_2(\phi_2^{Lim})\,\right]}{k_1} \right)^{\frac{1}{\gamma-1}}$. (13), (15), and (67) imply that under the specified conditions:

$$\phi_1^S \;\rightarrow\; \phi_1^{Lim} \;\text{ and }\; \phi_1^S \;\rightarrow\; \phi_1^{Lim} \;\text{ as } \delta \rightarrow 1 \quad \Rightarrow \quad \lim_{\delta \rightarrow 1} \left(\phi_1^I - \phi_1^S \right) \;=\; 0\,. \tag{68}$$

(67) and (68) imply:

$$\lim_{\delta \rightarrow 1} \left(\Phi^I - \Phi^S \right) \;=\; 0\,. \tag{69}$$

(64), (68), and (69) imply that $\lim_{\delta \rightarrow 1} \left(E_d\{W^S\} - E_d\{W^I\} \right)$ is:

$$\begin{aligned}
&\lim_{\delta \rightarrow 1} \delta^2 \left[\, \phi_1^S - \phi_1^I \,\right] \left[\, W_3(c_0 - \Delta) - W_3(c_0) \,\right] + \lim_{\delta \rightarrow 1} \delta^3 \left[\, \phi_1^S - \Phi^I \,\right] \left[\, W_4(c_0 - \Delta) - W_4(c_0) \,\right] \\
&+ \lim_{\delta \rightarrow 1} \delta^4 \left[\, \Phi^S - \Phi^I \,\right] \left[\, W_5(c_0 - \Delta) - W_5(c_0) \,\right] + \lim_{\delta \rightarrow 1} \delta^5 \left[\, \Phi^S - \Phi^I \,\right] \left[\, W_6(c_0 - \Delta) - W_6(c_0) \,\right] \\
&= \lim_{\delta \rightarrow 1} \delta^3 \left[\, \phi_1^S - \Phi^I \,\right] \left[\, W_4(c_0 - \Delta) - W_4(c_0) \,\right] \\
&< \lim_{\delta \rightarrow 1} \delta^3 \left[\, \Phi^S - \Phi^I \,\right] \left[\, W_4(c_0 - \Delta) - W_4(c_0) \,\right] \;=\; 0\,.
\end{aligned}$$

The inequality here holds because, from (4), $\Phi^S \;=\; \phi_1^S + \phi_2^S \left[\, 1 - \phi_1^I \,\right] > \phi_1^S$. ■

Proof of Proposition 8. (68) implies that under the specified conditions:

$$\lim_{k_1 \rightarrow \infty} \phi_1^S \;=\; 0 \;\text{ and }\; \lim_{k_1 \rightarrow \infty} \phi_1^I \;=\; 0\,. \tag{70}$$

(64) implies that under the specified conditions:

$$\begin{aligned}
E_d\{W^S\} - E_d\{W^I\} \;=\;& \delta^2 \left[\, \phi_1^S - \phi_1^I \,\right] D_{W3}(c_0, \Delta) + \delta^3 \left[\, \phi_1^S - \Phi^I \,\right] D_{W4}(c_0, \Delta) \\
&+ \delta^4 \left[\, \Phi^S - \Phi^I \,\right] D_{W5}(c_0, \Delta) + \delta^5 \left[\, \Phi^S - \Phi^I \,\right] D_{W6}(c_0, \Delta)
\end{aligned} \tag{71}$$

where $D_{Wt}(c_0, \Delta) \;=\; W_t(c_0 - \Delta) - W_t(c_0) \;=\; g^{t-1}\, D_{W1}(c_0, \Delta) \;>\; 0$ for $t \in \{1, ..., 6\}$. (71) implies:

$$E_d\{W^S\} - E_d\{W^I\} \;=\; A(k_1)\, D_{W1}(c_0, \Delta)$$

$$\text{where } \; A(k_1) \;\equiv\; \widetilde{\delta}^2 \left[\, \phi_1^S - \phi_1^I \,\right] + \widetilde{\delta}^3 \left[\, \phi_1^S - \Phi^I \,\right] + \widetilde{\delta}^4 \left[\, \Phi^S - \Phi^I \,\right] + \widetilde{\delta}^5 \left[\, \Phi^S - \Phi^I \,\right]. \tag{72}$$

$\frac{\partial D_{W1}(c_0, \Delta)}{\partial k_1} = 0$. Therefore, (72) implies:

$$\lim_{k_1 \rightarrow \infty} \left(E_d\{W^I\} - E_d\{W^S\} \right) \;>\; 0 \;\text{ if }\; \lim_{k_1 \rightarrow \infty} A(k_1) \;<\; 0\,.$$

(72) implies that $\lim_{k_1 \to \infty} A(k_1) < 0$ if: (i) $\lim_{k_1 \to \infty} \left(\phi_1^S - \phi_1^I \right) = 0$; (ii) $\lim_{k_1 \to \infty} \left(\phi_1^S - \Phi^I \right) < 0$; and (iii) $\lim_{k_1 \to \infty} \left(\Phi^S - \Phi^I \right) < 0$. We complete the proof by showing that (i), (ii), and (iii) hold.

(70) implies that $\lim_{k_1 \to \infty} \left(\phi_1^S - \phi_1^I \right) = 0$.

(4) and (70) imply:

$$\lim_{k_1 \to \infty} \left(\phi_1^S - \Phi^I \right) = \lim_{k_1 \to \infty} \left(\phi_1^S - \phi_1^I - \left[1 - \phi_1^I \right] \phi_2^I \right) = - \lim_{k_1 \to \infty} \phi_2^I = - \phi_2^I < 0;$$

$$\lim_{k_1 \to \infty} \left(\Phi^S - \Phi^I \right) = \lim_{k_1 \to \infty} \left(\phi_1^S + \left[1 - \phi_1^S \right] \phi_2^S - \phi_1^I - \left[1 - \phi_1^I \right] \phi_2^I \right)$$

$$= \lim_{k_1 \to \infty} \left(\phi_2^S - \phi_2^I \right) = \phi_2^S - \phi_2^I < 0.$$

The last inequality here reflects Proposition 1. ■

Proof of Proposition 9. Define $\widetilde{k}_2 \equiv \frac{k_2}{g}$, $x \equiv \frac{\Delta Q_0}{\widetilde{k}_2} = \frac{\Delta g Q_0}{k_2}$, and $\widetilde{\delta} \equiv g \delta$. (35) and (39) imply that under the maintained assumptions:

$$\phi_2^S = \frac{\Delta Q_0 g^2 \delta [1 + g \delta]}{k_2} = x g \delta [1 + g \delta] = x \widetilde{\delta} [1 + \widetilde{\delta}] \text{ and} \tag{73}$$

$$\phi_2^I = \frac{\Delta g Q_0 [1 + g \delta]}{k_2} = x [1 + \delta g] = x [1 + \widetilde{\delta}]. \tag{74}$$

