# Hidden Testing and Selective Disclosure of Evidence*

Claudia Herresthal†

November 12, 2018

For latest version click here.

**Abstract**

I consider a game with two players, a decision maker and an advisor, who are uncertain about the state of the world. The advisor can sequentially run informative tests and disclose (some or all) outcomes to the decision maker. The decision maker then faces a binary choice. Players agree on the optimal choice under certainty, but their preferences are misaligned under uncertainty in that players differ in how they trade off losses from wrong choices. I characterize equilibria of this game. In particular, I compare the case where testing is hidden and the advisor can choose which test outcomes to verifiably disclose to the case where testing is observable. I show that the decision maker is weakly better off when testing is hidden rather than observable if players' preferences are sufficiently misaligned. Otherwise, hidden testing can leave the decision maker strictly worse off. I identify conditions on preference parameters under which both players can be strictly better off when testing is hidden rather than observable.

Keywords: endogenous information acquisition, verifiable disclosure, transparency
JEL codes: D83, D82

*The latest version of this paper can be found at https://cherresthal.github.io/CHerresthal_HiddenTesting.pdf. I would like to thank Inga Deimen, Matt Elliott, Robert Evans, Pawel Gola, Ole Jann, Meg Meyer, Anna Sanktjohanser and conference and seminar audiences at Conference on Economic Design, Game Theory Conference at Stony Brook, EEA, University of Arizona, Toulouse, Bocconi, Birkbeck and Oxford for valuable comments and feedback.

†Queens' College, University of Cambridge, Silver Street, Cambridge, CB3 9ET, UK (email: cl.herresthal@gmail.com).

Pharmaceutical companies have recently come under scrutiny for selective reporting of clinical trial outcomes.[1] As a response to demands for greater transparency, several companies have pledged to register trials and report their outcomes in open online databases (Goldacre et al. (2017)). At first sight, it seems that such transparency would improve regulation because pharmaceutical companies can no longer hide trials with unfavorable outcomes. However, companies may also strategically respond by changing which trials they run and this could leave the regulator to base his approval decision on weaker evidence.

This paper asks when and why a decision maker (DM), e.g. a regulator, prefers hidden to observable information acquisition by an advisor, e.g. a company. The model is based on a statistical decision problem. While the advisor (he) has control over testing, the decision maker (she) has control over whether to accept or reject a given hypothesis, e.g. that a drug is safe. What makes players' strategic interaction interesting to study is that their preferences are misaligned when faced with uncertainty. In particular, players are assumed to trade off the loss from falsely accepting and falsely rejecting differently, e.g. a regulator is relatively more averse to the mistake of approving an unsafe drug than a company.

I first consider a setting with *hidden testing*, which corresponds to not having a trial registry. The advisor can, in private, sequentially run tests over a fixed number of periods.[2] Each test yields either a positive or a negative outcome. The advisor incurs an infinitesimal cost for each test, which implies that he stops testing if further tests yield no strict benefit.[3] In the final period, the advisor chooses which outcomes to disclose and then the DM takes an action. While each outcome is verifiable, the time at which an outcome was discovered is not verifiable.[4] This implies that, even in equilibrium, the DM cannot necessarily infer how many tests the advisor has run. In the context of drug licensing, even though the time at which evidence was acquired may be verifiable, the regulator may not be able to infer how much evidence has been acquired in total. This could be because he is uncertain about when the company started its investigation or how many trials it managed to run within a given time frame. I contrast this with a setting with *observable testing*, which corresponds to having a trial registry. In this case, the advisor has no private information, but can influence the DM's choice by strategically choosing when to stop testing.[5]

[1] For example, GlaxoSmithKline was accused to have withheld data that suggested that its antidepressant Paxil was linked to suicidal behavior in teenagers (Rabin (2017)), Merck was accused of downplaying evidence that its schizophrenia drug Zyprexa increased the risk of diabetes (Berenson (2006)), and the Cochrane Review concluded that Roche withheld trial data on its influenza drug Tamiflu to make the drug look more effective (Jefferson et al. (2014)).

