Anti-Herding and Strategic Consultation*

Gilat Levy†

Abstract

In this paper I analyze how careerist decision makers aggregate and use information provided by others. I find that decision makers who are motivated by reputation concerns tend to 'anti-herding', i.e., they excessively contradict public information such as the prior or others' recommendations. I also find that some decision makers may deliberately act unilaterally and not consult advisers although advice is costless. Moreover, advisers to the decision maker may not report their information truthfully. Even if the advisers care only about the outcome, they bias their recommendation since they anticipate inefficient anti-herding behavior by the decision maker.

JEL Classification Numbers: D82; D83.

Keywords: Reputation; Anti-herding.

1 Introduction

In the biblical story about Abshalom and David, Abshalom consults Ahitophel and asks for his advice how to capture the throne from David - Abshalom's father. Ahitophel offers Abshalom to lead a battalion of 12,000 soldiers in order to force David to surrender. God, according to the story, intervenes and prevents Abshalom from following the sound advice, so that David eventually prevails. The sub-text is provided by the bible's interpreters, who

*I thank Oriana Bandiera, Patrick Bolton, Allesandro Lizzeri, Meg Meyer, Marco Ottaviani, Ady Pauzner, Wolfgang Pesendorfer, Michele Piccione, Yossi Rapoport, Ran Spiegler, seminar participants and two anonymous referees, and especially Ronny Razin and my adviser Gene Grossman, for many valuable comments.

†Department of Economics, LSE, and Berglas School of Economics, Tel Aviv University. Current address: Department of Economics, London School of Economics, Houghton St., London, WC2A 2AE, UK. Tel.: +44-20-7955-6652; fax: +44-20-7831-1840. E-mail address: g.levy1@lse.ac.uk

claim that Abshalom did not follow Ahitophel's advice because he feared that the glory of victory would go to Ahitophel rather than to himself.[1] There are more recent examples of decision makers who choose to ignore advice or go against public opinion. President Truman's decision to fire General Douglas MacArthur during the Korean War was in defiance of public opinion.[2] The decision of President Ford in 1976, to intervene (diplomatically) in the internal affairs in Rhodesia was a surprising decision; not only was it against public opinion, but also in contradiction to Ford's previous actions and announcements, that suggested that he did not believe that it was in the national interest of Americans.[3]

In this paper I offer a career concerns explanation to the behavior of decision makers who ignore or contradict advice, go against the herd or even against their own belief. Career concerns have been recognized as a cause for inefficient decision making, since Fama (1980), Lazear and Rosen (1982) and Holmström (1982). More recently, several papers focused on the link between career concerns and the tendency of managers to 'herd', that is, to follow others' advice or actions while ignoring and suppressing their own information (see for example Scharfstein and Stein, 1990, Ottaviani and Sørensen, 2002, Zwiebel, 1995, Effinger and Polborn, 2001 and Prendergast and Stole, 1996). But, as Zwiebel (1995) notes, in many situations it seems more likely that the reputation gained from *outperforming* the competition would exceed the reputation gained from equal success across the board. As a result, decision makers may 'anti-herd' and reject advice. The aim of this paper is to demonstrate the conditions for anti-herding behavior and its implications.

I analyze the following model. A decision maker has a private signal about the state of the world, and the accuracy of this signal depends on her ability. There is a public prior, which determines what is the likely state of the world.[4] The decision maker takes an action based on the prior and her own private information. An evaluator observes the state of the world, the prior, and the action taken by the decision maker. On the basis of these observations, the evaluator can assess the quality of the decision maker's private information, which reflects her respective ability. The decision maker is concerned about the evaluator's assessment of her ability. A high assessment can be translated into higher wages or promotion.[5]

[1] The Interpreter's Bible (New York: Abingdon Press, 1953), vol. 2.

[2] This is probably the most famous example of a President taking a decision against public opinion. See Dennis D. Weinstock, *Truman, MacArthur and the Korean War*, Greenwood Press, 1999, chapters 9 and 10.

[3] William G. Hyland, *Mortal Rivals: Superpower Relations from Nixon to Reagan*, Random House, 1987.

[4] The prior can be formed for example given others' public recommendations.

[5] We can therefore think of the evaluator as representing firm owners in the market, who search for talented

I find that in the unique equilibrium of the model, the decision maker excessively contradicts the prior. Even when she believes that the prior is correct, she may go against it, a behavior which I term 'anti-herding'.[6] Intuitively, if the decision maker takes an action that follows the public prior, she reveals that her own information may not be accurate enough. If, however, the decision maker's action goes against the prior, or against the herd, she suggests that her own information is at least as accurate as the prior. Thus, going against public information can serve as a signal about ability, which implies that decision makers have an excessive incentive to use it. Thereby in equilibrium they distort their actions in this direction.

In the model, I also allow the decision maker to decide how much information to aggregate before taking her decision.[7] That is, she can choose to consult an adviser who provides public advice. I find that the most able decision makers choose not to consult, inefficiently. In particular, the signal of anti-herding can be substituted by the signal of not gathering information at all. Intuitively, if the decision maker chooses not to check the content of available information, she shows even a greater confidence in her ability. I then compare between the two types of signals, that of not soliciting information and that of contradicting available information. I find that decision makers behave even less efficiently when they use the signal of not soliciting information.

Finally, I ask whether advisers to the decision maker indeed report their information truthfully. If such an adviser has career concerns himself, he may wish to pose as a talented type and hence manipulate his recommendation. I show that advisers bias their recommendations even if they do not have career concerns but care only about the outcome. It is the anticipated anti-herding behavior of the decision maker which induces them to distort their recommendations. Thus, career concerns of either agent, be it the adviser or the decision maker, are sufficient to induce sub-optimal sharing of information.

Anti-herding results are derived in the literature in several contexts and under a variety of assumptions. I find that in terms of the decision maker's objectives, the decision maker

managers. This accords with the career concerns view of Holmström and Ricart i Costa (1986), in which the wage is a function of the market's perception of the ability of the manager. A different type of career concerns arises when agents are induced to signal their preferences, instead of their ability. This type of career concerns is analyzed for example by Morris (2000).

[6] Herding is defined as the behavior of decision makers who follow others' recommendations while ignoring their own information (Scharfstein and Stein, 1990, Banerjee, 1992).

[7] As far as I know, no other paper analyzes whether managers solicit advice or information in the context of career concerns.

must be motivated by career concerns. I illustrate though that anti-herding arises even if the decision maker has only reputation concerns and does not care about the outcome at all. For example, Prendergast and Stole (1996) analyze a dynamic model in which managers initially overreact to their own information compared to a common market prior, and then 'herd' relative to their own previous decisions later in their career. In their model they assume sufficiently high outcome concerns. I show that reputation concerns alone are sufficient.

In my setting, career concerns are individualistic. That is, the decision maker cares only about the perception of the evaluator about her type. There is no direct competition with other managers in the market. Given that career concerns are individualistic, I show that a necessary condition for anti-herding is that the decision maker knows her type. Intuitively, when the decision maker does not know her type, she can only signal it by taking the correct decision, which accords with efficiency. Hence, anti-herding cannot arise.[8]

Indeed, Scharfstein and Stein (1990) who initiated the reputational herding literature, assume that decision makers do not know their type and find that herding may arise.[9] Avery and Chevalier (2001), however, show that when the decision maker has sufficient information about her type, the resulting behavior is anti-herding. In their paper they assume that the information of able agents is correlated. This assumption, as my model illustrates, is not important for the derivation of anti-herding. Trueman (1994) was the first to set up a model in which an expert knows her type and tries to signal it to the market. His model is essentially a two-type model and as a result he derives anti-herding only for some parameters.[10] Here I generalize anti-herding results found in Trueman (1994), Avery and Chevalier (2001) and Prendergast and Stole (1996) for any distribution over types and for various assumptions about the information held by the players.[11] In addition, I extend this literature by analyzing

[8] The literature on reputation is split on the assumption whether agents know their type or not. Holmström (1982) and Holmström and Ricart i Costa (1986) assume that agents do not know their type, whereas Kreps and Wilson (1982) and Milgrom and Roberts (1982) assume that agents do know their type.

[9] They assume that information among smart managers is correlated as opposed to untalented managers. Ottaviani and Sørensen (2000), however, prove that the assumption of correlated information among talented managers is not necessary for the derivation of herd behavior; decision makers who do not know their type, can signal it only by taking the right decision. Thus, herd behavior arises in the presence of career concerns, exactly as in the models in which decision makers simply care about taking the correct action (see Banerjee, 1992, Bikhchandani, Hirshleifer and Welch, 1992).