(41) and (74) imply:

$$\phi_1^I = \frac{\Delta [1 + g \delta] Q_0 - \delta \left[g (1 + g \delta) \Delta Q_0 \phi_2^I - K_2(\phi_2^I) \right]}{k_1}$$

$$= \frac{\Delta [1 + \widetilde{\delta}] Q_0}{\widetilde{k}_2} \frac{\widetilde{k}_2}{k_1} - \widetilde{\delta} \frac{\widetilde{k}_2}{k_1} \left[\frac{\left(\phi_2^I \right)^2}{2} \right]. \tag{75}$$

(74) and (75) imply:

$$\phi_1^I = \phi_2^I \frac{\widetilde{k}_2}{k_1} - \widetilde{\delta} \frac{\widetilde{k}_2}{k_1} \frac{\left(\phi_2^I \right)^2}{2} = x \frac{\widetilde{k}_2}{k_1} [1 + \widetilde{\delta}] \left[1 - \frac{\widetilde{\delta}}{2} x \left(1 + \widetilde{\delta} \right) \right]. \tag{76}$$

(36) – (38) imply:

$$\phi_1^S = \frac{\widetilde{k}_2}{k_1} x [1 + \widetilde{\delta}] \left[1 - \frac{\widetilde{\delta}^3 \left(1 + \widetilde{\delta} \right)}{2} x \right]. \tag{77}$$

(4), (73), (74), (76), and (77) imply:

$$\Phi^I = \phi_1^I + \left[1 - \phi_1^I \right] \phi_2^I = \frac{\widetilde{k}_2}{k_1} x [1 + \widetilde{\delta}] \left[1 - \frac{\widetilde{\delta}}{2} x (1 + \widetilde{\delta}) \right]$$

$$+ \left[1 - \frac{\widetilde{k}_2}{k_1} \, x \left(1 + \widetilde{\delta} \right) \left(1 - \frac{\widetilde{\delta}}{2} \, x \, [\, 1 + \widetilde{\delta} \,] \right) \right] x \, [\, 1 + \widetilde{\delta} \,] \, ; \tag{78}$$

$$\Phi^S \;=\; \phi_1^S + \left[\, 1 - \phi_1^S \,\right] \phi_2^S \;=\; \frac{\widetilde{k}_2}{k_1} \, x \, [\, 1 + \widetilde{\delta} \,] \left[1 - \frac{\widetilde{\delta}^3 \, (\, 1 + \widetilde{\delta} \,)}{2} \, x \right]$$

$$+ \left[1 - \frac{\widetilde{k}_2}{k_1} \, x \left(1 + \widetilde{\delta} \right) \left(1 - \frac{\widetilde{\delta}^3 \, [\, 1 + \widetilde{\delta} \,]}{2} \, x \right) \right] [\, 1 + \widetilde{\delta} \,] \, \widetilde{\delta} \, x \, . \tag{79}$$

(64) implies that in the ID setting:

$$E_d\{W^I\} - \, E_d\{W^S\} \;\stackrel{s}{=}\; \left[\, \phi_1^I - \phi_1^S \,\right] D_{W3} + \delta \left[\, \Phi^I - \phi_1^S \,\right] D_{W4}$$

$$+ \;\; \delta^2 \left[\, \Phi^I - \Phi^S \,\right] D_{W5} + \delta^3 \left[\, \Phi^I - \Phi^S \,\right] D_{W6} \tag{80}$$

where $D_{Wt} \;\equiv\; W_t(c_0 - \Delta) - W_t(c_0) \;>\; 0$. Assumption G implies that $D_{Wt} \;=\; g \, D_{W(t-1)}$. Therefore, (80) implies:

$$E_d\{W^I\} - \, E_d\{W^S\} \;\stackrel{s}{=}\; \left[\, \phi_1^I - \phi_1^S \,\right] D_{W3} + g \, \delta \left[\, \Phi^I - \phi_1^S \,\right] D_{W3}$$

$$+ \;\; [\, g \, \delta \,]^2 \left[\, \Phi^I - \Phi^S \,\right] D_{W3} + [\, g \, \delta \,]^3 \left[\, \Phi^I - \Phi^S \,\right] D_{W3}$$

$$\stackrel{s}{=}\; \phi_1^I - \phi_1^S \, + \widetilde{\delta} \left[\, \Phi^I - \phi_1^S \,\right] + \widetilde{\delta}^{\,2} \left[\, \Phi^I - \Phi^S \,\right] + \widetilde{\delta}^{\,3} \left[\, \Phi^I - \Phi^S \,\right] \tag{81}$$

where $\widetilde{\delta} \;\equiv\; \delta \, g$. (81) implies:

$$E_d\{W^I\} \;>\; E_d\{W^S\} \;\; \text{if}$$

$$\phi_1^I - \phi_1^S + \widetilde{\delta} \left[\, \Phi^I - \phi_1^S \,\right] + \widetilde{\delta}^2 \left[\, \Phi^I - \Phi^S \,\right] + \widetilde{\delta}^3 \left[\, \Phi^I - \Phi^S \,\right] \;>\; 0 \, . \tag{82}$$

(76) – (79) and (82) imply that $E_d\{W^I\} > E_d\{W^S\}$ if:

$$\frac{\widetilde{k}_2}{k_1} \, x \, [\, 1 + \widetilde{\delta} \,] \left[1 - \frac{\widetilde{\delta}}{2} \, x \left(1 + \widetilde{\delta} \right) \right] - \frac{\widetilde{k}_2}{k_1} \, x \, [\, 1 + \widetilde{\delta} \,] \left[1 - \frac{\widetilde{\delta}^3 \left(1 + \widetilde{\delta} \right)}{2} \, x \right]$$

$$+ \, \widetilde{\delta} \left\{ \frac{\widetilde{k}_2}{k_1} \, x \, [\, 1 + \widetilde{\delta} \,] \left[1 - \frac{\widetilde{\delta}}{2} \, x \left(1 + \widetilde{\delta} \right) \right] \right.$$

