Performance Cycles*

David Lagziel[†] and Ehud Lehrer[‡]

June 19, 2018

ABSTRACT:

A decision maker repeatedly exerts effort to produce output. His past average performance defines his reward. We show that the decision maker's optimal strategy dictates a cyclic, oscillatory performance throughout the stages. Our model applies to a wide-range of economic settings where agents are subjected to history-based payoffs, including an R&D investment problem, the delegated portfolio-managers problem, and a dynamic advertising problem.

Journal of Economic Literature classification numbers: C61, C62, C79.

Keywords: Performance cycles; Performance persistence; Dynamic Optimization.

---

*The authors wish to thank David Schmeidler, Eddie Dekel, Avi Wohl, Zvika Neeman, Eilon Solan, and Dotan Persitz for their valuable comments; as well as the participants of the ninth edition of the UECE Lisbon Meetings, the Bar-Ilan University Economics Seminar, and the 10th annual Israeli Game-Theory conference. Lehrer acknowledges the support of the Israel Science Foundation, Grant #963/15.

†Ben-Gurion University of the Negev, Beer-Sheva 8410501, Israel. e-mail: davidlag@bgu.ac.il.

‡Tel Aviv University, Tel Aviv 69978, Israel and INSEAD, Bd. de Constance, 77305 Fontainebleau Cedex, France. e-mail: lehrer@post.tau.ac.il.

1 Introduction

Every person, firm, or economy are measured through current and past performance. Whether it is an academy professor designated for promotion, an investment firm with the standard assets-under-management fees, or an elected administration at mid-term elections, all possess a certain track record, used to determine their objective value and relevant reward, accordingly. But how does our past performance affects our incentives? Does recent over-performance carry some long- or short-term assurances to whatever follows?

In this paper, we study the basic tension between past performance and incentives. We pursue this goal through a dynamic model where a strategic decision maker (DM) repeatedly exerts effort to produce output, while being compensated according to his past and current production levels. We show that such evaluation process dictates a cyclic performance. Namely, the constant evaluation of the DM's performance triggers a strategic counter-reaction, by the DM, to level his efforts accordingly.

Our analysis spotlights two balancing forces that lead the evaluation-incentives interaction. On the one hand, past and current production levels are observed to generate an assessment over the DM's abilities. Evaluations are translated into reward-allocation decisions that define the DM's payoff. These *history-based payoffs* are the first key element behind our result. The DM, however, is subjected to convex operating costs. He maintains his positive evaluation by exerting convexly-increasing per-period "effort" that translates into output. The *marginally-decreasing output* is the second key element, to balance the first, in the dynamic setting. In other words, the history-dependant mechanism dictates the DM's reward, and requires the DM to balance current costs with future earnings, thus inducing a cyclic performance.

1.1 Main results

In practice, any DM can be measured in numerous ways, ranging from the sole last performance to his accumulated infinite track record. We capture these alternatives through two distinct mechanisms that represent short- and long-term averaging alike. The first mechanism is the *Transient model*, where the evaluation is based on the average of the last two-periods performance. The second mechanism is the *Persistent model*, where a discounted sum of all past performance is used to generate a performance evaluation. We present the main results derived from each, separately.

Under the Transient model, the DM is evaluated and compensated based on his last 2-stage average output. Namely, at any given stage, the DM is assessed by his current and last levels of production. This assessment ensures that the DM constantly monitors his recent performance. Any cost cut at a given stage, immediately limits the DM's payoff at the adjacent stage, since one period's output is next period's benchmark position. This intuition establishes our first main result where, even in a deterministic set-up, the DM's optimal policy dictates that production oscillates around a certain stable level, while converging towards it asymptotically. To put it differently, high-evaluation DMs have weaker incentives to produce high levels of output, compared to low-evaluation ones, since their benchmark high-level position entails a lower marginal payoff.

In the Persistent model the DM is evaluated by a discounted sum of all past production levels. Specifically, at any given stage, all past single-stage outputs are weighed-in while focusing more heavily on recent performance rather than earlier one. Given the updated setting, we show that a high-valuation position carries similar adverse-effects over incentives. Although consistent with the results of the Transient model, it is not the main focus of the current set-up.

The importance of the Persistent model follows from wide-perspective analysis. We study how changes in the evaluation process, as more weight is given to a recent performance rather than to an older one, affects the DM's payoff and his level of production. Non surprisingly, it appears that a high-evaluation DM profits from a higher weight on past performance maintaining his elite status at lower costs, while a low-evaluation DM benefits from myopic assessments of past production for the opposite reasons. On the other hand, we also prove the existence of a *basic tension between incentives and screening*. We show that optimal incentives are reached only in case past performance is not valued in any way, ignoring past production completely. To put it differently, a history-based screening process cannot produce first-best incentives, as the DM's optimal strategy would rely, to some extent, on past performance. Therefore, whenever there exists an uncertainty regarding the DM's differential abilities, the evaluator needs to balance between the screening process and optimal incentives.

1.2 Related literature

Our model and analysis combine several well-documented models and policies in the economic world. First, the oscillatory optimal strategy carries some resemblance to the optimal (S,s)-policies in inventory problems, where an individual agent allows his inventory to fall until is reaches a low level s, only to be imminently and actively increased to a high level S. Such policies were vastly studied in the context of the Pricing Problem (price adjustments and inflation), the Technology-Update Problem, and the Capital Stock Adjustment Problem (see, e.g., Arrow et al. (1951); Dvoretzky et al. (1952, 1953); Bellman et al. (1955); Bailey (1956); Arrow et al. (1958); Scarf (1959); Barro (1972); Sheshinski and Weiss (1977), and Sheshinski and Weiss (1993) for a general survey). Second, our convergence results correspond to the study of global stability in discounted problems, as in Scheinkman (1976); Rockafellar (1976); Cass and Shell (1976); Brock and Scheinkman (1976), among many others. We combine the two research branches by depicting a systematic method of oscillating convergence towards a stable production level for any initial condition.

The generality of our framework and approach follows, to some extend, the work of Chassang (2013). Chassang derives a robust approach to approximate the properties of linear high-liability contracts. Both the current work and the work of Chassang (2013) make few assumptions over the underline stochastic process. To be clear, we neither tackle the problem of expert testing as in Foster (1998); Fudenberg and Levine (1999); and Lehrer (2001), nor do we deal with optimal incentives as Holmstrom and Milgrom (1987); Laffont and Tirole (1988); and DeMarzo and Fishman (2007), among many others. We, to differ, study the implications of natural assessments over incentives.

The problem of non-persistent performance was broadly discussed in the delegated portfolio-manager context. The rational model of Berk and Green (2004) contributed one of the key ingredients in that aspect. Berk and Green proposed a non-strategic model, with decreasing returns to scale, where firms that recently outperformed suffer from the positive inflow of funds. They showed that much of the non-persistence and other related phenomenons could be attributed to the increasing convex costs, with respect to fund sizes. Nevertheless, their solution does not account for the cyclic performance that emerges from recent empirical studies of Cornell et al. (2017) and Bessler et al. (2017). In addition, the motivating conjecture that funds outflow benefit poorly-preforming firms is incompatible with the later work of Coval and Stafford (2007) and Rakowski (2010), suggesting outflows could be just as harmful as inflows. In this context, the study of Amihud and Goyenko (2013) gives a limited empirical basis to our work. Amihud and Goyenko (2013), building on the work of Carhart (1997), used robust empirical methods to produce a measure of predicting a fund's future performance. They show that the ability to generate excess return over a risk-adjust benchmark is positively associated with funds' expenses, which we denote as *effort*.

As we deal with a cyclic performance, one should consider the rich literature on optimal growth, consumption smoothing, and business cycles. These problems date back to the seminal works of Ramsey (1928); Brumberg and Modigliani (1954); Cass (1965); Koopmans (1965); Modigliani (1966) and Hall (1978). The significant difference between our work and previous ones follows from our emphasis on production and effort, rather than consumption. Our results indicate that the DM's optimal strategy generates some production smoothing, since a higher evaluation level is followed by a balancing lower one, and vice versa. Yet, these oscillation are endogenously-derived in our model, and do not depend on exogenous shocks, as in previous work.

1.3 Main contribution

In light of previous studies, we can underline several leading contributions of the current work. First and most importantly, we account for a cyclic performance in a general framework, linking oscillatory production and incentives. Our results indicate that cycles could be attributed to incentives whenever payoffs are not instantaneous, but distributed throughout the dynamics, depending on the DM's past performance. Therefore there need not be an exogenous shock to produce a cyclic performance, as it naturally arises from basic economic forces. In some respects, we show that a *production-smoothing* phenomenon originates in case past and current output levels dictate the single-stage reward.

On a theoretical level, we identify the process and method by which the production converges to a stable state. We do not limit ourselves to proving the existence of a unique solution, but show how it converges systematically through the performance mechanism. That is, we go beyond the fundamental work of Blackwell (1965), which guarantees convergence to an optimal production policy, and show that the optimal policy itself has a fixed point. Then, we prove that the same optimal policy induces a cyclic convergence towards that fixed point.

Another contribution is attributed to the inclusion of information frictions between the DM and his evaluator. In our model, the evaluation process follows simple heuristics that dictate the DM's payoff. Thus, there is no need for a common prior, Bayesian updating, or even common knowledge of distributions over abilities. The stochastic elements in our model need not to be normally-distributed, i.i.d., or even ergodic. In contrast, we consider a general Markov Decision Process. This generality opens the door to a broad analysis of the problem from a designer's viewpoint.

Lastly, we emphasize that the stochastic part on our work is more than a mere technical extension, and could accommodate several interpretations. For example, it captures the performance of a DM relative to an exogenous economy and other agents, alike. A broader discussion on this subject is given after Theorem 2 and in Section 2.2.1.

1.4 Structure of the paper

The paper is organized as follows. Section 2 concerns the Transient model divided into two parts: the deterministic case and the stochastic one. In Section 3 we revert to the Persistent model, and in Section 4 we focus on the effect that changes in the evaluation process have over incentives. Final conclusions and remarks are given in Section 5.

2 The Transient model

Consider a decision maker (DM) in an infinitely-repeated set-up. At every stage, the DM strategically exerts effort that translates into output. Then, the DM is rewarded a certain payoff, determined by a fixed allocation rule, and based on his aggregated past and current output. His main goal is to maximize his profit given by an infinite discounted-sum of the effort-deducted payoffs. Therefore, the process continues indefinitely as newly-realized output is observed and the payoff-allocation process repeats.

