A Theory of Self-Prospection*

Polina Borisova† Nikhil Vellodi‡

[Current Draft]: May 15, 2024

Abstract

A present-biased decision maker (DM) faces a two-armed bandit problem whose risky arm generates random payoffs at exponentially distributed times. The DM cannot perfectly observe payoffs but receives informative feedback. Our main finding is that, in the unique stationary Markov perfect equilibrium of the multi-self game, positive feedback supports greater equilibrium welfare than both negative and transparent feedback. It does so by encouraging the DM to *self-prospect* — imagine one's future goals and outcomes when evaluating the present. We relate our results to findings in psychology promoting the motivational effects of positive feedback, as well as more recent findings regarding self-prospection theory.

JEL Classification: D83, D91, C73

Keywords: Present-bias, strategic experimentation, motivational feedback.

1 Introduction

'The only limit to our realization of tomorrow will be our doubts of today.'

(Franklin D. Roosevelt)

*We would like to thank Abhijit Banerjee, Olivier Compte, Johannes Hörner, Navin Kartik and Sven Rady for their thoughtful comments, and especially Glenn Ellison for incredibly insightful conversations at an early stage in the project. This work has benefited from the support of the Agence Nationale de la Recherche through the program Investissements d'Avenir ANR-17-EURE-0001. Vellodi acknowledges Columbia Graduate School of Business and Economics faculty, where part of this research was performed.

†Paris School of Economics, Université Paris 1 Panthéon Sorbonne (polina.borisova@psemail.eu).

‡Paris School of Economics (nikhil.vellodi@psemail.eu).

How does feedback motivate a person to experiment when they lack willpower? Consider a person who can choose to engage in a risky activity, such as taking up a new hobby or employment in a skilled sector. Not only might this process of self-discovery take time and patience, but the person may also struggle with limited willpower. For instance, when learning an instrument or language or training for a marathon, people might give up too quickly if they doubt their talent and over-weigh their short-run struggle over their long-run reward. Researchers might give up on great ideas too soon if their impatience to see early results gets the better of them. Such behavior is ubiquitous and empirically well-established.[1]

Common to these scenarios is the arrival over time of informative feedback that helps the DM evaluate the risks involved. For instance, students receive feedback from their teachers, and researchers from their peers. More generally, the market for personal coaches in various aspects of life, ranging from physical trainers to professional and life coaches, is built around the provision of feedback, oftentimes specifically to compensate for a lack of willpower, and has seen enormous growth over the past decade.[2]

While limited willpower is well-known to generate a demand for "commitment devices" that help the DM overcome their lack of self-control (Ashraf et al., 2006; Bryan et al., 2010), the feedback received by the DM might also perform a similar role, working as a motivational tool to help bolster their willpower (Bénabou and Tirole, 2003). Contrasting the effects of positive and negative feedback on personal motivation, especially in the absence of personal willpower, is of great practical relevance in the market for personal coaching. It is also a topic of central interest in behavioral psychology (Hattie and Timperley, 2007; Fong et al., 2018), where positive feedback is often cited as providing motivation through positive self-affirmation. We contribute to this ongoing and important debate by developing a theoretical model that allows us to formalize precisely how positive feedback can enhance motivation.

[1] See O'Donoghue and Rabin (2001) for a survey.

[2] The International Coaching Federation's (ICF) 2023 Global Coaching Study reported that worldwide growth of professional coaching services was around 54 percent over 2019-2022. See https://tinyurl.com/mphdb7sy. The U.S. life coaching market is set to grow by almost 5 percent over 2023-2030. See https://tinyurl.com/yer43yv8.

In our infinite horizon, continuous time model, a single decision-maker (DM) faces a two-armed bandit problem, comprising a safe arm (S) and a risky arm (R). S generates a deterministic flow payoff, while R generates lump-sum rewards, taking two possible values $h_1 > h_0$, that arrive at exponentially distributed times (Keller et al., 2005; Keller and Rady, 2015). The size of these rewards is determined by a hidden state $\theta \in \{0, 1\}$, representing the underlying trait the DM wishes to learn about. However, the DM cannot perfectly observe rewards; they receive information through a *feedback structure* which hides certain reward arrivals from the DM. Crucially, the DM has *time-inconsistent preferences*, which we model through the continuous-time pseudo-exponential specification introduced in Harris and Laibson (2013). (See also Ekeland and Lazrak (2006); Karp (2007); Tan et al. (2021).) This specification is flexible and in particular nests exponential and quasi-hyperbolic preferences as special cases.

To capture limited willpower, we focus on the *sophisticated* problem, wherein the DM is modeled as a sequence of different selves, and solve for subgame perfect equilibria of the resultant multi-self game (Strotz, 1955; Phelps and Pollak, 1968). We further restrict our attention to stationary Markov perfect equilibria (Bernheim and Ray, 1987; Harris and Laibson, 2013, 2001), where the state variable is the DM's (self-)belief regarding θ. In this way, we are then able to determine to what extent the dynamics of feedback can substitute for the DM's lack of willpower by shaping their beliefs.

Our main finding is that in the unique equilibrium, positive feedback provides greater welfare than either negative or transparent feedback. Specifically, positive feedback does just as well as negative and transparent feedback for any prior belief, while there exist priors at which it strictly dominates both. The result stems from how positive feedback enables the DM to *self-prospect*. Self-prospection is the process of imagining one's future events, goals and life outcomes when evaluating the present.[3] The defining property of positive feedback — that it induces a positive trend in self-beliefs while the DM is actively learning — allows the DM to consider a future in which they continue to experiment as their beliefs

[3] See https://tinyurl.com/4b8vu5yh for an overview.

improve. In contrast, negative feedback generates a downward spiral of beliefs, and thus at the point of quitting, the DM is entirely consumed by the outcome of experimentation over the next instant. Simply put, positive (negative) feedback causes beliefs to drift away (into) from the quitting region. We develop this intuition formally in Section 3.

Our connection to self-prospection goes further still. Stephan and Sedikides (2023) argue how self-prospection can help motivate people through positive self-affirmation via three distinct channels: *self-esteem* – the extent to which a person values themselves –, *coherence* – the ability to create consistent goals and narratives over time – and *self-control* – the ability to take actions in the present that serve long-term goals. In Sections 5.3-5.6, we decompose our main characterization (Theorem 1) into these separate channels, showing how positive feedback enhances the DM's ability to self-prospect through each of these channels, subsequently motivating them to experiment more in line with their long-term goals. Thus, we provide a clear mechanism through which positive feedback enhances motivation; it promotes self-affirmation by allowing the DM to self-prospect.

Beyond this direct contribution, the tractability of our model allows us to derive further results. For instance, we show how regardless of the form of feedback, present-bias causes the DM to exhibit *indecision*; they split their time between experimenting and not for a positive time duration. Indecision is a result of the DM's desire to *procrastinate*; knowing that their future selves will bear the cost of experimenting, the current self would rather play safe and avoid such costs themselves. This behavior cannot occur under exponential discounting, and thus directly links lack of willpower to procrastination and indecision. This link is well-documented empirically (Ferrari et al., 1995; Ferrari and Pychyl, 2007). We also show how, as internal conflict disappears, the form of feedback becomes irrelevant (they each induce the same behavior), whereas when it is severe, positive feedback is optimal. This finding is consistent with evidence in the "goal orientation" literature (Dahling and Ruppel, 2016).

Related Literature and Contribution – Our work contributes to a nascent literature studying the effects of information on incentives under present-bias. Carrillo and

Mariotti (2000); Brocas and Carrillo (2000) study models of "strategic ignorance", whereby a present-bias DM is better off avoiding freely-available information, in order to manipulate self-beliefs and motivate further experimentation. Auster et al. (forthcoming) study an optimal-stopping model where time-inconsistency derives from model uncertainty. As in our Theorem 1, they find a role for randomized stopping, in contrast to the analogous time-consistent setting, while again abstracting from a comparison of feedback structures. In Ali (2011), the DM learns about their degree of self-control, rather than a parameter relevant to flow-payoffs directly. Bénabou and Tirole (2002, 2003) are most closely related, asking whether the DM can benefit from selectively avoiding or forgetting certain information. Their three-period models are however not long enough to nest certain phenomena captured in our infinite horizon setting, for instance, our "self-control" effect (Lemma 2), while none of these papers ask how the nature of the learning process itself can shape incentives.

By casting our model as a sequential game, we offer a new model of strategic experimentation (Bolton and Harris, 1999; Keller et al., 2005; Keller and Rady, 2015). We highlight several conceptual and technical parallels to these previous works, documented in Section 7.3. Despite the fact that ours is a sequential game with direct payoff externalities — both distinct from previous works — we leverage these parallels by adapting existing proof techniques to our setting when solving for equilibrium. Finally, our work shares similarities to work studying "incentivized exploration" wherein information is distorted by a planner facing a sequence of myopic agents (Che and Hörner, 2018; Kremer et al., 2014). Our "self-control" (Lemma 2) sees the DM act explicitly in order to drive beliefs up and induce further experimentation, knowing that they will be too impatient in the future.

The paper is structured as follows. We begin by studying the efficient benchmark in Section 3, turning to an alternative, "single-player" benchmark in Section 4. These initial insights begin to demonstrate the power of positive feedback in motivating the DM through enhanced "self-prospection". In Section 5 and 6, we offer the main findings of the paper by fully characterizing the sophisticated solution of the model under positive, negative and transparent feedback. We offer a brief discussion of our model in Section 7, and offer avenues for further research in Section 8.

2 Model

A decision-maker (DM) faces a two-armed bandit problem, comprising a safe (S) and risky (R) arm, over a continuous, infinite time-horizon $[0, \infty)$. If played, S yields a constant flow payoff of s, while R generates payoffs that depend on an unknown state of the world $\theta \in \{0, 1\}$. Specifically, lump-sum payoffs h_θ arrive in state θ according to a time-homogeneous Poisson process with parameter $\lambda \in (0, \infty)$. The DM holds a prior belief $p_0 \in (0, 1)$ that $\theta = 1$. We assume that expected flow payoffs are greater under R than S if and only if $\theta = 1$:

Assumption 1.

$$g \equiv \lambda h_1 > s > \lambda h_0 \equiv f.$$

The DM has a unit resource that can be perfectly split between the two arms at any instant. Let $\alpha_t \in [0, 1]$ denote the share of resource devoted to R at time $t \in [0, \infty)$.

Feedback, beliefs – The DM imperfectly observes lump-sum payoff arrivals when playing R, and receives information via a *feedback structure*. Formally, let $\mathbb{G} \equiv (\mathcal{G}_t)_{t \geqslant 0}$ denote the filtration generated by the payoff arrival process. A *feedback structure* $\varphi = (\varphi_1, \varphi_0)$ is a pair of $\mathbb{G}-$adapted $[0, 1]$-valued processes φ_1, φ_0 that denote the probability a lump-sum payoff is observed by the DM in state 1, 0 respectively, given that it arrives at time t. We will be entirely concerned with the following three feedback structures:

- Transparent feedback: $\varphi_{1,t} = \varphi_{0,t} = 1$ for all $t \in [0, \infty)$. All payoff arrivals are observed.

- Negative feedback: $\varphi_{1,t} = 1$, $\varphi_{0,t} = 0$ for all $t \in [0, \infty)$. Payoff arrivals are observed only when $\theta = 1$. Thus, in the absence of an arrival, the DM conflates no arrival with payoffs that arrive when $\theta = 0$.

- Positive feedback: $\varphi_{1,t} = 0$, $\varphi_{0,t} = 1$ for all $t \in [0, \infty)$. Payoff arrivals are observed only when $\theta = 0$.

We often refer to the arrival of payoff g as a "breakthrough" or "good news" and payoff f as a "breakdown" or "bad news". To this end, where appropriate, we use the subscript g for negative feedback (as good news is observed), b for positive feedback (bad news is observed) and f for transparent feedback (full news is observed).

It may seem somewhat perverse that we refer to a feedback structure that conceals breakthroughs as "positive". To understand why, let p_t denote the DM's subjective probability at time t that $\theta = 1$, based on all observable information up to that date.[4] Given a feedback structure, the evolution of beliefs follows a well-established law of motion: after observing the payoff h_θ, the DM's belief p_t jumps to θ, while absent observation of any payoff, the belief p_t evolves continuously according to

$$\dot{p}_t = \lambda \alpha_t (\varphi_{0,t} - \varphi_{1,t}) p_t (1 - p_t). \tag{1}$$

This equation reveals our choice of terminology. During the "active" phase of play, — when the DM is still uncertain regarding θ so that $p_t \neq 0, 1$ and $\alpha_t > 0$ — positive (negative) feedback induces a positive (negative) drift in beliefs. We discuss this interpretation further in Section 7.2.

Discount Function – The DM evaluates the present value of future payoffs according to a pseudo-exponential discounting function (Harris and Laibson, 2013; Karp, 2007; Ekeland and Lazrak, 2006). Specifically, for a given (measurable) sequence of actions $(\alpha_t)_{t \geq 0}$, the DM's time-0 expected present-discounted average value is

$$\mathbb{E}\left[\int_0^\infty \gamma \bar{d}(t) u(\alpha_t, p_t)\, dt\right], \tag{2}$$

[4] Formally, let $\mathbb{F} \equiv (\mathcal{F})_{t \geq 0}$ be the filtration generated by the observed payoff process. Then $p_t = \mathbb{E}(\theta \mid \mathcal{F}_t)$.

where $u(\alpha_t, p_t) \equiv (1-\alpha_t)s + \alpha_t(gp_t + f(1-p_t))$ is the expected flow payoff and

$$\bar{d}(t) \equiv e^{-(\eta+\gamma)t} + (1-e^{-\eta t})\beta e^{-\gamma t}. \tag{3}$$

for $\eta, \gamma \geq 0$ and $\beta \in [0,1]$. For simplicity, we shall refer to this form of discounting as (β, η)-*discounting.*

As discussed in Harris and Laibson (2013), (β, η)-discounting admits a natural, stochastic interpretation. The DM is comprised of multiple *selves* whose arrival times are random and follow a time-homogeneous Poisson process with parameter η. Each self scales down the payoffs generated after the next self arrives by a fixed factor β. Since arrivals are governed by a Poisson point process, we may index selves by $n \in \mathbb{N}$, with related arrival times $s_n \in [0,\infty)$ and inter-arrival times $\tau_n \equiv s_{n+1} - s_n \in [0,\infty)$. We say that the "future" arrives when the next self turns up. Formally, self n applies the discount factor $d_n(t)$ to the flow utility at time $s_n + t$, where

$$d_n(t) = \begin{cases} e^{-\gamma t}, & t \in [0, \tau_n) \\ \beta e^{-\gamma t}, & t \in [\tau_n, \infty). \end{cases} \tag{4}$$

We adopt this approach throughout our paper. To this end, fix a (measurable) strategy $\hat{\alpha} : [0,\infty) \to [0,1]$ and a feedback technology (φ_1, φ_0), and consider self n arriving at time s_n. We define the *continuation value* for self n as

$$v(p_{s_n+\tau_n}, \hat{\alpha}) = \mathbb{E}_{s_n+\tau_n}\left[\int_{s_n+\tau_n}^{\infty} \gamma e^{-\gamma(t-(s_n+\tau_n))} u(\hat{\alpha}_t, p_t)\, dt\right]. \tag{5}$$

Suppose that self n employs the strategy $\alpha : [0,\infty) \to [0,1]$. Then we define the *current value* for self n as

$$w(\alpha; p_{s_n}, \hat{\alpha}) = \mathbb{E}_{s_n}\left[\int_{s_n}^{s_n+\tau_n} \gamma e^{-\gamma(t-s_n)} u(\alpha_t, p_t)\, dt + \beta e^{-\gamma\tau_n} v(p_{s_n+\tau_n}, \hat{\alpha})\right] \tag{6}$$

as the combined value self n gets from payoffs accruing both before and after the future arrives.

Sophistication – As is typical in models with non-exponential discounting, there are various notions of optimal behavior within our framework. Our main results pertain to so-called *sophisticated* policies that treat each self as a separate player in a sequential, *intra-personal* game, and solve for its subgame perfect equilibria. This concept thus captures limited self-control; the DM perfectly understands their present-bias, but is unable to commit to actions that ameliorate it.

We focus on *stationary Markov perfect equilibria* (SMPE), wherein each self plays a best response to the strategies of future selves and use the same Markov strategy $\alpha : [0,1] \to [0,1]$ that depends only on the current belief $p \in [0,1]$.

Definition 1. For a Markov strategy $\tilde{\alpha} : [0,1] \to [0,1]$, let $\mathcal{B}(\tilde{\alpha})$ denote the set of Markov strategies α' such that $\alpha' \in \arg\max_{\alpha''} w(\alpha''; p, \tilde{\alpha})$ for all $p \in [0,1]$. A SMPE of the intra-personal game is then a function $\alpha : [0,1] \to [0,1]$ such that $\alpha \in \mathcal{B}(\alpha)$.

Definition 2. A strategy $\alpha : [0,1] \to [0,1]$ is a *threshold* strategy if $\alpha(p) = \mathbb{I}_{p \geq \underline{p}}$ for some $\underline{p} \in [0,1]$.

Naivité – In contrast, the *naive* solution to the problem facing the DM at time t is to best-respond to the belief that their future selves do not face the same self-control problem; specifically, they assume their future selves have $\beta = 1$ and thus will adopt the time-consistent policy α^*, where α^* maximizes (6) when $\beta = 1$.[5]

Definition 3. The *naive* solution of the intra-personal game is a function $\alpha^n : [0,1] \to [0,1]$ such that for all $p \in [0,1]$, $\alpha^n \in \arg\max_{\alpha'} w(\alpha'; p, \alpha^*)$.

3 Efficient (Exponential) Benchmark

To begin, we first characterize optimal experimentation in the exponential benchmark, as first introduced in Presman (1990). This is a limiting case of our model; set either $\beta = 1$

[5] In this case, the restriction to Markov policies is without loss, as a best-response to Markov strategies is necessarily Markovian.

or $\eta = 0$.[6]

Proposition 1. *With exponential discounting, the optimal policies are* $\alpha_b^*(p) = \alpha_f^*(p) = \mathbb{I}_{p \geq p^*}$ *and* $\alpha_g^*(p) = \mathbb{I}_{p > p^*}$*, where*

$$p^* = \frac{\gamma(s-f)}{\gamma(g-f) + \lambda(g-s)}. \tag{7}$$

The DM exclusively plays R so long as their belief is above p^*.[7] Once their belief drops below this level, they switch to exclusively playing S. The term $\lambda(g-s)$ captures the option value of learning; were $\varphi_0 = \varphi_1 = 0$, the option value effect disappears, and the DM switches to safe at the *myopic threshold*

$$p^m \equiv \frac{s-f}{g-f}.$$

3.1 Myopic versus Prospective Option Value

While Proposition 1 shows that the various feedback structures lead to identical policies for the DM, the result belies a fundamental difference in how these structures provide option value to the DM.