$$+ \left[1 - \frac{\widetilde{k}_2}{k_1} \, x \left(1 + \widetilde{\delta} \right) \left(1 - \frac{\widetilde{\delta}}{2} \, x \left[1 + \widetilde{\delta} \right] \right) \right] x \, [\, 1 + \widetilde{\delta} \,]$$

$$\left. - \, \frac{\widetilde{k}_2}{k_1} \, x \, [\, 1 + \widetilde{\delta} \,] \left[1 - \frac{\widetilde{\delta}^3 \left(1 + \widetilde{\delta} \right)}{2} \, x \right] \right\}$$

$$+ \left[\widetilde{\delta}^2 + \widetilde{\delta}^3 \right] \left\{ \frac{\widetilde{k}_2}{k_1} \, x \, [1 + \widetilde{\delta}\,] \left[1 - \frac{\widetilde{\delta}}{2} \, x \left(1 + \widetilde{\delta} \right) \right] \right.$$

$$+ \left[1 - \frac{\widetilde{k}_2}{k_1} \, x \left(1 + \widetilde{\delta} \right) \left(1 - \frac{\widetilde{\delta}}{2} \, x \left[1 + \widetilde{\delta} \right] \right) \right] x \, [1 + \widetilde{\delta}\,]$$

$$- \frac{\widetilde{k}_2}{k_1} \, x \, [1 + \widetilde{\delta}\,] \left[1 - \frac{\widetilde{\delta}^3 \left(1 + \widetilde{\delta} \right)}{2} \, x \right]$$

$$\left. - \left[1 - \frac{\widetilde{k}_2}{k_1} \, x \left(1 + \widetilde{\delta} \right) \left(1 - \frac{\widetilde{\delta}^3 \left[1 + \widetilde{\delta} \right]}{2} \, x \right) \right] [1 + \widetilde{\delta}\,] \, \widetilde{\delta} \, x \right\} \; > \; 0 \qquad (83)$$

$$\Leftrightarrow \quad \left[\frac{\widetilde{k}_2}{k_1} \right] \frac{\widetilde{\delta}^3 \, [1 + \widetilde{\delta}\,]}{2} \, x - \left[\frac{\widetilde{k}_2}{k_1} \right] \frac{\widetilde{\delta}}{2} \, x \, [1 + \widetilde{\delta}\,]$$

$$+ \, \widetilde{\delta} \left[- \frac{\widetilde{\delta}}{2} \frac{\widetilde{k}_2}{k_1} \, x \left(1 + \widetilde{\delta} \right) + 1 - \frac{\widetilde{k}_2}{k_1} \, x \left(1 + \widetilde{\delta} \right) \left(1 - \frac{\widetilde{\delta}}{2} \, x \left[1 + \widetilde{\delta} \right] \right) + \frac{\widetilde{k}_2}{k_1} \, \frac{\widetilde{\delta}^3 \left(1 + \widetilde{\delta} \right)}{2} \, x \right]$$

$$+ \left[\widetilde{\delta}^2 + \widetilde{\delta}^3 \right] \left\{ - \frac{\widetilde{k}_2}{k_1} \, \frac{\widetilde{\delta}}{2} \, x \, [1 + \widetilde{\delta}\,] + 1 - \frac{\widetilde{k}_2}{k_1} \, x \, [1 + \widetilde{\delta}\,] \left[1 - \frac{\widetilde{\delta}}{2} x \left(1 + \widetilde{\delta} \right) \right] \right.$$

$$\left. + \frac{\widetilde{k}_2}{k_1} \, \frac{\widetilde{\delta}^3 \, [1 + \widetilde{\delta}\,]}{2} \, x - \widetilde{\delta} \left[1 - \frac{\widetilde{k}_2}{k_1} \, x \left(1 + \widetilde{\delta} \right) \left(1 - \frac{\widetilde{\delta}^3 \left[1 + \widetilde{\delta} \right]}{2} \, x \right) \right] \right\} \; > \; 0 \, . \qquad (84)$$

As $x \equiv \frac{\Delta Q_0}{\widetilde{k}_2} \rightarrow 0$, the inequality in (84) becomes:

$$\widetilde{\delta} + \, [\, \widetilde{\delta}^2 + \widetilde{\delta}^3 \,] \, [\, 1 - \widetilde{\delta} \,] \; > \; 0 \, . \qquad (85)$$

The inequality in (85) holds because Assumption G implies that $\widetilde{\delta} < 1$. Therefore, $E_d\{W^I\} > E_d\{W^S\}$ when $\Delta \, Q_0$ is sufficiently small.

Finally, suppose $k_1 = k_2 \equiv k$, so $\frac{\widetilde{k}_2}{k_1} = \frac{1}{g}$. As $k \rightarrow \infty$, $x \equiv \frac{\Delta \, g \, Q_0}{k} \rightarrow 0$ and the inequality in (84) becomes the inequality in (85). Because this inequality holds, $E_d\{W^I\} > E_d\{W^S\}$ when k is sufficiently large. ■

<u>Proof of Proposition 10</u>. Recall from the proof of the Corollary to Lemma 2 that $\widehat{\widetilde{\delta}}$ is the value of $g \, \delta$ for which:

$$g \, \delta \, [\, 1 + g \, \delta \,] \; = \; (g \, \delta)^2 + \delta \, g \; = \; 1 \, . \qquad (86)$$

Initially suppose that $\delta \, g = \widehat{\widetilde{\delta}}$ and $\frac{\Delta \, Q_0 \, [\, 1 + \widehat{\widetilde{\delta}} \,]}{k_2} = 1$. Then (74) implies that under the specified conditions:

$$\phi_2^I \;=\; \frac{\Delta\, Q_0\, g\,[\,1+\widehat{\widehat{\delta}}\,]}{k_2} \;=\; 1\,. \tag{87}$$

(64) implies that under the specified conditions:

$$E_d\{W^S\} - E_d\{W^I\} \;=\; A_D\; D_W$$

$$\text{where}\;\; A_D \;\equiv\; [\,g\,\delta\,]^2 \left[\,\phi_1^S - \phi_1^I\,\right] + [\,g\,\delta\,]^3 \left[\,\phi_1^S - \Phi^I\,\right] + [\,g\,\delta\,]^4 \left[\,\Phi^S - \Phi^I\,\right]$$

$$+\; [\,g\,\delta\,]^5 \left[\,\Phi^S - \Phi^I\,\right] \;\; \text{and}$$

$$D_W \;\equiv\; W(c_0 - \Delta) - W(c_0) \;>\; 0\,. \tag{88}$$

We will show that $E_d\{W^S\} > E_d\{W^I\}$ by showing that $A_D > 0$ when $\frac{\Delta\, Q_0\, g\,[\,1+\widehat{\widehat{\delta}}\,]}{k_2} = 1$ and $\delta\, g = \widehat{\widehat{\delta}}$. The continuity of $E_d\{W^S\} - E_d\{W^I\}$ then ensures that $E_d\{W^S\} > E_d\{W^I\}$ when δ is sufficiently close to $\widehat{\widehat{\delta}}$ and $\frac{\Delta\, Q_0\, g\,[\,1+\widehat{\widehat{\delta}}\,]}{k}$ is less than, but sufficiently close to, 1.