[10] A version of his model can be found in Ottaviani and Sørensen (2002).

[11] In particular, I show that anti-herding arises also if the evaluator does not observe the state of the world, and when the evaluator or the decision maker do not observe the accuracy of the prior.

other implications of anti-herding, such as how it affects advisers to the decision maker and her inclination to consult at all.

Effinger and Polborn (2001) analyze a model in which experts do not know their type, as in Scharfstein and Stein (1990). They derive the benefit for the expert from fundamentals, and create a benefit function such that an expert is most valuable if he is the only smart expert. This induces an incentive for agents to differentiate themselves. Thus, since experts directly compete in their model (i.e., career concerns are not individualistic), they can derive anti-herding in a set up in which decision makers do not know their type.

The rest of the paper is organized as follows. Section 2 presents the model and a benchmark analysis of an efficient decision maker. Section 3 presents the main result; I derive the anti-herding result and its generalizations. Section 4 extends the model by allowing the decision maker to decide how much information to aggregate and also examines how much information advisers provide to the decision maker. Section 5 concludes and the appendix contains all proofs, which are not in the text.

2 The model

The players and the action space

There are two players in the game, a decision maker D, and an evaluator E. The decision maker must take an action, a. For simplicity, suppose there are only two possible actions, $a \in \{l, h\}$. D gains utility from advancing her career. That is, if E believes that D is an able decision maker, she can be rewarded with high wages or a promotion. Hence, the decision maker tries to signal to E that she is able. E, on the other hand, tries to guess the decision maker's true type or 'ability'. E's action is therefore to form beliefs on the type of D, as will be explained below.

The information structure

There is an underlying state of the world w, which is unknown and is realized only after the action a is taken. There are two possible states, which are also denoted by l or h. Let action l be appropriate in state l and action h be appropriate in state h. There is a common knowledge prior distribution on the states of the world. The prior can summarize previous public information, such as other's actions or recommendations. In particular, let $\Pr(w = h) = q$, for $q \in (.5, 1)$ whereas $\Pr(w = l) = 1 - q$.

(i) The information of D : The decision maker receives a private signal $s \in \{l, h\}$ about

the state of the world. The more able is D, the more likely it is that her signal is reliable. In particular, let p represent the ability of D and assume that $\Pr(s = w|w, p) = p$. D knows p, which is drawn from a continuous density function $f(p)$ on $[.5, 1]$.

Given the prior q, and her own information (s, p), the decision maker D forms the following beliefs, according to Bayes rule:

$$\Pr(w = h|s, p, q) = \left\{ \begin{array}{c} \frac{pq}{pq+(1-p)(1-q)} \text{ if } s = h, \\ \frac{(1-p)q}{(1-p)q+p(1-q)} \text{ if } s = l \end{array} \right\} \quad (1)$$

where $\Pr(w = l|s, p, q) = 1 - \Pr(w = h|s, p, q)$.

(ii) The information of E : The evaluator observes neither the signal s of the decision maker nor her type p, about which he forms beliefs. E knows that p is distributed on $[.5, 1]$ according to $f(p)$. I assume that E observes the action a chosen by D, the realized state of the world ω, and the prior q.

The objectives of the players, strategies, and equilibria

E updates his beliefs about p rationally. Denote $E's$ posterior expectations about p by π. That is, $\pi = E(p|q, w, a)$, where (q, w, a) is the information of E. The objective of E is to guess the type of D correctly. Define $I = 1$ if $w = a$ and $I = 0$ otherwise. *The objective of D is to maximize $\pi + \theta I$, for $\theta \geq 0$.* θ is a parameter that measures how strong are the outcome concerns of D.

The strategy of D is to pick a, that is, a function $\alpha : (s, p, q) \to \{l, h\}$. Similarly, E uses a belief updating function $\pi : (q, w, a) \to [.5, 1]$. I use the concept of a Perfect Bayesian Equilibrium to solve the model. I analyze only informative equilibria, i.e., when the strategy of the decision maker is responsive to her signal. Moreover, I ignore "mirror" equilibrium, i.e., an equilibrium that takes an original equilibrium and switches each action from l to h and vice versa.

The timing of the game:

1. q, the accuracy of the prior, is observed (where $q = \Pr(w = h)$ and $q > \frac{1}{2}$).
2. w is realized and D receives a private signal, (s, p), about w.
3. D takes an action a.
4. w becomes observable and E forms beliefs on p, given a, w, q and $f(p)$.

Benchmark: an efficient decision maker

As a benchmark, let us consider the behavior of an efficient decision maker, who cares only about taking the correct decision. This implies that she should take h if $\Pr(w = h|s, p, q) \geq$

$\Pr(w = l|s,p,q)$. By (1), this implies that if $s = h$, the decision maker takes h, whereas if $s = l$, the decision maker takes h as long as $p \leq q$.[12] Otherwise, if $p > q$, that is, her accuracy is greater than that of the prior, she takes l.

It is useful to describe such a cutoff point strategy in the following graph, which will accompany the analysis throughout. The right part of the graph describes the action of the decision maker when $s = l$, for p ranging from .5 to 1. The left part of the graph, describes the decision maker's action when $s = h$, and p ranges from .5 (in the middle) to 1 (in the left). Thus, as we go from left to right, $\Pr(w = l|s,p)$ increases. The cutoff point, which in the efficient case is at $(s = l, p = q)$, is such that for all information (s,p) to the right of it, D takes l, whereas for all information (s,p) to its left, D takes h :

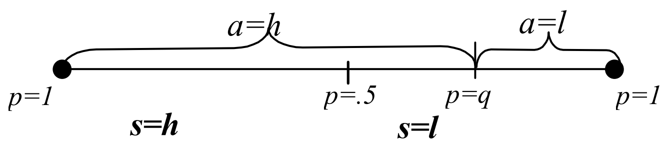

Figure 1: The efficient benchmark

3 Results

The incentives of the careerist decision maker are not necessarily aligned with efficient decision making, and hence she may distort her actions. I now analyze the behavior of the careerist decision maker in equilibrium and its sensitivity to the parameters of the model.

Let us define the following reputation function: $\pi(a,w,\alpha)$ is the expectations of E on $D's$ type, when E observes her action a, the state of the world w, and upon conjecturing some strategy α.[13] Essentially, E needs to update his beliefs about the density of each type, given the conjecture of α and the observation of a, w :

$$\pi(a,w,\alpha) = \int_{.5}^{1} p \frac{(\Pr(s = w|w,p)\Pr(\alpha(s,p,q) = s) + \Pr(s \neq w|w,p)\Pr(\alpha(s,p,q) \neq s))f(p)}{\int_{.5}^{1}(\Pr(s = w|w,p)\Pr(\alpha(s,p,q) = s) + \Pr(s \neq w|w,p)\Pr(\alpha(s,p,q) \neq s))f(p)dp} dp$$

Hence, $\pi(a,a,\alpha)$ denotes the reputation that D receives if she is correct (where $a = w$), whereas $\pi(a,a',\alpha)$ denotes the reputation that she receives if she is wrong $(a \neq a', a' = w)$.

Since D *knows* that E *knows* w, D believes that she receives the reputation $\pi(a,a,\alpha)$ if indeed she is correct. This occurs with $\Pr(w = a|s,p,q)$ which is defined in equation (1). We

[12] At $s = l$ and $p = q$, $\Pr(w = h|s,p,q) = \Pr(w = l|s,p,q)$.

[13] I omit the index q from the beliefs of E, since given the observation of w, q may affect the beliefs of the evaluator only through the decision maker's strategy.

can therefore describe the utility of D from taking an action a, in the following way:

$$\Pr(w=a|s,p,q)(\pi(a,a,\alpha)+\theta)+(1-\Pr(w=a|s,p,q))\pi(a,a',\alpha)$$

The first Lemma characterizes the strategy of the decision maker in any equilibrium.

