The Culture of Overconfidence*

V Bhaskar†　　Caroline Thomas‡

May 12, 2018

Abstract

Why do political leaders or managers persist with their pet projects despite bad news? When continuing the project is more informative than terminating it, a reputationally concerned leader is biased towards continuation, as it enables her to disclose her private information. Perceived overconfidence on the part of the leader aggravates this tendency, even when the leader is not overconfident. Higher-order beliefs regarding overconfidence can induce inefficient equilibrium selection even when it is "almost common knowledge" that the leader is not overconfident. A culture where leaders are expected to be overconfident has undesirable effects even upon leaders with correct beliefs.

JEL codes: C73, D82 , D72

Keywords: Overconfidence, Policy persistence, Mis-specified models, Non-common priors, Higher-order beliefs.

*We are grateful to Drew Fudenberg, François Lagarde, Stephen Morris and Muhamet Yildiz for helpful comments. Bhaskar thanks the National Science Foundation for its support via grant 201503942.

†Department of Economics, University of Texas at Austin. v.bhaskar@austin.utexas.edu.

‡Department of Economics, University of Texas at Austin. caroline.thomas@austin.utexas.edu.

To those waiting with bated breath for the 'U-turn', I have only one thing to say: 'You turn, if you want to. The lady's not for turning.'
Margaret Thatcher.

When the facts change, I change my mind. What do you do, Sir?
J.M. Keynes.

Why do managers and political leaders persist with their pet projects despite adverse information? Margaret Thatcher was "not for turning": not when UK unemployment hit two million, not when there was widespread opposition to the poll tax. The latter led to her being deposed as leader by her own party. Mao Tse-Tung intensified the Great Leap Forward, despite reports of widespread starvation in rural China. Over 30 million people died in the consequent famine (**?**). When Texas Instruments abandoned its foray into the home computer market, its stock price rose 22%. **?** find that a company's share price rises, on average, after a project termination that is preceded by public bad news.

To the ancient Greeks, the answer was hubris. In Aeschylus' *The Persians* and Sophocles' *Antigone*, arrogance leads the protagonist kings to ignore advice and omens, precipitating their demise. Historians invoke hubris to explain Napoleon's doomed march on Moscow (**?**) and Hitler's expansion (**?**). **?** argue that "hubris syndrome" afflicted many US presidents and British prime ministers. **?**, **?**, **?**, and **?** argue that managerial overconfidence lies behind corporate takeovers, and explains why they reduce the acquiring firm's value.

This paper shows that reputational concerns can induce a leader to continue an ineffectual project when she has correct beliefs regarding the quality of the project. Our main contribution is to show that overconfidence, perceptions of her overconfidence, or higher-order beliefs regarding overconfidence, magnify the resulting inefficiencies. Thus, a culture where leaders are expected to be overconfident can have adverse consequences even when the leader in question is not, in fact, overconfident.

The manager of a firm or a political leader receives inconclusive *private* information regarding the viability of her project, and must decide whether to continue the project to its conclusion, or abandon it. The leader cares about an outside observer's evaluation of the project's quality, since this is informative about her ability. Abandoning the project stops the arrival of further information; continuing reveals the underlying state publicly, resolving all uncertainty. This gives rise to a *disclosure value of continuation* — by continuing the project, the leader publicly verifies her private belief about its quality. Terminating it prevents any further learning, and only reveals that she deemed the project bad enough to

warrant termination. When she is moderately pessimistic about the project she therefore faces a reputational cost, inducing her to continue projects that are socially wasteful. There are strategic complementarities between the leader's continuation decision and the outside observer's inference; if the leader is more stubborn and continues the project at worse beliefs, the observer's inference is more adverse in the event that the project is terminated, increasing the reputational penalty for stopping. Consequently, there can be multiple equilibria, and an equilibrium where the project is stopped more often (at higher beliefs) is more efficient.

We examine the interaction between the leader's private information and overconfidence. We assume that the leader has a more optimistic prior on the project's chances of success than the observer. Because continuing the project results in more information revelation, it becomes even more appealing for an overconfident leader. Moreover, there is a second, more subtle effect, that induces excessive project continuation. Since the observer believes that the leader is overconfident, he draws even more negative inferences about her private information in the event that she terminates the project. This raises the reputational cost of termination. Indeed, this increase obtains even when the leader is not overconfident, but is merely *perceived* to be so, because the negative inference from termination is identical to that drawn when she is, in fact, overconfident.

Taking this argument further, suppose that the leader is not overconfident, and not perceived to be overconfident. But she believes that she is perceived to be overconfident. More generally, there may be mutual knowledge that the leader is *not* overconfident up to many levels, but not common knowledge: at some level, there is a perception that the leader is (believed to be) overconfident. Equilibria in such a game are anchored to equilibria in the game where the leader is *in fact* overconfident, exacerbating the leader's tendency to continue projects. This is true also in the limit, where it is "almost common knowledge" that the leader is not overconfident. If possible overconfidence is large enough, then there is a unique rationalizable outcome: the most inefficient equilibrium of the game with common knowledge of a common prior.

These results show that a culture where business and political leaders are expected to be substantially overconfident may have pernicious effects even for leaders who are not overconfident. Indeed, it can be mutual knowledge to a high degree that the leader in question has the right beliefs. Nonetheless, the lack of common knowledge entailed by the culture ensures that the most inefficient equilibrium is selected.

Our theory highlights a novel channel though which cultural stereotypes might determine outcomes: via higher-order beliefs. A culture where leaders are expected to be overconfident

plays a powerful role even when a leader is not in fact overconfident. This can be contrasted with a different view of the role of culture, namely that it coordinates expectations, and thereby selects among equilibria, much as history does.[1]

One possible normative implication of our results is that individuals who are not stereotypically expected to be overconfident — such as women — might prove better leaders, as they are under less pressure to persist with unprofitable projects.[2]

Related Literature: **?** and **?** argue that leaders are reluctant to abandon failing projects because this signals incompetence. They assume that competent leaders always select good projects, while incompetent leaders sometimes select bad ones, and subsequently receive information about the project. Cancelling a project reveals that a leader has changed her mind, and that she must be unskilled. **?** show that there is a unique equilibrium, with excessive continuation by the unskilled leader.

The main difference with our model arises from our ("Keynesian") assumption that no leader is infallible, and therefore it is wise to sometimes reverse one's decisions.Consequently, the inference made by the population when the leader cancels the project is not constant, but depends upon the leader's strategy. Thus, there may be multiple equilibria, and higher-order beliefs can play a role in selecting equilibria. This property is robust to assuming the leader has private information on her ability, provided she is not infallible. Also, we assume that information is fully revealed to the public when the project is continued, so that continuation is an informative experiment, analogous to disclosure with a verifiable type, as in **?** or **?**. This tempers inefficiency.

Majumdar and Mukand also discuss briefly the implications of the incompetent leader's prior on the project being larger than the observer's. They argue that policy persistence is likely to be reduced. Since overconfidence affects the decisions of both ability types in our setting, the findings are the opposite.