(75) and (87) imply that under the specified conditions:

$$\phi_1^I \;=\; \frac{\Delta\, Q_0\,[\,1+g\,\delta\,] - \delta\left[\,\Delta\, Q_0\, g\,(1+g\,\delta) - \frac{k_2}{2}\,\right]}{k_1} \;=\; \frac{k_2}{k_1}\left[\,\frac{1}{g} - \frac{\delta}{2}\,\right]. \tag{89}$$

(89) reflects the maintained assumption that $\delta\, g = \widehat{\widehat{\delta}}$.

(73) and (87) imply that when $\delta\, g \,=\, \widetilde{\delta}$ under the specified conditions:

$$\phi_2^S \;=\; \frac{\Delta\, Q_0\, \delta\, g^2\,[\,1+g\,\delta\,]}{k_2} \;=\; \delta\, g \;=\; \widetilde{\delta}\,. \tag{90}$$

(77), (87), and (90) imply that when $\delta\, g \,=\widehat{\widehat{\delta}}$ under the specified conditions:

$$\phi_1^S \;=\; \left[\frac{k_2}{k_1}\right] \frac{\Delta\, Q_0\,[\,1+g\,\delta\,] - \delta\left[\,\phi_2^S\,\Delta\, Q_0\,\delta\, g^2\,(1+g\,\delta) - K_2(\phi_2^S)\,\right]}{k_2}$$

$$=\; \frac{k_2}{k_1}\left[\,\frac{1}{g} - \frac{\delta\left[\,\delta\,\Delta\, Q_0\,\delta\, g^2\,(1+g\,\delta) - \frac{k_2}{2}\,\delta^2 g^2\,\right]}{k_2}\,\right]$$

$$=\; \frac{k_2}{k_1}\left[\,\frac{1}{g} - \delta^3\, g + \frac{1}{2}\,\delta^3\, g^2\,\right] \;=\; \frac{k_2}{g\,k_1}\left[\,1 - \frac{1}{2}\,(\,\widetilde{\delta}\,)^3\,\right] \;<\; 1\,. \tag{91}$$

The inequality in (91) holds because $\widetilde{\delta} < 1$ and $\frac{k_2}{g\,k_1} \leq 1$, by assumption.

(88) implies that when $\delta\, g = \widehat{\widehat{\delta}}$ and $\widetilde{\delta} = g\,\delta$:

$$A_D \;\stackrel{s}{=}\; \phi_1^S - \phi_1^I + \widetilde{\delta}\left[\,\phi_1^S - \Phi^I\,\right] + \widetilde{\delta}^{\,2}\left[\,\Phi^S - \Phi^I\,\right] + \widetilde{\delta}^{\,3}\left[\,\Phi^S - \Phi^I\,\right] \;>\; 0$$

$$\text{if}\;\; \phi_1^S - \phi_1^I + \widetilde{\delta}\left[\,\phi_1^S - \Phi^I\,\right] + \widetilde{\delta}^{\,2}\left[\,\phi_1^S - \Phi^I\,\right] + \widetilde{\delta}^{\,3}\left[\,\phi_1^S - \Phi^I\,\right] \;>\; 0$$

$$\Leftrightarrow \quad \phi_1^S - \phi_1^I + \widetilde{\delta}\left[\phi_1^S - 1\right] + \widetilde{\delta}^{\,2}\left[\phi_1^S - 1\right] + \widetilde{\delta}^{\,3}\left[\phi_1^S - 1\right] \;>\; 0\,. \tag{92}$$

The last equivalence here holds because $\Phi^I = 1$ when $\phi_2^I = 1$ (from (87)).

Observe that when $\delta\, g = \widehat{\widehat{\delta}}$:

$$\phi_1^S - \phi_1^I + \widetilde{\delta}\left[\phi_1^S - 1\right] + \widetilde{\delta}^2\left[\phi_1^S - 1\right] + \widetilde{\delta}^3\left[\phi_1^S - 1\right]$$

$$= \; \phi_1^S - \phi_1^I - \widetilde{\delta}\left[1 - \phi_1^S\right]\sum_{t=0}^{2}\widetilde{\delta}^{\,t} \;=\; \phi_1^S - \phi_1^I - \widetilde{\delta}\left[1 - \phi_1^S\right]\left[\frac{1 - \widetilde{\delta}^3}{1 - \widetilde{\delta}}\right]$$

$$= \; \phi_1^S - \phi_1^I - \widetilde{\delta}\left[1 - \phi_1^S\right]\left[\widetilde{\delta}^2 + \widetilde{\delta} + 1\right] \;=\; \phi_1^S - \phi_1^I - 2\,\widetilde{\delta}\left[1 - \phi_1^S\right]. \tag{93}$$

The last equality in (93) reflects (86). (92) and (93) imply:

$$A_D \;>\; 0 \;\text{ if }\; \phi_1^S - \phi_1^I \;>\; 2\,\widetilde{\delta}\,\left[1 - \phi_1^S\right]$$

$$\Leftrightarrow \;\; \frac{k_2}{k_1}\left[\frac{1}{g} - \frac{1}{2}\,\delta^3 g^2 - \left(\frac{1}{g} - \frac{\delta}{2}\right)\right] \;>\; 2\,\widetilde{\delta}\left[1 - \frac{k_2}{k_1}\left(\frac{1}{g} - \frac{1}{2}\,\delta^3 g^2\right)\right]$$

$$\Leftrightarrow \;\; \frac{\widetilde{\delta}}{2} - \frac{1}{2}\,\widetilde{\delta}^3 \;>\; 2\,\widetilde{\delta}\left[\frac{k_1\, g}{k_2} - \left(1 - \frac{1}{2}\,\widetilde{\delta}^3\right)\right]. \tag{94}$$