Lemma 1 *In any informative equilibrium, the decision maker uses a strategy* α^* *which is characterized by a cutoff point* (s^*,p^*)*, such that the decision maker takes* h *if* $\Pr(w=h|s,p,q)\geq \Pr(w=h|s^*,p^*,q)$*, and otherwise she takes* l*.*

Figure 2 describes an example of a cutoff point strategy for the decision maker, with $s^*=h$. Recall that the right part of the figure depicts $p\in[.5,1]$ for $s=l$ whereas the left part depicts $D's$ types with $s=h$, so that $\Pr(w=l|s,p)$ increases from left to right. The figure illustrates that given a cutoff point (s^*,p^*), the decision maker takes l for all (s,p) to the right of (s^*,p^*), and takes h for all (s,p) to its left.

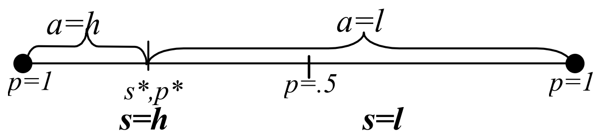

Figure 2: A cutoff point strategy with $s^*=h$.

Proof: In an informative equilibrium, some types of D take l whereas some types take h. This implies that there must be at least one type (s,p) who is indifferent between taking l and h. That is, the following condition must hold for some (s,p) :

$$\begin{aligned}&\Pr(w=l|s,p,q)(\pi(l,l,\alpha)+\theta)+\Pr(w=h|s,p,q))\pi(l,h,\alpha)\\ =&\Pr(w=h|s,p,q)(\pi(h,h,\alpha)+\theta)+\Pr(w=l|s,p,q))\pi(h,l,\alpha)\end{aligned}$$

re-arranging, we get:

$$\frac{\Pr(w=h|s,p,q)}{\Pr(w=l|s,p,q)}=\frac{\pi(l,l,\alpha)-\pi(h,l,\alpha)+\theta}{\pi(h,h,\alpha)-\pi(l,h,\alpha)+\theta} \tag{2}$$

The right-hand-side of (2) is fixed for all (s,p), since these are the beliefs of the evaluator. The evaluator does not know (s,p) and hence cannot condition his beliefs on this information. The left-hand-side of (2), on the other hand, changes with (s,p). In particular:

$$\frac{\Pr(w=h|s,p,q)}{\Pr(w=l|s,p,q)}=\left\{\begin{array}{c}\frac{pq}{(1-p)(1-q)}\text{ for } s=h\\ \frac{q(1-p)}{p(1-q)}\text{ for } s=l\end{array}\right\}. \tag{3}$$

For $s = h$, (3) is strictly increasing in p, and for $s = l$, (3) is strictly decreasing in p. Since both expressions are equal for $p = \frac{1}{2}$, there can be at most one pair (s^*, p^*) that satisfies equation (2). Thus, there is a unique cutoff point (s^*, p^*), such that if $\Pr(w = h|s, p, q) \geq \Pr(w = h|s^*, p^*, q)$ the decision maker takes h, and otherwise, she takes l.■

The strategy of the careerist decision maker is therefore similar to that of the efficient decision maker, who uses a cutoff strategy, with $s^* = l$ and $p^* = q$. We therefore have to check what cutoff point the reputational incentives of the careerist decision maker induce her to choose. As a first step, we can impose more structure on the beliefs of the evaluator E whenever he conjectures that D uses a cutoff point strategy (s^*, p^*):

Lemma 2 *(i) For any action a, $\pi(a, a, \alpha^*) > \pi(a, a', \alpha^*)$, (ii) If $s^* = l$, then $\pi(l, l, \alpha^*) > \pi(h, h, \alpha^*)$ and $\pi(l, h, \alpha^*) > \pi(h, l, \alpha^*)$, (iii) If $s^* = h$, then $\pi(h, h, \alpha^*) > \pi(l, l, \alpha^*)$ and $\pi(h, l, \alpha^*) > \pi(l, h, \alpha^*)$.*

Proof: see the appendix.■

The Lemma follows from Bayesian updating. The first part asserts that the reputation of D is higher if she takes the correct decision; this can arise as a signal on ability since D is more likely to receive the correct signal when she is able, and her strategy is responsive to her signal. In addition, the lemma asserts the following; if $s^* = l$, the reputation that E attributes to those who take l, whether they succeed or fail in taking the right decision, is higher than the reputation they receive when they take h. Intuitively, when $s^* = l$, D takes l only if $p > p^*$ (as in Figure 1, which describes an efficient decision maker). Hence, E knows that if $a = l$, it must be that $p > p^*$, whereas if $a = h$, D may admit a lower type, of $p < p^*$. The opposite happens when $s^* = h$ (as in the example described in Figure 2). In this case, higher reputation is attributed to those who take h.

Given Lemma 2, the next result helps us to focus our analysis:

Lemma 3 *In equilibrium, $s^* = l$.*

Proof: To find an equilibrium, we have to find the cutoff point that satisfies the following fixed point equation (that is, equation (2) with correct beliefs for the evaluator):

$$\frac{\Pr(w = h|s^*, p^*, q)}{\Pr(w = l|s^*, p^*, q)} = \frac{\pi(l, l, \alpha^*) - \pi(h, l, \alpha^*) + \theta}{\pi(h, h, \alpha^*) - \pi(l, h, \alpha^*) + \theta} \tag{4}$$

Assume now that $s^* = h$. Such a value implies that the right-hand-side of (4) is smaller than 1 by Lemma 2, whereas the left-hand-side of (4) is greater than 1 by equation (3). As a result, $s^* = h$ cannot solve (4).■

Intuitively, when the prior is contradicted by too many types of the decision maker, the evaluator may realize that a decision maker who contradicts is not necessarily an able one. Higher reputation would be attributed to those who follow the prior. Moreover, types with $s = h$, are more likely to take the right decision when they follow the prior and hence receive a high reputation (as well as θ). This implies that all types with $s = h$ would rather follow the prior, so no such type can be indifferent. A cutoff point with $s^* = h$ cannot be sustained.

Given that $s^* = l$ in equilibrium, we can now concentrate our analysis on finding p^*. Note that if in equilibrium, $p^* = q$, then it implies that the decision maker behaves efficiently. We can say that the decision maker behaves more (less) efficiently the closer (further) is p^* to q. If $p^* > q$, it means that the decision maker takes h even when she thinks that w is more likely to be l. This is what we call *excessively following the prior*, or '*herding*'. If, on the other hand, the equilibrium value admits $p^* < q$, the decision maker takes l even when she thinks that w is more likely to be h. This is what we call *excessively contradicting the prior*, or '*anti-herding*'.[14] The first Proposition establishes that anti-herding arises in equilibrium.

Proposition 1 *(the Anti-Herding result): For all $\theta \geq 0$, there exists a unique informative equilibrium. In this equilibrium, anti-herding arises and the decision maker excessively contradicts the prior. The decision maker behaves more efficiently when outcome concerns are stronger. Also, when the accuracy of the prior q increases, the decision maker follows the prior more often. However, when career concerns are strong enough, then no matter how high q is, a distortion always arises.*

Proof: see the appendix.■

Figure 3 describes the behavior of the decision maker in equilibrium. The figure focuses on the region of $s = l$ and shows the area (p^*, q) in which D contradicts the prior, inefficiently.

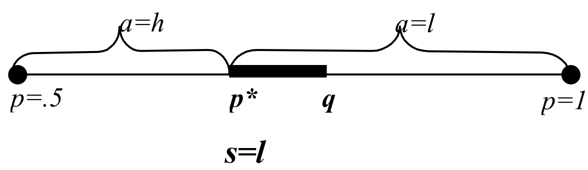

Figure 3: The equilibrium strategy of a careerist decision maker

In equilibrium, a trade-off is created between two types of signals: contradicting the prior and taking the right decision. In order to focus on the reputational trade-off, assume that

[14]My definition of herding and anti-herding is similar to that of Prendergast and Stole (1996).

$\theta = 0$, i.e., the decision maker has only career concerns. First, contradicting the prior is a signal for high ability. This is so because only the most able decision makers take this route. On the other hand, taking the correct action always enhances reputation because it indicates that the decision maker has received, on average, a correct signal.

Consider now a type with $p < q$. When this type follows the prior, she is more likely to take the correct decision, since she utilizes more information. As taking the correct decision is a signal for high ability, this increases her reputation. On the other hand, if she contradicts the prior she receives high reputation. But, she is more likely to be wrong and receive the low reputation attached to a failure. Formally, $\pi(l, l, \alpha) > \pi(h, h, \alpha)$ and $\pi(l, h, \alpha) > \pi(h, l, \alpha)$ in equilibrium, but when a type $p < q$ contradicts the prior she is more likely to receive $\pi(l, h, \alpha)$ whereas when she follows the prior she is more likely to receive $\pi(h, h, \alpha)$, where $\pi(h, h, \alpha) > \pi(l, h, \alpha)$. Thus, low types cannot fully mimic the able types who contradict the prior, because they are more likely to fail. The evaluator knows the state of the world and can 'punish' those that take the wrong decision by holding them, and rightly so, in lower regard. Thus, the tension between the two types of incentives determines the equilibrium level of contradiction.