Rational explanations for overconfidence include **?** and **?**. **?** argue that beliefs play a motivational role, so that individuals disregard or forget unfavourable information, thereby sustaining overconfidence. These considerations are likely to be important for CEOs or

[1]For instance, in a society where women defer to men, it is focal to coordinate on the men-preferred equilibrium in the battle of the sexes. **?** presents experimental evidence supporting this.

[2]**?** examine common stock investments, and find that men trade 45 percent more than women and earn lower returns, a difference they attribute to men's overconfidence. **?** show that girls perform worse than boys in a competitive contest, especially when they compete against boys. Their performance is no different in a non-competitive setup. If girls are less confident in their abilities, they will exert less effort in a winner-take-all contest.

political leaders, who must motivate their followers as well as themselves.

Models with heterogeneous priors are used to examine a range of issues: financial markets (?, ?, ?); bargaining (?); the exchange of opinions (?). ? argue against the use of Nash equilibrium as a solution concept in games with heterogeneous priors, and advocate the use of non-equilibrium notions, such as rationalizability. Our analysis takes on board this criticism — our main result on the negative effects of the culture of overconfidence extends to rationalizable strategies. ? provide a connection between Bayes Nash equilibria in games without a common prior and interim correlated rationalizability.

Our results on the effects of higher-order beliefs are reminiscent of the electronic mail game (?), ? and ?). The lack of common knowledge pertains to the state in these papers, rather than the prior. We discuss the connection in more detail after presenting our findings. The strategic implications of mis-specified beliefs are explored by ?.

1 The Baseline Model

We study the interaction between a decision maker (DM) — a manager or political leader — who undertakes a project, and an outside observer — the firm's shareholders/potential employers, or voters — who evaluates the DM's ability. In period zero, nature chooses the ability $\tau \in \{H, L\}$ of the DM and the quality $\omega \in \{G, B\}$ of her project, where a more able DM is more likely to be endowed with a good project. Neither the DM nor the observer observe the DM's ability or project quality, and share a common prior p that the project is good.

In period one, the DM privately observes a signal informative of project quality, and decides whether to continue or terminate the project. If she continues (action Y), a cost c is incurred, and the project's outcome is publicly realized in period two. The outcome is a success if the project is good, yielding a return v, and a failure if the project is bad, yielding zero. It resolves all uncertainty about the project quality. If the DM terminates (action N), this is publicly observed, and there is no further learning about the project's quality. The costs and returns accrue to the firm (in the manager example) or to society as large (in the politician context).

At the end of period two (after observing the project's outcome in the event that it was not cancelled at date one), the observer chooses an action in $[0, 1]$. His optimal action equals his posterior belief β about project quality.

The DM maximizes the sum of the social payoff from the project and of the observer's

action — interpreted as the DM's reputation. The former is weighted by $\theta > 0$, which measures the intensity of her social concerns relative to reputational ones. Appendix A.1 shows that our model is equivalent to one where the DM is concerned about perceptions of her ability, as in **?**, and demonstrates that θ is smaller the greater the correlation between the DM's ability and the project's quality. We assume linearity in both payoff dimensions for analytical clarity, to focus on the differential informational content of the two experiments, stopping and continuing. If the DM's evaluation of either payoff dimension were convex (resp. concave), this would automatically bias her towards continuing (resp. terminating) the project.

The signal privately observed by the DM at the beginning of period one induces a cumulative distribution F over the DM's posterior belief μ that the project is good. Let $\mu^{**} := \frac{c}{v}$. The socially optimal decision is to continue with the project if and only if $\mu > \mu^{**}$. We assume:

Assumption 1 *F has an atomless distribution, and $F(\mu^{**}) \in (0, 1)$.*

Our results do not depend on F being atomless (see **?**) but this simplifies exposition. Since $F(\mu^{**})$ is interior, the signal observed by the DM is decision-relevant, so that both decisions, stop and continue, are socially optimal for some beliefs.

We analyze Perfect Bayesian Equilibria. A pure strategy σ for the DM maps her private posterior belief $\mu \in [0, 1]$ about the project's quality to $\{Y, N\}$. A strategy for the observer specifies an action in $[0, 1]$ at date two for each public event: success, failure, and N (project termination). Since the observer must take action 1 following a success and 0 following a failure, it suffices to specify the observer's action when the project is cancelled, and we identify it with the observer's strategy ρ. Sequential rationality implies that $\rho = \beta$, where β is the observer's belief about project quality upon cancellation.

Fix an equilibrium strategy profile (σ, ρ). If the DM's private belief about the project is μ, then her expected payoff from continuing the project is

$$U(Y, \mu) := \mu [\theta v + 1] - \theta c. \tag{1}$$

With probability μ, the project succeeds, so that the social value rises by v and the observer's belief jumps to 1. With complementary probability the project fails and the observer's belief falls to 0. Thus the payoff from continuation is an increasing affine function of μ, and is *independent* of σ — it does not depend on the observer's beliefs about the strategy played by the DM. Since project continuation fully reveals the project's quality, it is analogous to

disclosure with a verifiable type.

However, the DM cannot verifiably disclose her belief on project quality when she terminates the project, and the reputational payoff is independent of her belief. The social component is zero, since no costs or returns arise. Thus any equilibrium is characterized by a threshold x: the DM cancels the project if $\mu < x$, and continues if $\mu > x$. The DM's payoff from cancelling under a strategy with threshold x equals:

$$U(N, x) := \mathbb{E}(\mu|\mu < x). \tag{2}$$

An equilibrium strategy for the DM is summarized by a threshold belief μ^* satisfying $U(Y, \mu^*) = U(N, \mu^*)$. An equilibrium is the pair (μ^*, ρ^*), where the observer's action upon cancellation is $\rho^* = \mathbb{E}(\mu|\mu < \mu^*)$. Equilibrium existence follows from standard arguments (see Appendix A.2). The threshold μ^* need not belong to the support of F, since there may be gaps.

The DM's optimal stopping threshold x is increasing in the observer's action ρ. Also, ρ is increasing in x. Thus our game is supermodular, and multiple equilibria, arise naturally, due to the strategic complementarity, and to the fact that the conditional expectation $\mathbb{E}(\mu|\mu < x)$ can increase rapidly with x. Mass points in the distribution F cause discontinuous upward jumps in the conditional expectation, but a rapid increase in the conditional expectation can arise without mass points. This is illustrated in Figure 1, which shows three intersections of the two curves, at $\underline{\mu}^*, \mu_2^*$ and $\bar{\mu}^*$, corresponding to three distinct equilibrium thresholds. Multiplicity can arise even when the DM knows her ability, as long as the high ability DM is fallible and sometimes chooses bad projects — see **?**.

