TI 2002-123/1
Tinbergen Institute Discussion Paper

Vanity in Politics

Guido Suurmond[1]
Otto H. Swank[2,3]
Bauke Visser[2]

[1] *University of Leiden,* [2] *Faculty of Economics, Erasmus University Rotterdam,* [3] *Tinbergen Institute.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Amsterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to wdriessen@few.eur.nl.

Vanity in Politics: A Problem?

Guido Suurmond[*]
University of Leiden

Otto H. Swank[†]
Erasmus University Rotterdam
Tinbergen Institute

Bauke Visser[‡]
Erasmus University Rotterdam

March 13, 2003

Abstract

Can vanity do any good? It may seem obvious to answer this question in the negative, as economists have shown how reputational concerns lead agents e.g. to ignore valuable information, to herd, and to become overly risk averse. We explore how proud agents may be a social blessing. An agent may exert effort to become informed about the uncertain benefits of a project. A smart agent's efforts make him better informed; a dumb agent's efforts are to no avail. If an agent does not know his type, pride is socially beneficial. If an agent knows his type, a dumb agent takes inefficient, unconventional decisions to mimick a smart agent. The latter exerts more effort in order not to be mistaken for a dumb. This holds whether or not project rejection is a save haven for the dumb.

Keywords: Reputation, information collection

JEL Classification: D72, D82, D83

[*]G.Suurmond@law.Leidenuniv.nl
[†]swank@few.eur.nl
[‡]bvisser@few.eur.nl

Vanity and pride are different things, though the words are often used synonymously. A person may be proud without being vain. Pride relates more to our opinion of ourselves, vanity to what we would have others think of us.
Jane Austin, 1813, Pride and Prejudice

1 Introduction

Can vanity do any good? It may seem obvious to answer this question in the negative. Vanity or pride is one of the Seven Deadly Sins. The "excessive belief in one's own abilities"[1], often combined with "the belief that [one is] better or more important than other people"[2] has been described by philosophers and depicted by many artists as the root of all sins. Although vanity and pride are sometimes used as synonyms, we follow the distinction Jane Austin makes: vanity refers to acting in a way such as to favourably influence other people's impression of oneself.

An incarnation of vanity also has gained a bad name with the worldly philosophers. The care for reputation typically leads a rational agent to ignore valuable, private information, which in turn gives rise to an outcome that is socially inferior to what would have been attained otherwise. Examples abound. In Holmström (1999), a manager who is concerned about his reputation may become more risk averse than is optimal for his firm, due to the negative effect on reputation that follows upon failure. When decisions are made sequentially by agents who care about their reputations, as is the case in Scharfstein and Stein (1990), this welfare loss stems from the conscious emulation of actions that were taken previously in an attempt to look smart and to share the blame if the decision works out to be wrong. In Prendergast and Stole (1996), an individual manager has to make decisions repeatedly and wants to be perceived as someone who is talented and dares to act upon his private information. Two distortions result from the eagerness of the manager to establish this reputation: exaggeration and conservatism. First, an impetuous young manager responds too much to new information about a project to show trust in his competence. Second, an experienced manager responds too little to new information as doing otherwise would upset the belief that his previous

[1] http://deadlysins.com/sins/pride
[2] 'Pride' in Cambridge International Dictionary of English.

actions were correct.[3]

Finally, it is often heard that the combination of elections and the power of the media in contemporary society cause politicians to care too little about content and too much about the way their decisions reflect upon them. The political campaign that preceded the elections in May 2002 in the Netherlands is a case in point. For fear of being made a scapegoat, many established politicians followed the herd. A general feeling of disenchantment with the existing body politic among the Dutch was the result. At the same time, in their eagerness to stand out certain aspirants maintained unconventional positions warranting 'clarification', rectification, or retraction the very next day. Many were left with the feeling that media attention and big egos were all that was at stake.

To examine whether vanity may lead to anything positive we analyse a model in which the decision to implement a project or to maintain the status quo has been delegated to an agent. The agent can exert effort with a view to becoming better informed about the uncertain, state dependent, consequences of the project. The costs of effort are incurred by the agent himself. The agent may be either smart or dumb: any effort exerted by a dumb agent is to no avail, while a smart agent may obtain valuable information. The agent cares both about the expected project benefits and about his reputation with the general public. We analyse both the case that the agent knows his type and the case that he does not know his type. In the base case, the state of nature remains unknown when the project is not implemented. This means that a smart agent can distinguish himself from a dumb agent only when he has implemented a project. We also assume that without further information acquisition, and absent pride, the agent deems the project to be 'bad', i.e., that implementation would on average lead to a loss.

We want to argue that this model can best be applied to the political realm for three reasons. First, just as in the managerial situations studied by e.g. Holmström, and Scharfstein and Stein, care for reputation may be a short cut for care for future income. A good reputation in the executive may strengthen the probability of re-election. Indeed, Rogoff (1990) and Sibert and Rogoff (1988) have already argued that incumbents' incentives to appear competent induce them to create electoral

[3] Other papers in which reputation is a negative force include Avery and Chevalier (1999), Biglaiser and Mezzetti (1997), Effinger and Polborn (2001), Zwiebel (1995).

cycles. Several studies have found evidence for the existence of electoral cyles in macroeconomic variables (for a recent survey see Drazen, 2000). Second, even if re-election is impossible incumbents care about the image the general public holds when leaving office. Such is at least the impression one gets when analysing the behaviour of re-elected presidents who have reached their final term. In several models of electoral competition, agents are assumed to maximize the value of exogenous ego rents (see e.g. Persson and Tabellini, 2000, chapter 3). These rents are not related to policies or budgets (as they are exogenous). This suggests that it is not only reputation, but also pride that is at stake. Third, our society also seems to stimulate and exploit pride of politicians by naming streets and squares after them, by inviting them to TV shows etc.

What emerges from our model is that the mechanisms that determine whether pride is beneficial or not depend crucially on whether the agent knows his type or not. In particular, if the agent does not know his type the amount of effort he exerts increases the more important becomes reputation. This is not because effort influences the posterior belief the public will eventually hold. Instead, it is because by exerting more effort he can improve the odds with which the public will hold a favourable belief. As a result, the average quality of implemented projects increases. Pride, then, may be beneficial.

If the agent knows his type, socially optimal behaviour would have been for the agent to base his decision on the information he has. A dumb agent, who does not exert any effort as it is costly and to no avail, should vote against implementation as projects are expected to be bad. A smart agent should exert some optimal level of effort and follow the signal he obtains. This will indeed be the behaviour if an agent cares little about reputation. The fact that a smart agent follows his signal and thus sometimes rejects projects, sometimes accepts them, points to what happens when agents have a concern with reputation. First, a dumb agent now wants to mimic a smart agent by both accepting and rejecting projects. As a dumb agent does not exert any effort this randomness is inefficient: some bad projects are implemented. The more pride, the more often the dumb agent will diverge from behaviour that is socially desirable, and the larger the inefficiency. The second effect is a reaction of the smart agent to the dumb agent's mimetic behaviour. With strong reputational concerns, the desire to mimic the smart agent induces a dumb agent to implement some bad projects. To distinguish himself from a dumb agent, a smart agent increases his efforts by obtaining more valuable information to convince the public of his qualities.

To gain further insight into the mechanisms behind our findings, we also analyse two variations on our base model. First, if the expected benefits of the project are positive (i.e., projects are 'good' instead of 'bad'), a dumb agent has a stronger incentive to behave in an unconventional, distortionary way. The reason is that the marginal costs of rejecting a good project are smaller than the marginal costs of accepting a bad project. The incidence of the dumb agent's mimetic behaviour is determined by equating the marginal revenue of doing so to its marginal costs. The revenues stem from the increase in reputation thanks to the emulation of the smart agent. The costs always include the opportunity costs arising from the implementation of a bad project or from the rejection of a 'good' project. Moreover, when a project is implemented the state of nature is revealed. Hence, when a bad project is implemented, there is the risk of negatively affecting reputation. This is costly. This risk is absent when a good project is rejected – maintaining the status quo acts as a save haven where the dumb can hide. More distortionary decisions result.