Under negative feedback, beliefs drift downward absent news according to $\dot{p}_t = -\lambda p_t(1-p_t)$. As such, the comparison facing the DM at time t is between playing S over $[t, \infty)$ and playing R over $[t, t+dt)$ and then S on $[t+dt, \infty)$ if no news arrives. (If a breakthrough occurs over $[t, t+dt)$, the DM clearly plays R thereafter.) Setting the value of these two policies to be equal generates the first-order condition satisfied at the threshold belief p^*:

$$(gp + f(1-p))\, dt + \Omega_g^*(\lambda p\, dt) = s\, dt, \tag{8}$$

where $\Omega_g^* = \int_0^\infty e^{-\gamma t}(g-s)\, dt = (g-s)/\gamma$. The term $\Omega_g^*(\lambda p\, dt)$ embodies what we call *myopic option value*, as it measures the net present value of information that arrives only

[6] While the result is stated as an optimal policy, the policy is of course a SMPE of the (time-consistent) intra-personal game when $\beta = 1$ but $\eta > 0$.

[7] The difference in the policies at p^* is due to the direction of belief drift.

within the next instant.

Under positive feedback, beliefs drift upward absent news according to $\dot{p}_t = \lambda p_t(1-p_t)$, and so the relevant comparison is between playing R until they observe a breakdown and playing S forever. The threshold belief p^* now satisfies

$$gp + f(1-p) + \Omega_b^*(g-s)p = s, \tag{9}$$

where

$$\Omega_b^* = \frac{\Lambda}{G-\Lambda} = \frac{\int_0^\infty (1-e^{-\lambda t})\gamma e^{-\gamma t}\,dt}{\int_0^\infty e^{-\lambda t}\gamma e^{-\gamma t}\,dt}.$$

The quantity $\frac{\Lambda}{G-\Lambda}$ measures the value of information under positive feedback. Crucially, unlike with negative feedback, where the option value term relates to news arriving over the next dt, positive feedback embodies *prospective option value*, as the term $\Omega_b^*(g-s)p$ measures the net present discounted value of news arrivals *over the entire future horizon*. Here then, we begin to see how positive feedback naturally induces the DM to *self-prospect*, by forcing them to consider outcomes occurring in the distant future when making choices in the present.

Under general discounting functions, the two quantities $\Omega_g^*(\lambda p\,dt)$ and $\Omega_b^*(g-s)p$ need not agree. However, under exponential discounting, they do, so that myopic and prospective option value are equivalent, and so positive and negative feedback entail the same informational value. This is peculiar to the fact that in this case, both discounting and news arrivals are stationary, and thus from today's perspective, the relative discounted likelihood of news arrivals over all future time, $\Lambda/(G-\Lambda)$, is identical to the analogous quantity defined over the next dt:

$$\frac{\Lambda}{G-\Lambda} = \frac{\lambda/(\gamma(\gamma+\lambda))}{1/(\gamma+\lambda)} = \frac{\lambda}{\gamma} = \frac{\lambda\,dt}{\gamma\,dt}.$$

Finally, note that under this benchmark, positive and transparent feedback not only deliver the same optimal switching time, but also induce identical optimization problems. To see this, fix a switching threshold p^*. If the current belief $p_t \geq p^*$, then both positive

Figure 1: Myopic and Prospective Option Value

Left panel: under negative feedback, information arriving over the next instant is relevant. Right panel: under positive feedback, information arriving over the entire future horizon is relevant.

and transparent feedback will dictate to play R until a breakdown is observed, whereas if $p_t < p^*$, they both dictate to play S forever. This equivalence, and the distinction with negative feedback, is best seen by computing the optimal values v_i^* for $i \in \{b, g, f\}$ that evaluate (5) at the optimal policy (note that v and w are equivalent in this setting) under each feedback structure.

Lemma 1. *For each $i \in \{b, g, f\}$ and $p \in [0, 1]$, let $v_i^*(p)$ denote (5) evaluated at α^* and $p_{s_n+\tau_n} = p$. Then*

1. *For all $p \in (p^*, 1)$, $v_b^*(p) = v_f^*(p) > v_g^*(p)$.*
2. *For all $p \in [0, p^*] \cup \{1\}$, $v_b^*(p) = v_f^*(p) = v_g^*(p)$.*

Thus, the value of experimenting coincides *precisely* at the cut-off belief p^*, but at all $p \in (p^*, 1)$, positive and transparent feedback provide higher value. This result gives a sense in which prospective option value might provide greater incentives for experimentation in general, an intuition that we will see play out in the remaining analysis.

4 Single-Player Problem

Before turning to the fully sophisticated solution of the (β, η)-model, we first introduce and characterize a useful *single-player benchmark*, wherein the DM is allowed to jointly maximize the strategies of each of their selves simultaneously, and thus relaxes the best-response requirement in Definition 1. They nevertheless use their own discounted value as their objective, and are still restricted to employ Markovian strategies.

Definition 4. The *single-player* solution of the intra-personal game is a function $\alpha : [0,1] \to [0,1]$ such that for all $p \in [0,1]$, $\alpha \in \arg\max_{\alpha'} w(\alpha'; p, \alpha')$.

This benchmark serves as a useful link between the exponential baseline and the equilibrium analysis that follows. In particular, the single-player solution is characterized by first-order conditions that are directly analogous to those in the exponential baseline.[8] Furthermore, by abstracting from strategic considerations, it will uncover any intrinsic differences that may exist between the various feedback technologies.

Proposition 2. *Let*

$$\Omega \equiv \frac{\gamma + \eta\beta}{\gamma + \eta}.$$

Under negative feedback, the single-player solution is $\alpha_g(p) = \mathbb{I}_{p > p_g^1}$*, where*

$$p_g^1 = \frac{\gamma(s-f)}{\gamma(g-f) + \lambda\Omega(g-s)}. \tag{10}$$

Notice that the threshold p_g^1 is identical in form to the exponential threshold p^*, with the additional factor Ω scaling the option value term. Indeed, since beliefs again drift down absent news, p_g^1 satisfies exactly the same first-order condition as in the exponential setting, but with (β, η)-discounting:

$$(gp + f(1-p))\,dt + \Omega\frac{g-s}{\gamma}(\lambda p\,dt) = s\,dt, \tag{11}$$

where

$$\Omega = \int_0^\infty \bar{d}(t)\,dt = \int_0^\infty e^{-(\eta+\gamma)t} + (1-e^{-\eta t})\beta e^{-\gamma t}\,dt = \frac{\gamma+\eta\beta}{\gamma+\eta}. \tag{12}$$

[8] Under exponential discounting, the Markovian restriction is without loss. This is not necessarily the case here, and thus it is an open question as to what the open-loop optimal control policy is when the restriction is relaxed.

Now the net present value of this instantaneous experiment is

$$\int_0^\infty \bar{d}(t)(g-s)\,dt = \Omega(g-s).$$

The term Ω exhibits various intuitive properties. First, it is strictly less than one, since the present-discounted value of a breakthrough occurring in the next instant is ameliorated by the factor $\beta < 1$. This, in turn, implies that $p_g^1 > p^*$, so that the DM stops earlier under (β, η)-discounting. Furthermore, $\lim_{\eta \to 0} \Omega = \lim_{\beta \to 1} \Omega = 1$, so that the exponential benchmark is recovered in both these time-consistent limits.

We next turn to positive feedback and show that in stark contrast to Proposition 1, the single-player solutions across feedback structures no longer coincide.

Proposition 3. *Let*

$$\Theta \equiv \frac{\gamma + \eta + \lambda}{\gamma + \eta\beta + \lambda}, \qquad \Omega_b^1 \equiv \frac{\gamma + \eta\beta\Theta}{\gamma + \eta}.$$

Under positive feedback, the single-player solution is $\alpha_b(p) = \mathbb{I}_{p \geq p_b^1}$, *where*

$$p_b^1 = \frac{\gamma(s-f)}{\gamma(g-f) + \lambda\Omega_b^1(g-s)}. \tag{13}$$

Here, the option value term has a different scaling parameter Ω_b^1, which itself is identical to Ω but for the constant Θ. It is readily verified that $\Theta > 1$, so that $\Omega < \Omega_b^1$. Thus, we arrive at our first key insight, namely that *under* (β, η)*-preferences, prospective option value is more powerful than myopic option value.*

Corollary 1. $p^* < p_b^1 < p_g^1 < p^m$.

To understand this result, we can again find p_b^1 through the same first-order approach as in Section 3:

$$gp + f(1-p) + \Omega_b^1(g-s)p = s, \tag{14}$$

where

$$\Omega_b^1 = \frac{\int_0^\infty (1-e^{-\lambda t})\bar{d}(t)\,dt}{\int_0^\infty e^{-\lambda t}\bar{d}(t)\,dt}$$
$$= \frac{\int_0^\infty (1-e^{-\lambda t})\left[e^{-(\eta+\gamma)t} + (1-e^{-\eta t})\beta e^{-\gamma t}\right]\,dt}{\int_0^\infty e^{-\lambda t}\left[e^{-(\eta+\gamma)t} + (1-e^{-\eta t})\beta e^{-\gamma t}\right]\,dt} \tag{15}$$
$$= \frac{\gamma + \eta\beta\Theta}{\gamma+\eta}.$$

Under positive feedback, the DM values news arrival even after the future arrives, as is evident from equation (15). However, whereas under exponential discounting, this made no difference, under (β, η)-preferences it is a crucial distinction.

How then does the arrival rate of the future affect both myopic and prospective option value? A first, direct effect is that more of the remaining time horizon is discounted by β. This direct effect dampens the present-value of learning and is present in both myopic and prospective option value. To see this, note that both Ω and Ω_b^1 *holding* Θ *fixed* are decreasing in η. However, a second counter-veiling effect is that, once the future arrives, the DM discounts exponentially thereafter, and thus the present-bias effect disappears. This effect is only present within prospective option value, as it relates to events occurring beyond the next instant. It is this effect that drives $\Omega < \Omega_b^1$ and hence $p_b^1 < p_g^1$.

4.1 Self-Esteem

We term the effect embodied by this additional factor Θ the **self-esteem effect**, inspired by Bénabou and Tirole (2002). Broadly speaking, self-esteem pertains to the evaluation of one's self (Smith and Mackie, 2007, p.107). The concept can be traced back to David Hume's seminal work, wherein he describes the synonymous concept of "self-love" as an important input into personal motivation (Hume, 1751, Appendix 2.1). In particular, he sees self-esteem as a driver of motivation, not a consequence. In this sense, our single-player problem divorces the DM from strategic considerations, so that Corollary 1 suggests that *positive feedback naturally fosters self-esteem*, in a way that negative feedback does not.

Self-esteem is one of the three pillars of *self-prospection* theory, as discussed in the introduction. In the equilibrium analysis that follows, we will see how the other two components — self-control and coherence — are also best supported by positive feedback.

5 Equilibrium Analysis

5.1 Best Responses

Toward a recursive representation of an SMPE, suppose first that all future selves use the strategy $\hat{\alpha}$, then $v(\hat{\alpha}, p) \equiv \hat{v}(p)$ must satisfy the differential equation

$$\hat{v}(p) = s + \hat{\alpha}(p) \cdot [b(p, \hat{v}) - c(p)], \tag{16}$$

where

$$\begin{aligned} b(p, \hat{v}) &= \frac{\lambda p \varphi_1}{\gamma} (\hat{v}(1) - \hat{v}(p)) + \frac{\lambda(1-p)\varphi_0}{\gamma} (\hat{v}(0) - \hat{v}(p)) \\ &\quad + \frac{\lambda(\varphi_0 - \varphi_1)}{\gamma} p(1-p)\hat{v}'(p), \\ c(p) &= s - (gp + f(1-p)), \end{aligned}$$

$v(1) = g$ and $v(0) = s$, while the current value function satisfies[9]

$$w(p) = s + \frac{\eta}{\gamma}[\beta \hat{v}(p) - w(p)] + \max_{\alpha \in [0,1]} \alpha \cdot [b(p, w) - c(p)], \tag{17}$$

$w(1) = \Omega g$ and $w(0) = \Omega s$. From the linearity of equation (17) in α, this requires the

[9]The reliance of w on $\hat{\alpha}$ via $\hat{v}(p)$ is suppressed in notation.

following best-response condition to hold:

$$\alpha(p) \begin{cases} = 1 & \text{if} \quad b(p,w) > c(p) \\ \in [0,1] & \text{if} \quad b(p,w) = c(p) \\ = 0 & \text{if} \quad b(p,w) < c(p), \end{cases} \tag{18}$$

where again, the reliance of w on $\hat{v}$ in (17) demonstrates how the current self's best-response depends on $\hat{\alpha}$. A SMPE comprises then of three functions: the strategy $\alpha : [0,1] \to [0,1]$, the current value function $w : [0,1] \to \mathbb{R}$ and the continuation value function $v : [0,1] \to \mathbb{R}$ such that w solves (17), v solves (16) and α is a maximizer of (17) for each $p \in [0,1]$ (or equivalently, satisfies (18)).

The final required ingredient is the necessary boundary conditions for the above differential equations. The only required conditions are that both w and v be bounded; it is readily seen that $\Omega s \leq w(p) \leq \Omega g$ and $s \leq v(p) \leq g$. With these arguments in hand, we can characterize a SMPE through the following definition:[10]

Definition 5 (Bellman System)**.** The function $\alpha : [0,1] \to [0,1]$ is an SMPE if there are functions $w, v : [0,1] \to \mathbb{R}$ such that for all $p \in [0,1]$, (α, w, v) jointly solve (16), (17) and (18), $\Omega s \leq w(p) \leq \Omega g$ and $s \leq v(p) \leq g$.

5.2 Main Result: Characterization of SMPE

Theorem 1.

1. *There exists a unique SMPE under each of positive, negative and transparent feedback, denoted by α_b, α_g and α_f respectively. Furthermore, for each $i \in \{b, g, f\}$, there exist*

[10] Formally, we say that w, v are solutions to their respective differential equations (17) and (16) in the *viscosity* sense (Crandall et al., 1992). In particular, the equilibrium strategy under positive feedback exhibits a discontinuity, and thus the usual Lipschitz conditions for uniqueness and existence of $\mathcal{C}^1([0,1])$-solutions to first-order ODEs do not apply. Nevertheless, we will adapt standard arguments (Kuvalekar and Lipnowski, 2020; Pham, 2009) to explicitly verify that at all regular states (Peskir and Shiryaev, 2006, §IV.9, VI.23), the candidate solution to w is in $\mathcal{C}^1([0,1])$. In the positive feedback case, w exhibits a kink (and is thus not $\mathcal{C}^1([0,1])$), but at an irregular state. These considerations are irrelevant in the transparent feedback case, wherein neither w nor v are described by differential equations.

$\underline{p}_i, \bar{p}_i$ *such that*

$$\alpha_i(p) \begin{cases} = 1 & \text{if} \quad p \geq \bar{p}_i \\ \in [0,1] & \text{if} \quad p \in [\underline{p}_i, \bar{p}_i) \\ = 0 & \text{if} \quad p < \underline{p}_i \end{cases} . \tag{19}$$

2. *For* $i \in \{b, g, f\}$, α_i *is continuous and strictly increasing on* $[\underline{p}_i, \bar{p}_i)$, *with* $\alpha_i(\underline{p}_i) = 0$. *Both* α_g *and* α_f *are Lipschitz continuous on* $[0,1]$. α_b *is discontinuous at* $\bar{p}_b$, *with* $\lim_{p \uparrow \bar{p}_b} \alpha_b(p) < \lim_{p \downarrow \bar{p}_b} \alpha_b(p) = 1$, *but is Lipschitz continuous everywhere else.*

3. $p^* < \underline{p}_b < \underline{p}_f = \underline{p}_g < \bar{p}_f = \bar{p}_b < \bar{p}_g < p^m$.

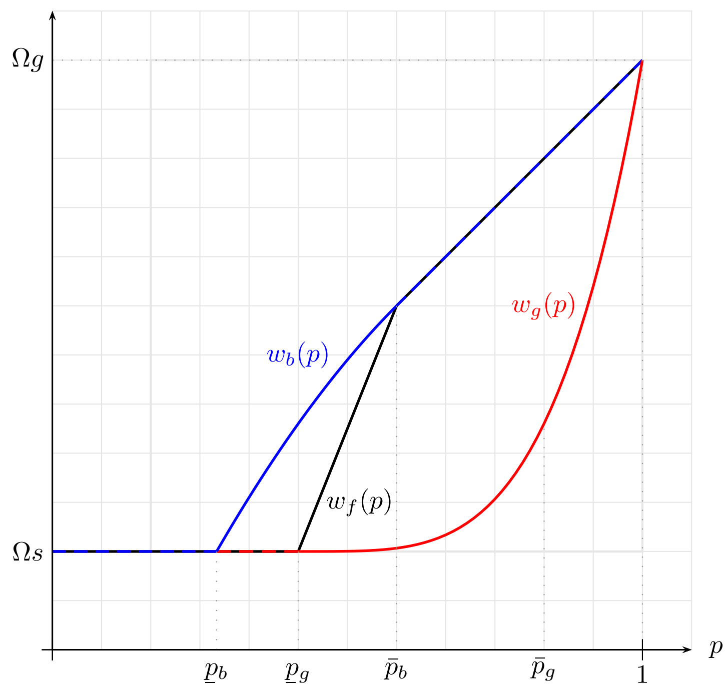
The remainder of the section is devoted to providing a detailed explanation of the various noteworthy features of Theorem 1, both in terms of their theoretical mechanisms as well as their psychological interpretations.[11] Figure 2 collects these insights and documents them graphically by plotting the equilibrium current values w.

5.3 Procrastination and Indecision

We begin by briefly describing why no equilibrium in simple threshold strategies exists. As an illustration, take the positive feedback case, and suppose that all future selves adopt the strategy $\alpha'(p) = \mathbb{I}_{p \geq p'}$. Clearly, it is a dominant strategy to play S at $p \leq p^*$, as the current self does not fully internalize future payoffs, and R at $p \geq p^m$, as this is a dominant strategy regardless. Thus, assume $p' \in (p^*, p^m)$. The continuation value then exhibits a discontinuity at p', with $\lim_{p \downarrow p'} v(p) > s$. To see this, note on the experimentation region, v is exactly the value function under exponential discounting; after the future arrives, discounting is exponential. As shown in Section 3, the *only* belief at which v can be continuous is p^*. While this is not an issue in the single-player problem — as v is optimized over simultaneously to w — in the equilibrium, v is taken as given by the current self, and thus such a discontinuity

[11] Through proving Theorem 1, we verify additional details that are omitted from the statement of the result for brevity. For instance, in all cases we fully solve for the equilibrium value functions w and v in closed form, and characterize the strategy on the region where $[\underline{p}_i, \bar{p}_i)$ as the solution to an implicit equation.

Figure 2: Equilibrium Current Value Functions



Blue lines: positive feedback. Red lines: negative feedback. Black lines: transparent feedback. For $p \geq \bar{p}_b$, the current value functions under both positive and transparent feedback coincide.