[2] The finite horizon represents the advisor's resource constraints. A finite horizon can be interpreted as a limiting case of an infinite horizon with a convex cost of testing, as discussed in Section 5. The assumption ensures that players never fully learn and, hence, that their preferences are not necessarily aligned.

[3] My results are independent of whether such a cost exists or not, as discussed in Section 3.2.

[4] This means that the advisor can hide but not forge outcomes.

[5] For my results, the important aspect of observable testing is that all realizations are disclosed. It is

I start with a two-period model and compare payoffs that arise in the set of Pareto-undominated equilibria. Theorem 1 characterizes under which conditions the DM is *strictly better* off in any equilibrium under hidden testing than in the unique equilibrium under observable testing. In addition, it characterizes when there exists an equilibrium under hidden testing in which the DM is *strictly worse* off than under observable testing. Based on these findings, Corollary 1 concludes that if players' preferences are sufficiently misaligned then the DM is weakly better off in any equilibrium under hidden testing than in the unique equilibrium under observable testing. Proposition 1 then compares the DM's payoff to a first-best benchmark, i.e. to a situation in which the DM has control over both testing and decision-making. It shows that the set of preference parameters for which the DM achieves her first-best expected payoff in any equilibrium under hidden testing is larger than the set of preference parameters for which she does so under observable testing. Proposition 2 compares the advisor's payoff across settings. Surprisingly, it shows that there exist conditions under which the advisor is strictly worse off in any equilibrium under hidden testing than in the unique equilibrium under observable testing, despite the fact that the advisor enjoys additional discretion when testing is hidden. Finally, Corollary 2 shows that both players can be strictly better off under hidden than observable testing.

My analysis identifies two distinct effects that cause the DM to benefit from hidden testing, and I refer to them as the insurance effect and the skepticism effect. The *insurance effect* gives a reason for why hidden testing encourages the advisor to test. In case the advisor finds evidence that would lead the DM to act against his interest, he has the option to hide it. Without a trial registry, a company may privately run additional trials on a drug that has already won approval and then withdraw the drug if and only if it finds strong evidence of adverse side effects. By contrast, with a trial registry, the company may be deterred from conducting additional studies. This is because, if these studies show that evidence of adverse side effects is weak, the company would prefer to continue selling the drug, but the regulator may withdraw it from the market. By contrast, the *skepticism effect* provides a reason why hidden testing allows the DM to credibly raise her acceptance threshold. With a trial registry, reported trial results can be taken at face value, but the regulator then also needs to see fewer trials with favorable outcomes to approve the drug.

The insurance and skepticism effects are fundamentally different. Under conditions for which the insurance effect exists, hidden testing can be compared to offering the advisor limited liability. The advisor has the upside from testing but no downside, because if outcomes lead players to disagree then the advisor's preferred action is chosen. Interestingly, offering this limited liability benefits the DM. Although she forgoes the option to choose her

irrelevant whether or not the time at which these outcomes were discovered is disclosed.

preferred action when players disagree, she benefits overall because the advisor acquires information. By contrast, the DM benefits from the skepticism effect because she can credibly threaten to act against the advisor's interest following certain outcomes. Under observable testing, the advisor can exploit the fact that, given sufficiently positive evidence, the DM chooses acceptance, although she would ideally like to make her choice of action contingent on what further tests show. However, under hidden testing, the DM can make a credible threat to reject unless additional positive outcomes are reported. This is because it is possible that the advisor is hiding contradictory evidence and, therefore, the evidence presented is weaker than its face value suggests. Since the DM's preferred action is chosen more often if the evidence leads the players to disagree, the advisor is made strictly worse off by hidden testing.