The set-up is mainly characterized by two functions: *the output function* and *the reward function*, defined as follows. Let $E = [e_{\min}, e_{\max}] \subseteq \mathbb{R}_+$ be a non-empty compact interval denoting the DM's single-period effort choice. For every $e \in E$, the production $Q(e)$ is determined by *the output function* $Q : E \to \mathbb{R}_+$. In other words, $Q(e)$ is the single-period output given the DM's chosen effort level $e \in E$. The DM's reward is defined by *the reward function* $R : Q(E) \to \mathbb{R}_+$. The reward function's input variable is the DM's aggregated (or average) recent performance.[1] Formally, last and current production are reduced to a single factor q, which we refer to as *past performance*, such that the DM receives the single-stage reward of $R(q)$. Due to notion of positive and diminishing marginal returns, both R and Q are assumed to be strictly-increasing, strictly-concave, and continuously-differentiable functions.[2]

The problem begins at stage $t = 1$, with an initial output of $Q_0 = Q(e_0)$. The DM chooses an effort e_1 to generate an output of $Q_1 = Q(e_1)$. Once Q_0 and Q_1 are realized, the DM is rewarded a payoff of $R(Q_0 + Q_1)$. That is, the past-performance assessment is given by $Q_0 + Q_1$. Continuing inductively, at every stage $t > 1$ and given past output $Q_0, Q_1, \ldots, Q_{t-1}$ where $Q_{t-1} = Q(e_{t-1})$, the DM exerts effort e_t to generate an output of $Q_t = Q(e_t)$, and collects the amount of $R(Q_{t-1} + Q_t)$. The DM's profit in the repeated process is given by the β-discounted sum

$$\pi(\underline{e}) = \sum_{t=1}^{\infty} \beta^{t-1}[-e_t + R(Q_{t-1} + Q_t)], \tag{1}$$

where $\underline{e} = (e_1, e_2, \ldots)$ is the DM's infinite-horizon realized actions and $\beta \in (0, 1)$. In words, the DM's payoff $\pi(\underline{e})$ is the discounted sum of its per-period payoffs, where at every stage t the DM loses an amount of e_t due to the exertion of effort, and collects a reward of $R(Q_{t-1} + Q_t)$ based on his output (and therefore, his effort) at stages t and $t - 1$.

[1] We alternatively use the terms 'output' and 'performance', since the former is used to generate an assessment over the DM.

[2] We could omit the reward function's concavity assumption, as long as the composition of the reward and output functions is concave with respect to the DM's effort. Nevertheless, for the sake of simplicity, both functions are considered concave henceforth.

A *strategy* σ is a function from all past realized outputs (histories) $\bigcup_{t\in\mathbb{N}} Q(E)^t$ to the effort set E. A *stationary strategy* σ is a function from the set of single-period output $Q(E)$ to the effort set, or equivalently, a function from the effort set to itself. Given any strategy σ and an initial effort level e_0, denote the DM's payoff by $\pi(e_0|\sigma)$, where all effort levels $\{e_t\}_{t\in\mathbb{N}}$ are determined according to σ. A strategy is considered *optimal* if it solves the optimization problem

$$\pi(e_0|\sigma) = \sup_{\underline{e}\in E^\infty} \sum_{t=1}^{\infty} \beta^{t-1}[-e_t + R(Q_{t-1} + Q_t)].$$

That is, a strategy is optimal if it produces the maximal payoff, denoted $\hat{\pi}(e_0)$, given an initial effort level of e_0.

The interior-solution property. To simplify the analysis, we require an additional technical assumption stating the the optimal solution is not trivial. Namely, we fix the parameters such that the extreme points $\{e_{\min}, e_{\max}\}$ cannot be the DM's optimal action, independently of Q_0. In other words, the DM can only gain from exerting non-minimal effort and cannot gain from the exertion of the maximal effort, thus choosing only interior points of E. One can weaken this assumption by restricting the initial condition to a subset that ensures the non-interior solutions are suboptimal.

2.0.1 Economic applications

Though we use a generic decision-problem terminology, our model applies to several economic scenarios. For example, consider the portfolio-management industry. Any investment firm typically exerts per-period "effort", either through the accumulation of information as in Stoughton (1993) and Admati and Pfleiderer (1997) or through managerial replacements as in Lynch and Musto (2003) and Dangl et al. (2008). This single-stage effort of e_t translates into the return $Q(e_t)$. Now investors observe average performance of $\frac{1}{2}(Q_{t-1} + Q_t)$ and allocate funds accordingly. The dependence on past performance is crucial since it enables a more accurate evaluation of one's abilities. Moreover, the use of running averages implies that current and recent performance are perfect substitutes. Due to assets-under-management fees, the recent average return becomes lucrative as funds flow towards the firm in stages to follow. The latter interaction is captured through the reward function $R(Q_{t-1} + Q_t)$. Interestingly, recent empirical studies (see Cornell et al. (2017) and Figure 2 of Bessler et al. (2017)) portrayed a limited cyclic performance at the firm level, whenever firms were evaluated by current and recent performance.

Another application of our model is any trading process involving credit, such that payments are distributed along sequential time periods. A firm constantly produces output, but payments are partitioned over two adjacent stages. Thus, the single-stage payoff partially depends on recent past performance, due to postponed payments, and partially depends on current performance (while the remaining compensation for current production will only be available on the next stage and so on). On the one hand, the services and costs are instantaneous and, on the other hand, payments are postponed. Given such interpretation, past performance is translated into direct monetary transfers, and the firm would strive to balance its recent average production, i.e., income and production smoothing.

One could also adapt our model to a semi optimal-growth model with post-generational transfers, as a portion of one's wealth is transferred to subsequent generations. By and large, any strategic interaction that combines the two previously-mentioned key components of marginally-decreasing output and history-based payoffs will be closely related to our framework and, therefore, to our conclusions *linking cyclic performance to incentives.*

2.1 Analysis and results - the deterministic case

The payoff function in Eq. (1) presents the basic tension under which the DM operates. On the one hand, the DM receives a history-based payoff, determined by last and current output, and prompted through a concave production function. On the other hand, costs are convexly increasing in production, thus the DM needs to balance his current and future efforts accordingly. For example, an over (under) investment in effort at a single stage, will generate a balancing counter-reaction at the subsequent stage to invest less (more) effort.

This balancing effect motivates the first result, presented in Theorem 1 that follows. First, it shows that there exists an *absorbing* effort level e^*. Once the absorbing effort level is reached, the DM will consistently exert the same level of effort throughout the stages. Second, the theorem proves that the DM balances his performance with respect to the absorbing level, at every two adjacent stages. Namely, in case the initial level is higher (lower) than the absorbing effort, the DM will invest less (more) effort relative to e^*, to level the performance at the subsequent stage. These alternating effort levels will continue to fluctuate around the absorbing level, while converging to it asymptotically.

Theorem 1. *There exists a unique, stationary and continuous optimal-strategy $\sigma : E \to E$. Given σ, the payoff function $\hat{\pi}(e_0) = \pi(e_0|\sigma)$ is a strictly-concave, and continuously-increasing function of e_0. In addition, if the interior-solution property holds, then:*

- *the optimal strategy σ is strictly decreasing with a single fixed point $e^* \in (e_{\min}, e_{\max})$;*
- *the sequences $\left(\sigma^{2n}(e_0)\right)_{n\in\mathbb{N}}$ and $\left(\sigma^{2n+1}(e_0)\right)_{n\in\mathbb{N}}$ monotonically converge to e^*;*
- *the fixed point e^* is bounded between $\sigma^n(e_0)$ and $\sigma^{n+1}(e_0)$ for every $n \in \mathbb{N}$.*

In other words, Theorem 1 suggests that a cyclic performance, monotonically and systematically converging to equilibrium, is natural when dealing with a DM concerned with current and recent performance. This outcome captures two important aspects of the current work. First, depicting a specific path and method of converging to a stable state in a dynamic-optimization problem. Second, linking aggregated performance, and therefore incentives, to a cyclic-performance phenomenon.

Remark 1. *Due to the technical nature of the analysis and proofs, we postpone them to* Appendix A. *However, we wish to refer the reader to* Lemma 1*, which studies the properties of the DM's optimal strategy, under the relevant Bellman equation. The generality of this lemma might be of some assistance in similar cases, specifically with the analysis of implicit optimal strategies in dynamic-programming problems.*

2.2 Analysis and results - the stochastic case

The first extension of the Transient model concerns the introduction of randomness to the output function. The randomness that we impose need not be i.i.d or even ergodic. Rather, we assume that the output function depends on a randomly-chosen state of the world, dictated by a Markov process, along with prior dependence on the DM's strategic effort. Though its general nature, this extension does not impairment previously-stated results. That is, in this subsection we prove that the conclusions of Theorem 1 still apply, in expectation, under the stochastic extension.

Formally, consider a finite[3] set Ω of states and denote by $P = (P_{ij})_{1 \leqslant i,j \leqslant |\Omega|}$ the transition matrix where P_{ij} is the probability of moving from state i to state j in a single time period. Given the states and transition function, consider a generalization of the output function such that $Q : \Omega \times E \to \mathbb{R}_+$ depends on the realized state $\omega \in \Omega$ and on the DM's effort. We assume that the output function maintains its basic properties independently of the realized state. Namely, for every $\omega \in \Omega$, the output function's ω-section $Q_\omega : E \to \mathbb{R}_+$ is a strictly-increasing, strictly-concave, continuously-differentiable function. Denote by S the convex hull of the compact set $Q(\Omega, E)$ of all possible realized outputs.

The stochastic decision problem evolves similarly to the deterministic one. At stage $t = 1$, with an initial state $\omega_0 \in \Omega$ and an initial output of Q_0, the DM chooses an effort level $e_1 \in E$. Next, a state ω_1 is realized according to P and ω_0, and the single-stage realized-output is $Q_1 = Q(\omega_1, e_1)$. Continuing inductively, at every stage $t > 1$ and given a *history* $h_{t-1} = (\omega_0, Q_0, \omega_1, Q_1, \ldots, \omega_{t-1}, Q_{t-1})$ of past realized outputs and states, the DM chooses an effort e_t. A state ω_t is realized according to P and

[3] In general, the use of a finite state space could be avoided by taking any countable set or any compact Borel set in $\mathbb{R}$. However, under any compact Borel set, the transition function must hold the Feller property (See Stokey et al. (1989), p. 220) roughly stating that every bounded continuous function is mapped, under the expectation operator and given the transition function, to a bounded continuous function.