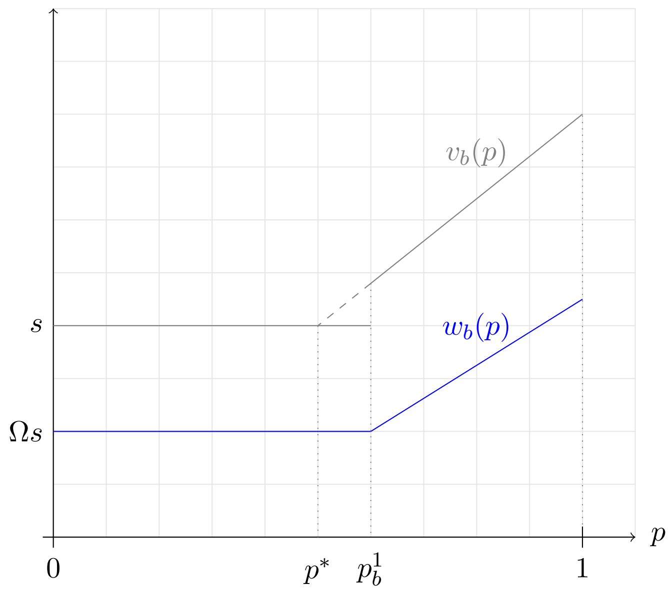
generates an incentive for randomization. Figure 3 demonstrates the argument graphically by plotting the single-player solution, which is the special case that $p' = p_b^1$.[12]

In essence, since the current self does not fully internalize the welfare of future selves, they are happy to free-ride on the experimentation of future selves and spare themselves the flow losses associated with experimentation today.[13] We thus term this effect the *procrastination effect*, inspired by O'Donoghue and Rabin (1999), as the DM delays costly experimentation when they shouldn't as it involves incurring costs in the immediate myopicrun.

[12]The argument is identical in the case of transparent feedback. Under negative feedback, the argument is qualitatively similar but slightly more subtle, and turns on a discontinuity in $v'(p)$ at $p = p'$, rather than $v(p)$.

[13]This effect is similar in spirit to the "free-rider" effect in experimentation games. See Section 7.3 for details.

Figure 3: Non-Existence of Threshold Equilibria



Blue line: current value under $\alpha'(p) = \mathbb{I}_{p \geq p'}$. Gray line: continuation value under $\alpha'(p)$. Gray dashed line: continuation under $\alpha(p) = \mathbb{I}_{p \geq p^*}$.

Remark 1 (Procrastination Effect). *No threshold equilibrium exists, regardless of the feedback structure.*

We interpret the DM playing $\alpha(p) \in (0, 1)$ over a range of beliefs as *indecision*.[14] Such indecision cannot obtain in the time-consistent version of the model (Proposition 1), nor in either the single-player or naive solutions (Propositions 2, 3 and Lemma 3), and thus is a sharp expression of the assumption of limited self-control within our framework. The relationship between lack of self-control, procrastination and indecision is well-documented (Ferrari et al., 1995; Ferrari and Pychyl, 2007). Indeed, the former relates indecision directly to *decisional procrastination*, the act of postponing a decision when faced with inner conflicts and choices (Janis and Mann, 1980). Furthermore, recent studies find that indecision is

[14] See (Bolton and Harris, 1999, §8) for a discussion of how the time-division problem we study is isomorphic to the problem where each self is allowed to randomize over $\alpha_t \in \{0, 1\}$. Our setting is less involved, as no two selves act simultaneously. Nevertheless, rigorously formulating mixed strategies in our continuous time setting is beyond the scope of the paper.

positively correlated with impulsiveness (Barkley-Levenson and Fox, 2016). Our results echo this finding, insofar as our equilibrium solution forces the DM to act in the present, rather than make longer-lasting plans.

5.4 Self-Control

The procrastination effect is present under all forms of feedback. In contrast, positive feedback exhibits a strategic externality that is not present in the other cases; by experimenting, the current self can push beliefs up, thereby encouraging their future selves to also experiment.[15] To see this, we prove an auxiliary result, namely that under negative feedback, the lower threshold $\underline{p}_g = p_g^1$, the analogous single-player solution. Intuitively, when the current self is at $\underline{p}_g$, they experiment in "isolation", as any experimentation they perform will terminate future experimentation, absent news. Thus, the first-order condition governing this lower threshold is identical to that governing p_g^1. By a similar token, under transparent feedback, if the current self is at $\underline{p}_f$ in a SMPE, any deviation to experiment over the next dt will be followed by playing R, since beliefs will not have drifted up into the region of greater experimentation. In contrast, under positive feedback, the current self can positively affect the experimentation rate through the induced upward belief drift absent news, thereby connecting them to their future selves.

Lemma 2 (Self-Control Effect). $\underline{p}_b < \underline{p}_f = \underline{p}_g = p_g^1$.

We term this effect the *self-control effect*. Control is the executive component of self-affirmation, and refers to confidence in one's ability to initiate actions that serve long-term goals. Stephan et al. (2017) find that by projecting further into the future, individuals attribute outcomes to their own willpower more, which in turn fortifies their commitment to achieving desired outcomes. In the present context, we argue that positive feedback engenders a sense of self-control. By inducing a positive trajectory of self-beliefs, positive feedback enables the DM to "look to the future" and partially mitigate present-bias. In

[15] This effect is similar in nature to the "encouragement" effect found in experimentation games. See Section 7.3.

contrast, negative and transparent feedback isolates the DM from their future.

5.5 Coherence and Self-Doubt

Theorem 1 exhibits further behavioral features that unlike the procrastination and self-control effects, have no direct analog in previous works on strategic experimentation. By inducing a path of improving beliefs, positive feedback allows the DM to envisage a future in which they are increasingly confident in their own abilities, which leads to behavior that aligns more closely with their long-run goals. This notion is best captured through another auxiliary result that compares the temptation to stop experimenting in equilibrium to that within the naive solution (Definition 3):

Lemma 3 (Coherence effect)**.** *For each $i \in \{b, g, f\}$, let α_i^n denote the solution to the naive problem. Then there exist p_i^n such that $\alpha_b^n(p) = \mathbb{I}_{p \geq p_b^n}$, $\alpha_f^n(p) = \mathbb{I}_{p \geq p_f^n}$ and $\alpha_g^n(p) = \mathbb{I}_{p > p_g^n}$. Furthermore, $\bar{p}_b = p_b^n$ and $\bar{p}_f = p_f^n$, while $\bar{p}_g > p_g^n$.*

Recall that the naive solution pitches the current self against the strategy wherein all future selves behave exponentially, and thus will experiment above p^* (Proposition 1). But note that under positive feedback, the upper threshold is best-responding to precisely the same strategy; since beliefs drift up, future play will coincide with the efficient benchmark. (The same logic holds under transparent feedback.) Hence, $\bar{p}_b$ and p_b^n must coincide.

Under positive feedback, the current and future selves are all perfectly aligned with their long-run goal at this belief. It is for this reason we call this effect the *coherence* effect. Coherence is another pillar of self-affirmation theory. It refers to the ability of individuals to create a consistent inner narrative, connecting current and future selves through a unified set of goals. Experimental evidence shows how by imagining their future selves, people make stronger connections between their present and future, reducing inner conflict (Stephan et al., 2015). In contrast, under negative feedback, at the upper threshold the DM is staring into a future plagued with indecision ($\alpha \in (0, 1)$) leading ultimately to the untimely termination of experimentation; beliefs will drift down into the region on which α gradually falls toward 0 at $\underline{p}_b > p^*$. Thus, the DM is trapped in a spiral of self-doubt, which unravels the current self's incentives to experiment, leading to $\bar{p}_g > p_g^n$.

5.6 Self-Esteem, Revisited

The final claim within Theorem 1 part 3) that demands explanation is why the upper thresholds $\bar{p}_g > \bar{p}_b$. Putting together the various previous claims, it would suffice to show that the naive thresholds $p_g^n > p_b^n$, which we verify now:

Lemma 4. $p_g^n > p_b^n$.

This result is directly analogous to why $p_g^1 > p_b^1$ in the single-player problem, and demonstrates that the self-esteem effect is still at play in equilibrium. The intuition is very similar to that developed in Section 4. At the threshold belief, the current self considers outcomes of experimentation only for an additional instant under negative feedback, whereas, under positive feedback, the current self knows that experimentation will continue (whether on their watch or during the future selves' tenure, who behave efficiently) unless a breakdown occurs.

5.7 Uniqueness

We now provide a brief sketch of how the SMPE identified in Theorem 1 are unique. From the form of equation (17), it is clear that there must be lowest and highest threshold beliefs at which the DM is willing to play R in any SMPE. The former is bounded below by p_i^1 for each $i \in \{b, g, f\}$, as the single-player problem is a relaxation of the equilibrium problem. The latter is clearly bounded above by the myopic threshold p^m, as playing R is strictly dominant at $p > p^m$.

Take first the negative feedback case. Any SMPE must have a lower threshold exactly equal to p_g^1. Suppose not, i.e. an SMPE exists such that $\underline{p}_g > p_g^1$. Note that at beliefs $p \in [p_g^1, \underline{p}_g)$, the current self's optimization problem is identical to the single-player counterpart; since beliefs drift down absent news, they experiment in isolation. Thus, they would have a strict incentive to continue experimenting on $[p_g^1, \underline{p}_g)$, a contradiction. Next, on the region where $\alpha_g \in (0,1)$, the indifference condition $b(p,w) - c(p) = 0$ must necessarily be satisfied. This equation generates an ODE governing w on this region, which admits a unique solution. By indifference, equation (17) then implies that v is also uniquely determined by the linear transformation $(\gamma + \eta)w(p) = \gamma s + \eta\beta v(p)$. The strategy α_g is then uniquely determined via (16). It is readily shown that $w(p)$ is continuous and strictly increasing on this region, and thus so too is $v(p)$ and hence $\alpha_g(p)$. Furthermore, we show there exists a $p^c < p^m$ such that $\lim_{p \uparrow p^c} \alpha_g(p) = \infty$, so that the upper threshold is then uniquely pinned down as the belief p at which $\alpha_g(p) = 1$.

The positive feedback case works in reverse. First, we argue that the upper threshold in any SMPE is necessarily the naive threshold p_b^n. This is because the upper threshold is precisely the belief at which the current self is indifferent between playing R and S, given that future play involves R being played at all higher beliefs, which is equivalent to the naive optimality condition. On the region where $\alpha_b \in (0,1)$, w is again the unique solution to an ODE, and is strictly increasing. The lower threshold is then uniquely determined as the belief at which $w(p) = \Omega s$, the present value of playing S forever. The transparent feedback is the simplest; since beliefs do not drift continuously, turning the differential equations (17) and (16) into a system of linear equations, one for each p, that can each be solved independently.

Finally, the discontinuity of the equilibrium strategy α_b under positive feedback is largely a technical feature, reflecting that the current value function is strictly concave on $[\underline{p}_b, \bar{p}_b)$ and linear on $[\bar{p}_b, 1]$, thus exhibiting a discontinuous second derivative at $\bar{p}_b$. Equation (17) then implies that the current value v has a kink and α_b a discontinuity. In general, it is not possible to rank experimentation rates across feedback structures. (See Section 6 for special cases where it is possible.)

6 The Benefits of Positive Feedback

Theorem 1 provides a sharp characterization of equilibrium outcomes, and allows us to contrast feedback structures through their effects on incentives and beliefs. But which feedback structure provides the greatest welfare?

As is standard in time-inconsistent models, there is no single welfare criterion, as each self has different intertemporal preferences. Nevertheless, we take the standard approach of using the DM's long-run discount factor as the appropriate weight on each instant (DellaVigna and Malmendier, 2004; O'Donoghue and Rabin, 1999, 2001; Gottlieb and Zhang, 2021). Thus, the welfare of a given strategy $(\alpha_t)_{t\geq 0}$ is evaluated according to

$$\mathcal{W}(\alpha; p) = \mathbb{E}\left[\int_0^\infty \gamma e^{-\gamma t} u(\alpha_t, p_t)\, dt \mid p_0 = p\right]. \tag{20}$$

Calculating equilibrium welfare is made straightforward by noticing that the long-run criterion (20) is precisely the equilibrium continuation value function $v(p)$, which was calculated in each case when proving Theorem 1:

Proposition 4 (The Benefits of Positive Feedback)**.**

1. *For $p \in [\bar{p}_b, 1)$: $\mathcal{W}(\alpha_b; p) = \mathcal{W}(\alpha_f; p) > \mathcal{W}(\alpha_g; p)$.*

2. *For $p \in (\underline{p}_g, \bar{p}_b)$: $\mathcal{W}(\alpha_b; p) > \mathcal{W}(\alpha_f; p) > \mathcal{W}(\alpha_g; p)$.*

3. *For $p \in (\underline{p}_b, \underline{p}_g]$: $\mathcal{W}(\alpha_b; p) > \mathcal{W}(\alpha_f; p) = \mathcal{W}(\alpha_g; p)$.*

4. *For $p \in [0, \underline{p}_b] \cup \{1\}$: $\mathcal{W}(\alpha_b; p) = \mathcal{W}(\alpha_f; p) = \mathcal{W}(\alpha_g; p)$.*

Thus, positive feedback is (weakly) preferred to transparent feedback, and at times strictly so, and is always strictly preferred to negative feedback. This result lies in stark contrast to all previous variants of the model; the DM was indifferent between positive and transparent feedback in the single-player problem, the naive problem and the exponential benchmark. The discrepancy must then stem from the strategic forces present exclusively in the equilibrium problem. (Indeed, in any single-player problem, full information must be (weakly) optimal.)

Motivation versus Information – Combining the insights acquired in Sections 5.3-5.6, we see that both transparent and positive feedback engender coherence (Lemma 3) and self-esteem (Lemma 4), but that only positive feedback promotes self-control (Lemma 2). Thus, the positive strategic value of linking the DM's selves in time through positively trending beliefs and bolstering self-control more than compensates for the informational loss in concealing breakthroughs.[16]

Of course, were feedback to be "babbling" and provide no information whatsoever, i.e. $\varphi_1 = \varphi_0 = 0$, the DM's belief would remain unmoved, which is easily seen to be dominated by.[17] Evidence shows that praise that is deemed excessive or unrealistic is ignored by children when forming beliefs about their abilities (Henderlong and Lepper, 2002). Relatedly, Finkelstein and Fishbach (2012) find that people who are certain of their abilities ("experts") respond to the informational content of negative feedback more effectively than "novices". More generally, a line of research emphasizes the role of negative feedback in providing informational content (Mayer, 1996; Locke and Latham, 2002).

While the role for "constructive feedback" is absent from our framework, we view these findings as echoing the fact that our DM benefits from being informed of breakdowns — in some sense, the harshest form of negative feedback possible — as the information therein allows them to better tailor their choices (they switch from R to S upon receiving this news). This is yet another asymmetry with negative feedback; observing a breakthrough does not alter the DM's current action.

Goal Orientation – Proposition 4 echoes a wealth of literature within psychology, both experimental and theoretical. The so-called "self-determination" theory suggests the beneficial/detrimental impact of positive/negative feedback on motivation (Ryan and Deci,

[16] Our work thus connects to the literatures on *strategic ignorance* (Bénabou and Tirole, 2002; Carrillo and Mariotti, 2000) and motivated beliefs (Bénabou and Tirole, 2004; Gottlieb, 2024) in economics.

[17] To see this, note that without information, the equilibrium strategy is $\alpha(p) = \mathbb{I}_{p \geq p^m}$, with $v(p) = \alpha(p)[gp + f(1-p) - s] + s$. For $p \geq p^m$, $\gamma[v_f(p) - v(p)] = \lambda gp + \lambda f(1-p) > 0$, while for $p < p^m$, $v_f(p) \geq s$, hence $v_f(p) - v(p) \geq 0$. Thus, transparent feedback always dominates no feedback.

2000). Ryan and Deci (2016) find strong evidence for this relationship in the context of school children performing ability-related tasks.

More specifically, a line of research relating to "goal orientation" argues that having a clearly defined set of goals can mitigate the detrimental impact of negative feedback, and that conversely, less goal-oriented give up on ability-based tasks more easily when presented with negative feedback (Dahling and Ruppel, 2016). As previously discussed, it is in general not possible to determine whether or not positive feedback induces greater experimentation at each $p \in [0,1]$, since on the region $[\underline{p}_g, \bar{p}_b]$, we cannot in general sign $\alpha_b(p) - \alpha_g(p)$ unambiguously. We can however do so when the degree of internal conflict is either very large or very small. We measure the degree of internal conflict by η, as the higher is η, the less time is each self in control, and thus the more present is internal conflict in the mind of the current self.

The following result shows that, in the quasi-hyperbolic case ($\eta \to \infty$), positive feedback exhibits higher rates of experimentation at all beliefs than both negative and transparent feedback, whereas when $\eta \to 0$, the manner of feedback becomes irrelevant. Figure 4 plots equilibrium strategies in the case where $\eta \to \infty$.

Lemma 5.

1. *For each $i \in \{b, g, f\}$, as $\eta \to \infty$, α_i converges pointwise to strategies α_i^∞ that take the form (19). Furthermore,*

   1) For $p \in [0, \underline{p}_b^\infty] \cup [\bar{p}_g^\infty, 1]$: $\alpha_b^\infty(p) = \alpha_f^\infty(p) = \alpha_g^\infty(p)$.

   2) For $p \in (\underline{p}_b^\infty, \underline{p}_g^\infty]$: $\alpha_b^\infty(p) > \alpha_f^\infty(p) = \alpha_g^\infty(p)$.

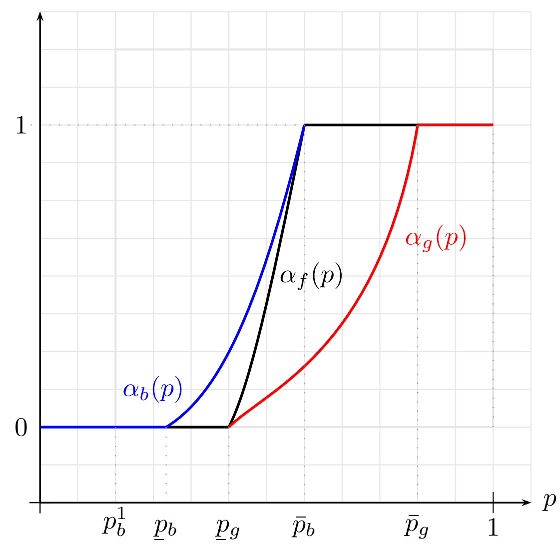
   3) For $p \in (\underline{p}_g^\infty, \bar{p}_b^\infty)$: $\alpha_b^\infty(p) > \alpha_f^\infty(p) > \alpha_g^\infty(p)$.

   4) For $p \in [\bar{p}_b^\infty, \bar{p}_g^\infty)$: $\alpha_b^\infty(p) = \alpha_f^\infty(p) > \alpha_g^\infty(p)$.

2. *For each $i \in \{b, g, f\}$, as $\eta \to 0$, $\bar{p}_i, \underline{p}_i \to p^*$.*

While our results on coherence (Lemma 3) suggest that positive feedback can help to mitigate conflicting personal goals, we view Lemma 5 above as providing a dual insight.

Figure 4: Quasi-Hyperbolic Equilibrium Strategies ($\eta \to \infty$)



Namely, in the absence of any internal conflict ($\eta = 0$), experimentation incentives are insensitive to the form of feedback provided, whereas when the internal conflict is strong (η high), the DM experiments more when provided with positive feedback.