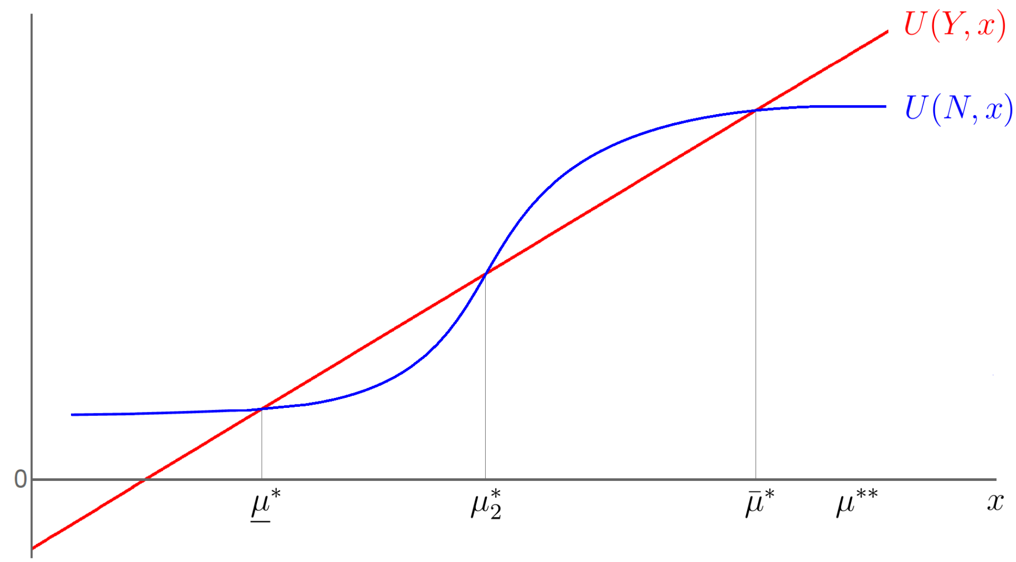

Figure 1: The DM's payoff $U(Y,x)$ from continuing when her private belief μ equals the threshold x, and her payoff $U(N,x)$ from stopping under the strategy with threshold belief x.

Definition 2 *An equilibrium with threshold μ^* is :*

- *Stable if there exists an open interval containing μ^* such that $\xi(x,\mu^*) := [U(N,x) - U(Y,x)][\mu^* - x] > 0$ for all $x \neq \mu^*$ in this interval.*
- *Unstable if there exists an open interval containing μ^* such that $\xi(x,\mu^*) < 0$ for all $x \neq \mu^*$ in this interval.*
- *Left-stable if there exists an open interval $(\mu^* - \epsilon, \mu^*)$ such that $\xi(x,\mu^*) > 0$ for all x in this interval.*

Graphically, stability of a threshold μ^* corresponds to the line depicting the payoff from continuing cutting the graph of the payoff from stopping from below. In Figure 1, the equilibria with the largest and smallest thresholds are stable, while the middle equilibrium is unstable. If an equilibrium is stable, then the comparative statics predictions are intuitive — for example, a small decrease in θ increases the equilibrium threshold. Comparative statics predictions are reversed for an unstable equilibrium.

1.1 Inefficient Continuation

In any equilibrium, the observer's posterior belief is a martingale, and thus its expectation at date zero must equal the prior p. Consequently, the DM's ex-ante payoff in any equilibrium, before she observes her private information, equals the payoff from the project under this equilibrium, plus her expected reputational payoff. Since the latter does not vary across equilibria, the equilibrium that maximizes the project's payoff is best for the DM.

Proposition 1

1. *There is inefficient continuation in every equilibrium if reputational concerns are sufficiently large, i.e. if θ is small enough.*
2. *When the distribution of beliefs has no gap immediately below μ^{**}, there is inefficient continuation in every equilibrium, no matter how small reputational concerns are.*
3. *Inefficiency is one-sided: the DM never terminates a profitable project.*
4. *Sufficiently bad projects are always terminated: for any degree of reputational concerns, there exists $\epsilon > 0$ such that the DM terminates the project for beliefs in $[\underline{\mu}, \underline{\mu}+\epsilon)$, where $\underline{\mu}$ denotes the lower bound of the support of F.*
5. *When there are multiple equilibria, efficiency increases with the equilibrium threshold.*
6. *The most inefficient equilibrium (i.e. smallest threshold) is left-stable.*

Note that sufficiently bad projects must be terminated, since continuing the project can only verify the DM's private beliefs. Indeed, if we consider a distribution F with only a single (mass) point in the support below μ^{**}, then one would have efficient decisions, no matter how large reputational concerns were. With mass points in the distribution of beliefs, there is inefficiency for large enough reputational concerns if and only if beliefs are rich, i.e. the conditional distribution of F below μ^{**} is not degenerate.[3]

By explicitly modelling the project initiation decision, one can show that this is conditionally efficient. Thus any initiated project is ex ante profitable, and terminating it must be, on average, bad news. However, *public* bad news can depress the average public belief below μ^{**}, and in this case, project terminations will increase the stock price, as found by **?**. A modification of our model can also explain why new managers scrap the projects of their predecessors. If a manager's contribution to firm value is persistent, a new manager will seek to minimize the reputation of her predecessor, thereby maximizing the observer's evaluation of her own contribution to firm value. She will terminate her predecessor's projects too

[3] See **?** for further elaboration of the points made in this and the next paragraph.

often, and these terminations must reduce the share price.[4] Thus our analysis suggests that empirical work on the announcement effects of project terminations needs to condition on whether these are preceded by changes in management, and also on the nature of preceding public news.

2 Overconfident Leaders

Suppose that DM and observer have different priors on the competence of the DM, and therefore, on the quality of the project, with the DM having a prior q on the quality of the project that is strictly greater than p, the prior of the observer. Assume that this is common knowledge. This defines G^0, a game with non-common priors.

If the DM observes a signal s with likelihood ratio ℓ,[5] her posterior belief equals

$$\pi(\ell) := \frac{q\,\ell}{q\,\ell + 1 - q}. \tag{3}$$

The posterior belief of the observer, were he to observe the same signal, would be

$$\mu(\ell) := \frac{p\,\ell}{p\,\ell + 1 - p}. \tag{4}$$

We now identify a signal with its associated likelihood ratio and define the above belief-updating functions for all $\ell \in [0, \infty)$, with $\pi(\infty) = \mu(\infty) = 1$. Define $\pi^\dagger(\mu)$ to be the posterior belief of the DM after a signal that would induce in the observer the posterior belief μ. That is, $\pi^\dagger : [0,1] \to [0,1]$ is obtained by using the inverse of (4) as the argument in (3):

$$\pi^\dagger(\mu) = \frac{q\,(1-p)\,\mu}{q\,(1-p)\,\mu + (1-q)\,p\,(1-\mu)}. \tag{5}$$

Lemma 1 *The function $\pi^\dagger$ is strictly increasing and strictly concave. For any $\mu \in (0,1)$, $\pi^\dagger(\mu) > \mu$.*

Arguing as before, any equilibrium must be in threshold strategies: there exists a likelihood ratio, denoted ℓ_0, such that the DM continues the project for $\ell \geq \ell_0$ and terminates

[4] If the DM seeks to *minimize* the observer's beliefs on project quality, there is strategic substitutability, and therefore, a unique equilibrium. The impact of differences in the prior is dampened, rather than magnified.