These observations naturally lead one to ask what role is played by the assumption that the state of nature is not revealed when the status quo is maintained. This is the second variation on the base model. A smart agent cannot distinguish himself from a dumb agent when projects are rejected. A dumb agent's possibly wrong decision is not revealed. Would it then be beneficial to invest in mechanisms revealing the state of nature in case the project is not implemented? In other words, is there a role for a Government audit office to evaluate the adequacy of the decision to reject projects and maintain the status quo? This question is analysed in section 6. In short, a dumb agent, losing his save haven, further distorts project choice. A smart agent will exert more effort as he can now also distinguish himself from a dumb in case of project rejection.

This paper is organized as follows. In the next section we present the model. In section 3 we analyse the case where the agent does not know his type, and in section 4 the alternative case is presented. We then turn to the first variation in section 5. Here, we pay special attention to the relationship between pride and unconventional decisions. Section 6 analyses what happens when the correctness of the decision to reject a project can be observed. Section 7 concludes by discussing how are findings are related to the existing literature. Proofs can be found in the Appendix.

2 The Model

On behalf of the public or the electorate an agent has to decide whether to implement a project ($X = 1$), or to maintain the status quo ($X = 0$). The payoff of an implemented project depends on the state of the world, μ. This state is good or bad $\mu \in \{G, B\}$ with equal prior probability. Before the agent decides on the project, he may exert effort to obtain a signal, $s \in \{g, b\}$ about the state of the world. A signal could be interpreted as the agent's verdict (view, opinion) on the project resulting from his investigations. A "smart" agent, by exerting more effort e, can improve the probability $\pi(e)$ with which his signal is informative. If a signal is informative, it accurately reveals the state of the world: $\Pr(G|g, \text{informative}) = \Pr(B|b, \text{informative}) = 1$, where $\Pr(G|g, \text{informative})$ denotes the probability that the state is G conditional on having observed an informative signal $s = g$. We assume that the marginal returns to effort are positive but declining, $\pi' > 0$ and $\pi'' < 0$. Exerting effort is costly, $c(e) \geq 0$, and the marginal costs of exerting effort are positive and increasing, $c' > 0$, $c'' > 0$.[4] With probability $1 - \pi(e)$ the signals remain uninformative even to a smart agent and does not add any new information to the initially held belief that both states are equally likely. That is, if signals are uninformative, $\Pr(G|g, \text{uninformative}) = \Pr(B|b, \text{uninformative}) = \frac{1}{2}$. To a "dumb" agent all signals remain uninformative, however much effort he exerts.[5]

The agent cares about the outcome of his decision for two reasons. First, the project may be intrinsically beneficial to undertake. The payoffs when G equal $p + h > 0$, whereas they equal $p - h < 0$ when B. Maintaining the status quo gives a normalised payoff of 0, independent of the state of the world. The ex ante expected benefits of project implementation are assumed smaller than zero, $p < 0$.[6]

Second, the agent cares about the beliefs the public holds about him, and may

[4] We assume that the functions $c(\cdot)$ and $\pi(\cdot)$ are such that any equilibrium effort level satisfies $\pi(e) < 1$.

[5] Assuming that a dumb agent's effort is less productive than that of a smart instead of not productive at all does not change the mechanisms that determine the results in this paper.

[6] In Section 5, we discuss the results for the alternative case that $p > 0$.

want to use the decision he makes to influence these beliefs. In particular, he cares whether he is thought to be a smart agent worth re-electing or a dumb one without any decision-making merits. The public attaches a prior probability α to the possibility that the agent is smart. Throughout the paper we assume that this prior probability is common knowledge. We define the reputation of the agent as the posterior belief that the agent is smart, $\widehat{\alpha}$. The public cannot observe the amount of effort e the agent exerts, and so the posterior belief can only be based on the decision to implement or to maintain the status quo, and, in the former case, whether the state of the world is good or bad. Hence, the agent's preferences are described by:

$$\begin{aligned} U(X=1 \mid G) &= p+h+\lambda\widehat{\alpha}(X=1,G)-c(e) \\ U(X=1 \mid B) &= p-h+\lambda\widehat{\alpha}(X=1,B)-c(e) \\ U(X=0) &= \lambda\widehat{\alpha}(X=0)-c(e) \end{aligned} \quad (1)$$

where λ is the weight the agent attributes to his reputation relative to the outcomes of the project and the costs of effort.

The timing in this model is as follows. First, Nature determines the state of the world μ and the type of the agent, dumb or smart. Next, the agent chooses effort e, and receives a signal s. He then decides whether to implement or to reject the project. If the project is implemented, μ becomes common knowledge, whereas if the status quo is maintained everyone remains ignorant about μ. The public revises its prior about the probability that the agent is smart, and payoffs are realized. All of the above is common knowledge.

In the next section we assume that it is common knowledge that the agent does not know whether he is smart or dumb. In section 4, it is common knowledge that the agent knows his type but the public does not.

3 The Decision Maker Does Not Know His Type

In this section we determine the optimal strategy for an agent who does not know his type. The equilibrium in this case consists of two parts. For the agent, it characterises a strategy, i.e., an effort level and a decision rule that determines whether to implement the project or not, such that his expected utility level is maximized given the posterior beliefs held by the public. For the public, it means that the posterior beliefs it holds are obtained using Bayes' rule, the equilibrium strategy of the agent, and the prior α, where possible.[7]

Suppose the agent decides to follow his signal, i.e., to implement the project if he receives a good signal, $s = g$, and to maintain the status quo in case of a bad signal, $s = b$. If the agent does not know his type the probability that $\mu = G$ when $s = g$ equals

$$\Pr(G|g) = \alpha \left(\pi(e) + (1 - \pi(e)) \frac{1}{2} \right) + (1 - \alpha) \frac{1}{2} = \frac{1 + \alpha\pi(e)}{2} \tag{2}$$

and similarly

$$\Pr(G|b) = \frac{1 - \alpha\pi(e)}{2} \tag{3}$$

Furthermore, let e^* be the equilibrium belief about the agent's choice of e.

With this strategy of the agent, the reputation of the agent held by the public once one of the three events $\{X = 0\}$, $\{X = 1, \mu = G\}$, or $\{X = 1, \mu = B\}$ is observed can be calculated. If $\{X = 0\}$, the state of the world is not known, and since μ remains unknown, no new information about the agent's type becomes available. Consequently, the posterior belief equals the prior:

$$\widehat{\alpha}(X = 0) = \alpha \tag{4}$$

If the project is accepted, the state of nature becomes known, and some additional information about the type of the agent is revealed. In fact, as only a smart agent may receive an informative signal, and as the postulated strategy of the agent is to follow his signal, accepting projects that work out to be good enhances the agent's reputation. Similarly, accepting projects that work out to be bad deteriorates his reputation. In fact, using Bayes' rule one obtains

$$\widehat{\alpha}(X = 1, G) = \frac{\alpha\pi(e^*) + \alpha\frac{1}{2}(1 - \pi(e^*))}{\alpha\pi(e^*) + \alpha\frac{1}{2}(1 - \pi(e^*)) + \frac{1}{2}(1 - \alpha)} = \frac{\alpha(1 + \pi(e^*))}{1 + \alpha\pi(e^*)} \geq \alpha \tag{5}$$

[7] In case out-of-equilibrium beliefs are required, these beliefs should be 'reasonable'. We specify what is reasonable when we discuss the relevant out-of-equilibrium beliefs.

and

$$\widehat{\alpha}(X=1,B)=\frac{\alpha\frac{1}{2}(1-\pi(e^{*}))}{\alpha\frac{1}{2}(1-\pi(e^{*}))+\frac{1}{2}(1-\alpha)}=\frac{\alpha(1-\pi(e^{*}))}{1-\alpha\pi(e^{*})}\leq\alpha \tag{6}$$

Equations (4)-(6) show the possible posterior probabilities under the assumption that the agent follows his signal.