7 Discussion

7.1 Model Discussion

We take a moment to discuss the various features of our model. Our pseudo-exponential discounting framework is taken from Harris and Laibson (2013). Its stationary, continuous-time specification is highly tractable, affording us the power of dynamic programming despite the inherent time-inconsistency, and delivering a unique SMPE. Furthermore, quasi-hyperbolic discounting are nested in this setting, by taking $\eta \to \infty$. We should note that our stochastic arrivals interpretation is not innocuous; Tan et al. (2021) demonstrate in an optimal-stopping setting with pseudo-exponential discounting that adopting the convention where a continuum of agents each operate instantaneously can expand the equilibrium set.

The assumption that payoff arrivals are imperfectly observed allows for direct comparison between positive and negative feedback without rewriting the model and separates payoffs from learning. Nevertheless, the analysis could be easily re-formulated to dispense with this assumption. For instance, an equivalent formulation is where the DM receives a flow payoff that is linear in the current belief p_t if they experiment, with news arrivals arriving independently from payoffs (Board and Meyer-ter-Vehn, 2013).

Our framework is built upon perfectly revealing Poisson learning processes. This is a well-established approach within the strategic experimentation literature (Keller et al., 2005; Keller and Rady, 2015) and beyond (Halac et al., 2017; Halac and Prat, 2016; Board and Meyer-ter-Vehn, 2013), with these papers finding stark differences in equilibrium behavior between good and bad news learning modes.[18] Furthermore, recent models of information choice based on good and bad news learning have delivered rich insights (Che and Mierendorff, 2019; Auster et al., forthcoming).

Finally, we restrict attention to stationary Markov perfect equilibria of the intra-personal game, clearly a strict subset of the set of all subgame perfect equilibria. Indeed, recent work by Hörner et al. (2021) demonstrates that the Markov restriction is with loss in games of strategic experimentation. Nevertheless, we take this approach for several reasons. First, it keeps our work in line with the majority of works on both time-inconsistency (Harris and Laibson, 2001, 2013; Luttmer and Mariotti, 2003; Karp, 2007) and strategic experimentation (Bolton and Harris, 1999; Keller et al., 2005; Keller and Rady, 2015). Second, since our motivation concerns the evolution of self-belief and the fostering of self-confidence, focusing on equilibria in which the DM's beliefs drive behavior is natural.

7.2 Feedback in Practice

As mentioned briefly in Section 2, we view a feedback structure that conceals breakthroughs and thus induces a positive trend in beliefs while the DM is still actively learning as providing positive feedback. In this paper, we do not motivate the origins of different feedback

[18] See Hörner and Skryzpacz (2016) for a survey.

structures. Nevertheless, to better relate our notions of feedback to those seen in practice, consider an alternative formulation of a feedback structure. Let $\vartheta : [0, \infty) \to \{0, 1, \emptyset\}$ be a $\mathbb{G}-$adapted "reporting function" where ϑ_t denotes an instantaneous 'report' at time t, observable to the DM, regarding payoff arrivals. Unlike in the baseline model, where arrivals could only be concealed, here we do not constrain ϑ in this way. This approach can replicate our three central feedback structures:

- Transparent feedback: Let $\vartheta_t = 0$ (1) if a breakdown (breakthrough) occurs at time t, and $\vartheta_t = \emptyset$ if no payoff arrives at time t.
- Positive feedback: Let $\vartheta_t = 0$ if a breakdown occurs at time t, and $\vartheta_t = 1$ otherwise.
- Negative feedback: Let $\vartheta_t = 1$ if a breakthrough occurs at time t, and $\vartheta_t = 0$ otherwise.

The reporting functions above give a clearer sense of how our definition of feedback capture practical settings. Take, for instance, positive feedback. The reporting function above can be interpreted as 'no news' being conflated with 'good news'. This is reminiscent of how for instance a personal trainer might give encouraging feedback even in the absence of supporting evidence.

7.3 Strategic Experimentation

Our analysis exhibits structural similarities to previous work on multi-player experimentation games (Bolton and Harris, 1999; Keller and Rady, 2015). We view these links as a strength; not only do we leverage them in proving Theorem 1, but we view them as harmonizing insights from disparate works. We take a moment to draw out these parallels in greater detail.

The procrastination effect is similar to the "free-rider" effect found in previous works on strategic experimentation. The free-rider effect describes a player's incentive not to experiment, instead letting other players experiment and generate valuable information that they can subsequently use themselves. Our procrastination effect also turns on an incentive to free-ride on other selves' experimentation and instead play safe. Heuristically, note that more experimentation by future selves increases v, which from equation (17) for the current value w can be seen to have the same effect as raising s, thus raising the value of playing S. However, the externality is qualitatively different, as the current self in our setting does not seek to procrastinate in order to use publicly-generated information themselves at a later time.

The self-control effect is reminiscent of the so-called *encouragement effect* in the strategic experimentation literature, which describes how an agent has an incentive to experiment in order to generate information that subsequently encourages other agents to experiment, which in turn generates valuable information for the agent.[19] Similar to our analysis, Keller and Rady (2015) identifies the presence of this effect with breakdown learning, while Keller et al. (2005) shows how the effect is absent under breakthrough learning. Mechanically, the reason is precisely the direction of belief drift absent arrivals, just as is here. Furthermore, in those papers, the misalignment between players disappears at extreme beliefs, as a dominant action exists.

However, there are some key differences. First, externalities are purely informational in these papers, whereas in the current setting payoff externalities also exist. Hence, the current agent seeks to encourage experimentation insofar as they care *directly* about the welfare of future selves. Second, our game is sequential, generating qualitatively different incentives, as discussed above. Beyond these underlying differences, our results differ in qualitative ways. The single-player benchmark is reminiscent of both the cooperative and single-player solutions studied in Keller et al. (2005) and Keller and Rady (2015), insofar as it strips away all strategic externalities that might arise in the equilibrium solution of the model. It is, however, distinct from both. In contrast to their cooperative solution, this benchmark is not efficient, as the DM still maximizes according to their non-exponential discount function. Furthermore, our single-player solution is a maximizer of (2) under the constraint

[19] Note that Bolton and Harris (1999) define the encouragement effect by examine how one player's best-response value is affected by another player's action. We thank Sven Rady for pointing out this distinction to us.

that feasible policies are Markovian, and hence may not coincide with the unconstrained open loop control maximizer of (2). Finally, equilibrium strategies are discontinuous under positive feedback, in contrast to the symmetric MPE in Keller and Rady (2015).

8 Conclusion and Future Work

We presented a model in which a present-biased decision maker (DM) with limited willpower faces an experimentation problem, and learns about its inherent risk through feedback. We showed that by enabling the DM to "self-prospect" – imagining one's self in the distant future when evaluating one's present –, positive feedback sustains greater welfare than either negative or transparent feedback.

While the tractability of our framework allowed us to derive several results, we conclude by highlighting potential avenues for future research within our broad agenda. First, while our primary focus was comparing positive and negative feedback, it is a natural question to ask more generally, what is the optimal form of feedback to provide such a person? We have also abstracted from the important question of where feedback originates from. In many instances, it is provided by a principal, whose interests may or may not be aligned with the DM, and who may be able to contract with the DM in more complex ways. We leave these interesting considerations for future work.

References

Ali, S. Nageeb, "Learning Self-Control," *The Quarterly Journal of Economics*, 2011, *126* (2), 857–893.

Ashraf, N., D. Karlan, and W. Yin, "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *The Quarterly Journal of Economics*, 2006, *121* (2), 635–672.

Auster, S., Y-K Che, and K. Mierendorff, "Prolonged Learning and Hasty Stopping: the Wald Problem with Ambiguity," *American Economic Review*, forthcoming.

Barkley-Levenson, E.E. and C.R. Fox, "The surprising relationship between indecisiveness and impulsivity," *Personality and Individual Differences*, 2016, *90*, 1–6.

Bénabou, R. and J. Tirole, "Self-Confidence and Personal Motivation," *The Quarterly Journal of Economics*, 2002, *117* (3), 871–915.

— **and** — , "Intrinsic and Extrinsic Motivation," *The Review of Economic Studies*, 2003, *70* (3), 489–520.

— **and** — , "Willpower and Personal Rules," *Journal of Political Economy*, 2004, *112* (4), 848–886.

Bernheim, B.D. and D. Ray, "Economic Growth with Intergenerational Altruism," *Review of Economic Studies*, 1987, *54* (2), 227–242.

Board, S. and M. Meyer-ter-Vehn, "Reputation for Quality," *Econometrica*, 2013, *81* (6).

Bolton, P. and C. Harris, "Strategic Experimentation," *Econometrica*, 1999, *67* (2), 349–374.

Brocas, I. and J. Carrillo, "The Value of Information when Preferences are Dynamically Inconsistent," *European Economic Review*, 05 2000, *44*, 1104–1115.

Bryan, G., D. Karlan, and S. Nelson, "Commitment Devices," *Annaul Review of Economics*, 2010, *2*, 671–698.

Carrillo, J. D. and T. Mariotti, "Strategic Ignorance as a Self-Disciplining Device," *The Review of Economic Studies*, 2000, *67* (3), 529–544.

Che, Y-K. and J. Hörner, "Recommender Systems as Mechanisms for Social Learning," *The Quarterly Journal of Economics*, 2018, *133* (2), 871–925.

Che, Y-K and K. Mierendorff, "Optimal Dynamic Allocation of Attention," *American Economic Review*, 2019, *109* (8).

Crandall, M., H. Ishii, and P-L. Lions, "User's Guide to Viscosity Solutions," *Bulletin of the American Mathematical Society*, 1992, *27* (1), 1–67.

Dahling, J.J. and C.L. Ruppel, "Learning goal orientation buffers the effects of negative normative feedback on test self-efficacy and reattempt interest," *Learning and Individual Differences*, 2016, *50*, 296–301.

DellaVigna, S. and U. Malmendier, "Contract Design and Self-Control: Theory and Evidence," *The Quarterly Journal of Economics*, 2004, *119* (2), 353–402.

Ekeland, I. and A. Lazrak, "Being serious about non-commitment: subgame perfect equilibrium in continuous time," *mimeo University of British Columbia*, 2006.

Ferrari, J., J. Johnson, and W. McCown, *Procrastination and Task Avoidance–Theory, Research and Treatment*, Springer Science & Business Media, 1 1995.

Ferrari, J.R. and T.A. Pychyl, "Regulating speed, accuracy and judgments by indecisives: Effects of frequent choices on self-regulation depletion," *Personality and Individual Differences*, 2007, *42* (4), 777–787.

Finkelstein, S.R. and A. Fishbach, "Tell Me What I Did Wrong: Experts Seek and Respond to Negative Feedback," *Journal of Consumer Research*, 2012, *39* (1), 22–38.

Fong, C.J., E.A. Patall, A.C. Vasquez, and S. Stautberg, "A Meta-Analysis of Negative Feedback on Intrinsic Motivation," *Educational Psychology Review*, 2018, *31*, 121–162.

Gottlieb, D., "Will You Never Learn? Self Deception and Biases in Information Processing," *working paper*, 2024.

_ and X. Zhang, "Long-term Contracting With Time-Inconsistent AgentS," *Econometrica*, 2021, *89* (2), 793–824.

Halac, M. and A. Prat, "Managerial Attention and Worker Performance," *American Economic Review*, 2016, *106* (10).

_ , N. Kartik, and Q. Liu, "Contests for Experimentation," *Journal of Political Economy*, 2017, *125* (5).

Harris, C. and D. Laibson, "Dynamic Choices of Hyperbolic Consumers," *Econometrica*, 2001, *69* (4), 935–957.

_ and _ , "Instantaneous Gratification," *Quarterly Journal of Economics*, 2013, *128* (1), 205–238.

Hattie, J. and H. Timperley, "The Power of Feedback," *Review of Educational Research*, 2007, *77* (1), 81–112.

Henderlong, J. and M.R. Lepper, "The effects of praise on children's intrinsic motivation: a review and synthesis," *Psychological bulletin*, 2002, *128* (5), 774–95.

Hörner, J. and A. Skryzpacz, "Learning, Experimentation and Informtion Design," *working paper, prepared for the Econometric Summer Meetings in Montreal*, 2016.

_ , N. Klein, and S. Rady, "Overcoming Free-Riding in Bandit Games," *The Review of Economic Studies*, 2021, *89* (4), 1948–1992.

Hume, D., *An Enquiry Concerning the Principles of Morals*, London: Andrew Millar, 1751.

Janis, I.L. and L. Mann, "Decision Making: A Psychological Analysis of Conflict, Choice, and Commitment," *The Annals of the American Academy of Political and Social Science*, 1980, *449* (1), 202–203.

Karp, L., "Non-constant discounting in continuous time," *Journal of Economic Theory*, 2007, *132*, 557—568.

Keller, G. and S. Rady, "Breakdowns," *Theoretical Economics*, 2015, *10*, 175–202.

_ , _ , and M. Cripps, "Strategic Experimentation with Exponential Bandits," *Econometrica*, 2005, *73* (1), 39–68.

Kremer, I., Y. Mansour, and M. Perry, "Implementing the "Wisdom of the Crowd"," *Journal of Political Economy*, 2014, *122* (5), 988–1012.

Kuvalekar, A. and E. Lipnowski, "Job Insecurity," *American Economic Journal: Microeconomics*, 2020, *12* (2), 188–229.

Locke, E. and G. Latham, "Building a practically useful theory of goal setting and task motivation - A 35-year odyssey," *The American psychologist*, 10 2002, *57*, 705–17.

Luttmer, Erzo G. J. and Thomas Mariotti, "Subjective Discounting in an Exchange Economy," *Journal of Political Economy*, 2003, *111* (5), 959–989.

Mayer, R.E., "Learners as information processors: Legacies and limitations of educational psychology's second metaphor," *Educational Psychologist*, 1996, *31* (3-4), 151–161.

O'Donoghue, T. and M. Rabin, "Doing It Now or Later," *The American Economic Review*, 1999, *89* (1), 103–124.

_ and _ , "Choice and Procrastination," *The Quarterly Journal of Economics*, 2001, *116* (1), 121–160.

Peskir, G. and A.N. Shiryaev, *Optimal Stopping and Free-Boundary Problems*, Vol. 10 of *Lectures in Mathematics. ETH Zürich*, Springer Verlag NY, 2006.

Pham, H., *Social Psychology*, Vol. 61, Continuous-time stochastic control and optimization with financial applications, 2009.

Phelps, E. S. and R. A. Pollak, "On Second-Best National Saving and Game-Equilibrium Growth," *The Review of Economic Studies*, 1968, *35* (2), 185–199.

Presman, E., "Poisson Version of the Two-Armed Bandit Problem with Discounting," *Theory of Probability and Its Applications*, 1990, *135.*

Ryan, R.M. and E.L. Deci, "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being," *The American psychologist*, 2000, *55*, 68–78.

— and — , "Optimizing Students' Motivation in the Era of Testing and Pressure: A Self-Determination Theory Perspective," in W.C. Liu, J.C.K. Wang, and R.M. Ryan, eds., *Building Autonomous Learners: Perspectives from Research and Practice using Self-Determination Theory*, Springer Singapore, 2016, pp. 9–29.

Smith, E.R. and D.M. Mackie, *Social Psychology*, Psychology Press, 2007.

Stephan, E. and C. Sedikides, "Mental Time Travel as Self-Affirmation," *Personality and Social Psychology Review*, 2023, pp. 1–28.

— , — , D. Heller, and D. Shidlovski, "My fair future self: The role of temporal distance and self-enhancement in prediction," *Social Cognition*, 2015, *33* (2), 149–168.

— , D. Shidlovski, and D. Heller, "Distant determination and near determinism: The role of temporal distance in prospective attributions to will," *Journal of Experimental Social Psychology*, 2017, *68*, 113–121.

Strotz, R.H., "Myopia and Inconsistency in Dynamic Utility Maximization," *The Review of Economic Studies*, 1955, *23* (3), 165–180.

Tan, K.S., W. Wei, and X.Y. Zhou, "Failure of Smooth Pasting Principle and Nonexistence of Equilibrium Stopping Rules under Time-Inconsistency," *SIAM Journal on Control and Optimization*, 2021, *59* (6), 4136–4154.

A Proofs

A.1 Proof of Proposition 1

Positive Feedback – The DM's problem is given by

$$v(p) = s + \max_{\alpha \in [0,1]} \alpha \cdot [b_b(p,v) - c(p)], \tag{A.1}$$

where the opportunity cost of playing R is

$$c(p) = s - [gp + f(1-p)], \tag{A.2}$$

and the discounted expected benefit of playing R is

$$b_b(p,v) = \frac{\lambda(1-p)}{\gamma}(s - v(p) + pv'(p)). \tag{A.3}$$

The boundary conditions are $s \leq v(p) \leq g$. The Bellman equation (A.1) is linear in α so admits a bang-bang solution. If $\alpha = 0$, $v(p) = s$, while if $\alpha = 1$,

$$(\gamma + \lambda(1-p))v(p) - \lambda p(1-p)v'(p) = \lambda(1-p)s + \gamma(gp + f(1-p)). \tag{A.4}$$

Since $v(p) \leq g$ for all $p \in [0,1]$, (A.4) can only admit the particular solution

$$v_b(p) = gp + \frac{\gamma f + \lambda s}{\gamma + \lambda}(1-p). \tag{A.5}$$

The optimal cutoff is uniquely pinned down by the boundary condition $v_b(p^*) = s$, implying the optimal threshold to be

$$p^* = \frac{\gamma(s-f)}{\gamma(g-f) + \lambda(g-s)}$$

and $\alpha_b(p) = \mathbb{I}_{p \geq p^*}$. The optimal value function exhibits continuity (however possessing a kink at p^*):

$$v_b^*(p) = \begin{cases} gp + \frac{\gamma f + \lambda s}{\gamma + \lambda}(1-p) & \text{if} \quad p \geq p^* \\ s & \text{if} \quad p < p^* \end{cases}. \tag{A.6}$$

Negative Feedback – The DM's problem is now given by

$$v(p) = s + \max_{\alpha \in [0,1]} \alpha \cdot [b_g(p, v) - c(p)], \tag{A.7}$$

where $c(p)$ is as in (A.2), and

$$b_g(p, v) = \frac{\lambda p}{\gamma}(g - v(p) - (1 - p)v'(p)), \tag{A.8}$$

subject to $s \leq v(p) \leq g$. Due to linearity in α, it is optimal to choose either $\alpha = 0$ (so that $v_g(p) = s$), or $\alpha = 1$ implying $v(p)$ satisfies the first-order ordinary differential equation

$$(\gamma + \lambda p)v(p) + \lambda p(1 - p)v'(p) = (\gamma + \lambda)gp + \gamma f(1 - p). \tag{A.9}$$

The differential equation (A.9) has a solution

$$v_g(p) = gp + f(1 - p) + C_g(1 - p)\left(\frac{1 - p}{p}\right)^{\frac{\gamma}{\lambda}} \tag{A.10}$$

with some constant C_g. Suppose the DM uses S below the cutoff belief p_g^* and R above. As is standard, smooth pasting and value matching conditions apply, confirming the equality $p_g^* = p^*$ (from straightforward algebra) and defining the constant of integration, providing

$$v_g^*(p) = \begin{cases} gp + f(1 - p) + \bar{C}_g(1 - p)\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}} & \text{if} \quad p > p^* \\ s & \text{if} \quad p \leq p^* \end{cases}, \tag{A.11}$$

where

$$\bar{C}_g = \frac{(s - f) - (g - f)p^*}{1 - p^*}\left(\frac{p^*}{1 - p^*}\right)^{\frac{\gamma}{\lambda}}. \tag{A.12}$$

On the entire interval of beliefs $[0, 1]$, the value function is smooth and convex (strictly convex on $(p^*, 1]$).