ω_{t-1}, and the single-stage output is $Q(\omega_t, e_t)$. Therefore, a strategy σ of the DM is a function from the set $\bigcup_{t\in\mathbb{N}}(\Omega \times S)^t$ of all finite histories to E, such that $\sigma(h_{t-1}) = e_t$ is the strategy's realized action at stage t.[4]

Given a strategy σ and initial conditions (ω_0, Q_0), the DM's expected β-discounted payoff is

$$\pi(\omega_0, Q_0|\sigma) = \mathbf{E}_{\sigma,\omega_0}\left[\sum_{t=1}^{\infty} \beta^{t-1}(-e_t + R(Q_{t-1} + Q_t))\right], \tag{2}$$

where $\mathbf{E}_{\sigma,\omega_0}[\cdot]$ is the expectation operator with respect to the probability measure induced by the transition probabilities P, the initial state ω_0, and the strategy σ. Note that the strategy is a random variable since it depends on realized states. Thus, the expectation operator also relates to the strategy-induced effort levels throughout the stages.

By the randomness of the process, the DM's realized output might not accurately follow the same cyclic performance as in Theorem 1. However, in the following Theorem 2 we prove that current output decreases, in expectation, with respect to previously-realized ones. Moreover, we show that the optimal strategy is a strictly-decreasing function of recently-realized output. Thus, the oscillating process presented in the deterministic case remains valid, in expectation, under the stochastic extension.

Theorem 2. *There exists a unique, stationary and continuous optimal-strategy $\sigma : \Omega \times S \to E$. Given σ and ω_0, the payoff function $\hat{\pi}(\omega_0, Q_0) = \pi(\omega_0, Q_0|\sigma)$ is a strictly-concave, continuously-increasing function of Q_0. In addition, if the interior-solution property holds, then $\boldsymbol{E}_{\sigma,\omega_{t-1}}[Q_t|Q_{t-1}]$ is a strictly-decreasing function of Q_{t-1}, and for every state ω visited infinitely many times, there exists an output level q_ω such that, w.p. 1, output oscillates around q_ω infinitely many times.*

An immediate question that follows from Theorem 2 relates to the role of the DM's strategy in inducing oscillation, versus the natural oscillations that occur due to randomness. In the proof, given in Appendix B, we not only show that infinitely many oscillations occur, but also prove that these oscillations occur infinitely many times between extreme output levels, related to extreme states. These levels are not necessarily close. In this sense, the strategic reaction of the DM intensifies the oscillatory trends, even under the mean-reversion phenomenon. Therefore, Theorem 2 joins Theorem 1 to support the results of various empirical studies suggesting that DMs have a cyclic performance. Our results and model explain such occurrences through a straightforward economic reasoning, which go beyond the probabilistic properties. Namely, when DMs' payoffs are history dependent, and production is increasingly costly, DMs' strategies level their performance accordingly.

[4] Note that the functions M and R are continuous, E is compact, and Ω is finite, therefore measurability requirements, given by Assumptions $9.1' - 9.3'$ in p.$256 - 258$ of Stokey et al. (1989), are met.

As a technical comment, we note that the monotonicity result in the last statement of Theorem 2 relates strictly to previously-realized output, and not to the different states. Therefore, we fix the state variable ω_{t-1}, and derive monotonicity through changes in the realized output R_{t-1}. Without further assumptions over transitions and states, the realized output in adjacent stages might actually increase in case, e.g., the fixed state variable is significantly worse (in terms of production) than all other states.

2.2.1 Accommodating for competition

The stochastic model analysed in Theorem 2 is more than just a technical extension of the deterministic set-up. In fact, it could accommodate several possible interpretations, specifically relating to competition between various DMs as in Lagziel and Lehrer (2018), where the interaction between DMs is mediated explicitly by performance-based relative compensation.

In a general strategic set-up, each DM should respond not only to the current state but also to other players' strategies. In the current work, an implicit interaction between DMs could occur through the reward function: the DM's performance can be interpreted as relative to, e.g., other DMs or some exogenous benchmark.

Our approach has an advantage and disadvantage. On one hand, our analysis does not require that a DM would know the strategies employed by all its rivals - the aggregate effect of all other players is encapsulated in an exogenous stochastic process which determines (along with his effort) the DM's payoff. On the other hand, our model does not treat situations where the DM is truly strategic and incorporates other players' behaviour in his decisions. Moreover, the DM's payoff in our model is independent of other opponents' performance. Thus, a full analysis of a general framework with multiple players is left for future research.

3 The Persistent model

Taking a general perspective on the problem, there are numerous modification to be made in the Transient process. One possibility is to condition evaluations and rewards on a longer history path with heterogeneous weights. In this section we study a process where the reward function depends on a discounted sum of all past levels of production. Though this set-up we address the decision problem through a wider scope. Namely, we study how changes in the evaluation process affect, in the long run, the DM's realized production and payoff. Therefore, instead of focusing on short-term effects as in Section 2, in this section we inspect long-term effects due to adjustments in the mechanism.

Formally, consider an optimization problem where the DM's performance, is evaluated by a λ-

discounted sum of past output, where $\lambda \in (0,1)$. That is, consider the following optimization problem

$$\hat{\pi}(Q_0) = \sup_{\underline{e} \in E^\infty} \sum_{t=1}^{\infty} \beta^{t-1}[-e_t + R(\widehat{Q_t})],$$

such that $\widehat{Q_t} = (1-\lambda)\widehat{Q_{t-1}} + \lambda Q(e_t)$ and $\widehat{Q_0} = Q_0$. In words, this optimization problem is similar to the original Transient model, other than the exchange of the two-stage evaluation $Q_{t-1} + Q_t$ with the discounted sum $(1-\lambda)^t Q_0 + \lambda \sum_{n=1}^{t} (1-\lambda)^{t-n} Q_n$. We refer to this problem as the *Persistent model.*

The evolution of the Persistent set-up is similar to the transient one. At the beginning of every stage t and given an evaluation of $\widehat{Q_{t-1}}$, the DM exerts effort e_t to generate an output of $Q(e_t)$. The DM's updated evaluation is set to $\widehat{Q_t} = (1-\lambda)\widehat{Q_{t-1}} + \lambda Q(e_t)$ and the DM collects a reward of $R(\widehat{Q_t})$. The process continues indefinitely.

A straightforward application of the persistent model relates to an R&D investment problem, where a firm decides how much to invest (in R&D) at every time period. The decision and costs are immediate, whereas the payoff is accumulated and spread throughout the stages. Any breakthrough, either a technological advantage or superior marketing abilities, would give the firm a long-term edge over the competition; an advantage which depreciates over time. Another application is attributed to a dynamic advertising problem, where performance is accumulated in the process of building a franchise. A firm typically invests in advertising, and its payoff depends on the accumulated investment with some form of depreciation.

A comparison of Section 2's results with the following Theorem 3 will certify that many previous results hold under the updated problem, while others change completely. Starting with the similarities, we show that there exists a unique, stationary and continuous optimal-strategy $\sigma : Q(E) \to Q(E)$, such that the optimal payoff function is a strictly-concave, continuously-increasing function. In addition, we prove that the optimal strategy is strictly decreasing, thus possessing a unique fixed point, as in the previous deterministic set-up.[5] On the other hand, the two models differ in the paths by which the systems converge to a stable production level. In particular, the Persistent model generates a monotonic convergence, rather than an oscillating one. We relate to this aspect, among several others, after formally presenting the results of Theorem 3 (proofs are given in Appendix C).

Theorem 3. *In the Persistent model, there exists a unique, stationary, and continuous optimal-strategy $\sigma : Q(E) \to Q(E)$. Given the optimal strategy, the payoff function $\pi(Q_0|\sigma)$ is a strictly-concave, continuously-increasing function of Q_0. In addition, if the interior-solution property holds,*

[5]Note that we now consider strategies as functions from the set of all possible outputs to itself, rather than functions from the set of efforts to itself. This alternation is a convenient technical modification of the strategy, that could also apply in the original Transient model.

then the optimal strategy is strictly decreasing with a single, interior, fixed point $Q^* \in Q(E)$, and the sequence $(\widehat{Q_t})_{t\in\mathbb{N}}$ of realized discounted performance, generated by σ and Q_0, monotonically converges to Q^*.

There are two important aspects that arise from the comparison of Theorem 3 with Theorem 1. First, the monotonic convergence of the discounted performance towards Q^* and, second, the optimal strategy's monotonicity.

The monotonic convergence, rather than an oscillating one, follows from the DM's need to minimize costs. The cost of moving from an initial evaluation of, e.g., $\widehat{Q_0} < Q^*$ to $\widehat{Q_1} > Q^*$ is much higher when the subsequent assessment is a convex combination of current and past outputs, rather than just the current one. Therefore, the need to minimize costs guarantees that the adjacent assessment tends towards the stable effort level, relative to the initial one, but does not cross it. In other words, it is cheaper and therefore more efficient for the DM to monotonically tend towards the stable level, thus discounting costs by exerting more effort in the future, instead of extorting a significant amount of effort in the present.

On the other hand, the evaluation's no-crossing statement does not hold when considering the realized single-stage output. The optimal strategy's monotonicity suggests that a below stable-level initial production generates an above stable-level output at the subsequent stage, and vice versa. Therefore, we do expect to see an over-performance by a DM, as long as the long-term performance is below its stable, fixed-point level.

Remark 2. *In* Appendix E *we extend the persistent model to a stochastic environment.* Theorem 5 *shows that the results of* Theorem 3 *still hold, in expectation, under the stochastic extension.*

4 Long-and short-term memory

In view of the two proposed evaluation mechanisms, one of the most interesting (and potentially most policy-related) question is: how much weight should be put on past performance? Therefore, in this section we analyze the ways in which changes in the evaluation process affect the DM's optimal payoff, as well as the realized output.

To simplify the analysis, we consider the Persistent model, where the evaluation at stage t is given by $\widehat{Q_t} = (1-\lambda)\widehat{Q_{t-1}} + \lambda Q(e_t)$. The advantage of this set-up is its ability to summarize the trade-off between past and current performance through a single factor, λ. Thus, it facilitates the analysis of changes in the evaluation process and their impact on payoffs and production. Namely, in Theorem 3 we showed that the DM's performance and performance converge to a stable level Q^*, so we can examine how changes in the evaluation process, through λ, affect the long term output, through the steady level Q^*. For example, if $\lambda = 0$, there is no value to future performance, and the steady level becomes Q_0 while the DM has no incentive to exert effort. That is, the system remains fixed to the initial condition and the DM produces the minimal feasible effort level. However, if $\lambda = 1$, then past performance is not taken into account during the valuation process, and the DM repeatedly solves the optimization problem $\max_{e \in E}\{-e + R(Q(e))\}$.

Before formally stating the results, a few preliminary explanations and notations are needed. For every parameter $\lambda \in [0, 1]$, let Q^*_λ be the limit production level in the λ-discounted model under the optimal behavior described in Theorem 3. We assume that the DM acts optimally, using the optimal stationary strategy given λ, and production converges to Q^*_λ. In addition, denote the DM's optimal payoff by $\hat{\pi}_\lambda(Q_0)$, where Q_0 is the initial evaluation.