We now identify the conditions under which it is optimal for the agent to follow his signal given posterior beliefs specified by (4)-(6). Combining these beliefs with the conditional probabilities (2) and (3), one directly sees that in equilibrium (i.e., when $e=e^{*}$) an agent's decision to implement or not does not affect the reputation he expects to hold. Obviously, if the agent is as ignorant as the public as to his type, he cannot outwit the public. Formally, in equilibrium

$$E_{\mu}\left[\widehat{\alpha}(X=1,\mu)|g\right]=\Pr(G|g)\,\widehat{\alpha}(X=1,G)+\Pr(B|g)\,\widehat{\alpha}(X=1,B)=\alpha$$

Therefore, the expected benefits of implementing a project when a good signal is received are:

$$\begin{aligned}U(X=1|g) &= (p+h)\Pr(G|g)+(p-h)\Pr(B|g)+\lambda\alpha-c(e^{*})\\ &= p+\alpha\pi(e^{*})h+\lambda\alpha-c(e^{*})\end{aligned}$$

Clearly, as the payoff of rejecting equals $U(X=0|g)=\lambda\alpha-c(e^{*})$, an agent follows a positive signal if $p+\alpha\pi(e^{*})h>0$. Similarly, an agent maintains the status quo in case of a bad signal, $s=b$, if $U(X=0|b)=\lambda\alpha-c(e^{*})$ exceeds

$$\begin{aligned}U(X=1|b) &= (p+h+\lambda\widehat{\alpha}(X=1,G))\Pr(G|b)+\\ &\quad (p-h+\lambda\widehat{\alpha}(X=1,B))\Pr(B|b)-c(e^{*})\\ &= p-\alpha\pi(e^{*})h+\lambda\alpha\frac{1-\alpha^{2}\pi(e^{*})^{2}-2\pi(e^{*})^{2}(\alpha-\alpha^{2})}{1-\alpha^{2}\pi(e^{*})^{2}}-c(e^{*})\end{aligned}$$

Clearly, $U(X=0|b)>U(X=1|b)$ as $0>p-\alpha\pi(e^{*})h$ and $\lambda\alpha>\lambda\alpha\frac{1-\alpha^{2}\pi(e^{*})^{2}-2\pi(e^{*})^{2}(\alpha-\alpha^{2})}{1-\alpha^{2}\pi(e^{*})^{2}}$.

Let us assume that $p+\alpha\pi(e^{*})h>0$ in fact holds, and derive the optimal value of e. The agent chooses e to maximize (1), given the equilibrium posterior expectations

(4-6). Though the agent cannot influence these expectations, he can influence their probabilities. The agent thus chooses e to maximize:

$$\begin{aligned} U(e) &= \frac{1}{2}(p+\alpha\pi(e)h)+\lambda\frac{1}{2}\alpha+ \\ &\lambda\left(\frac{1}{2}\frac{(1+\alpha\pi(e))}{2}\frac{\alpha(1+\pi(e^*))}{1+\alpha\pi(e^*)}+\frac{1}{2}\frac{(1-\alpha\pi(e))}{2}\frac{\alpha(1-\pi(e^*))}{1-\alpha\pi(e^*)}\right)-c(e) \end{aligned} \tag{7}$$

The optimal effort level e^* solves the first-order condition for a maximum

$$\frac{1}{2}\alpha\pi'(e^*)h+\frac{1}{2}\frac{\alpha^2(1-\alpha)\pi(e^*)}{1-\alpha^2(\pi(e^*))^2}\lambda\pi'(e^*)=c'(e^*) \tag{8}$$

For the agent to follow his signal, e^* should satisfy $p+\alpha\pi(e^*)h>0$. For the agent to exert any effort at all, $U(e^*)>\lambda\alpha$, or $\frac{1}{2}(p+\alpha\pi(e^*)h)-c(e^*)>0$. Equation (8) implicitly defines the optimal effort level e^* as a function of α, λ and h. Application of the implicit function theorem shows that effort increases with h and λ. The intuition behind the effect of h on e^* is clear. The larger the possible upside, the higher the value of information. To understand the effect of λ on e^*, first note that effort does not affect equilibrium beliefs about ability. Effort only affects the probability of events. By expending more effort, the agent increases the probability of the most favourable event (that is $\{X=1,\mu=G\}$) and decreases the probability of the least favourable event, $\{X=1,\mu=B\}$.

If e^* is such that $p+\alpha\pi(e^*)h<0$ holds, then the agent decides to ignore any good signal he may receive and maintains the status quo. Of course, he will then decide not to exert any effort at all. The agent has no reason to deviate from his strategy to always maintain the status quo if the out-of-equilibrium beliefs held by the public when either $\{X=1,\mu=G\}$ or $\{X=1,\mu=B\}$ occurs equal α. These beliefs seem reasonable as no effort is undertaken by the agent, and the agent does not know his type. Proposition 1 summarizes the main result of this section.

Proposition 1 *Suppose it is common knowledge that neither the agent nor the public knows whether the agent is dumb or smart. Let e^* be the solution to $\frac{1}{2}\alpha\pi'(e^*)h+\frac{1}{2}\frac{\alpha^2(1-\alpha)\pi(e^*)}{1-\alpha^2(\pi(e^*))^2}\lambda\pi'(e^*)=c'(e^*)$ and let it satisfy $\frac{1}{2}(p+\alpha\pi(e^*)h)-c(e^*)>0$. Then, the equilibrium strategy of the agent is to exert effort e^* and to follow his signal. The equilibrium posterior probabilities that the agent is smart are given by: $\widehat{\alpha}(X=0)=\alpha$, $\widehat{\alpha}(X=1,G)=\frac{\alpha(1+\pi(e^*))}{1+\alpha\pi(e^*)}\geq\alpha$, and $\widehat{\alpha}(X=1,B)=\frac{\alpha(1-\pi(e^*))}{1-\alpha\pi(e^*)}\leq\alpha$.*

(If instead e^* is such that $\frac{1}{2}(p+\alpha\pi(e^*)h) - c(e^*) \leq 0$, the equilibrium strategy of the agent is to exert no effort and to maintain the status quo. The equilibrium posterior belief that the agent is smart is given by: $\widehat{\alpha}(X=0) = \alpha$, while reasonable out-of-equilibrium beliefs that sustain this equilibrium are $\widehat{\alpha}(X=1,G) = \alpha$, and $\widehat{\alpha}(X=1,B) = \alpha$.)

4 The Agent Knows His Type

If it is common knowledge that the agent knows whether he is dumb or smart but the public does not, reputational concerns influence behaviour in two ways.[8] First, the weight the agent puts on reputation determines whether or not a dumb agent wants to mimic a smart agent by randomly implementing a project that is expected to lead to a loss. Second, if agents are quite sensitive to the way they are regarded by the public, a smart agent has an incentive to distinguish himself from a dumb agent. Reputational concerns induce a smart agent to exert more effort to reduce the risk of being mistaken for a dumb agent.

Proposition 2 *Suppose it is common knowledge that the agent knows his type but the public does not. Let $\lambda < \lambda^- := -\frac{(2-\alpha)p}{2(1-\alpha)}$ and let $e^*_{s,l}$ solve $\frac{1}{2}\pi'\left(e^*_{s,l}\right)h = c'\left(e^*_{s,l}\right)$ and satisfy $\frac{1}{2}\left(p+\pi\left(e^*_{s,l}\right)h\right) + \frac{1}{2}\lambda\left(1+\frac{\alpha}{2-\alpha}\right) - c\left(e^*_{s,l}\right) > \lambda\left(\frac{\alpha}{2-\alpha}\right)$. Then there exists an equilibrium in which (i) a dumb agent exerts no effort and maintains the status quo (ii) a smart agent selects $e^*_{s,l}$, and follows his signal, and (iii) posterior beliefs equal $\widehat{\alpha}(X=1,G) = \widehat{\alpha}(X=1,B) = 1$ and $\widehat{\alpha}(X=0) = \frac{\alpha}{2-\alpha}$.*

A comparison of propositions 1 and 2 shows that if it is common knowledge that the agent knows whether he is dumb or smart and the public does not, and if the agent is not very interested in his reputation, the decision-process is improved in two ways. First, the effort of a smart agent is increased because he has more certainty that his effort is productive and that he receives an informative signal. The effort of a dumb agent is decreased, but that does not matter because his effort is not productive. Secondly, in case the agent is not informed about his type, a dumb agent who follows his signal takes an inefficient decision half of the time. In the

[8] As before, it is possible that always maintaining the status quo is the optimal strategy for both a dumb and a smart agent. From now on we will ignore this equilibrium.

situation of proposition 2 a dumb agent always chooses in line with the expected benefits of the project. So, if reputational concerns are not very important and the agent knows his type while the public does not, then a higher expected pay-off on the project is realized. A drawback of the case in which the agent knows his type is that for low values of λ reputational concerns do not induce the agent to expend more effort.