To see that excessive contradiction must arise, i.e., that $p^* < q$, note that otherwise, if $p^* \geq q$, no trade-off arises. Both the reputation from contradicting the prior is higher at any state, and also types with $p \geq q$ are more likely to take the right decision when they contradict the prior, and hence receive high reputation for being correct. Thus, no type with $p \geq q$ can be indifferent, implying that in equilibrium $p^* < q$.

The proposition also characterizes the decision maker's behavior as a function of the parameters of the model, q and θ. Intuitively, when q increases, the benefit from following the prior, everything else being equal, is higher. This is because the terms of the reputational trade-off change; it becomes more likely to receive the (higher) reputation for taking the correct decision. Hence, more types are inclined to follow the prior, that is, the cutoff point p^* increases.[15]

However, the Proposition states that for small values of θ, efficiency, or a behavior which is close to efficient, cannot be achieved for all values of q. The implication of this result is that

[15] This only implies that more types follow the prior but not necessarily that the size of the distortion, i.e., $q - p^*$ is smaller. Note for example that when $q \to \frac{1}{2}$, the distortion is the smallest since the decision maker behaves efficiently; without almost any prior public knowledge about what is the right state, the only way to signal ability is by taking the correct decision.

the distortion due to career concerns is substantial; even if q admits high values, for example $q \to 1$, the cutoff point p^* is strictly below 1 and there is a positive measure of types who behave inefficiently. For example, when $\theta = 0$ and the density over types $f(p)$ is uniform, I find that p^* is bounded by 0.625. Thus, when $q \to 1$, all types in $(0.625, 1)$ take the wrong decision, consciously and inefficiently.

To see why p^* is bounded, consider the case of $\theta = 0$. Note that if the evaluator conjectures that p^* is very high, for example $p^* \to 1$, then $\pi(l, \cdot, \alpha^*) > \pi(h, \cdot, \alpha^*)$. That is, the reputation from contradicting the prior is higher than that from following the prior regardless of the state of the world, since those who contradict the prior are only the most able types, with $p \to 1$. In particular, $\pi(l, h, \alpha^*) > \pi(h, h, \alpha^*)$, i.e., even if the decision maker contradicts the prior and is found wrong, her reputation is higher compared to the scenario in which she follows the prior and is found correct. Thus, if these are the beliefs of the evaluator, any type of decision maker would rather contradict the prior. This implies that such beliefs for the evaluator cannot be sustained, for any q. Consequently, there is an upper bound on the cutoff point.[16]

Finally, note that in terms of the objectives of the decision maker, career concerns are an important condition; in particular, the distortion becomes smaller when the reputation motive decreases, and in the limit, when career concerns vanish, the decision maker behaves efficiently. Moreover, the Proposition illustrates that anti-herding arises when the decision maker is solely motivated by career concerns, i.e., even in the absence of outcome concerns.[17]

I now examine which of the assumptions about the information structure of the players is necessary for the derivation of anti-herding:

Proposition 2 *For all $\theta \geq 0$, anti-herding is an equilibrium phenomenon when the evaluator does not observe the state of the world, or when the evaluator or the decision maker do not observe the accuracy of the prior. On the other hand, anti-herding cannot arise in any equilibrium if the decision maker does not know her type or if the evaluator cannot observe whether her action contradicts or follows the prior.*

Proof: see the appendix.■

[16] The feature that p^* is bounded for all values of q will turn out to be important, since it implies that when q is high enough anti-herding must arise in an informative equilibrium, even if we change some of the assumptions of the model.

[17] Prendergast and Stole (1996) assume the existence of outcome concerns in order to derive a fully separating equilibrium. I show that at least for a semi-separating equilibrium, career concerns are enough.

I now explain the intuition of the result, which illustrates the robustness of anti-herding and the conditions under which it prevails. I find that anti-herding arises only if E observes the action of the decision maker, or more precisely, whether the decision maker follows or contradicts the prior. Otherwise, D can only signal her type by taking the correct decision and will therefore behave efficiently.

In terms of the information that D holds, I find that a crucial assumption is that D knows her type p. Intuitively, if the decision maker does not know p, she does not know whether her own information is more accurate than the prior and as a result cannot signal it. She only knows her signal s, and can only report information about s. This implies that the only way in which she can signal her type is by taking the correct decision.[18]

The Proposition shows that other assumptions are not important. For example, a common assumption in the literature is that both E and D observe the accuracy of the prior q.[19] Here, I show that the anti-herding result can be generalized beyond this assumption. Even if each of them does not know the exact accuracy of the prior, they both know whether D contradicts or follows the prior. Hence, the same trade-off described above can arise.

Finally, consider the scenario in which E does not observe w. This is highly relevant, since for many decisions, it is never known whether they are correct or not.[20] Two problems arise in this case. The first is to ensure that an informative equilibrium exists, for all $\theta \geq 0$.[21] The second is to ensure that anti-herding arises in equilibrium. But this is solved using the insight gained in Proposition 1. The result there established that the cutoff point p^* is bounded

[18] Effinger and Polborn (2001) derive anti-herding in a set up in which decision makers do not know their type. They assume that there is a direct competition between experts, i.e., that a decision maker is valuable only if she is considered the only smart decision maker. In this case, decision makers have an incentive to differentiate themselves. The implication is therefore that decision makers must know their type for anti-herding to arise, if career concerns are individualistic.

[19] I know of no paper in the herding/anti-herding literature that assumes differently.

[20] Consider for example judicial decision making. If there is no appeal, it is likely that we never find out whether the judge is right or wrong. See Levy (2002).

[21] When the evaluator does not observe w he cannot discipline the decision maker by 'providing' her incentives to take the right decision. This is an obstacle for an informative equilibrium only when $\theta = 0$, since when $\theta > 0$ the decision maker is inherently motivated by taking the correct decision. I show in the appendix that an informative equilibrium can exist even when $\theta = 0$. In the equilibrium, the decision maker behaves informatively but each of her types is indifferent between contradicting and following the prior. The only other paper that analyzes a career concerns model in which E does not observe w is Prendergast and Stole (1996). In their model they indeed substitute this assumption by assuming sufficiently high outcome concerns. I show that this is not necessary, neither for existence nor for anti-herding.

and cannot be 'too high', so that the interval of distortion (p^*, q) has a positive measure. If p^* is too high, then regardless of the state of the world, the reputation from contradicting must be higher than that from following, which cannot be sustained in equilibrium. Since this occurs regardless of the state of the world, it must be true also when E cannot observe w. The cutoff point is bounded in this case as well, for the same reason. As a result, when q is high enough, the cutoff is below q and anti-herding arises.

The 'anti-herding' effect may create implications above and beyond those described above. For example, if the prior is formed using reports of advisers, we can analyze whether the decision maker chooses to consult at all. Additionally, we can analyze the incentives of the advisers to transmit information to the decision maker, anticipating her contrarian behavior. In what follows, I analyze extensions of the model to explore some of these implications.

4 Extensions

So far, I assumed that there exists a known prior, which indicates that the state of the world is likely to be h, with an accuracy $q > \frac{1}{2}$. Alternatively, we can model the formation of this 'prior' by assuming that it is based on previous information or advice which is public and freely available to the decision maker. This allows us to endogenize some of the assumptions about the parameter q.

Assume that the initial prior is the symmetric one, i.e., that each state can occur with probability $\frac{1}{2}$. There is an adviser who receives a signal about the state of the world. The accuracy of the adviser's signal is known to be q. If the adviser transmits his signal truthfully (and publicly), then this structure is exactly identical to the structure assumed in the model. In the first extension, I analyze how much information the decision maker aggregates before she takes her decision. That is, I allow the decision maker to decide whether to consult the adviser or not and thus control the accuracy of her 'prior'.