[5] Here, $\ell = \frac{h(s|G)}{h(s|B)}$, where $h(s|\omega)$ denotes the value of the probability density function at s when the project quality is ω.

it for $\ell < \ell_0$. Let $\mu_0 := \mu(\ell_0)$, so that the DM's threshold belief is $\pi^\dagger(\mu_0)$. It is analytically convenient to identify the DM's strategy with the threshold μ_0.

Fix a strategy $x \in [0, 1]$ for the DM. If the DM terminates the project, the observer updates his belief regarding the project quality to $\mathbb{E}_F(\mu|\mu < x)$, where the expectation is taken with respect to F, the distribution of beliefs under the prior p.[6] Consequently, the DM's payoff from termination under the strategy x is given by

$$U(N, x) := \mathbb{E}_F(\mu|\mu < x). \tag{6}$$

The DM's private belief after observing the threshold signal ℓ_0 is $\pi^\dagger(\mu_0)$, and her payoff from continuing the project at that belief is

$$U(Y, \pi^\dagger(\mu_0)) = \pi^\dagger(\mu_0) [\theta v + 1] - \theta c. \tag{7}$$

The threshold μ_0 is an equilibrium strategy for the DM if

$$U(Y, \pi^\dagger(\mu_0)) = U(N, \mu_0), \tag{8}$$

where, by (6), the right-hand side equals ρ_0, the observer's equilibrium action.

Equilibrium existence follows from the same arguments as before, the only difference being that if overconfidence is sufficiently acute (i.e. the difference $q - p$ is sufficiently large), the DM may continue the project even at the lowest possible belief.

Figure 2 illustrates an equilibrium of the game G^0. The straight line, $U(Y, x)$, shows the payoff from continuation without overconfidence, i.e. when the posterior belief under the prior p equals the threshold x. The concave function $U(Y, \pi^\dagger(x))$ shows the DM's payoff from continuation with overconfidence, i.e. when the DM's posterior belief is $\pi^\dagger(x)$. By Lemma 1, this lies above $U(Y, x)$. The equilibrium threshold with overconfidence, μ_0, satisfies (8). It lies to the left of μ^*, the equilibrium threshold under a common prior. Thus, overconfidence leads to excessive continuation, over and above that arising under a common prior, for two reasons. The first is straightforward: since the DM's belief is $\pi^\dagger(\mu_0) > \mu_0$, the DM continues because she has more optimistic beliefs in G^0 than in the game with a common prior. The second reason is more subtle. Since in G^0 the observer knows that the DM is overconfident

[6]Let $H(s|\omega)$ denote the cumulative distribution function of the DM's private signals in the state ω. From the point of view of the observer, whose prior belief about the state $\omega = G$ is p, $H_p := p\,H(s|G) + (1 - p)\,H(s|B)$ is the distribution over the signals that the DM observes. We let F denote the associated distribution over the resulting posterior beliefs, $\mu(\ell(s))$.

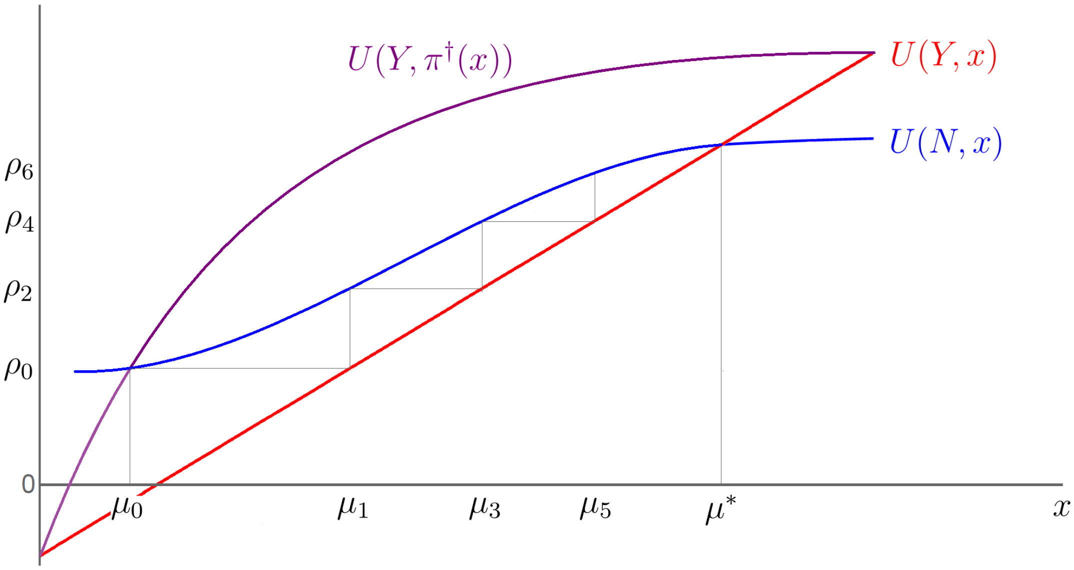

Figure 2: The threshold μ_0 and the observer action ρ_0 constitute an equilibrium of the game G^0. The threshold μ^* is the unique equilibrium threshold in the game with common knowledge of a common prior.

(as compared to his own prior), his inference upon project termination is more adverse. Thus the DM knows that she will be penalized more for terminating the project in G^0 than in the game with a common prior.

The argument, that greater confidence (or overconfidence) aggravates the tendency to continue bad projects, is more general. In Appendix A.5 we show that the game G^0 is supermodular: the observer's best response is increasing in the DM's threshold, and the DM's optimal threshold is increasing in the observer's strategy. Consequently, the results of **?** imply that the belief thresholds of the DM at the extremal equilibria are decreasing in q, the prior belief of the DM. Furthermore, at any stable equilibrium, the effect of a small increase in q is to reduce the equilibrium threshold.

2.1 Perceived Overconfidence

We now define G^1, a game where the DM and observer share the same prior, p, about the project. However, the outside observer *believes* that the DM is overconfident, i.e. that she has a prior $q > p$. We assume that the DM is aware of this belief. Specifically, assume:

- (**S1**) The DM and the observer share the prior p.

- (**T1**) The observer believes that the DM's prior is $q > p$.
- The observer's second-order belief is known to the DM, i.e. the DM knows T1.

Recalling that G^0 refers to the game with actual overconfidence, an alternative formalization of the game G^1 is:

- The DM and the observer share the prior p.
- The observer believes that the game G^0 is being played.
- The DM knows that the observer believes that G^0 is being played.

In any equilibrium of the game G^1, the observer's strategy must be the same as in an equilibrium of the game G^0, since he believes that G^0 is being played. In contrast, the DM's strategy differs across the two games. Thus, an equilibrium of the game G^1 is a triple $(\mu_0, \rho_0, \sigma_1)$ consisting of an equilibrium, (μ_0, ρ_0), of the game G^0, and σ_1, a best response to ρ_0 given the DM's prior, p.