This changes for higher values of λ for the following reason. A dumb agent, when contemplating whether to mimic the smart agent's behaviour or not, trades off the increase in reputation and the expected reduction in project benefits. For $\lambda < \lambda^{-}$, the dumb agent refrains from mimetic behaviour as it would hurt expected payoffs of the project more than it would add in terms of reputation. For values exceeding λ^{-}, however, the boost in reputation compensates sufficiently the expected loss in project payoffs. Clearly, the more important reputation, the more frequent inefficient implementation of harmful projects by the dumb. As a result, for a smart agent to convince the public of his qualities he has to exert more effort to obtain more valuable information than he would have in case the dumb agent had simply rejected all projects. The net effect of pride on the value of implemented projects is ambiguous.

Proposition 3 *Suppose it is common knowledge that the agent knows his type but the public does not. Let* $\lambda > \lambda^{-}$. *Then there exists an equilibrium in which (i) a dumb agent does not exert effort and implements the project with probability* γ^* *where* $0 < \gamma^* < \frac{1}{2}$; *(ii) a smart agent selects* $e^*_{s,h}$, *solving* $\frac{1}{2}\pi'\left(e^*_{s,h}\right)h + \frac{1}{4}\lambda\pi'\left(e^*_{s,h}\right)\left(\widehat{\alpha}\left(X=1,G\right) - \widehat{\alpha}\left(X=1,B\right)\right) = c'\left(e^*_{s,h}\right)$ *and follows his signal , and (iii)* $\widehat{\alpha}(X=0) = \frac{\alpha}{2-\alpha-2\gamma^*(1-\alpha)}$; $\widehat{\alpha}(X=1,G) = \frac{\alpha\left(1+\pi(e^*_{s,h})\right)}{\alpha\left(1+\pi(e^*_{s,h})\right)+2\gamma^*(1-\alpha)}$ *and* $\widehat{\alpha}\left(X=1,B\right) = \frac{\alpha\left(1-\pi(e^*_{s,h})\right)}{\alpha\left(1-\pi(e^*_{s,h})\right)+2\gamma^*(1-\alpha)}$.

Proposition 3 shows that when the agent knows his type reputational concerns have two effects. First, they distort the decision made by a dumb agent: although he expects the benefits to be negative, some projects are accepted. As a consequence, they induce a smart agent to exert more effort. If γ^* were equal to 0, reputational concerns would not affect effort (as we also know from Proposition 2). Effort increases with γ^*. The reason is that the more frequently a dumb agent chooses $X = 1$, the less information about the agent's type is contained in $X = 1$. Consequently, for a smart agent it becomes more important that the good state of the world occurs as γ^* rises.

It also follows from proposition 3 that a dumb agent never implements a project more often than a smart agent ($\gamma^* < \frac{1}{2}$). In particular, even when reputational concerns become of overriding importance, $\lambda \to \infty$, a dumb agent still does not want to implement a project with probability one half. The reason is that even when $\lambda \to \infty$ the fear of being found out to be dumb keeps the dumb agent from making the state revelating choice of implementation: the probability that an erroneously implemented project is due to a dumb agent is larger than the chances of such a mistake committed by a smart one. If project benefits matter, i.e. for finite λ, the dumb's expectation that projects are bad is a further reason not to choose implementation to often.

5 Reputation and unconventional decisions

So far we have assumed that the ex ante expected benefits are negative, $p < 0$. In this section, we examine the implications of the alternative assumption that $p > 0$. We restrict our attention to the case that the agent knows his ability, as it is easy to verify that the assumption about the sign of p is immaterial in case the agent does not know his ability[9].

As in case of $p < 0$, a dumb agent chooses efficiently only if he cares little about his reputation. When pride takes over, the dumb will mimic the smart by randomly rejecting projects.

Proposition 4 *Suppose it is common knowledge that the agent knows his ability. (1) Let* $\lambda < \lambda^+ := p\frac{4(1-\alpha)+\alpha^2\left(1-\pi\left(e^{**}_{s,l}\right)^2\right)}{4-6\alpha+2\alpha^2}p$, *and let* $e^{**}_{s,l}$ *solve*

$$\frac{1}{2}\pi'\left(e^{**}_{s,l}\right)h + \frac{1}{4}\lambda\pi'\left(e^{**}_{s,l}\right)\left[\widehat{\alpha}(X=1,G) - \widehat{\alpha}(X=1,B)\right] = c'\left(e^{**}_{s,l}\right)$$

and satisfy $0 < p < \pi\left(e^{**}_{s,l}\right)h$. *Then (i) a dumb agent chooses* $X = 1$*; (ii) a smart agent exerts effort level* $e^{**}_{s,l}$ *and follows his signal; (iii) posterior beliefs equal*

[9] The only implication of $p > 0$ rather than $p < 0$ when the agent does not know his ability is that for a wider range of parameters the agent will follow his signal.

$\widehat{\alpha}(X=0)=1$, $\widehat{\alpha}(X=1,G)=\frac{\alpha\left(1+\pi\left(e^{**}_{s,l}\right)\right)}{2+\alpha\pi\left(e^{**}_{s,l}\right)-\alpha}$, *and* $\widehat{\alpha}(X=1,B)=\frac{\alpha\left(1-\pi\left(e^{**}_{s,l}\right)\right)}{2-\alpha\pi\left(e^{**}_{s,l}\right)-\alpha}$. *(2) If instead* $\lambda>\lambda^{+}$, *and* $e^{**}_{s,h}$ *solves*

$$\frac{1}{2}\pi'\left(e^{**}_{s,h}\right)h+\frac{1}{4}\lambda\pi'\left(e^{**}_{s,h}\right)\left(\widehat{\alpha}(X=1,G)-\widehat{\alpha}(X=1,B)\right)=c'\left(e^{**}_{s,h}\right)$$

and let it satisfy $0<p<\pi\left(e^{**}_{s,h}\right)h$. *Then (i) a dumb agent chooses* $X=1$ *with probability* $\gamma^{**}\in[\underline{\gamma},1)$, *with* $0<\underline{\gamma}<\frac{1}{2}$; *(ii) a smart agent chooses* $e^{**}_{s,h}$ *and follows his signal; (iii) posterior beliefs equal* $\widehat{\alpha}(X=0)=\frac{\alpha}{2-\alpha-2\gamma^{**}(1-\alpha)}$, $\widehat{\alpha}(X=1,G)=\frac{\alpha\left(1+\pi(e^{**}_{s,h})\right)}{\alpha\left(1+\pi(e^{**}_{s,h})\right)+2\gamma^{**}(1-\alpha)}$ *and* $\widehat{\alpha}(X=1,B)=\frac{\alpha\left(1-\pi(e^{**}_{s,h})\right)}{\alpha\left(1-\pi(e^{**}_{s,h})\right)+2\gamma^{**}(1-\alpha)}$.

A comparison of propositions 4 and 3 shows that when $p>0$ a dumb agent is more inclined to distort policy than when $p<0$: $\lambda^{+}<\lambda^{-}$ for the same absolute value of p. The reason for this result is simple. In our model, maintaining the status quo, $X=0$, is a save haven in the sense that no information about the correctness of the decision becomes available. As a consequence, by choosing $X=0$ a dumb policy maker does not run the risk of making a mistake everybody observes. As we have shown in the previous section, when $p<0$ one cost of distorting policy is that mistakes are observed. This weakens the incentive of a dumb agent to choose $X=1$ if $p<0$.