The first result of Theorem 4 relates to the monotonicity of Q^*_λ with respect to λ. We prove that the optimal strategy's steady output-level, towards which production converges, is strictly increasing in λ. To put it another way, production increases as the evaluation is myopic (with respect to past output), and the assessments depends more heavily on current performance.

On the one hand, this result is quite intuitive from a strategic point of view. When agents cannot rely on past performance, they constantly need to re-justify their abilities at every stage to come, exerting more effort in the process. The accumulated performance generates a certain *inheritance effect* where the ability to transfer value from one stage to another, leads to less exertion of effort throughout the stages.

On the other hand, the same result also hints to an important economic observation. It implies that the first-best solution, where the DM exerts the maximal rational effort, is achievable only if the DM does not retain any past dependence. That is, the only possibility of exerting the optimal effort from the DM is by ignoring past results completely at any given stage. This may pose a problem is various scenarios. For example, when such a process comes into play, the ability to screen low-level DMs is eliminated. Therefore, whenever there exists an uncertainty regarding several DMs and their differential abilities, the market needs to balance between the screening process (putting more weight on past performance) and optimal incentives (putting more weight on current ones). In general, if screening is introduced then it is quite possible that output, as a function of λ, will take the shape of a *Laffer Curve*, which illustrates the concept of taxable income elasticity (see review in Laffer (2004)). It shows how, in theory, the government's revenue from taxation is eliminated when tax rates are either 0% or 100% (i.e., not collecting any income in the first case, and cutting incentives to produce income, in the other). If screening is introduced to our set-up, then DMs' average performance drops sharply as $\lambda = 0$ due to poor incentives, and once again drops sharply as $\lambda = 1$ due to poor screening.

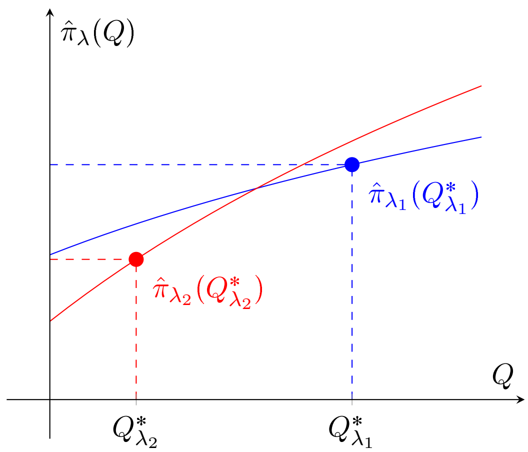

Figure 1: DM's expected payoff as a function of Q_0 and λ, where $\lambda_1 > \lambda_2$. Optimal-payoff functions are concave and continously increasing in the initial position.

Hence, the quest to achieve a first-best mechanism ends in an *Incentives-Screening Deadlock*, where one needs to balance between optimal incentives and optimal screening.

The second part of Theorem 4 concerns the DM's payoff as a function of λ and Q_0. These results are best exemplified in Figure 1, showing inverse effects between the discount factor and the initial condition. Namely, fix a discount factor λ_0 and the appropriate limit production level $Q^*_{\lambda_0}$. Whenever the initial output level is below $Q^*_{\lambda_0}$, the DM's payoff increases if the discount factor is above λ_0. However, whenever the initial position is above $Q^*_{\lambda_0}$, the DM's payoff increases if the discount factor is fixed below λ_0. Again, the economic intuition is clear. In case the initial position is high, the DM will prefer to maintain it as long as possible without exerting additional effort, while a low initial position can only harm the expected payoff if past performance is weighted heavily hereafter.

The combination of these two results generates a somewhat more surprising outcome. It shows that *any valuation factor* other than λ is preferable to the DM, given an initial position of $Q_0 = Q^*_\lambda$. That is, the valuation factor that generates the lowest expected output, given some initial position, is the one that imposes the same steady level. As it appears, once production converges to a stable level, the DM can only profit from either an increased or a decreased valuation factor, though the two generate inverse incentives.

Theorem 4. *Given the interior-solution property, the steady output level Q^*_λ strictly increases as a function of λ. Moreover, for every $\lambda_1 \neq \lambda_2$ and given an initial position of $Q^*_{\lambda_1}$, the DM's payoff is higher under the λ_2-valuation rather than the λ_1-valuation, i.e., $\hat{\pi}_{\lambda_2}\left(Q^*_{\lambda_1}\right) > \hat{\pi}_{\lambda_1}\left(Q^*_{\lambda_1}\right)$. In addition,* *If* $\lambda_1 > \lambda_2$*, then*

- $\hat{\pi}_{\lambda_2}(Q) > \hat{\pi}_{\lambda_1}(Q)$*, for every* $Q \geqslant Q^*_{\lambda_1}$*;*
- $\hat{\pi}_{\lambda_2}(Q) < \hat{\pi}_{\lambda_1}(Q)$*, for every* $Q \leqslant Q^*_{\lambda_2}$*;*
- $\hat{\pi}'_{\lambda_2}(Q) > \hat{\pi}'_{\lambda_1}(Q)$*, for every* $Q \leqslant Q^*_{\lambda_1}$*.*

Theorem 4 is best understood through Figure 1. First, if recent performance is weighted more heavily (i.e., $\lambda_1 > \lambda_2$), then incentives are sharpened such that: (i) output converges to a higher level of $Q^*_{\lambda_1} > Q^*_{\lambda_2}$; and (ii) dependence on the initial condition weakens, and the derivative w.r.t. Q_0 decreases. Next, in case the initial position Q_0 is low, the DM would prefer a lower evaluation of past production (blue line), quickly neglecting past performance, rather than a high evaluation of past production (red line). The opposite statement holds whenever the initial position is high. Moreover, once the initial position is a steady state of some λ_i evaluation (i.e., $Q_0 = Q^*_{\lambda_i}$), then any other λ_{-i} evaluation, *either low or high*, would be strictly preferable to the DM than λ_i.

5 Concluding remarks

In this paper we presented a model explaining the oscillatory performance issue arising from history-based payoffs. Though our setting is robust with respect to the governing output-generating process and underline probabilistic environment, there are several possible extensions to follow our model and analysis. First, one could model the interaction between multiple players to capture the oscillating performance in a general dynamic market. In light of the theoretical complexity behind such a model (as the dynamic interaction between DMs induces a stochastic game), the implementation of numerical analysis is imminent. Second, an extension of our Laffer-Curve idea, associating incentives and screening while taking into account the uncertainty regarding the agent's subjective abilities, is evident. We believe that the main obstacle lays in capturing the same phenomenon in a simple, yet general, static model. Lastly, a comprehensive analysis of our model while taking into account a wider range of evaluation processes, could produce important insights into the oscillating-performance problem.

References

Admati, Anat Ruth and Paul Pfleiderer, "Does it all add up? Benchmarks and the compensation of active portfolio managers," *The Journal of Business*, 1997, *70* (3), 323–50.

Amihud, Yakov and Ruslan Goyenko, "Mutual Fund's R2 as Predictor of Performance," *Review of Financial Studies*, mar 2013, *26* (3), 667–694.

Arrow, Kenneth J., Samuel Karlin, and Herbert Scarf, *Studies in the mathematical theory of inventory and production*, Stanford Univ. Pr, 1958.

_ , Theodore Harris, and Jacob Marschak, "Optimal Inventory Policy," *Econometrica*, 1951, *19* (3), 250–272.

Bailey, Martin J., "The Welfare Cost of Inflationary Finance," *Journal of Political Economy*, apr 1956, *64* (2), 93–110.

Barro, Robert J., "A theory of monopolistic price adjustment," *The Review of Economic Studies*, jan 1972, *39* (1), 17–26.

Bellman, Richard E., *Dynamic Programming*, Princeton University Press, 1957.

_ , Irving L. Glicksberg, and Oliver A. Gross, "On the Optimal Inventory Equation," *Management Science*, 1955, *2* (1), 83–104.

Berk, Jonathan B. and Richard C. Green, "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, 2004, *112* (6), 1269–1295.

Bessler, Wolfgang, David Blake, Peter Lückoff, and Ian Tonks, "Fund Flows, Manager Changes, and Performance Persistence," *Review of Finance*, may 2017, *43*, 741–768.

Blackwell, David, "Discounted Dynamic Programming," *The Annals of Mathematical Statistics*, 1965, *36* (1), 226–235.

Brock, William A and JoséA Scheinkman, "Global asymptotic stability of optimal control systems with applications to the theory of economic growth," *Journal of Economic Theory*, feb 1976, *12* (1), 164–190.

Brumberg, Richard and Franco Modigliani, "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data," in Kenneth K. Kurihara, ed., *Post Keynesian Economics*, Rutgers University Press, 1954, pp. 388–436.

Carhart, Mark M., "On Persistence in Mutual Fund Performance," *The Journal of Finance*, mar 1997, *52* (1), 57–82.

Cass, David, "Optimum Growth in an Aggregative Model of Capital Accumulation," *The Review of Economic Studies*, jul 1965, *32* (3), 233.

— **and Karl Shell**, "The structure and stability of competitive dynamical systems," *Journal of Economic Theory*, feb 1976, *12* (1), 31–70.

Chassang, Sylvain, "Calibrated Incentive Contracts," *Econometrica*, 2013, *81* (5), 1935–1971.

Cornell, Bradford, Jason Hsu, and David Nanigian, "Does Past Performance Matter in Investment Manager Selection?," *The Journal of Portfolio Management*, jul 2017, *43* (4), 33–43.

Coval, Joshua D. and Erik Stafford, "Asset Fire Sales (and Purchases) in Equity Markets," *Journal of Financial Economics*, 2007, *86*, 479–512.

Dangl, Thomas, Youchang Wu, and Josef Zechner, "Market Discipline and Internal Governance in the Mutual Fund Industry," *Review of Financial Studies*, sep 2008, *21* (5), 2307–2343.

DeMarzo, Peter M. and Michael J. Fishman, "Optimal Long-Term Financial Contracting," *Review of Financial Studies*, nov 2007, *20* (6), 2079–2128.

Dvoretzky, Aryeh, Jack Kiefer, and Jacob Wolfowitz, "The Inventory Problem: I. Case of Known Distributions of Demand," *Econometrica*, 1952, *20* (2), 187–222.

— , — , and — , "On the Optimal Character of the (s, S) Policy in Inventory Theory," *Econometrica*, 1953, *21* (4), 586–596.

Foster, D., "Asymptotic calibration," *Biometrika*, jun 1998, *85* (2), 379–390.