4.1 Strategic information aggregation

An adviser A receives a signal $s^A \in \{l, h\}$, with accuracy q. That is, $\Pr(s^A = w|w, q) = q > \frac{1}{2}$. The players E and D know the parameter q. In the first stage of the game, the decision maker can decide whether to consult her adviser or not, given q, her own information (s, p) and the initial symmetric prior on the state of the world. I assume that consultation is costless and that the adviser transmits his signal s^A truthfully upon consultation. I also assume that the decision whether to consult or not is observed by the evaluator E. The timing of the game is therefore the following:

1. It is common knowledge that $\Pr(w = h) = \frac{1}{2}$ and that $\Pr(s^A = w|w, q) = q$.
2. w is realized, A learns $s^A \in \{l, h\}$, and D receives a private signal (s, p) about w.
3. D may consult A and learn s^A.
4. D takes an action a.
5. E forms beliefs on p, given a, w, q, the decision to consult, s^A if she consults, and $f(p)$.

The efficient course of action is for D to consult the adviser and follow his advice when $p < q$ and follow her own signal s when $p > q$. But will D necessarily aggregate information? Intuitively, if D decides not to gather additional information, she reduces the likelihood of taking the appropriate action. But taking the correct action based only on her own information, is a more challenging test than doing so after collecting advice. The willingness of a decision maker to impose this test upon herself may show that she has faith in her own private information. To find if the decision not to consult can arise as a signal on ability, Proposition 3 identifies an equilibrium in which some types consult and some do not.[22]

Proposition 3 *(The Strategic Consultation Result): When θ is sufficiently high, there exists an equilibrium in which some types do not consult. In particular, the most able decision makers do not consult and follow their own signal whereas the less able decision makers consult and follow advice. Moreover, distortion arises since decision makers with types $p < q$ do not consult as well.*

Proof: see the appendix.■

Figure 4 depicts the equilibrium behavior of the decision maker, given some signal s and types $p \in [.5, 1]$. In equilibrium, all types with $p > p'$ do not consult and follow their own signal, i.e., take $a = s$. All other types, i.e., $p \in (.5, p')$, consult and follow advice ($a = s^A$). The area (p', q) describes the inefficiency created by career concerns.

[22] An equilibrium in the model is defined by an equilibrium in each of the continuation games calculated given some beliefs of E that a type p that consults belongs to the set $C \subseteq [.5, 1]$ and a type p that acts unilaterally belongs to the set $U = [.5, 1] \backslash C$, and by the condition that a decision maker of type $p \in U$ prefers to act unilaterally and a decision maker of type $p \in C$ prefers to consult. I focus on symmetric strategies, that is, the decision to consult depends on p but not on the content of s (since at the first decision node, the prior is symmetric).

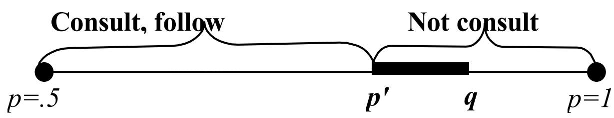

Figure 4: Not consulting as a signal of ability

In this equilibrium, the following trade-off arises. Consulting an adviser may increase the probability of taking the right decision (and receiving θ). However, choosing not to consult allows the decision maker to be perceived as a high ability type, since only able types choose to decide alone.[23] Obviously, able types have sufficiently accurate information so they can opt for high reputation, whereas the less able types need to rely on advice in order to receive θ and solve the trade-off by favouring consultation. Since the equilibrium separates between those who do not consult, the most able, and those who consult, the least able, it can exist only for high values of θ. Otherwise, the less able types are not 'willing' to distinguish themselves. Finally, the equilibrium reveals inefficient behavior on behalf of the decision maker; as in the previous result, the cutoff point is lower than q and hence some types with $p < q$ distort their actions by not consulting, in order to signal high ability.

Consider now an equilibrium in which all types consult.[24] Then, s^A becomes public. Suppose, without loss of generality, that $s^A = h$. The structure of the game is now identical to the model analyzed in the previous section, in which the prior (that $w = h$, with accuracy q) is public. Thus, the equilibrium described in Proposition 1 arises; there is a cutoff point $p^* < q$. If $s \neq s^A$, i.e., $s = l$, and $p > p^*$, D takes l and otherwise she takes h. In other words, if $p < p^*$, D follows advice whereas if $p > p^*$, D plans to ignore advice and to follow her own signal, s. If it turns out that $s^A = s$, then she appears to follow advice (as do able types with $s = h$ in our example); while if $s^A \neq s$, i.e., $s = l$, she appears to contradict advice. This equilibrium, as Proposition 1 implies, exists for all values of θ. For convenience, the next figure repeats the description of this equilibrium, this time for a general s and $p \in [.5, 1]$:

[23] Intuitively, if reputation for acting unilaterally would be low, then none of the decision maker's types would follow this route, which is inferior to consulting and following advice both in terms of career and outcome concerns.

[24] For all θ, such an equilibrium can be sustained with proper out-of-equilibrium beliefs for the evaluator.

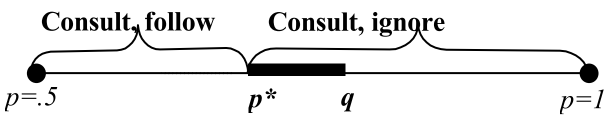

Figure 5: Contradicting advice as a signal of ability.

When D can choose whether to consult, there may exist two equilibria with a similar structure. In both, the most able decision makers act on the basis of their own information only (i.e., ignore information or do not collect it), whereas the less able decision makers follow advice. As a result, the signal of *not consulting* and the signal of *contradicting advice* may be substitutable; both D's decision whether to consult or not and her choice of how to use the advice she obtains can serve as signalling devices.

It is therefore interesting to investigate which of the equilibria described in Figures 4 and 5 is more efficient, i.e., which yields higher probability of taking the right decision:[25]

Proposition 4 *The probability of making the right decision is higher when the decision maker signals her type by contradicting advice.*

Proof: see the appendix.■

Intuitively, in the equilibrium in which the most able decision makers signal their type by not consulting (Figure 4) there is a high degree of separation between types. Since the decision to consult is public, all the types who do not consult distinguish themselves from less able types. On the other hand, the equilibrium in which contradicting advice is a signal of ability (Figure 5) is characterized by a high degree of pooling; the more able decision makers ignore advice but if the adviser's recommendation agrees with their signal, they appear as if they 'follow' advice. This allows the less able decision makers to mimic the most able decision makers more often.[26] Since the signal of not consulting is stronger, relatively more types are tempted to use it and distort their action. As a result, the equilibrium in which this signal is used is less efficient. In other words, $p' < p^* < q$.

4.2 Strategic advisers

So far I assumed that the adviser A transmits his information truthfully. I now endogenize the behavior of the adviser. In particular, the adviser may be careerist as well. Assume

[25] I thank a referee for suggesting this question.

[26] This is the reason why the equilibria in which decision makers ignore advice can be sustained for all values of θ, as opposed to the equilibrium in which some decision makers do not consult.

therefore that the accuracy of the adviser q is unknown to E and D, but they know its density function $g(q)$ over $[.5, 1]$. Let the utility function of A be $\pi^A + \theta^A I$, where π^A are the beliefs of E on the type of A and $\theta^A \geq 0$ measures how much A cares about the outcome. Finally, I allow the adviser A to transmit a message about (s^A, q) which is not necessarily truthful. The timing of the game is as follows:

1. It is common knowledge that $\Pr(w = h) = \frac{1}{2}$.
2. w is realized, A receives a signal (s^A, q) and D receives a signal (s, p) about w.
3. A transmits a message $m \in \{l, h\} \times [.5, 1]$ about (s^A, q).
4. D takes an action a given m, $g(q)$, and (s, p).
5. E forms beliefs π on p, and π^A on q, given $a, m, w, f(p)$ and $g(q)$.

It is then easy to prove the following:

Proposition 5 *(The Sub-optimal Sharing of Information Result): For all θ^A, there is no equilibrium in which the adviser always reveals his information (s^A, q).*

Proof: see the appendix.■

Intuitively, consider an adviser who reflects whether to exaggerate the accuracy of his report. If he poses as a more talented adviser, less types of the decision maker distort their action, since they face a more accurate prior or advice. The decision maker is therefore induced to behave more efficiently. This implies that if the adviser exaggerates his report, he increases both his reputation, and the probability of receiving θ^A. He therefore has an incentive to fool the decision maker and an equilibrium with full information revelation cannot exist. Thus, even if the adviser is only concerned about the outcome and has no concerns about his reputation, i.e., for $\theta^A \rightarrow \infty$, he cannot transmit his information fully, since he knows that the careerist decision maker distorts her actions and abuses his information.