Transparent Feedback – The DM solves

$$v(p) = s + \max_{\alpha \in [0,1]} \alpha \cdot [b_f(p, v) - c(p)], \tag{A.13}$$

with

$$b_f(p,v) = \frac{\lambda(1-p)}{\gamma}(v(0)-v(p)) + \frac{\lambda p}{\gamma}(v(1)-v(p)), \tag{A.14}$$

with $v(0) = s$, $v(1) = g$ and the boundary conditions $s \leq v(p) \leq g$. The optimization problem is linear in both α and $v(p)$ so leads to the bang-bang solution with two linear branches: $v(p) = s$ if $\alpha = 0$ and

$$v_f(p) = gp + \frac{\gamma f + \lambda s}{\gamma + \lambda}(1-p) \tag{A.15}$$

if $\alpha = 1$. Value matching implies the optimal threshold $p_f^* = p^*$ and the value function

$$v_f^*(p) = \begin{cases} gp + \frac{\gamma f + \lambda s}{\gamma + \lambda}(1-p) & \text{if} \quad p \geq p^* \\ s & \text{if} \quad p < p^* \end{cases}$$

to coincide with $v_b^*(p)$.

A.2 Proof of Lemma 1

1. The equality $v_f^*(p) = v_b^*(p)$ was verified explicitly in the proof of Proposition 1. Trivially, $v_g^*(p^*) = v_b^*(p^*)$, $v_g^*(1) = v_b^*(1)$. On $(p^*, 1)$, the optimal value function under negative feedback is strictly convex and thus strictly lower than $v_i^*(p)$, $i \in \{b, f\}$. Convexity follows from the direct calculation of the second derivative

$$(v_g^*(p))'' = \left(g - f - \bar{C}_g \frac{\gamma + \lambda p}{\lambda p}\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}}\right)' = \bar{C}_g \frac{\gamma(\gamma+\lambda)}{\lambda^2 p(1-p)}\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}}$$

   and the fact that $\bar{C}_g > 0$ (since $p^* < p^m$).

2. Below the efficient threshold p^*, the optimal strategy is $\alpha^* = 0$ and $v_b^*(p) = v_f^*(p) = v_g^*(p) = s$ regardless of the feedback mode. At $p = 1$, $v_b^*(1) = v_f^*(1) = v_g^*(1) = g$.

A.3 Proof of Proposition 2

The DM solves:

$$\begin{aligned} v(p,\alpha) &= s + \alpha \cdot [b_g(p,v) - c(p)], \\ w(p) &= s - \frac{\eta}{\gamma}w(p) + \max_{\alpha \in [0,1]} \alpha \cdot [b_g(p,w) - c(p) + \frac{\eta\beta}{\gamma}v(p,\alpha)], \end{aligned} \tag{A.16}$$

with boundary conditions $s \leq v(p, \alpha) \leq g$, $\Omega s \leq w(p) \leq \Omega g$, where $c(\cdot)$ and $b_g(\cdot, \cdot)$ satisfy (A.2) and (A.8), respectively, and $v(1) \equiv v(1, 1) = g$, $w(1) = \Omega g$. The optimization problem (A.16) in the Bellman system is linear in α, implying a threshold solution. Hereinafter, we omit the dependence of $v(p, \alpha)$ on α where clear. If $\alpha = 0$, then $v(p) = s$ and $w(p) = \Omega s$. If $\alpha = 1$, then the value functions satisfy the system of the differential equations

$$(\gamma + \lambda p)v(p) + \lambda p(1 - p)v'(p) = (\gamma + \lambda)gp + \gamma f(1 - p), \tag{A.17}$$

$$(\gamma + \lambda p + \eta)w(p) + \lambda p(1 - p)w'(p) = (\gamma + \lambda\Omega)gp + \gamma f(1 - p) + \eta\beta v(p). \tag{A.18}$$

(A.17) has a general solution (A.10) (with a constant C_g^1 that may differ from C_g), that helps us to identify a form of the particular solution for (A.18). The solution of (A.18) is

$$w(p) = \Omega gp + \Omega f(1 - p) + \beta C_g^1(1 - p)\left(\frac{1 - p}{p}\right)^{\frac{\gamma}{\lambda}} + K_g^1(1 - p)\left(\frac{1 - p}{p}\right)^{\frac{\gamma+\eta}{\lambda}}.$$

The constants C_g^1, K_g^1 and the threshold p_g^1 are obtained by imposing $w(p_g^1) = \Omega s$ (value matching for the current value), $w'(p_g^1) = 0$ (smooth pasting for the current value), $v(p_g^1) = s$ (value matching for the continuation value). The last condition arises from the fact that the beliefs move continuously in time through p_g^1, making it a regular boundary. Putting these calculations together, we arrive at explicit solutions for the value functions:

$$v_g^1(p) = \begin{cases} gp + f(1 - p) + \bar{C}_g(1 - p)\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}} & \text{if} \quad p > p_g^1 \\ s & \text{if} \quad p \leq p_g^1 \end{cases}$$

and

$$w_g^1(p) = \begin{cases} \Omega gp + \Omega f(1 - p) + \beta\bar{C}_g^1(1 - p)\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}} + \bar{K}_g^1(1 - p)\left(\frac{1-p}{p}\right)^{\frac{\gamma+\eta}{\lambda}} & \text{if} \quad p > p_g^1 \\ \Omega s & \text{if} \quad p \leq p_g^1 \end{cases},$$

where the constants $\bar{C}_g^1$, $\bar{K}_g^1$ are

$$\bar{C}_g^1 = \frac{(s-f)-(g-f)p_g^1}{1-p_g^1}\left(\frac{p_g^1}{1-p_g^1}\right)^{\frac{\gamma}{\lambda}},$$

$$\bar{K}_g^1 = \frac{\gamma(1-\beta)}{\gamma+\eta}\frac{(s-f)-(g-f)p_g^1}{1-p_g^1}\left(\frac{p_g^1}{1-p_g^1}\right)^{\frac{\gamma+\eta}{\lambda}}.$$

The threshold is given by

$$p_g^1 = \frac{\gamma(s-f)}{\gamma(g-f)+\lambda\Omega(g-s)}.$$

The continuation value $v_g^1(p)$ is continuous but possesses a convex kink at p_g^1. The current value function $w_g^1(p)$ is continuous and smooth on the entire interval $[0,1]$.

A.4 Proof of Proposition 3

The DM solves:

$$\begin{aligned} v(p,\alpha) &= s + \alpha\cdot[b_b(p,v)-c(p)], \\ w(p) &= s - \frac{\eta}{\gamma}w(p) + \max_{\alpha\in[0,1]}\alpha\cdot[b_b(p,w)-c(p)+\frac{\eta\beta}{\gamma}v(p,\alpha)], \end{aligned} \tag{A.19}$$

with $s \le v(p,\alpha) \le g$, $\Omega s \le w(p) \le \Omega g$, and where $c(p)$ and $b_b(p,\cdot)$ are given by (A.2) and (A.3). Due to linearity in α, this system admits a bang-bang solution with $v(p)=s$, $w(p)=\Omega s$, when $\alpha=0$ and $v(p)$, $w(p)$ satisfying the following system when $\alpha=1$:

$$(\gamma+\lambda(1-p))v(p) - \lambda p(1-p)v'(p) = \lambda(1-p)s + \gamma(gp+f(1-p)), \tag{A.20}$$

$$(\gamma+\lambda(1-p)+\eta)w(p) - \lambda p(1-p)w'(p) = \gamma(gp+f(1-p)) + \lambda\Omega(1-p)s + \eta\beta v(p). \tag{A.21}$$

The boundary condition $s \le v(p,\alpha) \le g$ ensures that (A.5) is a solution to (A.20), so that the right branch of the value function v_b^* in the exponential case aligns with the right branch of the continuation value function in the (β,η)-model. The differential equation (A.21) also admits only a particular solution under the presence of the bounds $\Omega s \le w(p) \le \Omega g$:

$$w(p) = \Omega gp + \bar{C}_b^1(1-p), \tag{A.22}$$

$$\bar{C}_b^1 = \frac{\gamma+\lambda+\eta\beta}{\gamma+\lambda+\eta}\frac{\gamma f}{\gamma+\lambda} + \frac{\gamma+\lambda+\eta\beta}{\gamma+\lambda+\eta}\frac{\gamma\lambda s}{(\gamma+\lambda)(\gamma+\eta)} + \frac{\lambda\eta\beta s}{(\gamma+\lambda)(\gamma+\eta)}. \tag{A.23}$$

Value matching of the solutions for $\alpha = 0$ and $\alpha = 1$ for the current value function yields the desired threshold value (13) and allows again for explicit solutions for the value functions:

$$v_b^1(p) = \begin{cases} gp + \frac{\gamma f + \lambda s}{\gamma + \lambda} & \text{if} \quad p \geq p_b^1 \\ s & \text{if} \quad p < p_b^1 \end{cases}$$

and

$$w_b^1(p) = \begin{cases} \Omega gp + \bar{C}_b^1(1-p) & \text{if} \quad p \geq p_b^1 \\ \Omega s & \text{if} \quad p < p_b^1 \end{cases}.$$

Both value functions are piecewise linear. The continuation value function exhibits a discontinuity at p_b^1, while the current value function exhibits a convex kink at p_b^1.

A.5 Proof of Theorem 1

A.5.1 Part 1

Positive Feedback – Fix $\alpha' : [0,1] \to [0,1]$, the future selves' strategy. The value functions then satisfy the Bellman system

$$v(p, \alpha') = s + \alpha'(p) \cdot [b_b(p, v) - c(p)], \tag{A.24}$$

$$w(p) = s + \frac{\eta}{\gamma}[\beta v(p, \alpha') - w(p)] + \max_{\alpha \in [0,1]} \alpha \cdot [b_b(p, w) - c(p)], \tag{A.25}$$

with the boundary conditions

$$s \leq v(p) \leq g, \tag{A.26}$$

$$\Omega s \leq w(p) \leq \Omega g, \tag{A.27}$$

and where $w(0) = \Omega s$, $v(0) = s$.

The current self's best response, denoted by α_b^*, is determined by comparing the opportunity cost of playing R with the expected private benefit:

$$\alpha_b^*(p) \begin{cases} = 1 & \text{if} \quad c(p) < b_b(w,p) \\ \in [0,1] & \text{if} \quad c(p) = b_b(w,p) \cdot \\ = 0 & \text{if} \quad c(p) > b_b(w,p) \end{cases} \tag{A.28}$$

Note that playing R is strictly dominant for $p > p^m$. Hence, in any SMPE, there must exist $\bar{p}_b \leq p^m$ such that current self is indifferent between S and R at $\bar{p}_b$, but plays R exclusively at all higher beliefs.

Lemma A.1. *In any SMPE, $\bar{p}_b = p_b^n$, where p_b^n is the solution of the naive problem.*

Proof. By definition 3, the naive problem is characterized by the Bellman system (A.24), (A.25), along with the boundary conditions (A.26), (A.27), with $\alpha'(p) = \mathbb{I}_{p \geq p^*}$. Thus, the continuation value function is given by (A.6). The resulting Bellman system is linear in α, hence it exhibits a bang-bang solution, with switching threshold p_b^n. Clearly $p_b^n > p^*$, thus beliefs fall into three regions: all selves play S ($p < p^*$), the current self plays S and the future selves play R ($p \in [p^*, p_b^n)$), and all selves play R ($p \geq p_b^n$).

When all selves play the same action, the values are known and coincide with the single-player solution: in the first region, the current value function is Ωs, in the third region, it is given by (A.22). When the future selves play R (with the value $v(p) = gp + \frac{\gamma f + \lambda s}{\gamma + \lambda}$) and the current self plays S, the current value function is given by:

$$\begin{aligned} w_2(p) &= s + \frac{\eta}{\gamma}[\beta v(p) - w_2(p)], \\ w_2(p) &= \frac{\gamma}{\gamma+\eta}\frac{\gamma s + \eta\beta f}{\gamma+\lambda} + \Omega\frac{\lambda s}{\gamma+\lambda} + \frac{\eta\beta}{\gamma+\eta}\frac{\gamma(g-f)+\lambda(g-s)}{\gamma+\lambda}p. \end{aligned} \tag{A.29}$$

The threshold that constitutes the best response to p^* balances the value from the experimentation, (A.22), and pulling S, (A.29):

$$w_2(p_b^n) = \Omega g p_b^n + \bar{C}_b^1(1 - p_b^n).$$

This equation holds at

$$p_b^n \equiv \frac{\gamma(s-f)}{\gamma(g-f) + \lambda\frac{\gamma(\gamma+\lambda+\eta\beta)}{\gamma(\gamma+\lambda+\eta)+\lambda\eta(1-\beta)}(g-s)},$$

defining the optimal naive threshold strategy $\alpha_b(p) = \mathbb{I}_{p \geq p_b^n}$. For the current self, experimentation yields higher value as belief increases, conditional on future selves experimenting at the belief p. Thus, the current self commences experimentation at $\bar{p}_b \leq p_b^n$.

We now argue that $\bar{p}_b = p_b^n$. Toward a contradiction, suppose future selves employs a strategy $\tilde{\alpha}$ such that $\tilde{\alpha}(p) = 1$ for $p \geq \tilde{p}$, where $\tilde{p} < p_b^n$. We show that deviation S at some belief $p_g^w \geq \tilde{p}$ is profitable. Consider the interval $[\tilde{p}, p_b^n)$. Here, the future selves play R, yielding $v(p) = gp + \frac{\gamma f + \lambda s}{\gamma + \lambda}(1-p)$. The present value to the current self from the putative deviation is then:

$$w_{dev}(p_g^w) = \frac{\gamma}{\gamma+\eta}s + \frac{\eta\beta}{\gamma+\eta}[gp_g^w + \frac{\gamma f + \lambda s}{\gamma+\lambda}(1-p_g^w)].$$

Comparing this value with the value from full experimentation (a linear function, given by (A.22)) gives $w_{dev}(p) > \Omega gp + \bar{C}_b^1(1-p)$ on the interval $[\tilde{p}, p_b^n)$, thus generating the desired contradiction. $\square$

Thus, on the interval $[\bar{p}_b, 1]$, both selves experiment. The differential equation (A.24) provides the continuation value function on this interval:

$$v_3(p) = gp + \frac{\gamma f + \lambda s}{\gamma + \lambda}(1-p).$$

From the Bellman equation (A.25), the current value $w_3(p)$ is the solution of the differential equation

$$w_3(p) = gp + f(1-p) + \frac{\eta}{\gamma}[\beta v_3(p) - w_3(p)] + \frac{\lambda(1-p)}{\gamma}(\Omega s - w_3(p) + pw_3'(p))$$

with the boundary conditions (A.27). The problem the current self solves aligns with the single-player problem (A.21) when both selves experiment. Thus, the solution coincides with (A.22) for $p \in [\bar{p}_b, 1]$. One can also verify that $c(p) < b(w_3, p)$ for $p \in (\bar{p}_b, 1]$, with equality at $\bar{p}_b$, so that on $(\bar{p}_b, 1]$,

$$w_3(p) > \frac{\gamma}{\gamma+\eta}s + \frac{\eta\beta}{\gamma+\eta}v_3(p)$$

and

$$w_3(\bar{p}_b) = \frac{\gamma}{\gamma+\eta}s + \frac{\eta\beta}{\gamma+\eta}v_3(\bar{p}_b). \tag{A.30}$$

Using the condition $c(p) = b(w, p)$ from (A.28), we obtain an explicit representation for w on the region of indifference:

$$\frac{\lambda(1-p)}{\gamma}\left[w(0)-w(p)+pw'(p)\right]=s-gp-f(1-p). \tag{A.31}$$

with the boundary condition $w(\bar{p}_b)=w_3(\bar{p}_b)$ (again, due to the indifference). The unique solution for $p<\bar{p}_b$ is given by

$$\begin{aligned} w_2(p) &= \frac{\gamma+\eta\beta}{\gamma+\eta}s-\frac{\gamma}{\lambda}(s-f)+\frac{\gamma}{\lambda}(g-s)p\ln\left(\frac{1-p}{p}\right)+\bar{C}_b^{eq}p, \\ \bar{C}_b^{eq} &= \Omega g-\bar{C}_b^1+\frac{\bar{C}_b^1-\Omega s+\frac{\gamma}{\lambda}(s-f)}{\bar{p}_b}-\frac{\gamma}{\lambda}(g-s)\ln\left(\frac{1-\bar{p}_b}{\bar{p}_b}\right). \end{aligned} \tag{A.32}$$

$w_2(p)$ is a concave and increasing function for $p<\bar{p}_b$.

Combining (A.28) and (A.25), we can relate $w_2(p)$ and $v_2(p)$ through a simple linear transform:

$$w_2(p)=s+\frac{\eta}{\gamma}[\beta v_2(p)-w_2(p)]. \tag{A.33}$$

The condition $w_2(\bar{p}_b)=w_3(\bar{p}_b)$ implies continuity of $v(p)$ at $\bar{p}_b$ ($v_2(\bar{p}_b)=v_3(\bar{p}_b)$). To see this, combine (A.30) and (A.33):

$$v_2(\bar{p}_b)=\frac{\gamma+\eta}{\eta\beta}w_2(\bar{p}_b)-\frac{\gamma}{\eta\beta}s=\frac{\gamma+\eta}{\eta\beta}w_3(\bar{p}_b)-\frac{\gamma}{\eta\beta}s=v_3(\bar{p}_b).$$

We can also verify that the smooth-pasting condition holds for the current value function at $\bar{p}_b$:

$$\begin{aligned} w_2(\bar{p}_b) &= s+\frac{\eta}{\gamma}[\beta v_2(\bar{p}_b)-w_2(\bar{p}_b)] \\ &= s+\frac{\eta}{\gamma}[\beta v_3(\bar{p}_b)-w_3(\bar{p}_b)] \\ &= s+\frac{\eta}{\gamma}[\beta v_3(\bar{p}_b)-w_3(\bar{p}_b)]+[b_b(\bar{p}_b,w_3)-c(\bar{p}_b)]. \end{aligned}$$

Hence, at the upper threshold $b_b(\bar{p}_b,w_3)=c(\bar{p}_b)=b_b(\bar{p}_b,w_2)$. Using the definition of the expected benefit function and the value matching for $w(p)$, we have

$$w_2'(\bar{p}_b)=\frac{\gamma c(\bar{p}_b)}{\lambda\bar{p}_b(1-\bar{p}_b)}-\frac{\Omega s}{\bar{p}_b}+\frac{w_2(\bar{p}_b)}{\bar{p}_b}=\frac{\gamma c(\bar{p}_b)}{\lambda\bar{p}_b(1-\bar{p}_b)}-\frac{\Omega s}{\bar{p}_b}+\frac{w_3(\bar{p}_b)}{\bar{p}_b}=w_3'(\bar{p}_b).$$

We may use the explicit formula for w_2 to derive the unique belief $\underline{p}_b$ at which $w_2(p)$ reaches the level Ωs, so that S is played at all $p < \underline{p}_b$. In particular, $\underline{p}_b$ satisfies the following equation

$$\Omega g - \bar{C}_b^1 + \frac{\bar{C}_b^1 - \Omega s + \frac{\gamma}{\lambda}(s-f)}{\bar{p}_b} - \frac{\gamma}{\lambda}(g-s)\ln\left(\frac{1-\bar{p}_b}{\bar{p}_b}\right) = \frac{\gamma}{\lambda}\frac{(s-f)}{\underline{p}_b} - \frac{\gamma}{\lambda}(g-s)\ln\left(\frac{1-\underline{p}_b}{\underline{p}_b}\right) \tag{A.34}$$

and defines the strategy:

$$\alpha_b(p) = \begin{cases} 1 & \text{if} \quad p \geq \bar{p}_b \\ \tilde{\alpha}_b(p) & \text{if} \quad p \in [\underline{p}_b, \bar{p}_b) \\ 0 & \text{if} \quad p < \underline{p}_b \end{cases}, \tag{A.35}$$

where $\tilde{\alpha}_b(p)$ is defined by (A.33) and

$$v_2(p) = s + \tilde{\alpha}_b(p) \cdot [b_b(p, v_2) - c(p)].$$

Lemma A.2. *The function $\tilde{\alpha}_b(p)$, determining the interior allocation, is continuous, strictly increasing, with a bounded derivative on $[\underline{p}_b, \bar{p}_b)$. At $\underline{p}_b$, $\tilde{\alpha}_b(\underline{p}_b) = 0$, and* $\lim_{p\uparrow\bar{p}_b} \alpha_b(p) < 1$.