Observe that:

$$\frac{\widetilde{\delta}}{2} - \frac{1}{2}\,\widetilde{\delta}^3 \;>\; 2\,\widetilde{\delta}\left[1 - \left(1 - \frac{1}{2}\,\widetilde{\delta}^3\right)\right] \;\;\Leftrightarrow\;\; \frac{\widetilde{\delta}}{2} - \frac{1}{2}\,\widetilde{\delta}^3 \;>\; 2\,\widetilde{\delta}\left[\frac{1}{2}\,\widetilde{\delta}^3\right]$$

$$\Leftrightarrow \;\; \frac{\widetilde{\delta}}{2} - \frac{1}{2}\,\widetilde{\delta}^3 \;>\; \widetilde{\delta}^4 \;\;\Leftrightarrow\;\; 1 - \widetilde{\delta}^2 \;>\; 2\,\widetilde{\delta}^3 \;\;\Leftrightarrow\;\; 1 \;>\; \widetilde{\delta}^2 + 2\,\widetilde{\delta}^3\,. \tag{95}$$

The last inequality in (95) holds because:

$$\widetilde{\delta}^2 + 2\,\widetilde{\delta}^3 \;=\; \widetilde{\delta}^2\,[\,1 + \widetilde{\delta}\,] + \widetilde{\delta}^3 \;=\; \widetilde{\delta} + \widetilde{\delta}^3 \;=\; \widetilde{\delta}\,[\,1 + \widetilde{\delta}^2\,] \;<\; \widetilde{\delta}\,[\,1 + \widetilde{\delta}\,] \;=\; 1\,.$$

The second and last equalities here reflect (86).

Because $k_1\, g \le k_2$ by assumption, (94) and (95) imply that $A_D > 0$ when $\phi_2^I = 1$ and $\delta\, g = \widehat{\widehat{\delta}}$. ■

Proof of Proposition 11. Define $\widetilde{k}_2 \equiv \frac{k_2}{g}$, $x \equiv \frac{\Delta\, g\, Q_0}{k_2} = \frac{\Delta\, Q_0}{\widetilde{k}_2}$, and $\widetilde{\delta} \equiv g\,\delta$. Recall that $E_d\{W^S\} > E_d\{W^I\}$ if the inequality in (84) is reversed. Because $x\,\widetilde{k}_2 = \Delta\, Q_0$, the inequality in (84) is reversed if:

$$\frac{\Delta\, Q_0}{k_1}\,\frac{\widetilde{\delta}^3\,[\,1 + \widetilde{\delta}\,]}{2} - \frac{\Delta\, Q_0}{k_1}\,\frac{\widetilde{\delta}}{2}\,[\,1 + \widetilde{\delta}\,]$$

$$+ \widetilde{\delta} \left[- \frac{\widetilde{\delta}}{2} \frac{\Delta Q_0}{k_1} \left(1 + \widetilde{\delta} \right) + 1 - \frac{\Delta Q_0}{k_1} \left(1 + \widetilde{\delta} \right) \left(1 - \frac{\widetilde{\delta}}{2} x [1 + \widetilde{\delta}] \right) + \frac{\Delta Q_0}{k_1} \frac{\widetilde{\delta}^3 \left(1 + \widetilde{\delta} \right)}{2} \right]$$

$$+ \left[\widetilde{\delta}^2 + \widetilde{\delta}^3 \right] \left\{ - \frac{\Delta Q_0}{k_1} \frac{\widetilde{\delta}}{2} [1 + \widetilde{\delta}] + 1 - \frac{\Delta Q_0}{k_1} [1 + \widetilde{\delta}] \left[1 - \frac{\widetilde{\delta}}{2} x \left(1 + \widetilde{\delta} \right) \right] \right.$$

$$\left. + \frac{\Delta Q_0}{k_1} \frac{\widetilde{\delta}^3 [1 + \widetilde{\delta}]}{2} - \widetilde{\delta} \left[1 - \frac{\Delta Q_0}{k_1} \left(1 + \widetilde{\delta} \right) \left(1 - \frac{\widetilde{\delta}^3 [1 + \widetilde{\delta}]}{2} x \right) \right] \right\} < 0 . \quad (96)$$

$x \equiv \frac{\Delta g Q_0}{k_2} \to 0$ as $k_2 \to \infty$. Therefore, as $k_2 \to \infty$, the inequality in (96) holds if:

$$\frac{\Delta Q_0}{k_1} \frac{\widetilde{\delta}^3 [1 + \widetilde{\delta}]}{2} - \frac{\Delta Q_0}{k_1} \frac{\widetilde{\delta}}{2} [1 + \widetilde{\delta}]$$

$$+ \widetilde{\delta} \left[- \frac{\widetilde{\delta}}{2} \frac{\Delta Q_0}{k_1} \left(1 + \widetilde{\delta} \right) + 1 - \frac{\Delta Q_0}{k_1} \left(1 + \widetilde{\delta} \right) + \frac{\Delta Q_0}{k_1} \frac{\widetilde{\delta}^3 \left(1 + \widetilde{\delta} \right)}{2} \right]$$

$$+ \left[\widetilde{\delta}^2 + \widetilde{\delta}^3 \right] \left\{ - \frac{\Delta Q_0}{k_1} \frac{\widetilde{\delta}}{2} [1 + \widetilde{\delta}] + 1 - \frac{\Delta Q_0}{k_1} [1 + \widetilde{\delta}] \right.$$

$$\left. + \frac{\Delta Q_0}{k_1} \frac{\widetilde{\delta}^3 [1 + \widetilde{\delta}]}{2} - \widetilde{\delta} \left[1 - \frac{\Delta Q_0}{k_1} \left(1 + \widetilde{\delta} \right) \right] \right\} < 0 . \quad (97)$$

Define $y \equiv \frac{\Delta Q_0}{k_1} [1 + \widetilde{\delta}]$. Then the inequality in (97) holds if:

$$\frac{\widetilde{\delta}}{2} y [\widetilde{\delta}^2 - 1] + \widetilde{\delta} \left[1 + \frac{\widetilde{\delta}}{2} y (\widetilde{\delta}^2 - 1) - y \right]$$

$$+ \left[\widetilde{\delta}^2 + \widetilde{\delta}^3 \right] \left[1 - \widetilde{\delta} + y \frac{\widetilde{\delta}}{2} (\widetilde{\delta}^2 - 1) - y (1 - \widetilde{\delta}) \right] < 0$$

$$\Leftrightarrow \quad \frac{\widetilde{\delta}}{2} y [1 - \widetilde{\delta}^2] \left[1 + \widetilde{\delta} + \widetilde{\delta}^2 + \widetilde{\delta}^3 \right] > \left\{ \widetilde{\delta} + [\widetilde{\delta}^2 + \widetilde{\delta}^3] [1 - \widetilde{\delta}] \right\} [1 - y] . \quad (98)$$

Assumption G requires $g < \frac{1}{\delta} \Rightarrow \widetilde{\delta} = \delta g < 1$. Therefore, the inequality in (98) holds as $y \equiv \frac{\Delta Q_0}{k_1} [1 + \widetilde{\delta}] \to 1$. ■

Proof of Proposition 12. Initially suppose the NID setting prevails, so the firm always implements an achieved cost reduction immediately.