To summarize, consultation in the presence of career concerns elicits low quality of information. Although the career concerns of the adviser and the decision maker are not conflicting, since they are not competing for the same job, the personal career concerns of either blocks truthful information revelation.[27]

[27] I do not fully characterize the equilibria in the case of strategic advisers. In a companion working paper (Levy, 2000) I show that contradicting an adviser still arises as a signal for high ability even in the presence of strategic advisers.

5 Conclusion

In this paper, I have shown how a decision maker who is concerned about her career distorts her decision by 'anti-herding'. Following several papers that have introduced anti-herding results in different contexts, the paper generalizes this result and illustrates the necessary conditions for its derivation. Career concerns are a necessary condition, and as long as career concerns are individualistic, that is, the decision maker does not care about the reputation of others, she must know her ability (type). Other assumptions are irrelevant. Anti-herding is the unique informative equilibrium behavior even if the decision maker cares only for reputation and has no outcome concerns. Anti-herding arises for any distribution over types, and even when the signals of smart decision makers are not correlated. Finally, it arises if the evaluator observes the environment only partially, i.e., he does not observe the state of the world or the accuracy of the public prior.

References

[1] Avery, C.N. and J. Chevalier, 1999, Herding over the career, Economic Letters 63, 327-333.

[2] Banerjee, A., 1992, A simple model of herd behavior, Quarterly Journal of Economics 107, 797-817.

[3] Bikhchandani, S., D. Hirshleifer and I. Welch, 1992, A theory of fads, fashion, custom, and cultural change as informational cascades, Journal of Political Economy 100, 992-1026.

[4] Effinger, M.R. and M.K. Polborn, 2001, Herding and anti-herding: A model of reputational differentiation, European Economic Review 45, 385-403.

[5] Fama, E.F., 1980, Agency problems and the theory of the firm, Journal of Political Economy 88, 288-307.

[6] Holmström, B., 1982, Managerial incentives problems: A dynamic perspective, in: Essays in economics and management in honor of Lars Wahlbeck (Helsinki: Swedish School Econ).

[7] Holmström, B. and J. Ricart i Costa, 1986, Managerial incentives and capital management, Quarterly Journal of Economics 101, 835-860.

[8] Kreps, D. and R. Wilson, 1982, Reputation and imperfect information, Journal of Economics Theory 27, 253-279.

[9] Lazear, E. and S. Rosen, 1982, Rank-order tournaments as optimum labor contracts, Journal of Political Economy 89, 841-864.

[10] Levy, G., 2002, Careerist judges, mimeo, LSE and TAU.

[11] Levy, G., 2000, Strategic consultation in the presence of career concerns, Sticerd Discussion Paper No. TE/00/404.

[12] Milgrom, P., 1982, Good news and bad news: Representation theorems and applications, Bell Journal of Economics 17, 380-391.

[13] Milgrom, P. and J. Roberts, 1982, Predation, reputation and entry deterrence, Journal of Economics Theory 27, 280-312.

[14] Morris, S., 2000, Political correctness, Journal of Political Economy 109, 231-265.

[15] Ottaviani, M. and P. Sørensen, 2002, Professional advice: The theory of reputational cheap talk, mimeo, LBS and Copenhagen University.

[16] Ottaviani, M. and P. Sørensen, 2001, Information aggregation in debate: Who should speak first?, Journal of Public Economics 81, 393-421.

[17] Ottaviani, M. and P. Sørensen, 2000, Herd behavior and investment: A Comment, American Economic Review 90, 695-704.

[18] Prendergast, C. and L. Stole, 1996, Impetuous youngsters and jaded old-timers: Acquiring a reputation for learning, Journal of Political Economy 104, 1105-1134.

[19] Scharfstein, D. and J. Stein, 1990, Herd behavior and investment, American Economic Review 80, 465-479.

[20] Trueman, B., 1994, Analyst forecasts and herding behavior, Review of financial studies 7, 97-124.

[21] Zwiebel, J., 1995, Corporate conservatism and relative compensation, Journal of Political Economy 103, 1-25.

Appendix

Proof of Lemma 2: (i) $\pi(a,a,\alpha^*) > \pi(a,a',\alpha^*)$:

$\pi(a,w,\alpha^*)$ is an expectation over p, using an updated density function given the observations of a and w, i.e.,

$$\pi(a,w,\alpha^*) = \int_{.5}^{1} pf(p|a,w,\alpha^*)dp$$

where $f(p|a,w,\alpha^*)$ is the updated density function on p given the observations of a, w and the knowledge of α^*. To show that $\int_{.5}^{1} pf(p|a,a,\alpha^*)dp > \int_{.5}^{1} pf(p|a,a',\alpha^*)dp$ we can use the MLRP property, i.e., show that $\frac{f(p|a,a,\alpha^*)}{f(p'|a,a,\alpha^*)} \geq \frac{f(p|a,a',\alpha^*)}{f(p'|a,a',\alpha^*)}$ for $p \geq p'$ with a strict inequality for at least one pair of values p and p'. We will show that the MLRP is satisfied in this case, since $\frac{f(p|a,a,\alpha^*)}{f(p|a,a',\alpha^*)}$ increases with p.

Given that the decision maker uses a cutoff strategy, then she takes l only when $s = l$ and $p > p^*$. Hence:

$$\frac{f(p|l,l,\alpha^*)}{f(p'|l,l,\alpha^*)} = \frac{p}{p'} \text{ for } p > p' > p^* \text{ and 0 otherwise,}$$
$$\frac{f(p|l,h,\alpha^*)}{f(p'|l,h,\alpha^*)} = \frac{1-p}{1-p'} \text{ for } p > p' > p^* \text{ and 0 otherwise.}$$

Hence, $\frac{f(p|l,l,\alpha^*)}{f(p|l,h,\alpha^*)}$ increases in p. A similar analysis holds for $a = h$.

(ii) If $s^* = l$, then $\pi(l,l,\alpha^*) > \pi(h,h,\alpha^*)$ and $\pi(l,h,\alpha^*) > \pi(h,l,\alpha^*)$:

Similarly to part (i), I show that when $s^* = l$, $\int_{.5}^1 pf(p|l,l,\alpha^*)dp > \int_{.5}^1 pf(p|h,h,\alpha^*)dp$ and $\int_{.5}^1 pf(p|l,h,\alpha^*)dp > \int_{.5}^1 pf(p|h,l,\alpha^*)dp$ by using the MLRP. To see that $\int_{.5}^1 pf(p|l,l,\alpha^*)dp > \int_{.5}^1 pf(p|h,h,\alpha^*)dp$, I have to show that $\frac{f(p|l,l,\alpha^*)}{f(p'|l,l,\alpha^*)} \geq \frac{f(p|h,h,\alpha^*)}{f(p'|h,h,\alpha^*)}$ for $p > p'$, or that $\frac{f(p|l,l,\alpha^*)}{f(p|h,h,\alpha^*)}$ increases with p. But $\frac{f(p|l,l,\alpha^*)}{f(p|h,h,\alpha^*)} = 1$ if $p > p^*$ and $\frac{f(p|l,l,\alpha^*)}{f(p|h,h,\alpha^*)} = 0$ if $p < p^*$, hence the result follows. Similarly, to see that $\int_{.5}^1 pf(p|l,h,\alpha^*)dp > \int_{.5}^1 pf(p|h,l,\alpha^*)dp$, we have to show that $\frac{f(p|l,h,\alpha^*)}{f(p|h,l,\alpha^*)}$ increases with p, but it is trivial since $\frac{f(p|l,h,\alpha^*)}{f(p|h,l,\alpha^*)} = 1$ if $p > p^*$ and $\frac{f(p|l,h,\alpha^*)}{f(p|h,l,\alpha^*)} = 0$ otherwise.

The analogous results for $s^* = h$, part (iii), follow from symmetry.■

Proof of Proposition 1: In particular, I will show that there is a unique cutoff point $p^*(q,\theta)$ which satisfies $p^*(q,\theta) < q$, that the cutoff $p^*(q,\theta)$ is increasing in θ and in q, and that when θ is sufficiently small, $p^*(q,\theta)$ is bounded for all q.