Arguing as before, the optimal strategy σ_1 for the DM must be a threshold strategy. Let ℓ_1 denote the equilibrium cut-off signal for the DM in the game G^1, and let $\mu_1 := \mu(\ell_1)$ be the corresponding threshold belief, derived according to (4). It must satisfy the equilibrium condition:

$$U(Y, \mu_1) = \rho_0. \tag{9}$$

Since the left-hand side of (9) is strictly increasing in μ_1, there is a unique solution. In other words, for any equilibrium (μ_0, ρ_0) of the game G^0 with overconfidence, there is a unique equilibrium (μ_0, ρ_0, μ_1) in the game G^1 with perceived overconfidence. This is illustrated in Figure 2. Here, the unique equilibrium of the game G^0 has threshold μ_0 satisfying (8). In G^1, the unique corresponding threshold, μ_1, satisfies (9), where ρ_0 is the observer's action in both equilibria. The equilibrium *outcome* in the game G^1 is given by the pair (μ_1, ρ_0). In the game G^0, the equilibrium outcome is (μ_0, ρ_0).

Observe from (8) and (9) that $\mu_1 = \pi^\dagger(\mu_0)$. Thus, the DM's *equilibrium cut-off beliefs are identical* in the two games, with actual overconfidence and perceived overconfidence. Lemma 1 then implies that $\mu_1 > \mu_0$, or equivalently $\ell_1 > \ell_0$, so a DM who is indeed overconfident requires more adverse news to cancel the project than a DM who is only perceived to be overconfident. Nonetheless, $\mu_1 < \mu^*$, so that, when compared to the game with a common prior analyzed in Section 1, the perception of overconfidence penalizes the reputation of a DM who is not, in fact, overconfident, and makes her more reluctant to cancel the project.

2.2 Higher-Order Beliefs about Overconfidence

Suppose that the DM is not overconfident and the observer knows this. However, the DM believes that the observer believes her to be overconfident. In this case, we can define the game G^2 with perceived perceived overconfidence. More generally, let us consider the game G^N, where the DM is not overconfident, but this is not common knowledge. The statements (**S1**) and (**T1**) have been defined in the context of the game G^1. Define the following statements for every integer $K > 1$:

- (**SK**) (K even) The DM believes that S(K-1) is true.
- (**SK**) (K odd) The observer believes that S(K-1) is true.
- (**TK**) (K even) The DM believes that T(K-1) is true.
- (**TK**) (K odd) The observer believes that T(K-1) is true.

Then, for every integer $N > 1$, the game G^N is defined by:

- The DM and the observer share the prior p.
- The statements S1 to S(N-1) are true.
- The statement TN is true.
- Both players know TN.

The games G^0 and G^1 have been defined. Let G denote the game with common priors analyzed in Section 1.

Consider the sequence of games, (G^N), $N \in \{0\} \cup \mathbb{N}$. Let $\mathcal{E}$ denote the set of even numbers, and let $\mathcal{O}$ denote the set of odd numbers. We define a sequence of strategies $(\mu_0, \rho_0), (\mu_n)_{n\in\mathcal{O}}, (\rho_n)_{n\in\mathcal{E}}$. An equilibrium of the game G^N consists of the sequence truncated at N, with the property that:

- (μ_0, ρ_0) is an equilibrium of the game G^0.
- For any $n \in \mathcal{O}, n \leq N$, μ_n is a best response to ρ_{n-1}.
- For any $n \in \mathcal{E}, n \leq N$, ρ_n is a best response to μ_{n-1}.

Thus for any odd n, μ_n is uniquely determined by ρ_{n-1} as follows:

$$U(Y, \mu_n) = \rho_{n-1}. \tag{10}$$

When n is an even number, ρ_n is uniquely determined by μ_{n-1} via:

$$\rho_n := U(N, \mu_{n-1}). \tag{11}$$

The equilibria in the different games are illustrated in Figure 2. The pair (μ_0, ρ_0) represents the equilibrium strategies in G^0, the game with overconfidence. In G^1, the game with perceived overconfidence, the observer believes that G_0 is being played, and therefore plays ρ_0. The DM's equilibrium strategy in G^1 is μ_1, the best response to ρ_0, satisfying (10). In G^2, the game with perceived perceived overconfidence, the DM believes that G^1 is being played, and therefore plays μ_1. The observer's equilibrium strategy is ρ_2, his best response to μ_1, satisfying (11). Thus in any game G^n, the equilibrium outcomes are (μ_n, ρ_{n-1}) if n is odd and (μ_{n-1}, ρ_n) if n is even.

Observe that the the equilibrium values μ_n for n odd and ρ_n for n even are defined iteratively and converge to (μ^*, ρ^*), the equilibrium the game with common priors. Also, even if there were additional equilibria to the right of (μ^*, ρ^*), the iterative process starting at (μ_0, ρ_0) cannot go to the right of (μ^*, ρ^*).

This last point can be made more precise as follows. Fix an equilibrium (μ_0, ρ_0) in the game G^0 with overconfidence. Let $\mu^*_+(\mu_0)$ denote the smallest equilibrium threshold in the game G that is larger than μ_0, i.e.

$$\mu^*_+(\mu_0) = \min\{x > \mu_0 : U(Y, x) = U(N, x)\}. \tag{12}$$

Let $\rho^*_+(\mu_0)$ denote the associated equilibrium strategy for the observer. Thus, $(\mu^*_+(\mu_0), \rho^*_+(\mu_0))$ is the smallest equilibrium in G that is larger than (μ_0, ρ_0) — this is necessarily left-stable.

Proposition 2 *Fix an equilibrium (μ_0, ρ_0) of G^0, the game with overconfidence. This induces sequences $(\mu_n)_{n\in\mathcal{O}}$ and $(\rho_n)_{n\in\mathcal{E}}$, such that for $n \in \mathcal{O}$, in any game G^n, the equilibrium outcome is (μ_n, ρ_{n-1}), and in any game G^{n+1}, the equilibrium outcome is (μ_n, ρ_{n+1}). The sequences (μ_n) and (ρ_n) are both increasing, and converge to $\mu^*_+(\mu_0)$ and $\rho^*_+(\mu_0)$ respectively, i.e. to a left-stable equilibrium of G, the game with common priors.*

Observe that, in the absence of common knowledge of common priors, no unstable equilibrium can be approximated. This is true even if we consider higher-order beliefs regarding underconfidence, i.e. regarding the DM having a prior q that is strictly less than the prior p of the observer. Only stable equilibria are selected if there is lack of common knowledge of common priors, regardless of how priors diverge.

2.2.1 Large Overconfidence

Hubris amongst leaders would not be a source of concern if it were not significant. The difference in priors, $q - p$, is therefore likely to be large. The direct consequences of large overconfidence, or large perceived overconfidence, are straightforward — the tendency to continue projects is aggravated. We now examine the implications of this assumption for the limit outcomes in the game G with common knowledge of common priors. Our first result is substantive — only the most inefficient equilibrium of G is selected. Furthermore, it is uniquely (limit) rationalizable. Thus, it is not essential to assume that players play a Nash equilibrium in any of the games where there is a lack of common knowledge of common priors. This addresses the concern raised by **?**, that there is a conceptual inconsistency in imposing equilibrium requirements in a setting where agents have different priors.