Also note that the probability with which the inefficient decision $X=0$ is taken may be larger than a half when a dumb agent expects projects to be good. Recall that the frequency of inefficient decisions when $p<0$ is strictly *smaller* than a half. The difference is due to the fact that choosing $X=0$ implies choosing the save haven in which the state of nature is not revealed. Hence, when $\lambda\to\infty$, and expected project benefits become immaterial, the dumb agent rejects the project more often than he implements it as the former does not expose him to possible public scorn.

Apart from influencing project choice, the assumption about the sign of p affects the effort a smart agent puts in investigation. To understand why, first recall that reputational concerns induce a smart agent to exert more effort in order to distinguish himself from a dumb agent. However, the public only observes whether the agent has made the correct decision if $X=1$. A dumb agent implements more projects when $p>0$ than when $p<0$. Consequently, for the same absolute value of p, a smart agent has stronger incentives to distinguish himself when $p>0$ than when $p<0$.

It is worth noting that when reputational concerns distort project choice, it is always in the *unconventional* direction ($X = 0$, if $p > 0$; $X = 1$, if $p < 0$). The reason is that we expect that a smart agent sometimes makes unconventional decisions. Ability and the probability of making unconventional decisions are positively related. This relationship induces dumb agents who care much about their reputation sometimes to adopt unconventional policies. This result of our model fits well with the observation made in the introduction that at the beginning of their careers politicians often launch unconventional proposals. It is at the beginning of their careers that the quality of a politician may not be known to the electorate. How this relationship between pride and the incidence of unconventional decisions compares to what is commonly found in the literature will be discussed in section 7.

6 Decision evaluation

In the model of Section 2 we have assumed that the consequences of the project are only revealed when the policy maker chooses implementation. The idea behind this assumption is that the proof of the pudding is in the eating. Learning requires doing. Sometimes, however, the full consequences of policy decisions are not easy to identify even in case of implemented projects. This provides a rationale for public agencies whose aim is to evaluate policy programmes. For example, the scope of various governmental accounting offices, like the United States General Accounting Office and the Netherlands Court of Audit, has gradually broadened over time. Initially, the court carried out audits. Today, it also evaluates the performance of programmes. It tries to determine whether or not policies have had the intended effects. As far as we know, these courts do not evaluate yet the performance of projects that are not implemented. One could ask whether it would be wise to further extend the tasks of these audit offices to also include the evaluation of the correctness of not implementing specific projects. Would this improve decision making? One way of checking the adequacy of rejecting specific projects is by looking at the performance of similar projects implemented abroad. In terms of our model, the implication of this additional task would be that the value of μ will be revealed even if the agent chooses $X = 0$. In this section we assume that such is the case.

It is not difficult to show that, if an agent does not know his own ability, the assumption about the observability of μ does not qualitatively change the results. In comparison with the case that μ is only revealed when projects are implemented, the agent exerts more effort. The reason is that if μ is also revealed when the status quo is maintained, the public always observes whether or not the agent has made the right decision. Hence, the marginal benefits of exerting effort increase. In section 7 we come back to this result as it contrasts sharply with the findings of Milbourn, Shockley and Thakor (2001). Formally,

Proposition 5 *Suppose μ is revealed when $X = 0$. Assume it is common knowledge that the agent does not know his type. Let $\hat{e}$ solve $\frac{1}{2}\alpha\pi'(\hat{e})h + \frac{\alpha^2(1-\alpha)\pi(\hat{e})}{1-\alpha^2(\pi(\hat{e}))^2}\lambda\pi'(\hat{e}) - c'(\hat{e}) = 0$ and satisfy $\frac{1}{2}(p + \alpha\pi(\hat{e})h) - c(\hat{e}) > 0$. Then the equilibrium strategy is for the agent to exert $\hat{e}$, and to follow his signal. The equilibrium posterior beliefs that the agent is smart are given by $\widehat{\alpha}(X = 1, G) = \widehat{\alpha}(X = 0, B) = \frac{\alpha(1+\alpha\pi(\hat{e}))}{1+\alpha\pi(\hat{e})}$ and $\widehat{\alpha}(X = 1, B) = \widehat{\alpha}(X = 0, G) = \frac{\alpha(1-\alpha\pi(\hat{e}))}{1-\alpha\pi(\hat{e})}$. Moreover, $\hat{e} > e^*$.*

Now suppose that the agent knows his ability. Furthermore, suppose that $p < 0$. In proposition 6, we show that in case the state of the world μ is always revealed, a dumb agent has *stronger* incentives to implement a project than in case μ is only revealed after implementation. To understand why, first note that with the assumption that μ is always revealed, rejection is no longer a save haven. The consequence is that rejection delivers a lower payoff to a dumb agent, and therefore it becomes less attractive. The cost of full revelation of μ is thus that a dumb agent becomes more inclined to distort project choice. There is also a benefit of full revelation, however. A smart agent will put more effort into investigation. The reason is that if the correctness of the decision is revealed irrespective of the decision made, a smart agent can also distinguish himself from a dumb when projects are rejected. As in section 3, without a further specification of our model, we cannot say whether or not courts of audit should be charged with the responsibility of evaluating projects which have been rejected. Important is that the result that reputational concerns may increase effort and distort project choice does not hinge on the assumption that the correctness of the decision is only revealed if projects are implemented. Formally,

Proposition 6 *Suppose μ is revealed when $X = 0$. Assume it is common knowledge that the agent knows his type. (1) Let $\lambda < \lambda_\mu^- := -p\frac{4(1-\alpha)+\alpha^2\left(1-\pi\left(\hat{e}_{s,l}\right)^2\right)}{4-6\alpha+2\alpha^2}$, where $\hat{e}_{s,l}$*

solves $\frac{1}{2}\pi'\left(\hat{e}_{s,l}\right)h+\frac{1}{4}\lambda\pi'\left(\hat{e}_{s,l}\right)\left(\widehat{\alpha}\left(X=0,B\right)-\widehat{\alpha}\left(X=0,G\right)\right)=c'\left(\hat{e}_{s,l}\right)$. *Then (i) the dumb agent chooses* $X=0$*; (ii) the smart agent's level of effort equals* $\hat{e}_{s,l}$*, and he follows his signal ; and (iii) the posteriors are* $\widehat{\alpha}\left(X=1,G\right)=\widehat{\alpha}\left(X=1,B\right)=1$, $\widehat{\alpha}\left(X=0,B\right)=\frac{\alpha\left(1+\pi\left(\hat{e}_{s,l}\right)\right)}{2+\alpha\pi\left(\hat{e}_{s,l}\right)-\alpha}$ *and* $\widehat{\alpha}\left(X=0,G\right)=\frac{\alpha\left(1+\pi\left(\hat{e}_{s,l}\right)\right)}{2-\alpha\pi\left(\hat{e}_{s,l}\right)-\alpha}$. *(2) Let* $\lambda>\lambda_{\mu}^{-}$*, and let* $\hat{e}_{s,h}$ *solve*

$$\frac{1}{2}\pi'\left(\hat{e}_{s,h}\right)h+\frac{1}{4}\lambda\pi'\left(\hat{e}_{s,h}\right)\left(\widehat{\alpha}\left(X=1,G\right)-\widehat{\alpha}\left(X=1,B\right)\right)$$
$$+\frac{1}{4}\lambda\pi'\left(\hat{e}_{s,h}\right)\left(\widehat{\alpha}\left(X=0,B\right)-\widehat{\alpha}\left(X=0,G\right)\right)=c'\left(\hat{e}_{s,h}\right)$$

Then (i) the dumb agent chooses $X=1$ *with probability* $\widehat{\gamma}\in(0,\frac{1}{2}]$*; (ii) the smart agent chooses* $\hat{e}_{s,h}$*, and follows his signal; (iii) the posteriors are* $\widehat{\alpha}\left(X=1,G\right)=\frac{\alpha\left(1+\pi\left(\hat{e}_{s,h}\right)\right)}{\alpha\left(1+\pi\left(\hat{e}_{s,h}\right)\right)+2\widehat{\gamma}(1-\alpha)}$, $\widehat{\alpha}\left(X=1,B\right)=\frac{\alpha\left(1-\pi\left(\hat{e}_{s,h}\right)\right)}{\alpha\left(1-\pi\left(\hat{e}_{s,h}\right)\right)+2\widehat{\gamma}(1-\alpha)}1$, $\widehat{\alpha}\left(X=0,B\right)=\frac{\alpha\left(1+\pi\left(\hat{e}_{s,h}\right)\right)}{\alpha\left(1+\pi\left(\hat{e}_{s,h}\right)\right)+2(1-\widehat{\gamma})(1-}$ *and* $\widehat{\alpha}\left(X=0,G\right)=\frac{\alpha\left(1-\pi\left(\hat{e}_{s,h}\right)\right)}{\alpha\left(1-\pi\left(\hat{e}_{s,h}\right)\right)+2(1-\widehat{\gamma})(1-\alpha)}$.