When the preferences of the DM and the advisor are sufficiently misaligned, it is possible that the DM is worse off under hidden testing. It is still the case that the advisor runs more tests when testing is hidden rather than observable, but the DM does not benefit from the additional information obtained. The skepticism effect no longer applies, because under hidden testing, the advisor searches for favorable outcomes to report those instead of unfavorable outcomes. Although the DM anticipates that the advisor may be hiding some unfavorable outcomes, on average, she is willing to act in line with his interest. In addition, the insurance effect no longer applies, because the advisor is willing to run some tests even when testing is observable. The additional tests he runs when testing is hidden contain information that is only valuable to him and not the DM, yet his selective reporting causes the DM's choice to depend on the additional information obtained.

To generalize these results, I turn to a model with an arbitrary number of periods and give sufficient conditions under which the payoff comparisons derived in the two-period model continue to apply (Theorem 2 and Proposition 3). I show that it is still true that the DM weakly benefits from hidden testing if preferences are sufficiently misaligned (Corollary 3). This model is also used to study payoff comparisons when the DM can delegate decision-making authority to the advisor, or when one of the two players can commit to their strategy ex ante, as discussed below.

Studying a situation in which the decision-making authority is delegated to the advisor further illustrates the differences between the skepticism and the insurance effect (Proposition 4). I first show that under hidden testing the DM is never strictly better off when delegating than when retaining decision rights. In addition, under observable testing, the DM is strictly better off when delegating than when retaining decision rights if the insurance effect applies, while the opposite is true if the skepticism effect applies.

If the DM can commit to a decision rule, she can achieve her first-best expected payoff, whether testing is hidden or observable. Without commitment power, the best the DM can do to compensate for her lack of options is not to monitor the advisor's testing, but instead to grant him the discretion to test in private, if either the insurance or the skepticism effect applies (Proposition 5). In addition, I study the DM's payoff comparison between hidden and observable testing if the advisor has the power to commit to a testing and a disclosure strategy and find that the skepticism effect ceases to exist, whereas the insurance effect persists (Proposition 6).

My paper is the first to identify the insurance and skepticism effects. For the insurance effect to arise, it is necessary that the two players agree on the optimal action under certainty, an assumption that has not been made in the related literature, e.g. Henry (2009), Che and Kartik (2009), Felgenhauer and Loerke (2017) and Di Tillio et al. (2018). In addition, the insurance effect shows that limited liability can increase the advisor's incentives to test when testing is sequential, contrary to findings by Mackowiak and Wiederholt (2012) in a different setting. The skepticism effect arises because the advisor can select how many outcomes to report, which has not been explicitly modeled in work comparing hidden to observable information acquisition with a given testing technology. In addition, contrary to existing findings, my paper shows that the DM can be strictly worse off under hidden testing. I give more detail on how my work compares to the existing literature in Section 1.

Although I use the drug approval process as my leading example, the insights also apply to scientific research. Scientists care about informing the public, but they are also under pressure to publish and therefore may be less averse to accepting a false hypothesis than the public. Recently, the fact that the results of many scientific studies cannot be replicated has strengthened demands for pre-analysis plans. However, if scientists have to register their experiments then it is harder for a journal editor to be credibly skeptical of the significance of their findings and, hence, harder for him to demand additional robustness checks before accepting a paper (skepticism effect). In addition, if scientists have to register their experiment then they may not include certain aspects in their research agenda out of fear that the additional findings might cast doubt on their existing conclusions and, hence, lead the editor to reject their work, whereas the scientists would still want to see their conclusions published (lack of insurance effect).

The rest of the paper is structured as follows. Section 1 compares my work to the existing literature. Section 2 introduces the model. Section 3 contains the key comparison of the DM's expected payoff under hidden versus observable testing, first for the two-period case and then for the case with an arbitrary finite number of periods. Section 4 analyzes this comparison when the DM delegates decision rights to the advisor and when one of the players has commitment power. Finally, Section 5 analyzes the comparison when the advisor has to commit to the number of tests ex ante and when the horizon is infinite. Section 6 concludes. All proofs can be found in the appendix.