The proof of Lemma 3 establishes that the equilibrium (s^*, p^*) solves

$$\frac{\Pr(w = h|s^*, p^*, q)}{\Pr(w = l|s^*, p^*, q)} = \frac{\pi(l,l,\alpha^*) - \pi(h,l,\alpha^*) + \theta}{\pi(h,h,\alpha^*) - \pi(l,h,\alpha^*) + \theta} \equiv k_\theta(\alpha^*) \tag{5}$$

and that the solution admits $s^* = l$. Consider now the case of $p^* \geq q$. By Lemma 2, the right-hand-side of (5) is greater than 1, whereas by (3) the left-hand-side of (5) is smaller than 1. Hence, a solution exists by continuity and it admits $s^* = l$ and $p^* < q$. It is left to show the properties of $p^*(q,\theta)$ and its uniqueness.

Step 1: Characterizing the upper bound on $p^*(q,\theta)$:

Let $s^* = l$. I will now show that there exists a unique $\tilde{p} > .5$, such that for all $p^* \geq (<) \tilde{p}$, $\pi(h,h,\alpha^*) \leq (>) \pi(l,h,\alpha^*)$ with equality only for $p = \tilde{p}$. This shows that in equilibrium, $p^*(q,\theta) < \tilde{p}$.

The expression for $\pi(h,h,\alpha^*)$ is

$$\pi(h,h,\alpha^*) = \int_{.5}^{p^*} p \frac{f(p)}{\int_{.5}^{p^*} f(p)dp + \int_{p^*}^1 pf(p)dp} dp + \int_{p^*}^1 p \frac{pf(p)}{\int_{.5}^{p^*} f(p)dp + \int_{p^*}^1 pf(p)dp} dp$$

Taking the derivative of $\pi(h,h,\alpha^*)$ w.r.t p^*, it is $\frac{(1-p^*)f(p^*)(p^*-\pi(h,h,\alpha^*))}{\int_{.5}^{p^*} f(p)dp + \int_{p^*}^1 pf(p)dp}$ and hence $\pi(h,h,\alpha^*)$ is a monotonically decreasing function as long as $p^* < \pi(h,h,\alpha^*)$ and a monotonically increasing function when $p^* > \pi(h,h,\alpha^*)$. When $p^* \to .5$, $p^* < \pi(h,h,\alpha^*)$ and when $p^* \to 1$, $\pi(h,h,\alpha^*) < p^*$. Therefore, there exists p^* such that $p^* = \pi(h,h,\alpha^*)$. Denote this value by p'. This value is unique since $\frac{d\pi(h,h,\alpha^*)}{dp^*}|_{p^*=\pi(h,h,\alpha^*)} = 0$.

On the other hand, $\pi(l,l,\alpha^*)$ and $\pi(l,h,\alpha^*)$ are averages over l for $l > p^*$ and thus increase with p^* for all $p^* > .5$. Also, since only values of $p > p^*$ are included in the computation of these averages, $\pi(h,w,\alpha^*) > p^*$ for all p^*. By the above, when $p^* \to 1$, $\pi(l,h,\alpha^*) > p^* > \pi(h,h,\alpha^*)$. When $p^* = .5$, $\pi(h,h,\alpha^*) = \pi(l,l,\alpha^*) > \pi(l,h,\alpha^*)$. Then, there must exist some $p^* \in (.5,1)$ satisfying $\pi(h,h,\alpha^*) = \pi(l,h,\alpha^*)$. Denote this value by $\tilde{p}$. Note that $\tilde{p} < p'$ and that $\tilde{p}$ is unique.

Step 2: Uniqueness: by step 1, when $p^* \in [.5,\tilde{p}]$, $\frac{d\pi(l,w,\alpha^*)}{dp^*} > 0$ and $\frac{d\pi(h,h,\alpha^*)}{dp^*} < 0$. A similar analysis as in step 1 holds for $\frac{d\pi(h,l,\alpha^*)}{dp^*} < 0$. This implies that $\frac{\partial}{\partial p^*}k_\theta(\alpha^*) > 0$, and hence the solution is unique.

Step 3: The cutoff point $p^*(q,\theta)$ is increasing in q and in θ :

To see that $p^*(q,\theta)$ is increasing in q, I take the total differentiation of the equilibrium condition:

$$\frac{\partial p^*(q,\theta)}{\partial q} = \frac{\frac{\partial}{\partial q}\frac{\Pr(w=h|s^*,p^*,q)}{\Pr(w=l|s^*,p^*,q)}}{\frac{\partial}{\partial p^*}k_\theta(\alpha^*) - \frac{\partial}{\partial p^*}\frac{\Pr(w=h|s^*,p^*,q)}{\Pr(w=l|s^*,p^*,q)}}$$

Thus, $\frac{\partial p^*(q,\theta)}{\partial q} > 0$ since $\frac{\partial}{\partial q}\frac{\Pr(w=h|s^*,p^*,q)}{\Pr(w=l|s^*,p^*,q)} > 0$, $\frac{\partial}{\partial p^*}k_\theta(\alpha^*) > 0$ by Step 2, and $\frac{\partial}{\partial p^*}\frac{\Pr(w=h|s^*,p^*,q)}{\Pr(w=l|s^*,p^*,q)} < 0$. To see that the solution increases in θ, note that

$$sign\frac{\partial}{\partial\theta}k_\theta(\alpha^*) = sign(\pi(h,h,\alpha^*) - \pi(l,h,\alpha^*) - (\pi(l,l,\alpha^*) - \pi(h,l,\alpha^*)) < 0$$

since $k_\theta(\alpha^*) > 1$ in equilibrium. Hence, the solution $p^*(q,\theta)$ induces a lower value for $\frac{\Pr(w=h|s^*,p^*,q)}{\Pr(w=l|s^*,p^*,q)}$ when θ is higher, implying that $p^*(q,\theta)$ is higher.■

Proof of Proposition 2: The proposition involves several claims which will be shown separately in different steps.

Step 1: If E does not observe w, there exists a unique informative equilibrium, in which the decision maker anti-herds whenever q is high enough.

Proof: Consider an equilibrium in which D behaves as in Proposition 1. That is, there exists a cutoff point (s',p') such that she takes h iff $\Pr(w=h|s,p,q) \geq \Pr(w=h|s',p',q)$. Assume also that $s'=l$. If E believes that this is her strategy, then given each action, he can updates his beliefs about the state of the world. In particular, denote by $q_a(\alpha')$, E's beliefs that $w=a$ given an action a. These believes can be updated by Bayes rule. Then an equilibrium is a solution to the fixed point equation:

$$q_h(\alpha')\pi(h,h,\alpha') + (1-q_h(\alpha'))\pi(h,h,\alpha') + \Pr(w=h|s,p,q)\theta =$$
$$q_l(\alpha')\pi(l,l,\alpha') + (1-q_l(\alpha'))\pi(l,h,\alpha') + \Pr(w=l|s,p,q)\theta$$

The right-hand-side (left-hand-side) is the expected utility of the decision maker from taking $l(h)$. Consider now $p' \to \frac{1}{2}$. When $p' = \frac{1}{2}$, I show in Levy (2002) that $q_h(\alpha') > q_l(\alpha')$. Since by Lemma 2 in this case the reputation for following and contradicting is equal, the left-hand-side is larger than the right-hand-side. On the other hand, for $p' \to 1$, and since in particular $p' > \tilde{p}$, where $\tilde{p}$ is the value of the cutoff point which satisfies $\pi(l,h,\alpha) = \pi(h,h,\alpha)$ and is defined in the proof of Proposition 1, the right-hand-side is higher than the left-hand-side. Thus, an equilibrium exists for all values of q. But also, for high enough values of q, if $p' > q > \tilde{p}$, the right-hand-side is larger than the left-hand-side, implying by continuity that in equilibrium $p' < q$. Hence, whenever $q > \tilde{p}$, this equilibrium must involve anti-herding.□

Step 2: Anti-herding arises disregarding whether E or D observe q or not.

Proof: Proposition 1 proves the case when both observe q. When D does not observe q, then she believes that she is found correct with probability $\frac{E(q)(1-p)}{E(q)(1-p)+p(1-E(q))}$ if she follows the prior and $\frac{p(1-E(q))}{E(q)(1-p)+p(1-E(q))}$ if she contradicts the prior. This is true no matter if E observes q or not since E observes w, and hence D has only to conjecture what is the probability that the state is h or l. Thus, the equilibrium is equivalent to the one with $\hat{q} = E(q)$. Finally, the case where E does not observe q but D does, is analyzed in Levy (2000) and I refer the reader to the proof there.□

Step 3: When D does not know p, in any informative equilibrium, she behaves efficiently, i.e., she follows her signal for $E(p) > q$ and follows the prior if $E(p) < q$.