The intuition is as follows. Consider $(\underline{\mu}^*, \underline{\rho}^*)$ the equilibrium with the smallest threshold in the game G with common priors. If q is sufficiently large, then $\pi^\dagger(\mu)$ becomes sufficiently large that the DM never wants to stop the project at any $\mu \geq \underline{\mu}^*$. Thus *every* equilibrium threshold in the game G^0 lies below $\underline{\mu}^*$. Since the game G^0 is supermodular, any rationalizable strategy ρ_0 for the observer has $\rho_0 \leq \underline{\rho}^*$. Thus any sequence of rationalizable thresholds, $(\mu_n)_{n \in \mathcal{O}}$, must start to the left of $\underline{\mu}^*$, and must therefore converge to $\underline{\mu}^*$. Similarly, any sequence of rationalizable actions, $(\rho_n)_{n \in \mathcal{E}}$, must converge to $\underline{\rho}^*$. Thus the limit points are unique, even if one considers rationalizable profiles, and does not require that players coordinate on an equilibrium.

Proposition 3 *If $q - p$ is sufficiently large, then any sequence of rationalizable DM strategies, $(\mu_n)_{n \in \mathcal{O}}$, converges to $\underline{\mu}^*$, and any sequence of rationalizable observer strategies, $(\rho_n)_{n \in \mathcal{E}}$, converges to $\underline{\rho}^*$.*

This result shows that a culture where business and political leaders are expected to be substantially overconfident may have pernicious effects even for leaders who are not overconfident. Indeed, it can be mutual knowledge to a high degree that the leader in question has the right beliefs. Nonetheless, the lack of common knowledge entailed by the culture ensures that the most inefficient equilibrium is selected. Our theory highlights a novel and important channel though which cultural stereotypes might determine outcomes: via higher-order beliefs. It may also have normative implications. In many environments, women are *perceived* as being less confident than men. Indeed, there is evidence that they are, in fact, less confident.[7] Thus, a female CEO may benefit from the stereotyping and may

[7]See, for example, the papers by **?** and **?** discussed in the introduction.

feel less pressure to pursue unprofitable projects than her male counterparts, since she will be penalized less for cancellations.

The implications of large underconfidence (i.e. q substantially smaller than p) are in the opposite direction: the equilibrium selected is the most efficient equilibrium in G. This result is partially reminiscent of **?**, who show that any rationalizable action in a complete information game can be made uniquely rationalizable by perturbing higher-order beliefs. The class of perturbations we consider can uniquely select only the extremal equilibria. It remains an open question whether and how one might uniquely select intermediate equilibria, especially unstable ones.

Our focus on overconfidence (rather than underconfidence) is motivated by the context. Literature, politics and the world of business are replete with examples where leaders are infected with hubris. This may influence outcomes even for the most clear-headed of leaders.

A Appendix: Proofs

A.1 When the DM is concerned about beliefs about her ability

In **?** and much subsequent work on career concerns, the DM's reputational payoff is a linear function of her expected ability. We now show that this assumption yields a model that is formally equivalent to the one analyzed in the text of the paper, where the DM's payoff is linear in the observer's belief about project quality. Recall that the DM's ability is $\tau \in \{H, L\}$, and the project's quality is $\omega \in \{G, B\}$. Let λ denote the prior probability that the DM is of type H. Let $p_\tau := \Pr(\omega = G|\tau)$ denote the probability with which type τ has a good project. Thus, the common prior on project quality is $p := \lambda\, p_H + (1-\lambda)\, p_L$. Recall that β is the observer's posterior belief, at date two, regarding the project's quality. Let ν denote the observer's posterior belief, at date two, that the DM is of type G. The relation between ν and β is as follows. When the project succeeds, so that $\beta = 1$,

$$\nu(1) = \frac{\lambda\, p_H}{\lambda\, p_H + (1-\lambda)\, p_L}.$$

When the project fails, $\beta = 0$ and

$$\nu(0) = \frac{\lambda\,(1-p_H)}{\lambda\,(1-p_H) + (1-\lambda)\,(1-p_L)}.$$

Both ν and β must satisfy the martingale property for a (hypothetical) experiment which perfectly reveals the project's quality, so that

$$\nu(\beta) = \beta\, \nu(1) + (1-\beta)\nu(0).$$

Let $\gamma := \nu(1) - \nu(0) < 1$. Then,

$$\nu(\beta) = \nu(0) + \gamma\beta.$$

Suppose that the observer takes an action in $[0, 1]$ to match ν, and the DM's total payoff equals $\tilde{\theta}V + \nu$, where V denotes the social payoff from the project, and $\tilde{\theta} > 0$ is a constant parameter reflecting the intensity of the DM's social concerns. Thus the DM's payoff equals

$$\tilde{\theta}V + \nu(0) + \gamma\beta.$$

If we let $\theta := \frac{\tilde{\theta}}{\gamma}$, then the DM's payoff is identical to the one analyzed in the text, except for

a constant term, $\nu(0)$, which accrues to both actions, stop and continue, and therefore does not affect the analysis.

A.2 Proof of Proposition 1

Let $\underline{\mu}$ denote the lower bound of the support of F. Define the function $g : [\underline{\mu}, 1] \to \mathbb{R}$ by

$$g(x) = U(Y, x) - U(N, x). \tag{A.13}$$

Since F has no mass points, g is continuous. At belief $\underline{\mu}$, there is no reputational loss from cancelling the project since $\mathbb{E}(\mu|\mu \le \underline{\mu}) = \underline{\mu}$. However, there is a social cost from continuing, since $F(\mu^{**}) > 0$ (see Assumption 1), so that $\underline{\mu} < \mu^{**}$. Therefore, $g(\underline{\mu}) < 0$. At any $x \ge \mu^{**}$, $g(x) > 0$. This follows because the social payoff from continuing is no less than that from stopping, and there is a reputational loss from stopping since $F(x) \ge F(\mu^{**}) > 0$, which implies $\mathbb{E}(\mu|\mu < x) < x$. Thus, there exists μ^* such that $g(\mu^*) = 0$, and any equilibrium threshold μ^* satisfies $\underline{\mu} < \mu^* < \mu^{**}$, proving parts 3 and 5 of Proposition 1.

Let $\mathcal{C}(F)$ denote the support of F. If there is no gap in $\mathcal{C}(F)$ immediately below μ^{**}, then for any equilibrium threshold μ^*, the interval (μ^*, μ^{**}) has positive F-measure, and there is inefficient continuation on this interval, thereby proving part 2 of Proposition 1.

If there is a gap in $\mathcal{C}(F)$ immediately below μ^{**}, let $\hat{\mu} := \sup\{\mu \in \mathcal{C}(F), \mu < \mu^{**}\}$. Then, $\hat{\mu}$ satisfies the following properties:

(i) $F(\hat{\mu}) = F(\mu^{**}) > 0$.