Under standard rebasing (SR) in this setting, the firm retains the full benefit of a cost reduction that is achieved in period 2 only for that period. Therefore, the firm's problem in period 2, given that it implemented first-period success probability ϕ_1^S but did not achieve a cost reduction in period 1, is:

$$\underset{\phi_2}{\text{Maximize}} \quad \left[\phi_2 + \alpha \, \phi_1^S \right] \Delta \, Q_2(c_0) - K_2(\phi_2)$$

$$\Rightarrow \quad K_2'(\phi_2^S) \;=\; \Delta \, Q_2(c_0) \;\; \text{at an interior optimum.} \tag{99}$$

Under IRIS in this setting, the firm retains the full benefit of a cost reduction achieved in period 2 during both period 2 and period 3. Therefore, the firm's problem in period 2, given that it implemented first-period success probability ϕ_1^I but did not achieve a cost reduction in period 1, is:

$$\underset{\phi_2}{\text{Maximize}} \quad \left[\phi_2 + \alpha \, \phi_1^I \right] \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] - K_2(\phi_2)$$

$$\Rightarrow \quad K_2'(\phi_2^I) \;=\; \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] \;\; \text{at an interior optimum.} \tag{100}$$

Under SR, the firm retains the full benefit of a cost reduction that is achieved in period 1 both in period 1 and in period 2. Therefore, the firm's problem in period 1 under SR in the NID setting is:

$$\underset{\phi_1}{\text{Maximize}} \quad \phi_1 \, \Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right]$$
$$+ \left[1 - \phi_1 \right] \delta \left[\left(\phi_2^S + \alpha \, \phi_1 \right) \Delta \, Q_2(c_0) - K_2(\phi_2^S) \right] - K_1(\phi_1) \, . \tag{101}$$

(101) implies that at an interior solution to this problem:

$$K_1'(\phi_1^S) \;=\; \Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right] - \delta \left[\left(\phi_2^S + \alpha \, \phi_1^S \right) \Delta \, Q_2(c_0) - K_2(\phi_2^S) \right]$$
$$+ \, \delta \, \alpha \left[1 - \phi_1^S \right] \Delta \, Q_2(c_0) \, . \tag{102}$$

Under IRIS, the firm retains for two periods the full benefit of an achieved cost reduction, whether the reduction is achieved in period 1 or period 2. Therefore, the firm's problem in period 1 under IRIS is:

$$\underset{\phi_1}{\text{Maximize}} \quad \phi_1 \, \Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right]$$
$$+ \left[1 - \phi_1 \right] \delta \left\{ \left[\phi_2^I + \alpha \, \phi_1 \right] \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] - K_2(\phi_2^I) \right\} - K_1(\phi_1) \, . \tag{103}$$

(103) implies that at an interior solution to this problem:

$$K_1'(\phi_1^I) \;=\; \Delta \left[Q_1(c_0) + \delta \, Q_2(c_0) \right] - \delta \left\{ \left[\phi_2^I + \alpha \, \phi_1^I \right] \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] - K_2(\phi_2^I) \right\}$$
$$+ \, \delta \, \alpha \left[1 - \phi_1^I \right] \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] . \tag{104}$$

(99) and (100) imply:

$$K_2'(\phi_2^I) \;=\; \Delta \left[Q_2(c_0) + \delta \, Q_3(c_0) \right] \;>\; \Delta \, Q_2(c_0) \;=\; K_2'(\phi_2^S) \quad \Rightarrow \quad \phi_2^I \;>\; \phi_2^S \, . \tag{105}$$

The implication ($\Rightarrow$) in (105) reflects the convexity of $K_2(\cdot)$.

To prove that $\phi_1^S > \phi_1^I$, suppose that $\phi_1^I \geq \phi_1^S$. Then:

$$\begin{aligned}
& \left[\, \phi_2^I + \alpha\, \phi_1^I \,\right] \Delta \left[\, Q_2(c_0) + \delta\, Q_3(c_0) \,\right] - K_2(\phi_2^I) \\
& \quad = \max_{\phi_2} \left\{ \left[\, \phi_2 + \alpha\, \phi_1^I \,\right] \Delta \left[\, Q_2(c_0) + \delta\, Q_3(c_0) \,\right] - K_2(\phi_2) \right\} \\
& \quad > \left[\, \phi_2^S + \alpha\, \phi_1^I \,\right] \Delta \left[\, Q_2(c_0) + \delta\, Q_3(c_0) \,\right] - K_2(\phi_2^S) \\
& \quad \geq \left[\, \phi_2^S + \alpha\, \phi_1^S \,\right] \Delta \left[\, Q_2(c_0) + \delta\, Q_3(c_0) \,\right] - K_2(\phi_2^S) \\
& \quad > \left[\, \phi_2^S + \alpha\, \phi_1^S \,\right] \Delta\, Q_2(c_0) - K_2(\phi_2^S)\,.
\end{aligned} \tag{106}$$

The equality in (106) reflects (100). The first inequality in (106) reflects (105). The second inequality in (106) reflects the maintained assumption that $\phi_1^I \geq \phi_1^S$.

Observe that:

$$\begin{aligned}
\phi_1^I \;=\; & \arg\max_{\phi_1} \left\{ \phi_1\, \Delta \left[\, Q_1(c_0) + \delta\, Q_2(c_0) \,\right] \right. \\
& \qquad \left. + \left[\, 1 - \phi_1 \,\right] \delta \left\{ \left[\, \phi_2^I + \alpha\, \phi_1 \,\right] \Delta \left[\, Q_2(c_0) + \delta\, Q_3(c_0) \,\right] - K_2(\phi_2^I) \right\} - K_1(\phi_1) \right\} \\
< \; & \arg\max_{\phi_1} \left\{ \phi_1\, \Delta \left[\, Q_1(c_0) + \delta\, Q_2(c_0) \,\right] \right. \\
& \qquad \left. + \left[\, 1 - \phi_1 \,\right] \delta \left\{ \left[\, \phi_2^S + \alpha\, \phi_1 \,\right] \Delta\, Q_2(c_0) - K_2(\phi_2^S) \right\} - K_1(\phi_1) \right\} \;=\; \phi_1^S\,.
\end{aligned} \tag{107}$$

The first equality in (107) reflects (103). The inequality in (107) follows from (106) because the value of ϕ_1 that maximizes the PDV of the firm's expected profit increases as the firm's expected profit following first-period failure declines.[62] The final equality in (107) reflects (101).