Fudenberg, Drew and David K. Levine, "An Easier Way to Calibrate," *Games and Economic Behavior*, oct 1999, *29* (1-2), 131–137.

Hall, Robert E., "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, dec 1978, *86* (6), 971–987.

Holmstrom, Bengt and Paul Milgrom, "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 1987, *55* (2), 303–28.

Koopmans, Tjalling C., "On the Concept of Optimal Economic Growth," *Cowles Foundation Discussion Papers*, 1965.

Laffer, Arthur B., "The Laffer Curve: Past, Present, and Future," in "Backgrounder," Vol. 1765, Washington: The Heritage Foundation, 2004.

Laffont, Jean-Jacques and Jean Tirole, "The Dynamics of Incentive Contracts," *Econometrica*, sep 1988, *56* (5), 1153.

Lagziel, David and Ehud Lehrer, "Reward schemes," *Games and Economic Behavior*, jan 2018, *107*, 21–40.

Lehrer, Ehud, "Any inspection is manipulable," *Econometrica*, 2001, *69* (5), 1333–1347.

Lynch, Anthony W. and David K. Musto, "How Investors Interpret Past Fund Returns," *The Journal of Finance*, 2003, *58* (5), 2033–2058.

Modigliani, Franco, "The Life Cycle Hypothesis of Saving, The Demand for Wealth and The Supply of Capital," *Social Research*, 1966, *33*, 160–217.

Rakowski, David, "Fund Flow Volatility and Performance," *The Journal of Financial and Quantitative Analysis*, 2010, *45*, 223–237.

Ramsey, F. P., "A Mathematical Theory of Saving," *The Economic Journal*, dec 1928, *38* (152), 543.

Rockafellar, Tyrrell R., "Saddle points of Hamiltonian systems in convex Lagrange problems having a nonzero discount rate," *Journal of Economic Theory*, feb 1976, *12* (1), 71–113.

Scarf, Herbert, "The optimality of (5, 5) policies in the dynamic inventory problem," in Kenneth J. Arrow, Samuel Karlin, and Patrick Suppes, eds., *Mathematical Models in Social Sciences*, Stanford University Press, 1959.

Scheinkman, Alexandre J., "On optimal steady states of n-sector growth models when utility is discounted," *Journal of Economic Theory*, feb 1976, *12* (1), 11–30.

Sheshinski, Eytan and Yoram Weiss, "Inflation and Costs of Price Adjustment," *The Review of Economic Studies*, jun 1977, *44* (2), 287–303.

Sheshinski, Eytan. and Yoram. Weiss, *Optimal pricing, inflation, and the cost of price adjustment*, MIT Press, 1993.

Stokey, Nancy L., Robert E. Lucas, and Edward C. Prescott, *Recursive Methods in Economic Dynamics*, Harvard University Press, 1989.

Stoughton, Neal M., "Moral Hazard and the Portfolio Management Problem," *The Journal of Finance*, 1993, *48* (5), 2009.

Appendices

A The deterministic case - analysis and proofs

Eq. (1) produces the following sequential optimization problem

$$\hat{\pi}(e_0) = \sup_{\underline{e} \in E^{\infty}} \sum_{t=1}^{\infty} \beta^{t-1}[-e_t + R(Q_{t-1} + Q_t)], \tag{3}$$

where $\hat{\pi}$ is a optimal payoff given initial effort level $e_0 \in E$. For the simplification of Eq. (3), Bellman (1957) defines the *Principle of Optimality* suggesting the analysis of the corresponding functional equation,

$$\hat{\pi}(e_0) = \sup_{e \in E} \left[-e + R(Q_0 + Q(e)) + \beta\hat{\pi}(e)\right], \tag{4}$$

where one can focus on a single-period problem instead of an infinite one. Note that the use of $\hat{\pi}$ in Eq. (4) is not trivial.[6] The properties of E (non-empty, compact, and convex) and the continuity of Q and R, suggest that the single-period payoffs are bounded. Thus, we can apply Theorems 4.2, 4.3, and 4.6 from chapter 4 of Stokey et al. (1989) to prove the existence, uniqueness, and continuity of the same solution $\hat{\pi}$ to both Eq. (3) and Eq. (4). In simple terms, we can use the customary approach in discounted dynamic optimization and restrict our analysis to the unique solution of Eq. (4).

Define the correspondence $\sigma : E \to 2^E$ such that

$$\sigma(e_0) = \{e \in E \mid \hat{\pi}(e) = F(e_0, e) + \beta\hat{\pi}(e)\},$$

where $F(e_0, e) = -e + R(Q(e_0) + Q(e))$ and for every $e_0 \in E$. Since all functions are bounded and continuous, σ is well defined. Moreover, Theorems[7] SLP-4.4, SLP-4.5, and SLP-4.6 prove that σ is a compact-valued, upper hemi-continuous correspondence, that generates the DM's optimal strategy either in the functional equation, Eq. (4), or in the sequential problem, Eq. (3).

To show that σ is a single-valued continuous function, we need to prove that $F(e_0, e)$ is concave w.r.t. e_0 and e, and strictly concave w.r.t. e_0 (by SLP-4.8). Fix $\delta \in (0, 1)$ and $(e_0, e), (e_0', e') \in E^2$ such that $e_0 \neq e_0'$. By the strict concavity of Q if follows that

$$Q(\delta e_0 + (1-\delta)e_0') + Q(\delta e + (1-\delta)e') > \delta[Q(e_0) + Q(e)] + (1-\delta)[Q(e_0') + Q(e')].$$

Since R is monotonic and concave,

$$R(Q(\delta e_0 + (1-\delta)e_0') + Q(\delta e + (1-\delta)e')) > \delta R(Q(e_0) + Q(e)) + (1-\delta)R(Q(e_0') + Q(e')),$$

[6] We also refer to $\hat{\pi}$ as the *value function* of the corresponding problems.

[7] Hereafter, we refer to Theorems 4.2 – 4.11 in p.71 – 85 of Stokey et al. (1989) as SLP-4.XX. A significant number of these theorems are based on the pioneering work of Blackwell (1965).

and

$$F(\delta(e_0, e) + (1-\delta)(e_0', e')) > \delta F(e_0, e) + (1-\delta)F(e_0', e').$$

Thus, σ is a single-valued continuous function and, by SLP-4.8, the value function $\hat{\pi}$ is strictly concave. In addition, the fact that F is strictly increasing in e_0 implies that the value function $\hat{\pi}$ is strictly increasing (see SLP-4.7).

Proving that $\hat{\pi}$ is differentiable using SLP-4.11 requires $\sigma(E)$ to be interior points of E (which holds by the interior-solution property). In such cases, the value function is continuously differentiable, and we can effectively use the envelope theorem: in the FOC of Eq. (4) we plug-in the optimal solution $\sigma(e_0)$ to obtain,

$$R'(Q(e_0) + Q(\sigma(e_0)))Q'(\sigma(e_0)) + \beta\hat{\pi}'(\sigma(e_0)) = 1. \tag{5}$$

Eq. (5) enables us to study the properties of σ. We start with monotonicity. Consider a small increase of e_0 to $e_0 + \varepsilon > e_0$. If $\sigma(e_0 + \varepsilon) \geqslant \sigma(e_0)$, then the LHS of Eq. (5) decrease, violating the equality, since R', Q', and $\hat{\pi}'$ are (non-negative) decreasing functions due to the strict concavity of R, Q, and $\hat{\pi}$. It implies that σ is a strictly-decreasing continuous function from E to E, thus it has a unique fixed point $e^* \in E$ (also an interior point of E) such that

$$R'(Q(e^*) + Q(e^*))Q'(e^*) + \beta\hat{\pi}'(e^*) = 1.$$

The next step of our analysis shows that, for every $e_0 \in E$, the sequences $\left(\sigma^{2n}(e_0)\right)_{n\in\mathbb{N}}$ and $\left(\sigma^{2n+1}(e_0)\right)_{n\in N}$ monotonically converge to e^*, as n tends to infinity. In addition, the uniqueness of e^*, along with the continuity of σ, imply that e^* is bounded between $\sigma^n(e_0)$ and $\sigma^{n+1}(e_0)$ for every n. That is, we show that the repeated use of σ tends monotonically to the fixed point e^*, and $\sigma^n(e_0)$ oscillates around e_0 as a function of n, with an amplitude converging to 0.

Now define $H(x, y) = R'(Q(x) + Q(y))Q'(y) + \beta\hat{\pi}'(y)$ and note that the concavity of Q and $\hat{\pi}$ imply that $H(x, y) > H(y, x)$ for every $(y, x) \subseteq E$. In addition, H is continuous and strictly-decreasing in both coordinates. Therefore, H satisfies the conditions of the following Lemma 1.

Lemma 1. *Let $H : E^2 \to \mathbb{R}$ be a continuous function, strictly-decreasing in both coordinates, such that*

$$H(x, y) > H(y, x), \tag{6}$$

for every $(y, x) \subseteq E$. Then,

1. *there exists $c \in \mathbb{R}$ such that for every $x \in E$, the equation $H(x, y) = c$ has a unique solution, $y_c(x)$;*
2. *the function y_c is a continuous, strictly decreasing function with a unique fixed point x_c;*

3. *for every* $x \in E$*, the sequences* $\left(y_c^{2n}(x)\right)_{n\in\mathbb{N}}$ *and* $\left(y_c^{2n+1}(x)\right)_{n\in\mathbb{N}}$ *monotonically tend to* x_c *as* $n \to \infty$*;*

4. *the fixed point* x_c *is bounded between* $y_c^n(x)$ *and* $y_c^{n+1}(x)$ *for every* $n \in \mathbb{N}$ *and every* $x \in E$*.*

Proof. Ineq. (6) implies that $H(e_{\max}, e_{\min}) > H(e_{\min}, e_{\max})$ and we can fix c between these two values of H. It follows from the strict monotonicity of H that $H(x, e_{\min}) > c > H(x, e_{\max})$ for every $x \in E$. By continuity, for every $x \in E$ there exists a solution $y_c(x)$ for $H(x, y) = c$, and strict monotonicity suggests $y_c(x)$ is unique. In addition, the same two properties of H imply that y_c is continuous and strictly-decreasing. Moreover, y_c is defined from E to E, therefore it has a unique fixed point, denoted x_c. We conclude that $H(x_c, y_c(x_c)) = H(x_c, x_c) = c$.