(ii) There exists $\epsilon > 0$ such that the interval $(\hat{\mu} - \epsilon, \hat{\mu}]$ has positive F-measure.

Both these properties follow immediately from the definition of $\hat{\mu}$ and Assumption 1. Property (i) implies that $\hat{\mu} > \mathbb{E}(\mu|\mu < \hat{\mu})$, so that the DM suffers a reputational loss from stopping when her belief equals $\hat{\mu}$. Thus, there exists $\hat{\theta}$ such that for all smaller values of θ, $U(Y, \hat{\mu}) > U(N, \hat{\mu})$. Since the payoffs are continuous, this inequality is also satisfied in a neighborhood of $\hat{\mu}$. That is, for every $\theta < \hat{\theta}$ there exists $\delta > 0$ such that $U(Y, x) > U(N, x)$ for all $x \in (\hat{\mu} - \delta, \hat{\mu}]$. Thus $\mu^* < \hat{\mu} - \delta$, and by property (ii) above, the interval (μ^*, μ^{**}) has positive F-measure, proving part 1 of the Proposition.

Part 4 of the proposition follows from part 3, and the fact that the social payoff from continuing is negative for all $x < \mu^{**}$.

We show next that the equilibrium with lowest threshold, $\underline{\mu}^*$, is left-stable. Consider a strictly increasing sequence, (x_n), with $x_1 = \underline{\mu}$ and which converges to $\underline{\mu}^*$. We have just

argued that $g(\underline{\mu}) < 0$. We now argue that this inequality must hold all along the sequence: $g(x_n) < 0$ for every n. Suppose instead that for some n, $g(x_n) = 0$. Then, there exists an equilibrium with a threshold lower than $\underline{\mu}^*$, a contradiction. Suppose that for some n, $g(x_n) > 0$. Then, since g is continuous, there exists $x < x_n$ such that $g(x) = 0$, again contradicting the definition of $\underline{\mu}^*$. This proves left-stability. Similarly, it can be shown that the equilibrium with threshold $\bar{\mu}^*$ is right-stable, where the definition of right-stability mirrors that of left-stability.

A.3 Proof of Lemma 1

The derivative of $\pi^\dagger(\mu)$ is

$$\frac{q(1-p)p(1-q)}{[q(1-p)\mu + (1-q)p(1-\mu)]^2} > 0.$$

The numerator in the above expression does not depend on μ, and the denominator is decreasing in μ, since the derivative of $q(1-p)\mu + (1-q)p(1-\mu)$ equals $q - p > 0$. Thus the derivative of $\pi^\dagger$ is strictly decreasing in μ. Since $\pi^\dagger$ is strictly concave in μ, with $\pi^\dagger(0) = 0$ and $\pi^\dagger(1) = 1$, it follows that $\pi^\dagger(\mu) > \mu$ for every $\mu \in (0, 1)$.

A.4 Proof of Proposition 2

In the game G^0, an equilibrium consists of a pair (μ_0, ρ_0) that satisfies

$$\rho_0 = U(N, \mu_0) = U(Y, \pi^\dagger(\mu_0)). \tag{A.14}$$

We determine the values of the equilibrium sequences $(\mu_n)_{n\in\mathcal{O}}$ and $(\rho_n)_{n\in\mathcal{E}}$, given the equilibrium (μ_0, ρ_0) in G^0.

Lemma 1 shows that $\mu_0 < \pi^\dagger(\mu_0)$, which, together with (A.14), implies that $U(N, \mu_0) > U(Y, \mu_0)$. Since $U(Y, x)$ is a strictly increasing (affine) function of x, it has a strictly increasing inverse, and we let $U_Y^{-1}(.)$ denote this inverse. Define μ_1 to be the unique value satisfying $U(Y, \mu_1) = \rho_0$, or equivalently, $\mu_1 = U_Y^{-1}(\rho_0)$. Since $\rho_0 = U(N, \mu_0) > U(Y, \mu_0)$, this implies that $\mu_1 > \mu_0$. Now, any value μ_1 generates a unique ρ_2 in G^2, satisfying the equation:

$$\rho_2 = U(N, \mu_1).$$

By the same logic, equilibrium thresholds and actions in higher-order games are defined

recursively as follows. For any $n \in \mathcal{O}$, $n > 1$, given any ρ_{n-1}, the threshold μ_n is defined by

$$U(Y, \mu_n) = \rho_{n-1}. \tag{A.15}$$

Observe that the above defines a unique μ_n, since $U(Y, x)$ is a strictly increasing function. Conversely, for any $n \in \mathcal{O}$, given any μ_n, the action ρ_{n+1} is uniquely defined by

$$\rho_{n+1} = U(N, \mu_n). \tag{A.16}$$

Given any equilibrium (μ_0, ρ_0) in G^0, equations (A.15) and (A.16) therefore define a unique sequence

$$\Big((\mu_0, \rho_0), (\mu_n)_{n\in\mathcal{O}}, (\rho_n)_{n\in\mathcal{E}}\Big).$$

The sequence satisfies the inequalities

$$U(N, \mu_{n+2}) \geq U(Y, \mu_{n+2}) = U(N, \mu_n) \geq U(Y, \mu_n), \quad n \in \{-1\} \cup \mathcal{O}, \tag{A.17}$$

where we let $\mu_{-1} := \mu_0$.

We show that the sequence $(\mu_n)_{n\in\mathcal{O}}$ is (weakly) increasing. We have already established that $\mu_1 > \mu_{-1}$. Now consider $n \in \mathcal{O}$, $n \geq 3$, and assume, by the induction hypothesis, that the sequence (μ_m) is weakly increasing for all $m < n, m \in \{-1\} \cup \mathcal{O}$. Then $\mu_{n-2} \geq \mu_{n-4}$. Since, $U(N, x)$ is increasing in x, being the conditional expectation, equation A.16 implies that $\rho_{n-1} \geq \rho_{n-3}$. Then, A.15 establishes that $\mu_n \geq \mu_{n-2}$.

We have established that (μ_n) is an increasing sequence. Since it is bounded, it must converge to some value, denoted μ^∞. There are two possibilities:

(i) either $\mu_n = \mu^\infty$ for some n, or

(ii) $\mu_n < \mu^\infty$, $\forall n$.

We will show that μ^∞ is an equilibrium threshold for the DM in the game G with common priors in both the cases above, by showing that $U(Y, \mu^\infty) = U(N, \mu^\infty)$.

Consider case (i). Let m denote the smallest value of n such that $\mu_n = \mu^\infty$. By the definition of m, and using the inequalities in A.17, we deduce that

$$U(N, \mu_m) = U(Y, \mu_m) = U(N, \mu_{m-2}) > U(Y, \mu_{m-2}).$$

The first equality implies that μ_m is an equilibrium threshold in G. $U(N, \mu_m) = U(N, \mu_{m-2})$

implies that $U(N,x)$ is constant on the interval (μ_{m-2},μ_m). (This is possible only if this interval has zero F-measure.) Thus μ^∞ is an equilibrium threshold in G, and we have also proved that it is left-stable, since $U(N,x)$ is constant on the interval (μ_{m-2},μ^∞), while $U(Y,x)$ is strictly increasing.