# 1 Related Literature

My work builds on the extensive literature on persuasion when the advisor strategically acquires information.[6] In contrast to the literature on Bayesian persuasion, I assume the testing technology is exogenously given, but how the advisor selects which outcomes to report when testing is hidden is related to the concavification approach by Kamenica and Gentzkow (2011). The part of the literature which takes the testing technology as given can be divided into two strands. The first assumes that information acquisition is observable and asks how the advisor can manipulate the DM through strategically stopping the flow of information.[7] The second assumes information acquisition is hidden and asks how the advisor can manipulate the DM by disclosing evidence selectively.[8]

My paper compares the DM's payoff under hidden and observable information acquisition. Some of the existing work on this question restricts the advisor's information acquisition to a choice of whether or not to acquire a single signal. Matthews and Postlewaite (1985) studies a seller who faces a choice of whether or not to acquire a single costless signal about product quality.[9] When disclosure is compulsory, the seller does not test and buyers take his claimed ignorance at face value. But when disclosure is voluntary, the seller tests and reveals the signal if the quality is good, which allows buyers to learn about quality. Dahm et al. (2009) study disclosure rules for medical trials, assuming a pharmaceutical company can decide whether or not to run a single trial whose outcome is either positive, negative or inconclusive. They find that compulsory registries combined with a voluntary results database can result in full transparency, but reduce the company's incentive to test.[10]

The most closely related papers share the assumption that the advisor can acquire multiple signals, which expands the advisor's scope for strategically influencing the DM's choice. However, none of these papers identify the insurance effects as a reason why the DM can benefit from hidden testing, because they assume that players do not always agree on the optimal action under certainty.[11] In addition, they do not identify the skepticism effect

[6] A large literature studies persuasion when the advisor is exogenously informed, starting with Milgrom (1981), Milgrom and Roberts (1986). For recent work on selective disclosure from multiple exogenously given signals, see e.g. Dziuda (2011), Hart et al. (2017).

[7] Brocas and Carrillo (2007) were the first to study this type of manipulation in a discrete time model, while Henry and Ottaviani (2018) and McClellan (2017) study a continuous time version.

[8] In most models, the advisor chooses once and for all how much information to acquire, while in Felgenhauer and Schulte (2014) information acquisition is sequential, just like in my framework.

[9] See also Farrell and Sobel (1983) and Shavell (1994).

[10] Di Tillio et al. (2017) study manipulation in both the design and reporting of medical trials.

[11] Either the advisor prefers acceptance irrespective of the state so that players can only ever agree on acceptance (Brocas and Carrillo (2007), Felgenhauer and Loerke (2017), Di Tillio et al. (2018)), or, in models with continuous action choice, the advisor's ideal action always differs from the DM's ideal action by a constant independent of the state (Henry (2009), Che and Kartik (2009)). A more detailed explanation for what is necessary to generate the insurance effect can be found in Section 3.3.

because they do not explicitly model the advisor as facing a fixed testing technology and choosing how many of the acquired outcomes to report.

Di Tillio et al. (2018) compare the DM's payoff under two scenarios, one in which she allows a biased advisor to collect a sample of size $n$ in private and report his preferred draw and one in which she restricts him to collect a single draw in public. They show that depending on the distribution, selective disclosure may leave the DM either better or worse off. My focus lies on identifying how the conflict of interest between players affects the DM's benefit from hidden testing when the advisor can disclose as many outcomes as desired.

Henry (2009) assumes that the advisor commits ex ante to a quantity of costly research, which maps into a state-dependent number of infinitesimal positive and negative signals. The advisor chooses a higher quantity of research if his choice is hidden and the DM perfectly infers his findings in equilibrium, due to unraveling in the vein of Milgrom (1981) and Grossman (1981). Unraveling does not occur in my paper since the DM faces a discrete choice and the advisor acquires test outcomes sequentially, as explained in Section 3.2.