Fix $x \in E$ such that $x > x_c$. Since $H(x_c, y_c(x_c)) = c$ where H is strictly decreasing, we deduce that $y_c(x) < y_c(x_c) = x_c$. Assume, contrary to the stated lemma, that $y_c^2(x) = x$. Then,

$$c = H(y_c(x), y_c^2(x)) = H(y_c(x), x) < H(x, y_c(x)) = c,$$

where the inequality follows from Ineq. (6). A contradiction. Since $H(y_c(x), x) < c$, we conclude that $x_c < y_c^2(x) < x$, as needed. A similar proof holds for $x > x_c$. Hence, we can now consider the sequences $\left(y_c^{2n}(x)\right)_{n\in\mathbb{N}}$ and $\left(y_c^{2n+1}(x)\right)_{n\in\mathbb{N}}$ bounding x_c. Each of the two sequences tends closer to x_c as n grows. Since both are monotonic and bounded, they converge. Assume, by contradiction, that a sequence, e.g., $\left(y_c^{2n+1}(x)\right)_{n\in\mathbb{N}}$, converges to $x'_c \neq x_c$. Then,

$$x'_c = \lim_{n\to\infty} y_c^{2n+3}(x) = \lim_{n\to\infty} y_c^2\left(y_c^{2n+1}(x)\right) = y_c^2\left(\lim_{n\to\infty} y_c^{2n+1}(x)\right) = y_c^2(x'_c),$$

contradicting the strict monotonicity of the sequences when $x \neq x_c$, and concluding the proof. ■

To conclude, consider the previously-defined function $H(x, y) = R'(Q(x) + Q(y))Q'(y) + \beta\hat{\pi}'(y)$, and note that it sustains all the conditions of Lemma 1. Fix $c = 1$ and substitute x and y by e_0 and σ, respectively. The previous analysis, prior to Lemma 1, shows that σ has a unique fixed point e^* given $c = 1$, and Lemma 1 ensures that an iteration of σ cyclically converges to the fixed point e^*. Thus, we conclude the proof of Theorem 1. ■

B The stochastic case - analysis and proof

Similarly to the previous analysis, we transform the optimization problem derived from Eq. (2) to the following Bellman equation

$$\pi(\omega_0, Q_0) = \sup_{e\in E} \mathbf{E}_{\omega_0}\left[-e + R(Q_0 + Q(\tilde{\omega}, e)) + \beta\pi(\tilde{\omega}, Q(\tilde{\omega}, e))\right], \tag{7}$$

where the expectation relates to the random variable $\tilde{\omega}$, drawn according to P and ω_0.

Proof of Theorem 2. To prove Theorem 2, we follow Chapter 9 of Stokey et al. (1989), and specifically, Assumptions $9.4-9.12$ and $9.16-9.17$ along with Exercise 9.7. Formally, Assumption 9.4 follows from the definition of S; Assumption 9.5 follows from the assumptions on Ω and P; Assumptions $9.6, 9.9, 9.11$ and 9.16 hold since E is a fixed convex interval; Assumptions $9.7, 9.8, 9.10$ and 9.12 follow from Assumptions $4.4, 4.5, 4.7$, and 4.9 respectively, mentioned in Appendix A along with the linearity of the expectation operator; and Assumption 9.17 holds by the differentiability of R_ω. In addition, the stage-t output depends solely on the realized action and the realized state at stage $t-1$, thus the condition given in Exercise 9.7-f is met, and the result follows by the interior-solution property. To sum-up, there exists a continuously-increasing, strictly-concave, differentiable payoff function $\hat{\pi}(\omega_0, Q_0)$ (all w.r.t. Q_0), and there exists a unique, stationary, and continuous optimal-strategy $\sigma : \Omega \times S \to E$.

By the differentiability of the RHS of Eq. (7) w.r.t. e, we use the envelope theorem and plug-in $\sigma(\omega_0, Q_0)$ after taking the FOC, to get

$$0 = \mathbf{E}_{\omega_0}\left[-1 + \left(R'(Q_0 + Q(\tilde{\omega}, e)) + \beta \frac{\partial \pi(\tilde{\omega}, Q(\tilde{\omega}, e))}{\partial Q(\tilde{\omega}, e)}\right) \frac{\partial Q(\tilde{\omega}, e)}{\partial e}\right]_{e=\sigma(\omega_0, Q_0)}. \tag{8}$$

The monotonicity and concavity of the output function, the reward function, and the payoff function imply that the derivatives on the RHS decrease when either Q_0 or $\sigma(\omega_0, Q_0)$ increase. Thus, an increase in Q_0 must follow a decrease in $\sigma(\omega_0, Q_0)$ to maintain Eq. (8). Hence, $\mathbf{E}_{\sigma,\omega_{t-1}}[Q_t|Q_{t-1}] = \mathbf{E}_{\omega_{t-1}}[Q(\tilde{\omega}, \sigma(\omega_{t-1}, Q_{t-1})]$, and $\sigma(\omega_{t-1}, Q_{t-1})$ decrease w.r.t Q_{t-1}.

To prove a cyclic performance, we start with the simple case where the state variable is absorbed, w.p. 1, to some fixed state $\omega \in \Omega$. In such a case, the proof of Theorem 1 holds and a cyclic performance follows. Otherwise, assume w.l.o.g. that the chain is irreducible. For every $\omega \in \Omega$, consider the function $\psi_\omega(q) = \mathbf{E}_\omega\left[Q(\tilde{\omega}, \sigma(\omega, q))\right]$. By the continuity and monotonicity of σ along with the compactness assumption over $Q(E)$, there exists a unique fixed point q_ω such that $\psi_\omega(q_\omega) = q_\omega$. Since the Markov chain is finite and irreducible, we can take the stationary distribution μ and define $Q^* = \mathbf{E}_\mu[q_{\tilde{\omega}}]$, where the expectation is taken w.r.t. μ.

Fix $\overline{\omega}, \underline{\omega} \in \Omega$ such that $q_{\overline{\omega}} > q_\omega > q_{\underline{\omega}}$, for every $\omega \in \Omega \backslash \{\overline{\omega}, \underline{\omega}\}$. We will show that for every $\varepsilon > 0$, w.p. 1, every trajectory visits the two intervals $(-\infty, q_{\underline{\omega}} + \varepsilon]$ and $[q_{\overline{\omega}} - \varepsilon, \infty)$ infinitely many times, thus oscillating around Q^* as needed. The idea behind this statement is that both $\overline{\omega}$ and $\underline{\omega}$ are visited infinitely many times, and whenever the realized output is within $[q_{\underline{\omega}}, q_{\overline{\omega}}]$, then the expected production in the subsequent period is outside $[q_{\underline{\omega}}, q_{\overline{\omega}}]$. Namely, the monotonicity of σ implies that for every state ω, the inequality $\psi_\omega(q) > q_\omega$ holds if and only if $q < q_\omega$. Meaning, a realized production below (above) the fixed point q_ω ensures next-stage's expected output is above (below, resp.) the fixed point. In other words, output *oscillates in expectation.*

Fix a small $\varepsilon > 0$ such that $Q^* \in (q_{\underline{\omega}} + \varepsilon, q_{\overline{\omega}} - \varepsilon)$. The compactness of $Q(E)$ along with the oscillation-in-expectation property guarantees that there exists $\delta > 0$ such that $\Pr(Q(\tilde{\omega}, \sigma(\overline{\omega}, q)) > q_{\overline{\omega}}) > \delta$ for every $q < q_{\overline{\omega}}$ and, equivalently, $\Pr(Q(\tilde{\omega}, \sigma(\underline{\omega}, q)) < q_{\underline{\omega}}) > \delta$ for every $q > q_{\underline{\omega}}$. We will now turn to a proof by contradiction.

Denote $I = [q_{\underline{\omega}} + \varepsilon, q_{\overline{\omega}} - \varepsilon]$ and assume there is a positive probability event $D = \bigcup_{t \in \mathbb{N}} D_t$ where D_t includes all histories such that the realized output from stage t onwards is in I. Since D has positive probability, there exists $D_T \subseteq D$ with positive probability, and a positive-probability finite history h, of length greater than T stages, such that $\Pr(D_T|h) > 1 - \delta$. Now consider all continuations of h. Each continuation h' settles in $\overline{\omega}$ infinitely often. Let $\tau[h']$ be the first stage, after h, where $\overline{\omega}$ is the state variable according to a continuation h'. The construction implies that $Q_{\tau[h']} \in I$, and specifically $Q_{\tau[h']} < q_{\overline{\omega}}$. By the previous statement, we know that for every $Q_{\tau[h']} < q_{\overline{\omega}}$,

$$\Pr\left(Q\left(\tilde{\omega}, \sigma\left(\omega_{\tau[h']}, Q_{\tau[h']}\right)\right) > q_{\overline{\omega}} \middle| \, h, \omega_{\tau[h']} = \overline{\omega}\right) > \delta.$$

Summing over all stages $\tau(h')$, we get that $\Pr(\overline{D_T}|h) > \delta$, contradicting the initial assumption that $\Pr(D_T|h) > 1 - \delta$ and concluding the proof. ∎

C Proof of Theorem 3

Proof. In this proof we follow the same analysis present in Appendix A. However, to simplify the notation, we use the set $Q(E)$ instead of E to denote the DM's actions, and denote the initial position by $Q_0 \in Q(E)$. Therefore, the equivalent functional equation to Eq. (4) becomes

$$\hat{\pi}(Q_0) = \sup_{q \in Q(E)} \left[-Q^{-1}(q) + R((1-\lambda)Q_0 + \lambda q) + \beta \hat{\pi}((1-\lambda)Q_0 + \lambda r)\right], \tag{9}$$

where the DM chooses a production level q, receives a payoff of $-Q^{-1}(q) + R((1-\lambda)Q_0 + \lambda q)$, and moves on to the next stage with evaluation $(1-\lambda)Q_0 + \lambda q$. Note that Q^{-1} is the inverse function of Q, and therefore strictly-increasing, strictly-convex and continuously-differentiable.

By the properties of E, Q, and R we can use SLP-4.2, SLP-4.3, and SLP-4.6 (similarly to Theorem 1) to prove the existence, uniqueness, and continuity of $\hat{\pi}$. Re-define the correspondence $\sigma : Q(E) \to 2^{Q(E)}$ such that

$$\sigma(Q_0) = \{q \in Q(E) \mid \hat{\pi}(q) = F(Q_0, q) + \beta \hat{\pi}((1-\lambda)Q_0 + \lambda q)\},$$

where $F(Q_0, q) = -Q^{-1}(q) + R((1-\lambda)Q_0 + \lambda q)$. Theorems SLP-4.4, SLP-4.5, and SLP-4.6 prove that σ is a compact-valued, upper hemi-continuous correspondence, that generates the DM's optimal strategy.