The conclusion in (107) that $\phi_1^I < \phi_1^S$ contradicts the maintained assumption that $\phi_1^I \geq \phi_1^S$. Therefore, by contradiction:

$$\phi_1^S \;>\; \phi_1^I\,. \tag{108}$$

The proof that $\phi_1^S > \phi_1^I$ in the ID setting is analogous. ■

<u>Proof of Proposition 13</u>. (99) and (100) imply that in the NID setting:

$$k_2 \left(\phi_2^S\right)^{\gamma - 1} \;=\; \Delta\; g\, Q_0 \;\;\Rightarrow\;\; \phi_2^S \;=\; \left[\, \frac{\Delta\; g\, Q_0}{k_2} \,\right]^{\frac{1}{\gamma-1}} \;=\; \left[\, \frac{\Delta\; Q_0}{\widetilde{k}_2} \,\right]^{\frac{1}{\gamma-1}} \quad \text{and}$$

$$k_2 \left(\phi_2^I\right)^{\gamma - 1} \;=\; \Delta \left[\, g\, Q_0 + g^2\, \delta\, Q_0 \,\right]$$

[62] Formally, if $\phi_1^I \in (0,1) = \arg\max_{\phi_1} \{\, \phi_1\, A + [\, 1 - \phi_1 \,]\, B - K_1(\phi_1) \,\}$, then $A - B = K_1'(\phi_1^I) \Rightarrow \frac{d\phi_1^I}{dB} = -\frac{1}{K_1''(\phi_1^I)} < 0$.

$$\Rightarrow \quad \phi_2^I \ = \ \left[\frac{\Delta\, Q_0\, g \left(1 + \widetilde{\delta} \right)}{k_2} \right]^{\frac{1}{\gamma - 1}} \ = \ [\, 1 + \widetilde{\delta}\,]^{\frac{1}{\gamma-1}} \left[\frac{\Delta\, Q_0}{\widetilde{k}_2} \right]^{\frac{1}{\gamma-1}} \ = \ \phi_2^S\, [\, 1 + \widetilde{\delta}\,]^{\frac{1}{\gamma-1}}. \tag{109}$$

(4) implies that in the setting with innovation persistence:

$$\Phi^j \ = \ \phi_1^S + \left[1 - \phi_1^S \right] \left[\phi_2^S + \alpha\, \phi_1^S \right] \quad \text{for } j \in \{ S, I \}. \tag{110}$$

(57) implies that under the specified conditions:

$E_d \left\{ W^S \right\} \ > \ E_d \left\{ W^I \right\}$ if

$$\widetilde{\delta}^2 \left[\Phi^S - \phi_1^I \right] [\, S_0(c_0 - \Delta) - S_0(c_0)\,]$$

$$+ \left[\Phi^S - \Phi^I \right] [\, S_0(c_0 - \Delta) - S_0(c_0)\,] [\, \widetilde{\delta}^3 + \widetilde{\delta}^4 + \widetilde{\delta}^5\,] \ > \ 0$$

$$\Leftrightarrow \quad \Phi^S - \phi_1^I + \left[\Phi^S - \Phi^I \right] [\, \widetilde{\delta} + \widetilde{\delta}^2 + \widetilde{\delta}^3\,] \ > \ 0. \tag{111}$$

First suppose that $\Phi^S \geq \Phi^I$. $\Phi^S > \phi_1^I$ because $\phi_1^S > \phi_1^I$, from Proposition 12. Therefore, (111) implies that $E_d \left\{ W^S \right\} \ > \ E_d \left\{ W^I \right\}$ when $\Phi^S \geq \Phi^I$.

Now suppose that $\Phi^S < \Phi^I$. (110) implies that (111) holds in this case if:

$$\Phi^S - \phi_1^I + \left[\Phi^S - \Phi^I \right] [\, \widetilde{\delta} + \widetilde{\delta}^2 + \widetilde{\delta}^3 + \widetilde{\delta}^4 + \ldots\,] \ > \ 0$$

$$\Leftrightarrow \quad \Phi^S - \phi_1^I + \left[\Phi^S - \Phi^I \right] \frac{\widetilde{\delta}}{1 - \widetilde{\delta}} \ > \ 0 \quad \Leftrightarrow \quad \left[\Phi^S - \phi_1^I \right] \left[1 - \widetilde{\delta} \right] + \widetilde{\delta} \left[\Phi^S - \Phi^I \right] \ > \ 0$$

$$\Leftrightarrow \quad \phi_1^S + \phi_2^S \left[1 - \phi_1^S \right] + \alpha\, \phi_1^S \left[1 - \phi_1^S \right] - \phi_1^I - \widetilde{\delta}\, \phi_2^I \left[1 - \phi_1^I \right] - \ \alpha\, \widetilde{\delta}\, \phi_1^I \left[1 - \phi_1^I \right] \ > \ 0$$

$$\Leftrightarrow \quad \left[1 - \phi_1^I \right] \left[1 - \widetilde{\delta}\, \phi_2^I - \alpha\, \widetilde{\delta}\, \phi_1^I \right] - \left[1 - \phi_1^S \right] \left[1 - \phi_2^S - \alpha\, \phi_1^S \right] \ > \ 0. \tag{112}$$

Proposition 12 implies:

$$1 - \phi_1^I \ > \ 1 - \phi_1^S. \tag{113}$$

Furthermore, (109) implies that when $\widetilde{\delta}\, [\, 1 + \widetilde{\delta}\,]^{\frac{1}{\gamma-1}} < 1$:

$$\phi_2^S - \widetilde{\delta}\, \phi_2^I \ = \ \phi_2^S - \widetilde{\delta}\, \phi_2^S\, [\, 1 + \widetilde{\delta}\,]^{\frac{1}{\gamma-1}} \ = \ \phi_2^S \left[1 - \widetilde{\delta} \left(1 + \widetilde{\delta} \right)^{\frac{1}{\gamma-1}} \right] \ > \ 0. \tag{114}$$