Proof. The continuation value function $v_2(p)$ admits two representations on the interval $[\underline{p}_b, \bar{p}_b)$. First, (A.33) links the current and the continuation value functions

$$w_2(p) = \frac{\gamma}{\gamma+\eta}s + \frac{\eta\beta}{\gamma+\eta}v_2(p). \tag{A.36}$$

Second, (A.24) implies that

$$v_2(p) = s + \tilde{\alpha}_b(p)\left[\frac{\lambda(1-p)}{\gamma}[s - v_2(p) + pv_2'(p)] - [s - gp - f(1-p)]\right].$$

Together with (A.36), this implies

$$w_2(p) - \frac{\gamma+\eta\beta}{\gamma+\eta}s = \tilde{\alpha}_b(p)\frac{\gamma+\eta(1-\beta)}{\gamma+\eta}[(s-f)(1-p)-(g-s)p]. \tag{A.37}$$

The left-hand side of (A.37) is a non-negative (zero at $p = \underline{p}_b$ and positive on $(\underline{p}_b, \bar{p}_b)$), continuous, strictly increasing function on the interval $[\underline{p}_b, \bar{p}_b)$. The coefficient $\frac{\gamma+\eta(1-\beta)}{\gamma+\eta}[(s-f)(1-p)-(g-s)p]$ is a positive, strictly decreasing function on $[\underline{p}_b, \bar{p}_b)$, so $\tilde{\alpha}_b(p)$ is a non-negative, continuous, strictly increasing function, and $\tilde{\alpha}_b(\underline{p}_b) = 0$.

At $\bar{p}_b$, value matching for the current value function holds: $\lim_{p\uparrow\bar{p}_b} w_2(p) = \lim_{p\downarrow\bar{p}_b} w_3(p)$, making (A.37) equivalent to

$$\lim_{p\uparrow\bar{p}_b} \tilde{\alpha}_b(p) = \lim_{p\downarrow\bar{p}_b} \frac{\gamma+\eta}{\gamma+\eta(1-\beta)} \cdot \frac{w_3(p)-\Omega s}{(s-f)-(g-f)p} = \frac{\eta^2\beta(1-\beta)}{(\gamma+\eta(1-\beta))(\gamma+\lambda+\eta\beta)} < 1.$$

The boundedness of the derivative follows from the positivity, boundedness, and bounded derivative of the right-hand side of expression (A.37), along with the boundedness and bounded derivative of the left-hand side, all on the interval $[\underline{p}_b, \bar{p}_b)$. □

The continuation value function

$$v_b(p) = \begin{cases} v_3(p) & \text{if} \quad p \geq \bar{p}_b \\ v_2(p) & \text{if} \quad p \in [\underline{p}_b, \bar{p}_b) \\ s & \text{if} \quad p < \underline{p}_b \end{cases},$$

and the current value function

$$w_b(p) = \begin{cases} w_3(p) & \text{if} \quad p \geq \bar{p}_b \\ w_2(p) & \text{if} \quad p \in [\underline{p}_b, \bar{p}_b) \\ \Omega s & \text{if} \quad p < \underline{p}_b \end{cases}$$

satisfy the system of the Bellman equations (A.24), (A.25) with the boundary conditions (A.26), (A.27), with the maximum of $w(p)$ achieved at $\alpha_b(p)$, given by (A.35). The best-response analysis demonstrates optimality for all p where $w(p)$ is smooth. Due to smooth pasting obtaining at $\bar{p}_b$, $w(p)$ is smooth on the whole interval $(\underline{p}_b, 1]$. Equilibrium uniqueness is then a result of the uniqueness of $\bar{p}_b$, $\underline{p}_b$ and the interior allocation $\alpha_b(p)$ for $p \in [\underline{p}_b, \bar{p}_b)$ as given in (A.37).

Negative Feedback – Fix $\alpha' : [0,1] \to [0,1]$, the future selves' strategy. The value functions then satisfy the Bellman system

$$v(p, \alpha') = s + \alpha'(p) \cdot [b_g(p, v) - c(p)], \tag{A.38}$$

$$w(p) = s + \frac{\eta}{\gamma}[\beta v(p, \alpha') - w(p)] + \max_{\alpha\in[0,1]} \alpha \cdot [b_g(p, w) - c(p)], \tag{A.39}$$

with the boundary conditions (A.26), (A.27) and with $w(1) = \Omega g$, $v(1) = g$. Denote the current self's best response by α_g^*. We again arrive at the best-response system:

$$\alpha_g^*(p) \begin{cases} = 1 & \text{if} \quad c(p) < b_g(w,p) \\ \in [0,1] & \text{if} \quad c(p) = b_g(w,p) \\ = 0 & \text{if} \quad c(p) > b_g(w,p) \end{cases} . \tag{A.40}$$

Note that playing S is strictly dominant at $p < p^*$. Hence, in any SMPE, there must exist $\underline{p}_g \geq p^*$ such that the current self is indifferent between S and R at $\underline{p}_g$, but plays S exclusively at all lower beliefs.

Lemma A.3. *In any SMPE,* $\underline{p}_g = p_g^1$.

Proof. In any SMPE, the current value can never exceed w_g^1, the solution of the single-player problem, since in the single-player problem, optimization takes place over both the present and the future. Therefore, $\underline{p}_g \geq p_g^1$. Suppose future selves plays a strategy $\tilde{\alpha}$ such that $\tilde{\alpha}(p) = 0$ for all $p \leq \tilde{p}$, where $\tilde{p} > p_g^1$. We show that a deviation by the current self to switching at a lower belief $p_g^w \in (p_g^1, \tilde{p})$ is profitable. This follows from the fact that playing R exclusively on the interval $(p_g^1, \tilde{p})$ will generate a current value $w(p) > s$. On the interval $(p_g^1, \tilde{p})$, $v(p) = s$ and the current value function under full experimentation satisfies

$$(\gamma + \lambda p + \eta)w(p) + \lambda p(1-p)w'(p) = (\gamma + \lambda\Omega)gp + \gamma f(1-p) + \eta\beta s.$$

The solution is

$$w_{dev}(p) = \frac{\gamma f + \eta\beta s}{\gamma + \eta} + \left[\frac{\gamma(g-f)}{\gamma+\eta} + \lambda\frac{\eta\beta(g-s)}{(\gamma+\eta)(\gamma+\lambda+\eta)}\right]p + \bar{C}_{dev}(1-p)\left(\frac{1-p}{p}\right)^{\frac{\gamma+\eta}{\lambda}},$$

where $\bar{C}_{dev}$ is a constant of integration, defined by the value matching condition at p_g^1: $w_{dev}(p_g^1) = \Omega s$. Moreover, at this point, the right derivative $(w_{dev})'_+(p_g^1) = 0$. The function $w_{dev}(p)$ is convex and strictly increasing, and hence $w_{dev}(p_g^w) > s$ for all $p_g^w \in (p_g^1, \tilde{p})$. □

Below $\underline{p}_g$, the continuation value function is $v(p) = s$ and the current self has the value function $w(p) = \Omega s$ (since $p < \underline{p}_g < p^m$, $c(p) > b_g(w,p) = 0$).

Next, on the indifference region where $c(p) = b_g(w,p)$, the value function $w_2(p)$ is the unique solution of the differential equation

$$\frac{\lambda p}{\gamma}\left[w(1)-w(p)-(1-p)w'(p)\right]=s-gp-f(1-p). \tag{A.41}$$

subject to the boundary condition $w_2(\underline{p}_g)=\Omega s$. This pins down the constant of integration:

$$w_2(p)=\frac{\gamma}{\lambda}(g-s)+\Omega g+\frac{\gamma}{\lambda}(s-f)(1-p)\ln\left(\frac{1-p}{p}\right)+\bar{C}_g^{eq}(1-p), \tag{A.42}$$
$$\bar{C}_g^{eq}=-\frac{\Omega(g-s)+\frac{\gamma}{\lambda}(g-s)}{1-\underline{p}_g}-\frac{\gamma}{\lambda}(s-f)\ln\left(\frac{1-\underline{p}_g}{\underline{p}_g}\right).$$

Smooth pasting holds automatically at $\underline{p}_g$, which can be verified by estimating (A.41) at $\underline{p}_g$ and applying the value-matching condition. The Bellman equation (A.39) and the best response (A.40) determine the relation between $w_2(p)$ and $v_2(p)$ and so the strategy $\tilde{\alpha}_g$ in the indifference region is defined by:

$$w_2(p)=\frac{\gamma}{\gamma+\eta}s+\frac{\eta\beta}{\gamma+\eta}v_2(p), \tag{A.43}$$

$$v_2(p)=s+\tilde{\alpha}_g(p)\cdot[b_g(p,v_2)-c(p)]. \tag{A.44}$$

Lemma A.4. *The function $\tilde{\alpha}_g(p)$ (given by (A.43) and (A.44)) is a continuous, strictly increasing, convex function on $[\underline{p}_g,p^c)$, where*

$$p^c=\frac{\gamma(s-f)}{\gamma(g-f)+\lambda\frac{\gamma}{\gamma+\eta(1-\beta)}(g-s)}<p^m,$$

and where $\tilde{\alpha}_g(\underline{p}_g)=0$, $\lim_{p\uparrow p^c}\tilde{\alpha}_g(p)=+\infty$.

Proof. Substitute (A.42) and the expression for $v_2(p)$ (given by (A.43)) into (A.44):

$$w_2(p)-\frac{\gamma+\eta\beta}{\gamma+\eta}s=\tilde{\alpha}_g(p)\left[\frac{\gamma+\eta(1-\beta)}{\gamma+\eta}(s-f)(1-p)-\frac{\gamma+\lambda+\eta(1-\beta)}{\gamma+\eta}(g-s)p\right] \tag{A.45}$$

The left-hand side of (A.45) is a non-negative (it takes zero value at $\underline{p}_g$ and is positive on $(\underline{p}_g,p^c)$), continuous, and strictly increasing function on $[\underline{p}_g,p^c)$. The linear coefficient $\left[\frac{\gamma+\eta(1-\beta)}{\gamma+\eta}(s-f)(1-p)-\frac{\gamma+\lambda+\eta(1-\beta)}{\gamma+\eta}(g-s)p\right]$ is a positive, strictly decreasing function on $[\underline{p}_g,\bar{p}_g)$. Hence, $\tilde{\alpha}_g(p)$ is a non-negative, continuous, strictly increasing function on $[\underline{p}_g,p^c)$.

At $p=\underline{p}_g$, $\tilde{\alpha}(\underline{p}_g)=0$ as the left-hand side turns to zero, while the coefficient on the right-hand side does not. When p approaches p^c, the coefficient on the right-hand side tends to 0 (whereas $w_2(p) - \frac{\gamma+\eta\beta}{\gamma+\eta}s$ does not), thus proving $\lim_{p\uparrow p^c} \tilde{\alpha}_g(p) = +\infty$.

Convexity follows from the fact that $\tilde{\alpha}_g$ can be represented as a ratio of a non-negative, increasing, convex function divided by a positive, decreasing, linear function. □

By Lemma A.4 and the intermediate value theorem, there exists a unique cut-off $\bar{p}_g < p^c < p^m$ such that $\tilde{\alpha}_g(\bar{p}_g) = 1$. We now conclude the proof of equilibrium uniqueness, by demonstrating that the upper threshold in any equilibrium is necessarily equal to $\bar{p}_g$. Were it higher, indifference could clearly not be maintained as $\bar{p}_g$ is the highest point at which $\tilde{\alpha}_g(p) \leq 1$. Were it lower, we will explicitly construct a profitable deviation for the current self.

Lemma A.5. *The following conditions are equivalent:*

1. $\tilde{\alpha}_g(\bar{p}_g) = 1$*;*

2. $\bar{p}_g$ *satisfies the following equation:*

$$\frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f)(1-\bar{p}_g)\ln\left(\frac{1-\bar{p}_g}{\bar{p}_g}\right) + \frac{\gamma+\eta}{\eta\beta}\bar{C}_g^{eq}(1-\bar{p}_g) = \\ = \frac{\gamma+\eta}{\eta\beta}(s-f)(1-\bar{p}_g) - \frac{\gamma+\lambda+\eta}{\eta\beta}(g-s)\left(\frac{\gamma}{\lambda}+\bar{p}_g\right) - (g-f)(1-\bar{p}_g). \tag{A.46}$$

Proof. In the interior allocation region, $\tilde{\alpha}_g(p)$ is given by the equation (A.45). Consider this equation at the point $\bar{p}_g$. Then, by definition, $\tilde{\alpha}_g(\bar{p}_g) = 1$. Substituting the explicit expression for $w_2(\bar{p}_g)$ from (A.42) yields the result. □

Lemma A.6. *There exists a unique equilibrium of the form:*

$$\alpha_g(p) = \begin{cases} 1 & \text{if} \quad p \geq \tilde{p}_g \\ \tilde{\alpha}_g(p) & \text{if} \quad p \in [\underline{p}_g, \tilde{p}_g) \\ 0 & \text{if} \quad p < \underline{p}_g \end{cases}, \tag{A.47}$$

where $\tilde{p}_g = \bar{p}_g$.

Proof. Suppose that the strategy α_g^v taking the form (A.47) constitutes a SMPE. Consider a deviation of the form

$$\alpha_g^w(p) = \begin{cases} 1 & \text{if} \quad p \geq p_g^w \\ 0 & \text{if} \quad p \in [\tilde{p}_g, p_g^w) \\ \tilde{\alpha}_g(p) & \text{if} \quad p \in [\underline{p}_g, \tilde{p}_g) \\ 0 & \text{if} \quad p < \underline{p}_g \end{cases}. \tag{A.48}$$

We prove that if $\tilde{p}_g < \bar{p}_g$, then both the continuation value and current value under the deviation will exhibit a convex kink at $\tilde{p}_g$, and hence there exists $\varepsilon > 0$ such that for all $p \in [\tilde{p}_g, \tilde{p}_g + \varepsilon)$, deviating to α_g^w is profitable.

Under α_g^v, the current value consists of three pieces, one in each of the three regions. When $p < \underline{p}_g$, all selves use S and so $v_1^v(p) = s$, $w_1^v(p) = \Omega s$. When $p \in [\underline{p}_g, \tilde{p}_g)$, all selves use the interior allocation $\tilde{\alpha}_g(p)$, uniquely defined through (A.43) and (A.44). The continuation value function is given by

$$v_2^v(p) = \frac{\gamma}{\lambda}\frac{\gamma+\lambda+\eta}{\eta\beta}(g-s) + g + \frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f)(1-p)\ln\left(\frac{1-p}{p}\right) + \frac{\gamma+\eta}{\eta\beta}\bar{C}_g^{eq}(1-p)$$

and the current value function satisfies the equation (A.41). When $p \in [\tilde{p}_g, 1]$, all selves play R and hence the continuation value solves (A.38) with $\alpha'(p) = 1$. The solution is:

$$v_3^v(p) = gp + f(1-p) + C(1-p)\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}} \tag{A.49}$$

for some constant C. The current value function is determined by the equation:

$$w_3^v(p) = gp + f(1-p) + \frac{\eta}{\gamma}[\beta v_3^v(p) - w_3^v(p)] + \frac{\lambda p}{\gamma}(\Omega g - w_3^v(p) - (1-p)(w_3^v)'(p)).$$

The solution is

$$w_3^v(p) = \Omega gp + \Omega f(1-p) + \beta C(1-p)\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}} + K(1-p)\left(\frac{1-p}{p}\right)^{\frac{\gamma+\eta}{\lambda}}.$$

Under the putative SMPE α_g^v, the continuation value function is continuous, so that $v_1^v(\underline{p}_g) = v_2^v(\underline{p}_g)$ and $v_2^v(\tilde{p}_g) = v_3^v(\tilde{p}_g)$. The second value matching condition defines the constant of integration C in (A.49). Furthermore, the current value function is then necessarily smooth, satisfying both value matching and smooth pasting conditions at the same points: $w_1^v(\underline{p}_g) = w_2^v(\underline{p}_g)$, $w_2^v(\tilde{p}_g) = w_3^v(\tilde{p}_g)$,

$(w_1^v)'(\underline{p}_g) = (w_2^v)'(\underline{p}_g)$, and $(w_2^v)'(\tilde{p}_g) = (w_3^v)'(\tilde{p}_g)$.

Now suppose the current self employs the strategy $\alpha_g^w(p)$. This divides the interval $[0,1]$ into four distinct regions, given by (A.48). Examine the two intermediate regions: $[\underline{p}_g, \tilde{p}_g)$, where all selves play $\tilde{\alpha}_g(p)$, and $[\tilde{p}_g, p_g^w)$, where future selves play R, while the current self deviates to S. When $p \in [\underline{p}_g, \tilde{p}_g)$, the continuation and current value functions are given by α_g^v, so that $v_2^v(p) = v_2^w(p)$ and $w_2^v(p) = w_2^w(p)$. When $p \in [\tilde{p}_g, p_g^w)$, $v_3^v(p) = v_3^w(p)$, and from the Bellman equation (A.39), the current value function satisfies the equation

$$w_3^w(p) = s + \frac{\eta}{\gamma}[\beta v_3^w(p) - w_3^w(p)]. \tag{A.50}$$

To demonstrate that α_g^w constitutes a profitable deviation from α_g^v, we compare the right derivatives of the current value functions at point $\tilde{p}_g$. If $w'_+(p) \equiv (w_3^v)'(\tilde{p}_g) < (w_3^w)'(\tilde{p}_g)$, then there exists $\varepsilon > 0$ such that for all $p \in [\tilde{p}_g, \tilde{p}_g + \varepsilon)$, $w_3^w(p) > w_3^v(p)$ and so the deviation α_g^w is profitable.