To show that σ is a single-valued continuous function, we need to prove that $F(Q_0, q)$ is concave w.r.t. Q_0 and q, and strictly concave w.r.t. Q_0 (see SLP-4.8). By the strict convexity of Q^{-1} and by the same analysis as in Appendix A, the concavity condition of F holds and σ is a single-valued continuous function, while the value function $\hat{\pi}$ is strictly concave. In addition, the fact that F is strictly increasing in Q_0 implies that the value function $\hat{\pi}$ is also strictly increasing (see SLP-4.7).

The interior-solution property and SLP-4.11 prove that the value function is continuously differentiable, and by the envelope theorem we can follow the analysis of Chapter 4 in Stokey et al. (1989), to write down the following FOC of the Bellman equation,

$$0 = -\frac{1}{\lambda Q'(Q^{-1}(\sigma(Q_0)))} + R'((1-\lambda)Q_0 + \lambda\sigma(Q_0)) + \beta\hat{\pi}'((1-\lambda)Q_0 + \lambda\sigma(Q_0)),$$

or equivalently,

$$\lambda Q'(Q^{-1}(\sigma(Q_0)))\left[R'((1-\lambda)Q_0 + \lambda\sigma(Q_0)) + \beta\hat{\pi}'((1-\lambda)Q_0 + \lambda\sigma(Q_0))\right] = 1 \tag{10}$$

Next, consider a small increase of Q_0 to $Q_0 + \varepsilon > Q_0$. If $\sigma(Q_0 + \varepsilon) \geqslant \sigma(Q_0)$, then the LHS of the last equation decreases (since Q, R, and $\hat{\pi}$ are concave), violating the equality. Hence, we proved that σ is a strictly-decreasing continuous function from $Q(E)$ to $Q(E)$, thus it has a unique, interior, fixed point $Q^* \in Q(E)$ such that

$$\lambda Q'(Q^{-1}(Q^*))\left[R'(Q^*) + \beta\hat{\pi}'(Q^*)\right] = 1.$$

Combining the last two equations (the FOC equality and the last fixed-point equality) yields

$$\lambda Q'(Q^{-1}(Q^*))\left[R'(Q^*) + \beta\hat{\pi}'(Q^*)\right] = \lambda Q'(Q^{-1}(\sigma(Q_0)))\left[R'(\hat{Q}) + \beta\hat{\pi}'(\hat{Q})\right],$$

where $\hat{Q} = (1-\lambda)Q_0 + \lambda\sigma(Q_0)$.

Assume $Q_0 < Q^*$. The monotonicity of σ implies that $\sigma(Q_0) > Q^*$, and so $Q'(Q^{-1}(Q^*))) > Q'(Q^{-1}(\sigma(Q_0)))$. Therefore, it follows from the last equation that $R'(\hat{Q}) + \beta\hat{\pi}'(\hat{Q}) > R'(Q^*) + \beta\hat{\pi}'(Q^*)$, or equivalently $Q^* > (1-\lambda)Q_0 + \lambda\sigma(Q_0) > Q_0$. In words, we showed that an initial position of $Q_0 < Q^*$ imposes a production above Q^* in the subsequent stage, while maintaining the subsequent position below Q^*. By induction, the same result applies in every stage to follow. Symmetrically, one reaches a similar conclusion given $Q_0 > Q^*$, and we derive that the sequence $(\widehat{Q_t})_{n\in\mathbb{N}}$ generated by σ and Q_0, monotonically converges to Q^*. ∎

D Proof of Theorem 4

Proof. To simplify the proof, we use the same notation as in the proof of Theorem 3 where the relevant Bellman equation is given by Eq. (9),

$$\hat{\pi}_\lambda(Q_0) = \sup_{q \in Q(E)} \left[-Q^{-1}(r) + R((1-\lambda)Q_0 + \lambda q) + \beta \hat{\pi}((1-\lambda)Q_0 + \lambda q) \right],$$

such that Q^{-1} is the (strictly increasing and convex) inverse function of Q. To use Bellman's principle of optimality and Blackwell's Contraction Mapping Theorem, we need to find a contracting operator from the set of bounded functions to itself. Let B be the set of bounded real-valued functions over $Q(E)$. For every λ, define the operator $T_\lambda : B \to B$ such that

$$(T_\lambda f)(Q_0) = \max_{q \in Q(E)} \left[-Q^{-1}(r) + R((1-\lambda)Q_0 + \lambda q) + \beta f((1-\lambda)Q_0 + \lambda q) \right],$$

for every $Q_0 \in Q(E)$. This operator, along with the results of chapter 3 of Stokey et al. (1989), was used explicitly to prove Theorem 3, and will be similarly used in the current proof.

The proof is divided into five parts with respect to the different parts of the theorem:

Part I proves $\hat{\pi}_{\lambda_2}\left(Q^*_{\lambda_1}\right) \geqslant \hat{\pi}_{\lambda_1}\left(Q^*_{\lambda_1}\right)$, for every two discount factors $\lambda_1 \neq \lambda_2$ such that $Q_0 = Q^*_{\lambda_1}$.

Part II proves $\hat{\pi}_{\lambda_2}(Q_0) > \hat{\pi}_{\lambda_1}(Q_0)$, for every two discount factors $\lambda_1 > \lambda_2$ such that $Q_0 > Q^*_{\lambda_1}$.

Part III proves $\hat{\pi}_{\lambda_2}(Q_0) < \hat{\pi}_{\lambda_1}(Q_0)$, for every two discount factors $\lambda_1 > \lambda_2$ such that $Q_0 < Q^*_{\lambda_2}$.

Part IV proves Q^*_λ is strictly increasing in λ.

Part V proves Part II and Part III for the cases where $Q_0 = Q^*_{\lambda_1}$ and $Q_0 = Q^*_{\lambda_2}$, respectively. Applying Parts II, III, and V to any λ and with respect to higher and lower discount factors produces the desired result.

Part I. Since $Q_0 = Q^*_{\lambda_1}$ is a fixed point of the λ_1-valuation problem, the DM will repeatedly generate an output of $Q^*_{\lambda_1}$ and a payoff of $\hat{\pi}_{\lambda_1}(Q^*_{\lambda_1})$. Thus, for any function $f \in B$ such that $f(Q^*_{\lambda_1}) \geqslant \pi_{\lambda_1}(Q^*_{\lambda_1})$, it follows that

$$\begin{aligned} (T_{\lambda_2} f)(Q^*_{\lambda_1}) &\geqslant -Q^{-1}(Q^*_{\lambda_1}) + R(Q^*_{\lambda_1}) + \beta f(Q^*_{\lambda_1}) \\ &\geqslant -Q^{-1}(Q^*_{\lambda_1}) + R(Q^*_{\lambda_1}) + \beta \pi_{\lambda_1}(Q^*_{\lambda_1}) \\ &= \hat{\pi}_{\lambda_1}(Q^*_{\lambda_1}), \end{aligned}$$

where the first inequality follows from substituting the optimal q with $Q^*_{\lambda_1}$, and the second inequality follows from the assumption over f. By Bellman's principle of optimality and Banach's Contraction

Mapping Theorem along with the fact that the set of bounded functions that sustain the condition $f(Q^*_{\lambda_1}) \geqslant \pi_{\lambda_1}(Q^*_{\lambda_1})$ is closed, it follows that $\hat{\pi}_{\lambda_2}(Q^*_{\lambda_1}) \geqslant \hat{\pi}_{\lambda_1}(Q^*_{\lambda_1})$, as needed.

Part II. Fix a function $f \in B$ such that $f(q) \geqslant \hat{\pi}_{\lambda_1}(q)$ for every $q > Q^*_{\lambda_1}$. By Theorem 3, we know that $Q_0 > Q^*_{\lambda_1}$ implies $\sigma_{\lambda_1}(Q_0) < Q^*_{\lambda_1}$, where σ_{λ_1} is the optimal stationary strategy in the λ_1-valuation problem, such that the position in the next stage tends towards $Q^*_{\lambda_1}$ from above. Hence,

$$\begin{aligned}
(T_{\lambda_2} f)(Q_0) &\geqslant -Q^{-1}(\sigma_{\lambda_1}(Q_0)) + R((1-\lambda_2)Q_0 + \lambda_2\sigma_{\lambda_1}(Q_0)) + \beta f((1-\lambda_2)Q_0 + \lambda_2\sigma_{\lambda_1}(Q_0)) \\
&\geqslant -Q^{-1}(\sigma_{\lambda_1}(Q_0)) + R((1-\lambda_2)Q_0 + \lambda_2\sigma_{\lambda_1}(Q_0)) + \beta\hat{\pi}_{\lambda_1}((1-\lambda_2)Q_0 + \lambda_2\sigma_{\lambda_1}(Q_0)) \\
&> -Q^{-1}(\sigma_{\lambda_1}(Q_0)) + R((1-\lambda_1)Q_0 + \lambda_1\sigma_{\lambda_1}(Q_0)) + \beta\hat{\pi}_{\lambda_1}((1-\lambda_1)Q_0 + \lambda_1\sigma_{\lambda_1}(Q_0)) \\
&= \hat{\pi}_{\lambda_1}(Q_0),
\end{aligned}$$

where the first inequality follows from substituting the optimal q with $\sigma_{\lambda_1}(Q_0)$, the second inequality follows from the assumption over f, and the third inequality follows from the monotonicity of $\hat{\pi}_\lambda$ (w.r.t. λ) and of R. Since the set of functions f sustaining the required condition is closed, and by the Contraction Mapping Theorem, the result follows.

Part III. Similarly to Part II, fix a function $f \in B$ such that $f(q) \geqslant \hat{\pi}_{\lambda_2}(q)$ for every $q < Q^*_{\lambda_2}$. By Theorem 3, we know that $Q_0 < Q^*_{\lambda_2}$ implies $\sigma_{\lambda_2}(Q_0) > Q^*_{\lambda_2}$, where σ_{λ_2} is the optimal stationary strategy in the λ_2-valuation problem, such that the position in the next stage tends towards $Q^*_{\lambda_2}$ from below. Hence,

$$\begin{aligned}
(T_{\lambda_1} f)(Q_0) &\geqslant -Q^{-1}\left(\frac{\lambda_1-\lambda_2}{\lambda_1}Q_0 + \frac{\lambda_2}{\lambda_1}\sigma_{\lambda_2}(Q_0)\right) + R((1-\lambda_2)Q_0 + \lambda_2\sigma_{\lambda_2}(Q_0)) \\
&+ \beta f((1-\lambda_2)Q_0 + \lambda_2\sigma_{\lambda_2}(Q_0)) \\
&\geqslant -Q^{-1}\left(\frac{\lambda_1-\lambda_2}{\lambda_1}Q_0 + \frac{\lambda_2}{\lambda_1}\sigma_{\lambda_2}(Q_0)\right) + R((1-\lambda_2)Q_0 + \lambda_2\sigma_{\lambda_2}(Q_0)) \\
&+ \beta\hat{\pi}_{\lambda_2}((1-\lambda_2)Q_0 + \lambda_2\sigma_{\lambda_2}(Q_0)) \\
&> -Q^{-1}(\sigma_{\lambda_2}(Q_0)) + R((1-\lambda_2)Q_0 + \lambda_2\sigma_{\lambda_2}(Q_0)) + \beta\hat{\pi}_{\lambda_2}((1-\lambda_2)Q_0 + \lambda_2\sigma_{\lambda_2}(Q_0)) \\
&= \hat{\pi}_{\lambda_2}(Q_0),
\end{aligned}$$

where the first inequality follows from substituting the optimal q with $\frac{\lambda_1-\lambda_2}{\lambda_1}Q_0 + \frac{\lambda_2}{\lambda_1}\sigma_{\lambda_2}(Q_0)$, the second inequality follows from the assumption over f, and the third inequality follows from the monotonicity of Q^{-1}. Since the set of functions f sustaining the required condition is closed, and by the Contraction Mapping Theorem, the result follows.