Note that $\widehat{\gamma}=\frac{1}{2}$ is possible: maintaining the status quo no longer implies choosing a save haven, but implies equal public exposure. This takes away the reason for making an inefficient decision less often than or more often than half of the time when reputational concerns are all that matter.

7 Discussion

Agents who are concerned about their reputation may ignore valuable information, downplay it, or exaggerate it. These individually rational actions are socially harmful. The aim of this paper was to understand whether pride, particularly in the political realm, can do any good. Our model, in which an agent can exert effort that may provide him with information about the value of an investment project, shows that pride may indeed be beneficial when an agent does not know whether he is smart or dumb. Moreover, it shows the mechanisms determining whether the same is the case when an agent does know his type

Scharfstein and Stein (1990) analyse the use of private information in a sequential decision problem. They show that herd behaviour, or conventional decisions, may result from agents' desire to gain a strong reputation. They assume that agents do not know whether they are dumb or smart. We analyse a one person decision problem with an objective function identical to Scharfstein and Stein's, and assume that an agent can consciously decide how much private information to acquire. When an agent knows his type reputational concerns lead the dumb agent to take unconventional decisions. Our analysis suggests therefore that if information acquisition is endogenised in a sequential decision structure it may matter whether agents do or do not know their type.[10]

Our model is closely related to that of Milbourn, Shockley and Thakor (2001) (henceforth MST). Like MST, we are concerned with the effect of reputational concerns on the incentives to invest in information. Our paper differs from MST in two important respects. First, MST only consider the case that the agent does not know his type, while we also examine the case that the policy maker knows his type. We showed that relaxing the assumption that the agent does not know his type has important implications for the results and the mechanisms determining them. Indeed, just as in MST when the agent does not know his type reputational concerns do not influence project choice. But when the agent knows his type reputational concerns lead a dumb agent to distort project choice.

Second, they show that the degree of overinvestment depends on the intensity of reputational concerns only if the state of nature remains unobserved after project rejection. Were the state of nature always revealed, the influence of reputation on effort would vanish, and any overinvestment would disappear. As we showed in section 6, reputational concerns influence effort whether the correctness of the decision to maintain the status quo can be observed or not. In fact, if the state becomes known whatever the decision taken and if the agent does not know his type, the information collection effort is strengthened.

The reason for this difference can best be explained as follows. In MST a manager creates a pile of projects on his desk, the average quality of which exactly reflects the manager's ability. Any investment in information helps the agent in distinguishing good from bad projects. The larger the investment, the better the screening. This is beneficial from a reputational point of view if the public only observes the outcome of that part of the pile of projects that the manager implements. By exerting

[10] Effinger and Polborn (2001) show how reputational concerns can give rise to anti-herding. This depends critically on the agents' objective function. An agent wants to show he is the unique smart agent.

more effort, the agent can then increase the probability with which the best event $\{X = 1, G\}$ is observed and decrease the frequency of the worst possible event, $\{X = 1, B\}$. But if the quality of the project is revealed whether it is implemented or rejected, i.e., if the public eventually observes the quality of the whole pile of projects, the public can perfectly observe the quality of the manager. Screening no longer affects reputation. Hence, reputational concerns no longer influence the choice of effort.

In our model, the average quality of the pile of projects depends exclusively on the state of nature. The quality of the agent is the quality of the screening of these projects. To enhance comparability with the MST findings, consider the case where an agent does not know his type. When the consequences of his decision also become visible when a project is rejected, an extra opportunity to show one's quality is offered, increasing the marginal reputational benefits of exerting effort. Hence, more information is collected. Moreover, we also showed that reputation continues to influence effort when the agent knows his type.

Adam Smith (1776) argued that "[i]t is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest." Paraphrasing Adam Smith, we would argue that the electorate addresses itself at least partly, not to the humanity of the politicians but to their self-love. In this paper we have tried to convince the reader that the private vice of vanity may be a public virtue.

8 References

Avery, C.N., Chevalier, J.A., 1999. Herding over the career. *Economics Letters* 63 (3), 327-333.

Biglaiser, G., Mezzetti, C., 1997. Politicians' decision making with re-election concerns. *Journal of Public Economics* 66 (3), 425-447

Drazen, A., 2000. *Political economy in macroeconomics*, Princeton University Press, Princeton.

Effinger, M.R., Polborn, M.K., 2001. Herding and anti-herding: A model of reputational differentiation. *European Economic Review* 45 (3), 385-403.

Holmström, B., 1999. Managerial incentive problems: A dynamic perspective. *Re-*

view of Economic Studies 66 (1), 169-182.

Milbourn, T.T., Shockley, R.L., Thakor, A.V., 2001. Managerial career concerns and investments in information. *RAND Journal of Economics* 32 (2), 334-351.

Persson, T., Tabellini, G., 2000. *Political economics: explaining economic policy.* MIT Press, Cambrigde.

Prendergast, C., Stole, L., 1996. Impetuous youngsters and jaded old-timers: Acquiring a reputation for learning. *Journal of Political Economy* 104 (6), 1105-1134.

Rogoff, K, 1990. Equilibrium political budget cycles. *American Economic Review* 80 (1), 21-36.

Rogoff, K., Sibert, A., 1988. Elections and macroeconomic policy cycles. *Review of Economic Studies* 55 (1), 1-16.

Scharfstein, D.S., Stein, J.C., 1990. Herd behavior and investment. *American Economic Review* 80 (3), 465-479.

Smith, A. (1776), An Inquiry into the Nature and Causes of The Wealth of Nations. Everyman's Library edition (1964), Dent: London.

Zwiebel, J., 1995. Corporate conservatism and relative compensation. *Journal of Political Economy* 103 (1), 1-25.

9 Appendix

Proof of Proposition 2: We first show that given the strategies of either type of agent, posterior beliefs are obtained using Bayes' rule, the equilibrium strategies of the agent, and the prior belief α. Since a dumb agent always chooses $X = 0$, while a smart agent sometimes chooses $X = 1$, $X = 1$ reveals that the agent is smart, whether the state is good or bad, implying $\widehat{\alpha}(X = 1, G) = \widehat{\alpha}(X = 1, B) = 1$. A smart agent chooses $X = 0$ with probability $\frac{1}{2}$. Consequently, $\widehat{\alpha}(X = 0) = \frac{\frac{1}{2}\alpha}{\frac{1}{2}\alpha+(1-\alpha)} = \frac{\alpha}{2-\alpha}$.

Now consider the strategy of a dumb agent. First note that a dumb agent never exerts effort, because there are no benefits of effort to him. A dumb agent prefers $X = 0$ to $X = 1$ if $\lambda\widehat{\alpha}(X = 0) > p + \lambda\widehat{\alpha}(X = 1)$ or $\lambda\frac{\alpha}{2-\alpha} > p + \lambda$, implying $\lambda < -\frac{2-\alpha}{2(1-\alpha)}p$.