Next, consider case (ii). $U(N,\mu)$ is a continuous function of μ. Therefore, $U(Y,\mu_n)\leq U(N,\mu_n)$, and $\mu^\infty>\mu^n$ for all n, imply that $U(Y,\mu^\infty)\leq U(N,\mu^\infty)$. We now prove the reverse inequality. Since the sequence (μ_n) is Cauchy, and since $U(Y,\mu)$ is an affine function, the sequence $(U(Y,\mu_n))$ is a Cauchy sequence. Thus for any $\epsilon>0$, there exists $\bar{n}$ such that $U(Y,\mu_{n+2})-U(Y,\mu_n)<\epsilon$ if $n>\bar{n}$. By (A.15) and (A.16), $U(Y,\mu_{n+2})=U(N,\mu_n)$. Therefore, $U(N,\mu_n)-U(Y,\mu_n)<\epsilon$ if $n>\bar{n}$. Since ϵ was arbitrary, and both functions are continuous, $U(N,\mu^\infty)\leq U(Y,\mu^\infty)$. We conclude that $U(Y,\mu^\infty)=U(N,\mu^\infty)$, so that μ^∞ must be an equilibrium threshold of G, the game with common priors.

Having established convergence to an equilibrium threshold, we now show that this must correspond to the threshold μ_+^* defined in (12). The equilibrium (μ_+^*,ρ_+^*) is left-stable by definition: recall that $\mu_+^*(\mu_0)=\min\{x>\mu_0 : U(Y,x)=U(N,x)\}$, and that μ_0 satisfies (A.14) which implies $U(Y,\mu_0)<U(N,\mu_0)$. It follows that

$$U(Y,x)<U(N,x),\qquad x\in[\mu_0,\mu_+^*)\tag{A.18}$$

To prove convergence to μ_+^*, we consider two possible cases: either there is no gap in the support of F to the left of μ_+^*, or this condition fails.

Suppose that there is no gap, i.e. there exists an open interval $(\tilde{\mu},\mu_+^*)$ such that F is strictly increasing on this interval. This implies that $U(N,x)<U(N,\mu_+^*)$ if $x<\mu_+^*$. This, together with (A.18), implies that for any $x\in[\mu_0,\mu_+^*)$:

$$U(Y,x)<U(N,x)<U(N,\mu_+^*).$$

We now show that the sequence (μ_n) is *strictly* increasing and that $\mu_n<\mu_+^*$ for all $n\in\mathcal{O}$. Since $\mu_n=U_Y^{-1}[U(N,\mu_{n-2})]$ while $\mu_+^*=U_Y^{-1}[U(N,\mu_+^*)]$, and the function U_Y^{-1} is strictly increasing, it follows from $\mu_0<\mu_+^*$ that $\mu_{n-2}<\mu_n$ for every $n>1$, $n\in\mathcal{O}$, and that $\mu_n<\mu_+^*$ for every $n\in\mathcal{O}$.

Now, suppose that there is a gap, i.e. there exists an interval $(\tilde{\mu},\mu_+^*)$ on which F is constant. Then, for any $x\in[\mu_0,\mu_+^*)$:

$$U(Y,x)<U(N,x)\leq U(N,\mu_+^*).$$

It follows that if $\mu_n < \mu_+^*$, then $\mu_{n+2} \le \mu_+^*$. If $\mu_n = \mu_+^*$, then since $U(Y, \mu_+^*) = U(N, \mu_+^*)$, it follows that $\mu_{n+2} = \mu_+^*$, and thus the sequence becomes constant, and therefore $\mu_n \le \mu_+^*$ for every $n \in \mathcal{O}$.

Thus, we have established that the sequence (μ_n) is an increasing sequence that converges to an equilibrium threshold, and also established that $\mu_n \le \mu_+^*$ for all n. This completes the proof of the proposition.

A.5 Proof of Proposition 3

We show first that the game G^0 is supermodular. If the DM uses a threshold x, then the observer's belief upon cancellation is $\mathbb{E}_F(\mu|\mu < x)$, a weakly increasing function of x. Since the observer's action must match his posterior belief, the observer's best response correspondence $\hat{\rho}(x)$ is weakly increasing in x. Let the DM's best response threshold be denoted by $\hat{\mu}(\rho) \in [0, 1]$. By (8), it satisfies

$$U(Y, \pi^\dagger(\hat{\mu})) = \rho.$$

Then $\hat{\mu}(\rho)$ is uniquely defined and increasing since both $\pi^\dagger$ and $U(Y, .)$ are strictly increasing functions. Thus the game G^0 is supermodular.

Let $(\underline{\mu}_0, \underline{\rho}_0)$ denote the Nash equilibrium of G^0 with the smallest threshold, and let $(\bar{\mu}_0, \bar{\rho}_0)$ denote the Nash equilibrium with the largest threshold. Since the game is supermodular, any rationalizable strategy for the observer lies in the interval $[\underline{\rho}_0, \bar{\rho}_0]$.

From (5), observe that for any interior value of μ, $\pi^\dagger(\mu) \to 1$ as $q \to 1$. Thus, for q sufficiently large, every equilibrium threshold in the game G^0 is strictly less than $\underline{\mu}^*$, and therefore $\bar{\mu}_0 < \underline{\mu}^*$.

Consider a rationalizable strategy of the DM in the game G^1. This must be a best response to a probability distribution over the rationalizable strategies of the observer in G^0, i.e. to a distribution with support $[\underline{\rho}_0, \bar{\rho}_0]$. Since the DM's payoff from stopping is linear in ρ, her rationalizable strategies in G^1 are the best responses to elements of the set $[\underline{\rho}_0, \bar{\rho}_0]$. Thus the rationalizable thresholds μ_1 in G^1 satisfy

$$U(Y, \mu_1) = \rho, \quad \rho \in [\underline{\rho}_0, \bar{\rho}_0].$$

For $n \in \mathcal{O}$, let $\underline{\mu}_n$ denote the smallest equilibrium threshold for the DM in G^n, and let $\bar{\mu}_n$ denote the largest. Any rationalizable threshold, μ_n, must belong to the interval spanned

by these two thresholds. Similarly, for $n \in \mathcal{E}$, let $\underline{\rho}_n$ denote the smallest equilibrium action for the observer in G^n, and let $\bar{\rho}_n$ denote the largest. Again, any rationalizable action ρ_n belongs to the interval spanned by these two actions.

Consider the sequences (μ_n) and (ρ_n) induced by $(\underline{\mu}_0, \underline{\rho}_0)$. By Proposition 2, these must converge to $\underline{\mu}^*$ and $\underline{\rho}^*$ respectively. Similarly, the sequences induced by $(\bar{\mu}_0, \bar{\rho}_0)$ also converge to the same limits. Since any rationalizable sequence μ_n (resp. ρ_n) is sandwiched between the two induced equilibrium sequences, this proves the proposition.