Proof: I will show the claim for $\theta = 0$. The result then must hold for positive values of θ, since these values only increase the motivation to take the right decision. The expected utility from taking an action a is

$$\Pr(a = w|s,q)\pi(a,a,\alpha) + (1 - \Pr(a = w|s,q))\pi(a,a',\alpha) \tag{6}$$

where

$$\Pr(a = h|s,q) = \left\{ \begin{array}{l} \frac{E(p)q}{(1-E(p))(1-q)+E(p)q} \text{ if } s = h, \\ \frac{(1-E(p))q}{(1-E(p))q+E(p)(1-q)} \text{ if } s = l \end{array} \right\}$$

When D does not know her type, she can contingent her strategy only on her signal s, that is, whether $s = h$ or $s = l$. Hence, $\alpha(s,p,q) = \alpha(s,q)$ for all p. For brevity, let us denote the probability with which D follows her signal by α_s. Then in equilibrium,

$$\pi(h,h,\alpha) = \int_{.5}^{1} p \frac{[\alpha_h p + (1-\alpha_l)(1-p)]f(p)}{\int_{.5}^{1}[\alpha_h p + (1-\alpha_l)(1-p)]f(p)dp} dp$$

and analogously we can define the rest. We now consider equilibria in which either $\alpha_l < 1$ or $\alpha_h < 1$ (but not both since there cannot be an informative equilibrium such that a decision maker is indifferent given both signals). It is first easy to see using the MLRP that if $\alpha_l < (=)1$ and $\alpha_h = (<)1$, then $\pi(l,l,\alpha) > (<)\pi(h,h,\alpha)$ and $\pi(l,h,\alpha) > (<)\pi(h,l,\alpha)$. Also, for any a, $\pi(a,a,\alpha) > \pi(a,a',\alpha)$. I will now show $\alpha_h = 1$. Assume that $\alpha_h < 1$ and therefore $\alpha_l = 1$. This implies that the reputation from h is higher than the reputation from l. Then, for any q and $E(p)$, it is better to take h when $s = h$, implying that $\alpha_h = 1$. Now, when $E(p) > q$, it also has to be that $\alpha_l = 1$. Assume to the contrary that $\alpha_l < 1$. Then the reputation from l is higher then the reputation from h. If $E(p) > q$, then it is the case that when $s = l$, there is a higher probability that the state is l. Then, it has to be that D takes l, so $\alpha_l < 1$ is not an equilibrium. We conclude that when $E(p) > q$, it must be that $\alpha_l = \alpha_h = 1$, and hence the decision maker follows her signal. Finally, we need to show that when $E(p) < q$, $\alpha_l = 0$. The proof shows that for any $\alpha_l \leq 1$, the utility from taking h when $s = l$ is higher than from $s = l$.[28] This implies that the only possible equilibrium has $\alpha_l = 0$, implying that the decision maker always follows the prior.□

Finally, it is trivial that if E does not observe whether the action of D contradicts or follows the prior then the decision maker can only signal her type by taking the right decision, if at all. Thus, the conjunction of Steps 1, 2 and 3 completes the proof of Proposition 2.■

Proof of Proposition 3: In particular, I will show that the equilibrium is characterized by a cutoff point $p'(q,\theta) < q$. Let us conjecture the equilibrium described in the Proposition. When the decision maker does not consult for all types $p > p'(q,\theta)$ and follows her signal, then her reputation is (where the superscript N denotes not consulting):

$$\pi^N(a,a,\alpha') = \int_{p'(q,\theta)}^{1} p\frac{pf(p)}{\int_{p'(q,\theta)}^{1} pf(p)dp}dp,$$

where $\pi^N(a,a',\alpha')$ can be similarly defined. It is easy to see, by MLRP, that $\pi^N(a,a,\alpha') > \pi^N(a,a',\alpha')$.We can therefore express the expected utility from not consulting, given the conjecture of the equilibrium, by

$$p\pi^N(a,a,\alpha') + (1-p)\pi^N(a,a',\alpha') + p\theta$$

When the decision maker consults for all $p < p'(q,\theta)$ and follows advice, her reputation is (where the superscript A denotes consulting the adviser A):

$$\pi^A(a,a,\alpha') = \int_{.5}^{p'(q,\theta)} p\frac{f(p)}{\int_{p'(q,\theta)}^{1} f(p)dp}dp = \pi^A(a,a',\alpha')$$

In this case, the observation of w entails no information about the type of D since the decision is not responsive to her signals. Note that $\pi^N(\cdot,\cdot,\alpha') > \pi^A(\cdot,\cdot,\alpha')$ for any a and w, since the types who do not consult are known to be in $[p'(q,\theta),1]$ whereas those who consult are in $[.5,p'(q,\theta)]$. Thus, the expected utility from consulting and following advice, for all p, is:

$$\pi^A(a,a,\alpha') + q\theta$$

[28]Detailes available upon request. The proof is techinically easy but tedious and long.

In the conjectured equilibrium, we have to find a type $p'(q,\theta)$ that solves the fixed point equation:

$$\pi^A(a,a,\alpha') + q\theta = p'(q,\theta)\pi^N(a,a,\alpha') + (1-p'(q,\theta))\pi^N(a,a',\alpha') + p'(q,\theta)\theta$$

When $p'(q,\theta) \to \frac{1}{2}$, then for high enough θ, the left-hand-side is higher from the right-hand-side because $q > p'(q,\theta)$, implying that the solution $p'(q,\theta) > \frac{1}{2}$. On the other hand, when $p'(q,\theta) \geq q$, then the utility from not consulting is higher than that from consulting since both the reputation from not consulting is higher and the probability of receiving θ. Hence, an equilibrium exists and satisfies $p'(q,\theta) < q$.■

Proof of Proposition 4. Consider the equilibrium in which all types consult and the most able types ignore advice. It is trivial to show that this equilibrium is analogous to the one described in Proposition 1. Consider now the cutoff point $p^*(q,\theta)$ in this equilibrium. Re-arranging the equilibrium condition, it solves:

$$\frac{q(1-p^*)}{p^*(1-q)} = \frac{\pi^c(a,a,\alpha^*) - \pi^f(a,a',\alpha^*) + \theta}{\pi^f(a,a,\alpha^*) - \pi^c(a,a',\alpha^*) + \theta} \tag{7}$$

where $\pi^f(a,a,\alpha^*)$ is the reputation from following advice while taking the correct decision, $\pi^c(a,a,\alpha^*)$ is the reputation from contradicting advice and taking the correct decision, and so on. Now consider the equilibrium fixed point equation when some types do not consult, i.e., the equilibrium described in Proposition 3:

$$\pi^A(a,a,\bar{\alpha}) = \bar{p}(q,\theta)\pi^N(a,a,\bar{\alpha}) + (1-\bar{p}(q,\theta))\pi^N(a,a',\bar{\alpha}) + (\bar{p}(q,\theta)-q)\theta$$

Note that:

$$\pi^N(a,w,\alpha^*) = \pi^c(a,w,\alpha^*);\ \pi^A(a,w,\alpha^*) < \pi^f(a,w,\alpha^*) < \pi^c(a,w,\alpha^*)$$

Given the above, I now plug the solution of the fixed point equation $p^*(q,\theta)$ from (7) to the fixed point equation in $\bar{p}(q,\theta)$ above. It is then easy to show that:

$$\pi^A(a,a,\alpha^*) < p^*(q,\theta)\pi^N(a,a,\alpha^*) + (1-p^*(q,\theta))\pi^N(a,a',\alpha^*) + (p^*(q,\theta)-q)\theta$$

which implies that at $p^*(q,\theta)$, the decision maker rather not consult, and hence the solution to the fixed point equation must admit a lower level of p, i.e., $\bar{p}(q,\theta) < p^*(q,\theta)$.■

Proof of Proposition 5: If A reveals all his information truthfully, given a message (s^A, q), the equilibrium is as described in Proposition 1. But then a deviation of a type q to a report of $q+\varepsilon$ increases both the efficiency of decision making (and induces a higher probability to receive θ^A) and the beliefs of E. Thus, no type can report his type truthfully.■[29]

[29] The proof actually illustrates that the adviser cannot transmit truthfully any connected interval.