Finally consider the strategy of a smart agent. When the smart agent decides not to exert any effort, the best he can do is to maintain the status quo, which gives him

$\lambda\frac{\alpha}{2-\alpha}$. Let $e^*_{s.l}$ be the hypothetical optimal level of effort. When he does exert effort and receives a bad signal, maintaining the status quo leaves him with $\lambda\frac{\alpha}{2-\alpha}-c\left(e^*_{s,l}\right)$, whereas deviating to implementation reduces his payoff to $p-\pi\left(e^*_{s,l}\right)h+\lambda-c\left(e^*_{s,l}\right)$ (recall that $\lambda\frac{\alpha}{2-\alpha}>p+\lambda$ by assumption). In case of a good signal, the payoff of implementation equals $p+\pi\left(e^*_{s,l}\right)h+\lambda-c\left(e^*_{s,l}\right)$, whereas the payoff of rejection amounts to $\lambda\frac{\alpha}{2-\alpha}-c\left(e^*_{s,l}\right)$. A necessary condition for implementation to be optimal is $p+\pi\left(e^*_{s,l}\right)h>-\lambda\left(\frac{2(1-\alpha)}{2-\alpha}\right)$. If e is such that he follows his signal, the smart agent chooses e to maximize $\frac{1}{2}(p+\pi\left(e_{s,l}\right)h)+\frac{1}{2}\lambda\left(1+\frac{\alpha}{(2-\alpha)}\right)-c\left(e_{s,l}\right)$. The first-order condition for a maximum, $e^*_{s,l}$, equals $\frac{1}{2}\pi'\left(e^*_{s,l}\right)h=c'\left(e^*_{s,l}\right)$. For the smart agent to exert effort and to follow his signal instead of not exerting any effort at all and maintaining the status quo

$$\frac{1}{2}(p+\pi\left(e^*_{s,l}\right)h)+\frac{1}{2}\lambda\left(1+\frac{\alpha}{(2-\alpha)}\right)-c\left(e^*_{s,l}\right)>\lambda\frac{\alpha}{2-\alpha}$$

should hold. QED

Proof of Proposition 3: The posterior probabilities directly follow from Bayes' rule and the equilibrium strategies of the agent.

Consider the strategy followed by a dumb agent. A dumb agent exerts no effort, because effort is costly and there are no benefits of effort. When the dumb agent mixes, his payoffs of accepting the project (ex ante) should equal those of maintaining the status quo. Let $\gamma=\Pr\left(X=1\right)$ when an agent is dumb. In equilibrium, the following must hold

$$p+\frac{1}{2}\lambda\left(\widehat{\alpha}\left(X=1,G\right)+\widehat{\alpha}\left(X=1,B\right)\right)=\lambda\widehat{\alpha}\left(X=0\right) \tag{9}$$

or,

$$-\frac{p}{\lambda}=\frac{1}{2}\left(\widehat{\alpha}\left(X=1,G\right)+\widehat{\alpha}\left(X=1,B\right)\right)-\widehat{\alpha}\left(X=0\right) \tag{10}$$

where $\widehat{\alpha}(X=0)=\frac{\alpha}{2-\alpha-2\gamma^*(1-\alpha)}$, $\widehat{\alpha}(X=1,G)=\frac{\alpha\left(1+\pi\left(e^*_{s,h}\right)\right)}{\alpha\left(1+\pi\left(e^*_{s,h}\right)\right)+2\gamma^*(1-\alpha)}$ and $\widehat{\alpha}\left(X=1,B\right)=\frac{\alpha\left(1-\pi\left(e^*_{s,h}\right)\right)}{\alpha\left(1-\pi\left(e^*_{s,h}\right)\right)+2\gamma^*(1-\alpha)}$, and where γ^* denotes the equilibrium value of γ.

The question becomes whether for all values of (p,λ,α) satisfying $\lambda>-\frac{(2-\alpha)p}{2(1-\alpha)}$ or $0<-\frac{p}{\lambda}<\frac{2(1-\alpha)}{2-\alpha}\ (\leq 1)$, and for all values $\pi(e^*_{s,h})\in(0,1)$, there is a value of

$\gamma^* \in \left(0, \frac{1}{2}\right)$ such that equation (10) holds.

Note that the left hand side of (10) is strictly positive and smaller than $\frac{2(1-\alpha)}{2-\alpha}$, and that the expression on the right hand side is monotonically decreasing in γ^*. Moreover, for $\gamma^* = 0$ the expression reduces to $1 - \frac{\alpha}{2-\alpha} = \frac{2(1-\alpha)}{2-\alpha} > 0$ while for $\gamma^* = \frac{1}{2}$ it becomes $\alpha^2 \pi(e^*_{s,h})^2 \frac{1-\alpha}{\left(\alpha\pi(e^*_{s,h})+1\right)\left(\alpha\pi(e^*_{s,h})-1\right)} < 0$. Hence, for every value of $\pi\left(e^*_{s,h}\right)$ there is a value of $\gamma^* \in \left(0, \frac{1}{2}\right)$ such that condition (10) is satisfied.

Now consider the strategy of a smart agent. If he chooses not to exert any effort, his payoffs equal those of a dumb agent. If instead he exerts effort, the best he can do is to follow his signal. Suppose he has received a good signal. Then $\Pr(G|g) = \frac{1+\pi}{2}$. The condition that should hold for implementation to be prefered to maintaining the status quo is

$$\begin{aligned} & p + \pi\left(e^*_{s,h}\right) h + \lambda(\Pr(G \mid g)\,\widehat{\alpha}(X=1,G) \\ & + \Pr(B \mid g)\,\widehat{\alpha}(X=1,B)) \\ > \; & \lambda\widehat{\alpha}(X=0) \end{aligned}$$

That this condition in fact holds can be seen by using equation (9) in conjunction with $\Pr(G|g) > \frac{1}{2} > \Pr(B|g)$. A similar line of reasoning shows that maintaining the status quo when he has received a bad signal dominates implementation.

When a smart agent decides to exert effort he maximizes

$$\frac{1}{2}\left(p + \pi\left(e_{s,h}\right) h\right) + \lambda\frac{1}{2}\left(\frac{1+\pi\left(e_{s,h}\right)}{2}\widehat{\alpha}(X=1,G) + \frac{1-\pi\left(e_{s,h}\right)}{2}\widehat{\alpha}(X=1,B)\right)$$
$$+\frac{1}{2}\lambda\widehat{\alpha}(X=0) - c\left(e_{s,h}\right)$$

Therefore, $e^*_{s,h}$ is the solution to

$$\frac{1}{2}\pi'\left(e_{s,h}\right) h + \frac{1}{4}\lambda\pi'\left(e_{s,h}\right)\left(\widehat{\alpha}(X=1,G) - \widehat{\alpha}(X=1,B)\right) = c'\left(e_{s,h}\right) \qquad (11)$$

QED

Proof of Proposition 4: This proof follows closely the one of Proposition 3. For values of $\lambda < \lambda^+$, a dumb agent implements a project with probability one and a smart agent follows his signal. Hence, in equilibrium, $\widehat{\alpha}(X=0) = 1$, $\widehat{\alpha}(X=1,G) = \frac{\alpha\left(1+\pi\left(e^{**}_{s,l}\right)\right)}{2+\alpha\pi-\alpha}$, and $\widehat{\alpha}(X=1,B) = \frac{\alpha\left(1-\pi\left(e^{**}_{s,l}\right)\right)}{2-\alpha\pi-\alpha}$, where $e^{**}_{s,l}$ solves

$\frac{1}{2}\pi'\left(e_{s,l}^{**}\right)h+\frac{1}{4}\lambda\pi'\left(e_{s,l}^{**}\right)\left[\widehat{\alpha}(X=1,G)-\widehat{\alpha}(X=1,B)\right]=c'\left(e_{s,l}^{**}\right)$. The threshold value λ^{+} is obtained by equating the payoff for a dumb agent of implementation and rejection, or

$$p+\lambda\left[\frac{1}{2}\frac{\alpha\left(1+\pi\left(e_{s,l}^{**}\right)\right)}{2+\alpha\pi\left(e_{s,l}^{**}\right)-\alpha}+\frac{1}{2}\frac{\alpha\left(1-\pi\left(e_{s,l}^{**}\right)\right)}{2-\alpha\pi\left(e_{s,l}^{**}\right)-\alpha}\right]=\lambda$$