Part IV. Assume that $Q^*_{\lambda_1} < Q^*_{\lambda_2}$ for $0 < \lambda_2 < \lambda_1 < 1$ (where the result for the end points is trivial). Take $q \in (Q^*_{\lambda_1}, Q^*_{\lambda_2})$. According to Parts II and III, we get $\hat{\pi}_{\lambda_2}(q) > \hat{\pi}_{\lambda_1}(q) > \hat{\pi}_{\lambda_2}(q)$. A contradiction. Thus, $Q^*_{\lambda_1} \geqslant Q^*_{\lambda_2}$, for $\lambda_1 > \lambda_2$.

Now assume that $Q^*_{\lambda_1} = Q^*_{\lambda_2}$ for $\lambda_2 < \lambda_1$. We can take the FOC of the RHS of the stated Bellman equation (similarly to Theorem 3), along with the derivative of $\hat{\pi}_{\lambda_1}(Q_0)$ to get the two equations,

$$\lambda \left[R'((1-\lambda)Q_0 + \lambda\sigma(Q_0)) + \beta\hat{\pi}'((1-\lambda)Q_0 + \lambda\sigma(Q_0))\right] = \left(Q^{-1}\right)'(\sigma(Q_0))$$

and

$$\hat{\pi}'(Q_0) = (1-\lambda)R'((1-\lambda)Q_0 + \lambda\sigma(Q_0)),$$

where the second equality follows from the envelope theorem. Taking $\lambda = \lambda_1$, $Q_0 = Q^*_{\lambda_1}$, and plugging the second equation into the first yields

$$\lambda_1\left[1 + \beta(1-\lambda_1)\right] = \frac{\left(Q^{-1}\right)'\left(Q^*_{\lambda_1}\right)}{R'\left(Q^*_{\lambda_1}\right)}.$$

Since $\beta \in (0,1)$, the LHS is an increasing function of λ_1, subject to $0 \leqslant \lambda_1 \leqslant 1$. Thus, $Q^*_{\lambda_1} = Q^*_{\lambda_2}$ contradicts the last equality, implying $Q^*_{\lambda_1} > Q^*_{\lambda_2}$, as needed.

Part V. We only prove the relevant case of Part II where $Q_0 = Q^*_{\lambda_1}$, while a similar proof holds for $Q_0 = Q^*_{\lambda_2}$ of Part III. Consider $\lambda_1 > \lambda_2$ and fix $\lambda_3 \in (\lambda_2, \lambda_1)$. According to Part IV, $Q^*_{\lambda_1} > Q^*_{\lambda_3} > Q^*_{\lambda_2}$. Hence by Part II, $\hat{\pi}_{\lambda_2}(Q^*_{\lambda_1}) > \hat{\pi}_{\lambda_3}(Q^*_{\lambda_1})$, whereas by Part I, $\hat{\pi}_{\lambda_3}(Q^*_{\lambda_1}) \geqslant \hat{\pi}_{\lambda_1}(Q^*_{\lambda_1})$, which concludes Part V.

Next, we prove the last statement of the theorem regarding the derivatives. If $Q^*_{\lambda_2} \leqslant Q \leqslant Q^*_{\lambda_1}$, then $\sigma_{\lambda_2}(Q) < Q < \sigma_{\lambda_1}(Q)$, and $(1-\lambda)Q + \lambda\sigma_\lambda(Q)$ increase with λ. If $Q < Q^*_{\lambda_2}$, we consider two cases where either $\sigma_{\lambda_1}(Q) \leqslant \sigma_{\lambda_2}(Q)$ or $\sigma_{\lambda_1}(Q) > \sigma_{\lambda_2}(Q)$. Assume that $\sigma_{\lambda_1}(Q) \leqslant \sigma_{\lambda_2}(Q)$. Thus, $\lambda Q'(Q^{-1}(\sigma_\lambda(Q)))$ increase w.r.t. λ, and by Eq. 10 along with the concavity of R and $\hat{\pi}$, it follows that $(1-\lambda)Q + \lambda\sigma_\lambda(Q)$ increase in λ. Otherwise, $\sigma_{\lambda_1}(Q) > \sigma_{\lambda_2}(Q) > Q$ and, again, we get the same monotonicity of $(1-\lambda)(Q) + \lambda\sigma_\lambda(Q)$ w.r.t. λ. By the previously-stated equation $\hat{\pi}'_\lambda(Q) = (1-\lambda)R'((1-\lambda)Q + \lambda\sigma(Q))$, along with the concavity of R, it follows that $\hat{\pi}'_{\lambda_2}(Q) > \hat{\pi}'_{\lambda_1}(Q)$, as stated.

■

E The stochastic persistent extension

The stochastic extension of the persistent model is a straightforward one. Following the same notation and analysis of Section 2.2, we assume that the output function depends on the effort level e and a state $\omega \in \Omega$, chosen according to the transition matrix P. Recall that S is the convex hull of the compact set $Q(\Omega, E)$ of all possible realized outputs. At every stage t, the DM exerts effort e_t given the realized state ω_{t-1} and position $\widehat{Q_{t-1}}$, where $\widehat{Q_t} = (1-\lambda)\widehat{Q_{t-1}} + \lambda Q(\omega_t, e_t)$ and $\widehat{Q_0} = Q_0$. Hence, for every strategy σ the DM's expected β-discounted payoff is

$$\pi(\omega_0, Q_0|\sigma) = \mathbf{E}_{\sigma,\omega_0}\left[\sum_{t=1}^{\infty} \beta^{t-1}(-e_t + R(\widehat{Q_t})\right].$$

The following theorem extends Theorem 3 to the stochastic set-up. The results of Theorem 3 hold, in expectation, similarly to the transient-stochastic extension. Namely, there exist a strictly-concave, continuously-increasing payoff function $\hat{\pi}(\omega_0, Q_0)$, and a unique, stationary, and continuous optimal-strategy σ, while the subsequent expected-production decreases as a function of past production.

Theorem 5. *There exists a unique, stationary, and continuous optimal-strategy $\sigma : \Omega \times S \to E$. Given σ and ω_0, the payoff function $\hat{\pi}(\omega_0, Q_0) = \pi(\omega_0, Q_0|\sigma)$ is a strictly-concave, continuously-increasing function of Q_0. In addition, if the interior-solution property holds, then $\boldsymbol{E}_{\sigma,\omega_{t-1}}[Q_t|\widehat{Q_{t-1}}]$ is a strictly-decreasing function of $\widehat{Q_{t-1}}$.*

Proof. The respective Bellman equation is

$$\pi(\omega_0, Q_0) = \sup_{e \in E} \mathbf{E}_{\omega_0}\left[-e + R((1-\lambda)Q_0 + \lambda Q(\tilde{\omega}, e)) + \beta\pi(\tilde{\omega}, (1-\lambda)Q_0 + \lambda Q(\tilde{\omega}, e))\right],$$

where the expectation relates to the random variable $\tilde{\omega}$, drawn according to P and ω_0. By the same reasoning as in the proof of Theorem 2, we follow Chapter 9 of Stokey et al. (1989), specifically using Assumptions 9.4 – 9.12 and 9.16 – 9.17, along with Exercise 9.7. The assumptions and statements hold as in the stochastic transient model, with the exception of Exercise 9.7-f that does not hold. Specifically, the condition of 9.7-f requires the position in the subsequent stage to be independent of the evaluation in the previous one. However, in the stochastic persistent model the dependence is straightforward by $(1-\lambda)Q_0 + \lambda Q(\tilde{\omega}, e)$. To sum-up, results 9.7-a through 9.7-d ensure the existence of a continuously-increasing, strictly-concave payoff function $\hat{\pi}(\omega_0, Q_0)$, all w.r.t. Q_0, and ensure the existence of a unique, stationary, and continuous optimal-strategy $\sigma : \Omega \times S \to E$. Thus, we can reformulate the previous Bellman equation as

$$\hat{\pi}(\omega_0, Q_0) = -e + \mathbf{E}_{\omega_0}\left[M((1-\lambda)Q_0 + \lambda Q(\tilde{\omega}, e)) + \beta\hat{\pi}(\tilde{\omega}, (1-\lambda)Q_0 + \lambda Q(\tilde{\omega}, e))\right],$$

where $e = \sigma(\omega_0, Q_0)$.

We wish to prove that e decreases as a function of Q_0, while ω_0 is fixed. Following a proof by contradiction, assume there exist $Q_0' > Q_0$ such that $e' = \sigma(\omega_0, Q_0') > \sigma(\omega_0, Q_0) = e$. We show that a deviation from e to e' is profitable given (ω_0, Q_0), contradicting the optimality of e. To simplify the notation, define the functions G by

$$G(x, y) = \mathbf{E}_{\omega_0}\left[R((1-\lambda)x + \lambda Q(\tilde{\omega}, y)) + \beta\hat{\pi}(\tilde{\omega}, (1-\lambda)x + \lambda Q(\tilde{\omega}, y))\right].$$

Note that G is concave as the sum of two concave functions. Thus,

$$G(Q_0', e) - G(Q_0, e) > G(Q_0', e') - G(Q_0, e').$$

In words, the concavity of G implies that an upwards deviation of the x-variable from Q_0 to Q_0' is more profitable in case $y = e$, rather than $y = e' > e$. Therefore,

$$G(Q_0, e') - G(Q_0, e) > G(Q_0', e') - G(Q_0', e) > 0,$$

where the second inequality follows from the optimality of $e' = \sigma(\omega_0, Q_0')$. However, the last inequality suggests that $G(Q_0, e') - G(Q_0, e) > 0$, contradicting the optimality of $e = \sigma(\omega_0, Q_0)$ and concluding the proof. ■