In the interior allocation region $[\underline{p}_g, \tilde{p}_g)$, the current value functions for the two strategies α_g^v and α_g^w coincide. Moreover, they are linear transformations of the continuation value function, leading to:

$$(w_2^i)'(p) = \frac{\eta\beta}{\gamma+\eta}(v_2^i)'(p), \quad i \in \{v, w\}.$$

Hence, we can express $(w_3^v)'(\tilde{p}_g) = (w_2^v)'(\tilde{p}_g) = \frac{\eta\beta}{\gamma+\eta}(v_2^v)'(\tilde{p}_g)$. From (A.50), $(w_3^w)'(p) = \frac{\eta\beta}{\gamma+\eta}(v_3^w)'(p)$ and so at $\tilde{p}_g$, $(w_3^w)'(\tilde{p}_g) = \frac{\eta\beta}{\gamma+\eta}(v_3^v)'(\tilde{p}_g)$. Therefore,

$$(w_3^v)'(\tilde{p}_g) < (w_3^w)'(\tilde{p}_g) \Longleftrightarrow (v_2^v)'(p) < (v_3^v)'(\tilde{p}_g).$$

To prove this inequality, consider the difference

$$\begin{aligned}
&(v_2^v)'(p) - (v_3^v)'(\tilde{p}_g) \\
&\quad = -\frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f)\frac{1}{\tilde{p}_g} - \frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f)\ln\left(\frac{1-\tilde{p}_g}{\tilde{p}_g}\right) - \frac{\gamma+\eta}{\eta\beta}\bar{C}_g^{eq} - (g-f) + C\frac{\gamma+\lambda\tilde{p}_g}{\lambda\tilde{p}_g}\left(\frac{1-\tilde{p}_g}{\tilde{p}_g}\right)^{\frac{\gamma}{\lambda}} \\
&\quad = \frac{\gamma}{\eta\beta}\frac{\gamma+\lambda+\eta}{\lambda}(g-s)(\frac{\gamma}{\lambda}+\tilde{p}_g) + (g-f)(1-\tilde{p}_g) + \frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f)(1-\tilde{p}_g)\ln\left(\frac{1-\tilde{p}_g}{\tilde{p}_g}\right) - \\
&\qquad - \frac{\gamma+\eta}{\eta\beta}(s-f)(1-\tilde{p}_g) + \frac{\gamma+\eta}{\eta\beta}\bar{C}_g^{eq}(1-\tilde{p}_g) \equiv F(\tilde{p}_g).
\end{aligned}$$

By (A.46), $F(\bar{p}_g) = 0$. To prove that $F(\tilde{p}_g) < 0$ for $\tilde{p}_g < \bar{p}_g$, we demonstrate that $F(\tilde{p}_g)$ is an increasing function of $\tilde{p}_g$.

$$(F(\tilde{p}_g))' = \underbrace{\frac{\gamma(\gamma+\lambda+\eta)+\lambda\eta(1-\beta)}{\lambda\eta\beta}(g-f) + \frac{\gamma+\lambda+\eta\beta}{\eta\beta}(g-s) - \frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f)\frac{1}{\tilde{p}_g}}_{>0} +$$
$$+ \frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f)\underbrace{\ln\left(\frac{\tilde{p}_g}{\underline{p}_g}\frac{1-\underline{p}_g}{1-\tilde{p}_g}\right)}_{\geq 0} > 0,$$

where both inequalities are direct consequences of $\tilde{p}_g \geq \underline{p}_g$. Hence, for $\tilde{p}_g < \bar{p}_g$, $(w_3^v)'(\tilde{p}_g) < (w_3^w)'(\tilde{p}_g)$, leading to a contradiction. Thus, $\tilde{p}_g = \bar{p}_g$ in a SMPE, which completes the proof of uniqueness.

□

Given the equilibrium strategy (A.47), we can now define the value functions for the experimentation region ($p \geq \bar{p}_g$). In this region, the value functions $v_3(p)$ and $w_3(p)$ satisfy the system of the differential equations (A.17), (A.18) with the boundary conditions $v_2(\bar{p}_g) = v_3(\bar{p}_g)$, $w_2(\bar{p}_g) = w_3(\bar{p}_g)$.

$$v_3(p) = gp + f(1-p) + \bar{\bar{C}}_g^{eq}(1-p)\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}},$$
$$w_3(p) = \Omega gp + \Omega f(1-p) + \beta\bar{\bar{C}}_g^{eq}(1-p)\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}} + \bar{\bar{K}}_g^{eq}(1-p)\left(\frac{1-p}{p}\right)^{\frac{\gamma+\eta}{\lambda}},$$

where the constants are given by

$$\bar{\bar{C}}_g^{eq} = \frac{\gamma+\eta}{\eta\beta}\frac{(s-f) - \left[(g-f) + \frac{\lambda}{\gamma+\eta}(g-s)\right]\bar{p}_g}{1-\bar{p}_g}\left(\frac{\bar{p}_g}{1-\bar{p}_g}\right)^{\frac{\gamma}{\lambda}} \tag{A.51}$$
$$\bar{\bar{K}}_g^{eq} = \Bigg[-\frac{\gamma(1-\beta)}{\gamma+\eta}[g\bar{p}_g + f(1-\bar{p}_g)] + \frac{\gamma}{\eta}s -$$
$$-\frac{\gamma}{\eta}\left[\frac{\gamma}{\lambda}(g-s) + \Omega g + \frac{\gamma}{\lambda}(s-f)(1-\bar{p}_g)\ln\left(\frac{1-\bar{p}_g}{\bar{p}_g}\right) + \bar{C}_g^{eq}(1-\bar{p}_g)\right]\Bigg]\frac{1}{\bar{p}_g}\left(\frac{\bar{p}_g}{1-\bar{p}_g}\right)^{\frac{\gamma+\lambda+\eta}{\lambda}}.$$

The first derivative of the current value function is continuous at the upper threshold $\bar{p}_g$. The argument resembles the one for the positive feedback case. Two value-matching conditions and the Bellman equation (A.39) provide us with $b_g(\bar{p}_g, w_3) = c(\bar{p}_g) = b_g(\bar{p}_g, w_2)$. This equality, together with $w_2(\bar{p}_g) = w_3(\bar{p}_g)$, implies that the smooth pasting condition $w_2'(\bar{p}_g) = w_3'(\bar{p}_g)$ holds.

Lemma A.7. *Smooth pasting holds for the continuation value function at the upper threshold:* $v_2'(\bar{p}_g) = v_3'(\bar{p}_g)$.

Proof. Due to the continuity of $v(p)$ and $\alpha_g(p)$ at the upper threshold $\bar{p}_g$:

$$\begin{aligned} v_2(\bar{p}_b) &= s + [b_g(\bar{p}_g, v_2) - c(\bar{p}_g)] \\ &= g\bar{p}_g + f(1-\bar{p}_g) + \frac{\lambda \bar{p}_g}{\gamma}(g - v_2(\bar{p}_g) - (1-\bar{p}_g)v_3'(\bar{p}_g)) \\ &= g\bar{p}_g + f(1-\bar{p}_g) + \frac{\lambda \bar{p}_g}{\gamma}(g - v_3(\bar{p}_g) - (1-\bar{p}_g)v_2'(\bar{p}_g)). \end{aligned}$$

Again, from the continuity

$$\begin{aligned} v_2(\bar{p}_b) &= v_3(\bar{p}_b) \\ &= s + [b_g(\bar{p}_g, v_3) - c(\bar{p}_g)] \\ &= g\bar{p}_g + f(1-\bar{p}_g) + \frac{\lambda \bar{p}_g}{\gamma}(g - v_3(\bar{p}_g) - (1-\bar{p}_g)v_3'(\bar{p}_g)). \end{aligned}$$

Hence, $v_2'(\bar{p}_g) = v_3'(\bar{p}_g)$. □

Therefore, the continuation value function is

$$v_g(p) = \begin{cases} v_3(p) & \text{if} \quad p \geq \bar{p}_g \\ v_2(p) & \text{if} \quad p \in [\underline{p}_g, \bar{p}_g) \\ s & \text{if} \quad p < \underline{p}_g \end{cases},$$

the current value function has a form

$$w_g(p) = \begin{cases} w_3(p) & \text{if} \quad p \geq \bar{p}_g \\ w_2(p) & \text{if} \quad p \in [\underline{p}_g, \bar{p}_g) \\ \Omega s & \text{if} \quad p < \underline{p}_g \end{cases}.$$

They satisfy the Bellman system of equations (A.38), (A.39) with the boundary conditions (A.26), (A.27). $w(p)$ attains its maximum at $\alpha_g(p)$, given by (A.47). The optimality follows from the best-response analysis, and the uniqueness is a result of Lemma A.6.

Transparent Feedback – The Bellman system that defines the equilibrium solution under transparent feedback is

$$v(p, \alpha') = s + \alpha'(p) \cdot [b_f(p, v) - c(p)], \tag{A.52}$$

$$w(p) = s + \frac{\eta}{\gamma}[\beta v(p, \alpha') - w(p)] + \max_{\alpha \in [0,1]} \alpha \cdot [b_f(p, w) - c(p)], \tag{A.53}$$

with the boundary conditions (A.26), (A.27) and with $w(1) = \Omega g$, $w(0) = \Omega s$, $v(1) = g$, $v(0) = s$. The absence of the belief's drift makes the system linear in $w(p)$. The best-response is again given by

$$\alpha_f^*(p) \begin{cases} = 1 & \text{if} \quad c(p) < b_f(w, p) \\ \in [0, 1] & \text{if} \quad c(p) = b_f(w, p) \cdot \\ = 0 & \text{if} \quad c(p) > b_f(w, p) \end{cases} \tag{A.54}$$

Due to linearity, in the indifference region $(c(p) = b_f(w, p))$, the current value function is easily solved for:

$$w_2(p) = \Omega s - \frac{\gamma}{\lambda}(s - f) + \left[\frac{\gamma}{\lambda}(g - f) + \Omega(g - s)\right] p. \tag{A.55}$$

This gives us the expression for $v_2(p)$ on the interior region, which is unique due to the uniqueness of $w_2(p)$:

$$\begin{aligned} w_2(p) &= s + \frac{\eta}{\gamma}[\beta v_2(p) - w_2(p)], \\ v_2(p) &= s - \frac{\gamma}{\lambda}\frac{\gamma + \eta}{\eta\beta}(s - f) + \frac{\gamma + \eta}{\eta\beta}\left[\frac{\gamma}{\lambda}(g - f) + \Omega(g - s)\right] p. \end{aligned}$$

In a SMPE, the optimal action of the current self must coincide with the action α' chosen by the future selves. If all selves exclusively play S, the current value function achieves its minimum of Ωs, and similarly, the continuation value is s. The threshold $\underline{p}_f$ is determined thus by the condition $w_2(p) = \Omega s$. More specifically,

$$\underline{p}_f = \frac{\gamma(s - f)}{\gamma(g - f) + \lambda\Omega(g - s)} \equiv \underline{p}_g.$$

If all selves exclusively play R, the current value function $w_3(p)$ aligns with (A.22). The continuation value function $v_3(p)$ is given by the equation (A.15). Simple algebra confirms that the point of indifference between an interior allocation and full experimentation $(w_2(p) = w_3(p))$ is reached at

$$\bar{p}_f = \frac{\gamma(s - f)}{\gamma(g - f) + \lambda\frac{\gamma(\gamma+\lambda+\eta\beta)}{\gamma(\gamma+\lambda+\eta)+\lambda\eta(1-\beta)}(g - s)} \equiv \bar{p}_b, \tag{A.56}$$

completing the definition of the strategy:

$$\alpha_f(p) = \begin{cases} 1 & \text{if} \quad p \geq \bar{p}_b \\ \tilde{\alpha}_f(p) & \text{if} \quad p \in [\underline{p}_g, \bar{p}_b) \\ 0 & \text{if} \quad p < \underline{p}_g \end{cases}, \tag{A.57}$$

where $\tilde{\alpha}_f(p)$ satisfies

$$v_2(p) = s + \tilde{\alpha}_f(p) \cdot [b_f(p, v_2) - c(p)].$$

Lemma A.8. *The function $\tilde{\alpha}_f(p)$ is a continuous, strictly increasing, convex function on $[\underline{p}_g, p^c)$ such that $\tilde{\alpha}_f(\underline{p}_g) = 0$, $\lim_{p\uparrow p^c} \tilde{\alpha}_f(p) = +\infty$, $\tilde{\alpha}_f(\bar{p}_g) = 1$.*

Proof. From the equivalence of two representations of $v_2(p)$, the characterization of $\tilde{\alpha}_f$ is given by

$$w_2(p) - \frac{\gamma + \eta\beta}{\gamma + \eta} s = \tilde{\alpha}_f(p) \left[\frac{\gamma + \eta(1-\beta)}{\gamma + \eta}(s - f)(1 - p) - \frac{\gamma + \lambda + \eta(1-\beta)}{\gamma + \eta}(g - s)p \right]. \tag{A.58}$$

The left-hand side of (A.58) is non-negative (specifically, it is zero at $p = \underline{p}_g$ and positive for all $p \in (\underline{p}_g, p^c)$), linear, and strictly increasing. The coefficient on the right-hand side is a positive decreasing function that reaches 0 at p^c. Hence, $\tilde{\alpha}_f(p)$ is a non-negative, strictly increasing function on $[\underline{p}_g, p^c)$, attaining 0 at $p = \underline{p}_g$. In the limit $p \uparrow p^c$, the coefficient on the right-hand side tends to 0, while the function on the left-hand side remains positive, ensuring the limit $\lim_{p\uparrow p^c} \tilde{\alpha}_f(p) = +\infty$. The result $\tilde{\alpha}_f(\bar{p}_b) = 1$ follows from straightforward algebra (evaluating (A.58) at $\bar{p}_b$).

Convexity follows from the fact that we can represent $\tilde{\alpha}_f$ as a ratio of a non-negative, increasing, linear function divided by a positive, decreasing, linear function. □

The continuation value function

$$v_f(p) = \begin{cases} v_3(p) & \text{if} \quad p \geq \bar{p}_b \\ v_2(p) & \text{if} \quad p \in [\underline{p}_g, \bar{p}_b) \\ s & \text{if} \quad p < \underline{p}_g \end{cases},$$

the current value function

$$w_f(p) = \begin{cases} w_3(p) & \text{if} \quad p \geq \bar{p}_b \\ w_2(p) & \text{if} \quad p \in [\underline{p}_g, \bar{p}_b) \\ \Omega s & \text{if} \quad p < \underline{p}_g \end{cases},$$

satisfy the system of the Bellman equations (A.52), (A.53) with the boundary conditions (A.26), (A.27), achieving the maximum at $\alpha_f(p)$.

A.5.2 Part 2

Consider the strategy $\alpha_g(p)$. From Lemma A.4, it follows that $\alpha_g(p)$ is continuous on the interval $[0, 1]$, strictly increasing on the interval $[\underline{p}_g, \bar{p}_g)$ (since, by construction, $\bar{p}_g < p^c$), and $\alpha_g(\underline{p}_g) = 0$. Within the intervals $[0, \bar{p}_g) \cup (\bar{p}_g, 1]$, $\alpha_g(p)$ remains constant, and so have a bounded derivative. Also by Lemma A.4, within the interval $[\underline{p}_g, \bar{p}_g)$, the derivative increases and approaches infinity only as $p \uparrow p^c > \bar{p}_g$, remaining bounded within $[\underline{p}_g, \bar{p}_g)$. Hence, $\alpha_g(p)$ is Lipschitz continuous on $[0, 1]$.

The argument for $\alpha_f(p)$ follows the same principle, applying Lemma A.8 instead of Lemma A.4.

According to Lemma A.2, $\alpha_b(p)$ is continuous on $[\underline{p}_b, \bar{p}_b)$ with $\alpha_b(\underline{p}_b) = 0$, and $\alpha_b(\bar{p}_b) < 1$, thereby exhibiting a discontinuity at $p = \bar{p}_b$. On the interval $(\bar{p}_b, 1]$, the function $\alpha_b(p)$ is constant, thus Lipschitz continuous. Within the interval $[0, \bar{p}_b)$, $\alpha_b(p)$ remains continuous, with a bounded derivative (Lemma A.2). Therefore, $\alpha_b(p)$ is also Lipschitz continuous on $[0, \bar{p}_b)$.

A.5.3 Part 3

Using (A.30) and the fact that

$$\frac{\gamma + \eta\beta}{\gamma + \eta} > \frac{\gamma(\gamma + \lambda + \eta\beta)}{\gamma(\gamma + \lambda + \eta) + \lambda\eta(1 - \beta)} > 0,$$

we obtain the order $\underline{p}_g < \bar{p}_b < p^m$. By construction, $\underline{p}_i \leq \bar{p}_i$, $i \in \{b, g\}$. Moreover, $\bar{p}_i \leq p^m$ and $\underline{p}_i \geq p^*$, leaving two inequalities for the verification. To show that $\underline{p}_b < \underline{p}_g$, introduce the auxiliary functions:

$$G_1(p) \equiv \Omega g - \bar{C}_b^1 + \frac{\bar{C}_b^1 - \Omega s + \frac{\gamma}{\lambda}(s - f)}{p} - \frac{\gamma}{\lambda}(g - s) \ln\left(\frac{1 - p}{p}\right)$$
$$G_2(p) \equiv \frac{\gamma}{\lambda} \frac{(s - f)}{p} - \frac{\gamma}{\lambda}(g - s) \ln\left(\frac{1 - p}{p}\right).$$

Then (A.34), the expression that defines the lower threshold for the positive feedback, transforms into $G_1(\bar{p}_b) = G_2(\underline{p}_b)$. $G_2(p)$ is a decreasing function on the interval $[p^*, p^m)$:

$$(G_2(p))' = \frac{\gamma}{\lambda} \frac{[(s-f)+(g-s)]\,p-(s-f)}{(1-p)p^2} < 0.$$

The explicit comparison $G_2(\underline{p}_g) < G_1(\bar{p}_b) = G_2(\underline{p}_b)$ leads to the desired inequality $\underline{p}_g > \underline{p}_b$. To show $\bar{p}_b < \bar{p}_g$, introduce

$$\begin{aligned} H_1(p) &\equiv \frac{\gamma+\eta}{\eta\beta} \frac{\Omega(g-s)+\frac{\gamma}{\lambda}(g-s)}{1-p} + \frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f)\ln\left(\frac{1-p}{p}\right) \\ H_2(p) &\equiv \frac{\gamma+\lambda+\eta}{\eta\beta}(g-s)\frac{\gamma+\lambda p}{\lambda(1-p)} + (g-f) + \frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f)\ln\left(\frac{1-p}{p}\right) - \frac{\gamma+\eta}{\eta\beta}(s-f) \end{aligned}$$

and observe that $H_2(p)$ is an increasing function on the interval $[p^*, p^m]$:

$$(H_2(p))' = \frac{\gamma+\eta}{\eta\beta} \frac{\left[(\gamma+\lambda)\frac{\gamma+\lambda+\eta}{\gamma+\eta}(g-s)+\gamma(s-f)\right]p-\gamma(s-f)}{p(1-p)^2} > 0.$$

The expression for the upper boundary of the interior region under negative feedback (A.46) is equivalent to $H_1(\underline{p}_g) = H_2(\bar{p}_g)$. From straightforward algebra, it follows that $H_2(\bar{p}_b) < H_1(\underline{p}_g) = H_2(\bar{p}_g)$ and $\bar{p}_b < \bar{p}_g$.