$\widetilde{\delta} \leq 1$ because $\widetilde{\delta}\, [\, 1 + \widetilde{\delta}\,]^{\frac{1}{\gamma-1}} < 1$. Therefore, from Proposition 12:

$$\phi_1^I \ < \ \phi_1^S \quad \Rightarrow \quad \alpha\, \phi_1^I \ < \ \alpha\, \phi_1^S \quad \Rightarrow \quad \alpha\, \widetilde{\delta}\, \phi_1^I \ < \ \alpha\, \phi_1^S. \tag{115}$$

Because $\phi_2^S + \alpha\, \phi_1^S < 1$, by assumption, (114) and(115) imply:

$$1 - \widetilde{\delta}\, \phi_2^I - \alpha\, \widetilde{\delta}\, \phi_1^I \ > \ 1 - \phi_2^S - \alpha\, \phi_1^S \ > \ 0. \tag{116}$$

(113) and (116) imply that the inequality in (112) holds. ■

Proof of the Corollary to Proposition 13. The Corollary follows directly from Proposition 13 because $\widetilde{\delta}\,[\,1+\widetilde{\delta}\,]^{\frac{1}{\gamma-1}} < 1$ under the specified conditions. (See the proof of the Corollary to Proposition 6.) ■

References

Armstrong, Mark, Ray Rees, and John Vickers, "Optimal Regulatory Lag under Price-Cap Regulation," *Revista Espanola de Economia*, 1995, 93-116.

Armstrong, Mark and David Sappington, "Recent Developments in the Theory of Regulation," in Mark Armstrong and Robert Porter (eds), *Handbook of Industrial Organization, Vol. 3.* Elsevier Science Publishers, 2007, pp. 1557-1700.

Arrow, Kenneth, "Economic Welfare and the Allocation of Resources for Invention," in National Bureau of Economic Research, *The Rate and Direction of Inventive Activity: Economic and Social Factors.* Princeton: Princeton University Press, 1962.

Australian Energy Regulator, "Electricity Distribution Network Service Providers: Efficiency Benefit Sharing Scheme," June 2008.

Australian Energy Regulator, "Better Regulation: Expenditure Incentives," November 2013.

Baumol, William and Alvin Klevorick, "Input Choices and Rate-of-Return Regulation: An Overview of the Discussion," *Bell Journal of Economics and Management Science*, 1(2), Autumn 1970, 162-190.

Biglaiser, Gary and Michael Riordan, "Dynamics of Price Regulation," *RAND Journal of Economics*, 31(4), Winter 2000, 744-767.

Coco, Giuseppe and Claudio De Vincenti, "Can Price Regulation Increase Cost-Efficiency?" *Research in Economics*, 58(4), December 2004, 303-317.

Coco, Giuseppe and Claudio De Vincenti, "Optimal Rate Base Reviews Under Price-Cap Regulation," Università di Bari Discussion Paper, December 2005.

Council of Economic Advisors, "Discounting for Public Policy: Theory and Recent Evidence on the Merits of Updating the Discount Rate," Issue Brief, January 2017 (https://obamawhitehouse.archives.gov/sites/default/files/page/files/201701_cea_discounting_issue_brief.pdf).

Duma, Daniel, Michael Pollitt, and Andrei Covatariu, "Regulatory Learning in the Face of Net Zero Climate Policy: The Case of the UK," *Review of Industrial Organization*, 65(2), 2024 (forthcoming).

Frontier Economics, "Application of Symmetric IRIS: A Report Prepared for Transpower," March 2015.

Joskow Paul, "Incentive Regulation in Theory and Practice: Electric Distribution and Transmission Networks," in Nancy Rose (ed.) *Economic Regulation and Its Reform.* Chicago: University of Chicago Press, 2014, chapter 5, pp. 291-343.

Joskow Paul, "The Expansion of Incentive (Performance-Based) Regulation of Electricity Distribution and Transmission in the United States," *Review of Industrial Organization*, 65(2), 2024, (forthcoming).

New Zealand Commerce Commission, "Introduction to the DPP for Stakeholders: 2020 Reset of the DPP for EDBs," 5 November 2018.

Ofgem, "Regulating Energy Networks for the Future: RPI-X@20 History of Energy Network Regulation," Supporting Paper 13b/09, February 27, 2009 (https://www.ofgem.gov.uk/ofgem-publications/51984/supporting-paper-history-energynetwork-regulation-finalpdf).

Office of Management and Budget, *Circular No. A-4*, November 9, 2023 (https://www.whitehouse.gov/wp-content/uploads/2023/11/CircularA-4.pdf).

Oxera, "Methodology Review for a Regulatory Framework Based on a Total Expenditure Approach," December 2021 (https://www.arera.it/fileadmin/allegati/docs/21/615-21oxera.pdf).

Pint, Ellen, "Price-Cap versus Rate-of-Return Regulation in a Stochastic-Cost Model," *RAND Journal of Economics*, 23(4), Winter 1992, 564-578.

Sappington, David and Dennis Weisman, "Price Cap Regulation: What Have We Learned from Twenty-Five Years of Experience in the Telecommunications Industry?" *Journal of Regulatory Economics*, 38, 2010, 227-257.

Sappington, David and Dennis Weisman, "The Disparate Adoption of Price Cap Regulation in the U.S. Telecommunications and Electricity Sectors," *Journal of Regulatory Economics*, 49(2), 2016, 250-264.

Sappington, David and Dennis Weisman, "40 Years of Incentive Regulation: What Have We Learned and What Questions Remain?," *Review of Industrial Organization*, 65(2), 2024, *forthcoming*.

Scalise, Joseph and Stephan Zech, "Sustained Cost Reduction for Utilities," Bain & Company *Brief*, June 19, 2013 (https://www.bain.com/insights/sustained-cost-reduction-for-utilities).

Turner, Douglas and David Sappington, "Technical Appendix to Accompany 'Enhancing Incentives for Cost Reduction in Regulated Industries'," October 2024 (https://people.clas.ufl.edu/sapping).

United States Postal Service, "First-Class Mail Volume Since 1926: Number of Pieces Mailed, to the Nearest Million," *Postal History*, accessed August 24, 2024 (https://about.usps.com/who/profile/history/first-class-mail-since-1926.htm).

Wilson, Gennelle, Cory Felder, and Rachel Gold, "States Move Quickly to Performance Based Regulation to Achieve Policy Priorities," Rocky Mountain Institute Report, March 31, 2022 (https://rmi.org/states-move-swiftly-on-performance-based-regulation-to-achieve-policy-priorities).