Solving for λ gives $\lambda^{+}=p\frac{4(1-\alpha)+\alpha^{2}\left(1-\pi\left(e_{s,l}^{**}\right)^{2}\right)}{4-6\alpha+2\alpha^{2}}$. For $\lambda>\lambda^{+}$, the dumb agent mimics the smart agent by implementing projects with a probability strictly smaller than one, $\gamma<1$. Let γ^{**} denote the equilibrium value. When the dumb agent mixes his payoffs of accepting the project (ex ante) should equal those of maintaining the status quo:

$$p+\frac{1}{2}\lambda\left(\widehat{\alpha}\left(X=1,G\right)+\widehat{\alpha}\left(X=1,B\right)\right)=\lambda\widehat{\alpha}\left(X=0\right) \tag{12}$$

or,

$$\frac{p}{\lambda}=\widehat{\alpha}\left(X=0\right)-\frac{1}{2}\left(\widehat{\alpha}\left(X=1,G\right)+\widehat{\alpha}\left(X=1,B\right)\right) \tag{13}$$

In equilibrium $\widehat{\alpha}(X=0)=\frac{\alpha}{2-\alpha-2\gamma^{**}(1-\alpha)}$, $\widehat{\alpha}(X=1,G)=\frac{\alpha\left(1+\pi(e_{s,h}^{**})\right)}{\alpha\left(1+\pi(e_{s,h}^{**})\right)+2\gamma^{**}(1-\alpha)}$ and $\widehat{\alpha}\left(X=1,B\right)=\frac{\alpha\left(1-\pi(e_{s,h}^{**})\right)}{\alpha\left(1-\pi(e_{s,h}^{**})\right)+2\gamma^{**}(1-\alpha)}$. Of course, the right hand side of equation (13) is monotonically increasing in γ^{**}. For $\gamma^{**}=1$, it equals $\frac{p}{\lambda^{+}}$, whereas for $\gamma^{**}=\frac{1}{2}$ it is strictly larger than zero. As for $\gamma^{**}=0$ the right hand side is negative, there exists a value of γ that depends on the parameters of the model, say $\underline{\gamma}\in\left(0,\frac{1}{2}\right)$ such that the right hand side equals zero for $\gamma^{**}=\underline{\gamma}$.

Comparing equation (9) which should hold for $p<0$ and (12) for $p>0$, one notes that for the same absolute value of p, and in particular for a fixed value of $\pi\left(e\right)$, if γ^{-} solves equation (9), then the solution γ^{+} that solves equation (12) satisfies $\gamma^{+}>\gamma^{-}$. This, together with the first order condition that determines the optimal level of effort of the smart agent, equation (11), shows that the marginal benefit of effort have increased relative to the case of $p<0$. Hence, $e_{s,h}^{**}>e_{s,h}^{*}$. QED

Proof of Proposition 5: The analysis is similar to that leading up to equation 7. Because of the observability of μ in case of $X=0$, the expression for the expected reputation when a project is rejected becomes

$$\Pr(B|b)\,\widehat{\alpha}(X=0,B)+\Pr(G|b)\,\widehat{\alpha}(X=0,G) \tag{14}$$

instead of α. Note that the expression in (14) is equal to the expected reputation when a project is implemented following a good signal. Hence, the maximand becomes

$$\begin{aligned} U(e) &= \frac{1}{2}(p+\alpha\pi(e)h)+ \\ &\quad 2\lambda\left(\frac{1}{2}\frac{(1+\alpha\pi(e))}{2}\frac{\alpha(1+\pi(\hat{e}))}{1+\alpha\pi(\hat{e})}+\frac{1}{2}\frac{(1-\alpha\pi(e))}{2}\frac{\alpha(1-\pi(\hat{e}))}{1-\alpha\pi(\hat{e})}\right)-c(e) \end{aligned} \tag{15}$$

and the first order condition amounts to

$$\frac{1}{2}\alpha\pi'(\hat{e})\,h+\frac{\alpha^2(1-\alpha)\pi(\hat{e})}{1-\alpha^2(\pi(\hat{e}))^2}\lambda\pi'(\hat{e})=c'(\hat{e}) \tag{16}$$

As the second term on the left hand side is twice as large for any value of e as compared to the same term in equation 8, $\hat{e}>e^*$. QED

Proof of Proposition 6:(1) The posterior beliefs follow from the equilibrium strategies. The threshold value λ_μ^- is calculated as in the previous propositions. The smart agent's objective function becomes $U(e_{s,l})=\frac{1}{2}(p+\pi(e_{s,l})h)+\frac{1}{2}\lambda+\frac{1}{2}\lambda\left(\frac{1-\pi(e_{s,l})}{2}\widehat{\alpha}(X=0,G)+\frac{1+\pi(e_{s,l})}{2}\widehat{\alpha}(X=0,B)\right)-c(e_{s,l})$, and so the first order condition becomes

$$\frac{1}{2}\pi'(\widehat{e}_{s,l})\,h+\frac{1}{4}\lambda\pi'(\widehat{e}_{s,l})\left(\widehat{\alpha}(X=0,B)-\widehat{\alpha}(X=0,G)\right)=c'(\widehat{e}_{s,l})$$

In (2), $\lambda>\lambda_\mu^-$. The posterior beliefs can be derived from the equilibrium strategies. The dumb chooses $X=1$ with probability γ. In equilibrium, he is indifferent between implementation and maintaining the status quo or

$$\frac{1}{2}\lambda\left[\widehat{\alpha}(X=0,B)+\widehat{\alpha}(X=0,G)\right]=p+\frac{1}{2}\lambda\left[\widehat{\alpha}(X=1,B)+\widehat{\alpha}(X=1,G)\right]$$

Compared to the case where $X=0$ is a save haven, the expected reputation of choosing $X=0$ goes down for fixed γ. Hence, there is a larger incentive to choose

$X = 1$. Rewriting this equilibrium condition, one gets

$$-\frac{p}{\lambda} = \frac{1}{2}\left[\widehat{\alpha}\left(X=1,B\right) + \widehat{\alpha}\left(X=1,G\right) - \left(X=0,B\right) - \widehat{\alpha}\left(X=0,G\right)\right]$$

The left hand side satisfies $0 < -\frac{p}{\lambda} < -\frac{p}{\lambda_{\mu}^{-}}$, whereas the right hand side monotonically declines from $-\frac{p}{\lambda_{\mu}^{-}}$ for $\gamma = 0$ to 0 for $\gamma = \frac{1}{2}$. Hence, for all values of $\pi\left(e\right)$ there is a value of γ, $\gamma_{\mu}^{-} \in (0, \frac{1}{2}]$, such that the equilibrium condition is satisfied. The smart agent's level of effort is obtained by maximizing

$$\begin{aligned} U\left(e_{s,h}\right) &= \frac{1}{2}\left(p + \pi\left(e_{s,h}\right)h\right) + \frac{1}{2}\lambda\left(\frac{1+\pi\left(e_{s,h}\right)}{2}\widehat{\alpha}\left(X=1,G\right) + \frac{1-\pi\left(e_{s,h}\right)}{2}\widehat{\alpha}\left(X=1,B\right)\right) \\ &\quad + \frac{1}{2}\lambda\left(\frac{1-\pi\left(e_{s,h}\right)}{2}\widehat{\alpha}\left(X=0,G\right) + \frac{1+\pi\left(e_{s,h}\right)}{2}\widehat{\alpha}\left(X=0,B\right)\right) - c\left(e_{s,h}\right), \end{aligned}$$

and so the first order condition becomes

$$\begin{aligned} &\frac{1}{2}\pi'\left(\widehat{e}_{s,h}\right)h + \frac{1}{4}\lambda\pi'\left(\widehat{e}_{s,h}\right)\left(\widehat{\alpha}\left(X=1,G\right) - \widehat{\alpha}\left(X=1,B\right)\right) + \\ &\quad \frac{1}{4}\lambda\pi'\left(\widehat{e}_{s,h}\right)\left(\widehat{\alpha}\left(X=0,B\right) - \widehat{\alpha}\left(X=0,G\right)\right) = c'\left(\widehat{e}_{s,h}\right) \end{aligned}$$

The extra marginal benefit in case of rejection leads to an increase in effort relative to the case where $X = 0$ was a save haven.