A.6 Proof of Lemma 2

$\underline{p}_b < \underline{p}_f = \underline{p}_g$ follows from Theorem 1, part 3. The equality $\underline{p}_g = p_g^1$ follows from Lemma A.3.

A.7 Proof of Lemma 3

The naive solution assumes future selves play the threshold strategy $\hat{\alpha}_i(p) = \mathbb{I}_{p \geq p^*}$, $i \in \{b, f\}$, $\hat{\alpha}_g(p) = \mathbb{I}_{p > p^*}$ (depending on the monitoring mode) and the current self responds optimally. The current self solves the Bellman system that consists of (16), (17), (A.26), (A.27), where the strategy of future selves is $\alpha'(p) = \hat{\alpha}_i(p)$, $i \in \{b, g, f\}$. Due to the linearity of the system in α, the system admits a threshold solution. The naive threshold p_i^n is not below the efficient one $p_i^n \geq p^*$, splitting the interval of beliefs into three regions: all selves play S, the current self plays S and the future selves experiment, all selves play R.

Positive feedback – Follows immediately from Lemma A.1.

Negative feedback – The future selves' value function is given by (A.11). In the first region ($p \leq p^*$), the current self has the value Ωs. In the third region, where all selves experiment ($p > p_g^n$), the current value function satisfies the differential equation

$$(\gamma + \lambda p + \eta)w(p) + \lambda p(1-p)w'(p) = \gamma g p + \gamma f(1-p) + \eta\beta v(p) + \lambda\Omega g p,$$
$$v(p) = v(p) = gp + f(1-p) + C_g^{br}(1-p)\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}}. \quad \text{(A.59)}$$

The solution has a form

$$w_3(p) = \Omega g p + \Omega f(1-p) + \beta C_g^{br}(1-p)\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}} + K_g^{br}(1-p)\left(\frac{1-p}{p}\right)^{\frac{\gamma+\eta}{\lambda}},$$

where C_g^{br} and K_g^{br} are some constants. In the region, where the current self plays S and the future selves experiment, $p \in (p^*, p_g^n]$, the continuation value function remains (A.59), whereas the current value changes:

$$w_2(p) = s + \frac{\eta}{\gamma}[\beta v(p) - w_2(p)]$$
$$w_2(p) = \frac{\eta\beta}{\gamma+\eta}(g-f)p + \frac{\gamma s + \eta\beta f}{\gamma+\eta} + \frac{\eta\beta}{\gamma+\eta}C_g^{br}(1-p)\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}}.$$

The optimal threshold is determined by the standard value matching and smooth pasting conditions:

$$w_2(p_g^n) = w_3(p_g^n),$$
$$(w_2)'(p_g^n) = (w_3)'(p_g^n).$$

This generates the following implicit expression for the naive threshold:

$$\frac{\gamma+\eta}{\eta\beta}\frac{(s-f) - [(g-f) + \frac{\lambda}{\gamma+\eta}(g-s)]p_g^n}{1-p_g^n}\left(\frac{p_g^n}{1-p_g^n}\right)^{\frac{\gamma}{\lambda}} = \frac{(s-f)-(g-f)p^*}{1-p^*}\left(\frac{p^*}{1-p^*}\right)^{\frac{\gamma}{\lambda}}. \quad \text{(A.60)}$$

Therefore, the optimal strategy for the current self is to play S for $p \leq p_g^n$ and to play R when $p > p_g^n$, where p_g^n is defined by (A.60).

Corollary A.1. *Represent the coefficients $\bar{\bar{C}}_g^{eq}$ (from (A.51)) and $\bar{C}_g$ (from (A.12)) as the functions of $\bar{p}_g$ and p^* respectively. Then (A.60) is equivalent to $\bar{\bar{C}}_g^{eq}(p_g^n) = \bar{C}_g(p^*)$.*

To verify the inequality $\bar{p}_g > p_g^n$, consider the functions $\bar{\bar{C}}_g^{eq}(p)$ and $\bar{C}_g(p)$. $\bar{\bar{C}}_g^{eq}(p)$ is decreasing on the interval $[p^*, p^m]$:

$$\text{sign}(\bar{\bar{C}}_g^{eq}(p))' = \text{sign}\left[\frac{\gamma+\eta}{\eta\beta}\frac{\frac{\gamma}{\lambda}(s-f) - [\frac{\gamma}{\lambda}(g-f) + \left(1+\frac{\gamma+\lambda}{\gamma+\eta}\right)(g-s)]p}{p(1-p)^2}\left(\frac{p}{1-p}\right)^{\frac{\gamma}{\lambda}}\right]$$
$$= \text{sign}\left[\frac{\gamma(s-f)}{\gamma(g-f) + \lambda\left(1+\frac{\gamma+\lambda}{\gamma+\eta}\right)(g-s)} - p\right] < 0,$$

where the last inequality is derived from the fact that

$$p \geq p^* > \frac{\gamma(s-f)}{\gamma(g-f) + \lambda\left(1+\frac{\gamma+\lambda}{\gamma+\eta}\right)(g-s)}.$$

$\bar{p}_g > p_g^n$ is equivalent to $\bar{\bar{C}}_g^{eq}(\bar{p}_g) < \bar{C}_g(p^*) = \bar{\bar{C}}_g^{eq}(p_g^n)$, which follows from $v_g(p) < v^*(p)$ on the interval $(\bar{p}_g, 1)$ (efficiency of the exponential benchmark).

Transparent feedback – In the transparent feedback mode, the current value functions for $p < p^*$ and $p \geq p_f^n$ admit the same representation as in the positive feedback case. Moreover, since $v_b^*(p) = v_f^*(p)$, we also have in the region $[p^*, p_f^n)$, the current value function coincides with that for the positive feedback. Together with the same optimality conditions, it leads to $p_f^n = p_b^n$ and so $p_f^n = \bar{p}_f$.

A.8 Proof of Lemma 4

In the proof of Lemma 3, we introduced the functions $\bar{\bar{C}}_g^{eq}(p)$ and $\bar{C}_g(p)$ such that $\bar{C}_g^1(p_g^n) = \bar{C}_g(p^*)$, which is equivalent to the equation (A.60). From straightforward algebra, we find that $\bar{\bar{C}}_g^{eq}(p_b^n) > \bar{C}_g(p^*) = \bar{\bar{C}}_g^{eq}(p_g^n)$. Since $\bar{\bar{C}}_g^{eq}(p)$ is a decreasing function on the interval $[p^*, p^m]$, we conclude that $p_b^n < p_g^n$.

A.9 Proof of Proposition 4

When $p \in [0, \underline{p}_b]$, regardless of the feedback mode, the continuation value function equals s. At $p = 1$, $v_b(1) = v_f(1) = v_g(1) = g$, confirming the equality $\mathcal{W}(\alpha_b; p) = \mathcal{W}(\alpha_f; p) = \mathcal{W}(\alpha_g; p)$ (proving part 4).

We prove the remaining statements in several steps. First, we show that on the interval $[\bar{p}_b, 1)$, $\mathcal{W}(\alpha_b; p) = \mathcal{W}(\alpha_f; p)$. Second, we demonstrate that the strict inequality $\mathcal{W}(\alpha_b; p) > \mathcal{W}(\alpha_f; p)$ holds for $p \in (\underline{p}_b, \bar{p}_b)$. In the final step we compare the welfare for transparent feedback and negative feedback, showing that $\mathcal{W}(\alpha_f; p) = \mathcal{W}(\alpha_g; p)$ on $(\underline{p}_b, \underline{p}_g]$ and $\mathcal{W}(\alpha_f; p) > \mathcal{W}(\alpha_g; p)$ on $(\underline{p}_g, 1)$.

Consider the interval $[\bar{p}_b, 1)$. As verified, $\bar{p}_f = \bar{p}_b$, and above this threshold, the value functions in both modes exhibit the same form $v_i(p) = gp + \frac{\gamma f + \lambda s}{\gamma + \lambda}(1-p)$, $i \in \{b, f\}$. Hence, both positive and transparent feedback deliver the same welfare: $\mathcal{W}(\alpha_b; p) = \mathcal{W}(\alpha_f; p)$.

Next, consider the interval $(\underline{p}_b, \bar{p}_b)$. We show that v_b is strictly concave on this interval, while v_f is weakly convex, hence proving the desired result.

Lemma A.9. *$v_b(p)$ is a strictly concave function on the interval $(\underline{p}_b, \bar{p}_b)$.*

Proof. Take the second derivative of the value function and verify the sign on the target interval

$$\begin{aligned}(v_b(p))'' &= \left(\frac{\gamma+\eta}{\eta\beta}\frac{\gamma}{\lambda}(g-s)\ln\left(\frac{1-p}{p}\right) - \frac{\gamma+\eta}{\eta\beta}\frac{\gamma}{\lambda}(g-s)\frac{1}{1-p} + \frac{\gamma+\eta}{\eta\beta}\bar{C}_b^{eq}\right)' \\ &= -\frac{\gamma+\eta}{\eta\beta}\frac{\frac{\gamma}{\lambda}(g-s)}{p(1-p)^2} < 0.\end{aligned}$$

□

Lemma A.10. *The value function for transparent feedback is piecewise linear, weakly convex on the interval $(\underline{p}_b, \bar{p}_b)$, and weakly concave on the interval $(\underline{p}_g, 1)$.*

Proof. The value function for transparent feedback is as follows:

$$v_f(p) = \begin{cases} gp + \frac{\gamma f + \lambda s}{\gamma + \lambda}(1-p) & \text{if} \quad p \geq \bar{p}_b \\ s - \frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f) + \frac{\gamma+\eta}{\eta\beta}\left[\frac{\gamma}{\lambda}(g-f) + \Omega(g-s)\right]p & \text{if} \quad p \in [\underline{p}_g, \bar{p}_b) \\ s & \text{if} \quad p < \underline{p}_g \end{cases}.$$

This confirms the result. □

Given that the value functions match at the boundary points of the target interval, $v_f(\underline{p}_b) = v_b(\underline{p}_b)$ and $v_f(\bar{p}_b) = v_b(\bar{p}_b)$, the inequality $v_f(p) < v_b(p)$ holds for interior points $p \in (\underline{p}_b, \bar{p}_b)$ (as per Lemma A.9 and Lemma A.10). $v_f(p) = v_g(p) = s$ within the interval $[\underline{p}_b, \underline{p}_g]$, proving $\mathcal{W}(\alpha_f; p) = \mathcal{W}(\alpha_g; p)$ (completing the proof for part 3). Finally, consider the interval $(\underline{p}_g, 1)$.

Lemma A.11. *$v_g(p)$ is a convex function on the interval $(\underline{p}_g, 1)$.*

Proof. Consider first $v_g(p)$ on the interval $(\underline{p}_g, \bar{p}_g]$ and take the second derivative.

$$(v(p))'' = \left(-\frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f)\frac{1}{p} - \frac{\gamma}{\lambda}\frac{\gamma+\eta}{\eta\beta}(s-f)\ln\left(\frac{1-p}{p}\right) - \frac{\gamma+\eta}{\eta\beta}\bar{C}_g^{eq}\right)' =$$
$$= \frac{\gamma+\eta}{\eta\beta}\frac{\frac{\gamma}{\lambda}(s-f)}{(1-p)p^2} > 0.$$

so the function $v_g(p)$ is convex on this interval. Now consider $v_g(p)$ on the interval $(\bar{p}_g, 1)$ and show that it is also convex:

$$(v(p))'' = \left(g - f - \bar{\bar{C}}_g^{eq}\frac{\gamma+\lambda p}{\lambda p}\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}}\right)' = \bar{\bar{C}}_g^{eq}\frac{\gamma(\gamma+\lambda)}{\lambda^2 p(1-p)}\left(\frac{1-p}{p}\right)^{\frac{\gamma}{\lambda}}$$

so the sign of the derivative depends on the sign of the constant $\bar{\bar{C}}_g^{eq}$, defined by (A.51).

The positivity of the coefficient $\bar{\bar{C}}_g^{eq} > 0$ is equivalent to $\bar{p}_g < p^c$. This inequality follows from the construction of the upper threshold $\bar{p}_g$ (Lemma A.4, Lemma A.5). On $(\underline{p}_g, 1)$, $v_g(p)$ is smooth. Smoothness on each interval $(\underline{p}_g, \bar{p}_g)$ and $(\bar{p}_g, 1)$ separately follows from the explicit form of the functions, smoothness at $\bar{p}_g$ follows from Lemma A.7. This ensures convexity on $(\underline{p}_g, 1)$. □

At the boundary points the value of the value functions coincide $v_f(\underline{p}_g) = v_g(\underline{p}_g)$, $v_f(1) = v_g(1)$, so by Lemma A.10 and by Lemma A.11, $v_f(p) > v_g(p)$ on $(\underline{p}_g, 1)$. This proves $\mathcal{W}(\alpha_f; p) > \mathcal{W}(\alpha_g; p)$ for $p \in (\underline{p}_g, 1)$.

A.10 Proof of Lemma 5

Part 1 – For a fixed η, $\alpha_b(p)$ is given by (A.35), with $\tilde{\alpha}_b(p)$ on the interval $(\underline{p}_b, \bar{p}_b)$ defined by (A.37), $\alpha_g(p)$ is given by (A.47) with $\tilde{\alpha}_g(p)$ on the interval $(\underline{p}_g, \bar{p}_g)$, defined by (A.45), and $\alpha_f(p)$ is given by (A.57) with $\tilde{\alpha}_f(p)$ on the interval $(\underline{p}_g, \bar{p}_b)$, defined by (A.58). As we take the limit, they converge to:

$$\alpha_i^\infty(p) = \begin{cases} 1 & \text{if} \quad p \geq \bar{p}_i^\infty \\ \tilde{\alpha}_i^\infty(p) & \text{if} \quad p \in [\underline{p}_i^\infty, \bar{p}_i^\infty)\ , \\ 0 & \text{if} \quad p < \underline{p}_i^\infty \end{cases} \tag{A.61}$$

where $\lim_{\eta\to\infty} \underline{p}_i = \underline{p}_i^\infty$, $\lim_{\eta\to\infty} \bar{p}_i = \bar{p}_i^\infty$, $i \in \{b, g, f\}$ (where $\underline{p}_f^\infty = \underline{p}_g^\infty$, $\bar{p}_f^\infty = \bar{p}_b^\infty$). The order of the thresholds stated in Theorem 1 is preserved in the limit, resulting in:

$$\underline{p}_b^\infty < \underline{p}_g^\infty < \bar{p}_b^\infty < \bar{p}_g^\infty. \tag{A.62}$$

From (A.62) together with (A.61), it immediately follows that $\alpha_b^\infty(p) = \alpha_f^\infty(p) = \alpha_g^\infty(p)$ on the union of the intervals $[0, \underline{p}_b^\infty] \cup [\bar{p}_g^\infty, 1]$, $\alpha_b^\infty(p) > \alpha_f^\infty(p) = \alpha_g^\infty(p)$ on $(\underline{p}_b^\infty, \underline{p}_g^\infty]$ and $\alpha_b^\infty(p) = \alpha_f^\infty(p) > \alpha_g^\infty(p)$ on $[\bar{p}_b^\infty, \bar{p}_g^\infty)$.

On the remaining interval $(\underline{p}_g, \bar{p}_b)$, all $\alpha_i^\infty(p)$ are determined by $\tilde{\alpha}_i^\infty(p)$. For the fixed η, $i \in \{b, g, f\}$, these strategies are given by (A.37), (A.45) and (A.58), exhibiting the general form of $\tilde{\alpha}_i(p) = f_1(w_i(p))/f_2^i(p)$, where $f_1(w_i(p)) \equiv w_i(p) - \Omega s$ is a linear and increasing function of $w_i(p)$, and $f_2^i(p)$ is a linear and decreasing function. $(\underline{p}_g, \bar{p}_b)$ is the region of the interior allocation, providing

$$w_i(p) = \frac{\gamma}{\gamma+\eta}s + \frac{\eta\beta}{\gamma+\eta}v_i(p).$$

Along with $v_b(p) > v_f(p) > v_g(p)$ (Lemma A.9, Lemma A.10, Lemma A.11), this leads to $w_b(p) > w_f(p) > w_g(p)$, and consequently, $f_1(w_b(p)) > f_1(w_f(p)) > f_1(w_g(p))$. At the same time

$$f_2^b(p) \equiv \frac{\gamma+\eta(1-\beta)}{\gamma+\eta}(s-f)(1-p) - \frac{\gamma+\eta(1-\beta)}{\gamma+\eta}(g-s)p,$$
$$f_2^f(p) = f_2^g(p) \equiv \frac{\gamma+\eta(1-\beta)}{\gamma+\eta}(s-f)(1-p) - \frac{\gamma+\lambda+\eta(1-\beta)}{\gamma+\eta}(g-s)p,$$

so that $\lim_{\eta\to\infty} f_2^b(p) = \lim_{\eta\to\infty} f_2^f(p) = \lim_{\eta\to\infty} f_2^g(p)$, completing the proof:

$$\tilde{\alpha}_b^\infty(p) = \lim_{\eta\to\infty} \frac{f_1(w_b(p))}{f_2^b(p)} = \lim_{\eta\to\infty} \frac{f_1(w_b(p))}{f_2^f(p)} > \lim_{\eta\to\infty} \frac{f_1(w_f(p))}{f_2^f(p)} = \tilde{\alpha}_f^\infty(p),$$
$$\tilde{\alpha}_f^\infty(p) = \lim_{\eta\to\infty} \frac{f_1(w_f(p))}{f_2^f(p)} = \lim_{\eta\to\infty} \frac{f_1(w_f(p))}{f_2^g(p)} > \lim_{\eta\to\infty} \frac{f_1(w_g(p))}{f_2^g(p)} = \tilde{\alpha}_g^\infty(p)$$

on the interval $(\underline{p}_g^\infty, \bar{p}_b^\infty)$.

Part 2 – The result for $\underline{p}_g = \underline{p}_f = p_g^1$ (Lemma 2), which is given by (10), and $\bar{p}_b = \bar{p}_f$ (Theorem 1, part 3), which is given by (A.56), is immediate. The thresholds $\underline{p}_b$ and $\bar{p}_g$ are given by the implicit expressions (A.34) and (A.46) respectively. The result that $\lim_{\eta\to 0} \underline{p}_b = \lim_{\eta\to 0} \bar{p}_g = p^*$ follows from straightforward algebra.