In Felgenhauer and Loerke (2017)'s framework, an advisor tests sequentially and can decide how informative each test will be. Interestingly, they find that the advisor runs only a single test in any Pareto-undominated equilibrium, whether testing is observable or hidden. However, if testing is hidden the advisor runs a more informative test, because this makes it credible that he will not run further tests even if the outcome is unfavorable. Comparing our work illustrates that whether or not hidden testing is beneficial for the DM depends on how much flexibility the advisor has when designing tests.[12]

Che and Kartik (2009) study the ideal bias of the advisor, taking the informational environment as given.[13] The advisor exerts costly effort to increase the chances to observe a single signal which is normally distributed about the true state. They show that the DM prefers a biased to an unbiased advisor provided his bias is sufficiently large. A more biased advisor does not report his signal for a larger range of realizations, but has greater incentives to exert effort. In my setting, an advisor whose preferences are less aligned runs weakly more tests when testing is hidden. However, this is not necessarily beneficial for the DM as the additional test outcomes can be used to substitute for unfavorable outcome in earlier tests.

In addition, my work relates to the literature on delegation. In Li and Suen (2004) the advisor is exogenously informed and preferences are as in this paper. They show that the DM weakly benefits from delegating decision-making authority to a more reluctant advisor. I show that their result no longer applies when information acquisition is endogenous.

[12] In a related paper, Liebgober (2017) studies how a scientist's choice of an experiment is affected by whether the experimental "protocol" is observable or not. Without transparency, the scientist has incentives to invest in a more informative experiment which is less susceptible to manipulation.

[13] See also Gerardi and Yariv (2008) and Dur and Swank (2005).

Argenziano et al. (2016) and Deimen and Szalay (2018) show that endogenous information acquisition causes the DM to have additional benefits from cheap-talk communication compared to delegation. In my setting with verifiable disclosure, I show that the DM weakly prefers communication to delegation.

## 2 Model

I first outline the game under hidden testing, denoted by $\Gamma$, and then the game under observable testing, denoted by $\Gamma^o$, which serves as a benchmark.

***Setting:*** Time is discrete and there are finitely many periods, $n = 0, 1, ..., N$. There are two players, a decision maker (DM) and an advisor (A).

***Timing and Information:*** In period $n = 0$, Nature draws a state $\omega \in \{false, true\}$, which determines whether a given hypothesis is true or false, where $Pr\,(true) = q \in [0, 1]$. The realized state is unobserved by both players. In each period $n = 1, ..., N$, the advisor first privately chooses whether to test, $\tau = 1$, or not to test, $\tau = 0$. If he tests, Nature draws a test outcome $s_n \in \{+, -\}$, with state-dependent accuracy given by

$$Pr\,(-|false) \equiv p_F \tag{1}$$

$$Pr\,(+|true) \equiv p_T \tag{2}$$

where $\frac{1}{2} < p_j < 1$ for $j = F, T$ and the realizations of outcomes are independent conditional on the state. In addition, the advisor incurs some infinitesimal cost of testing, $c > 0$, such that he does not test if he is indifferent between testing or not.[14] If the advisor does not test, Nature does not draw a test outcome. The test outcome is privately observed by the advisor. At the end of period $N$, the advisor sends a message $m \in \mathcal{M}$ to the DM, where the message space $\mathcal{M}$ is defined below. Then the DM chooses an action $a \in \{accept, reject\}$. Finally, payoffs are realized. It is assumed that players do not have commitment power.

***Payoffs:*** A player $i = DM, A$ incurs the following state-dependent payoff, denoted by $u_i\,(a, \omega)$:[15]

| $u_i\,(a, \omega)$ | $false$ | $true$ |
|---|---|---|
| $reject$ | $0$ | $-1$ |
| $accept$ | $-\lambda_i$ | $0$ |

where $\lambda_i \geq 0$. Players agree on the optimal action if the state of the world were known. In particular, if $\omega = true$ then both players prefer acceptance since $-1 < 0$ and if $\omega = false$

[14] This corresponds to a situation in which testing is costly and I consider the limit as this cost becomes small. My results are independent of whether such a cost exists or not, as discussed in Section 3.2.

[15] This framework is based on DeGroot (1970)'s framework for optimal statistical decisions.

then both players prefer rejection since $-\lambda_i < 0$ for any $i = A, DM$. However, if there is uncertainty about the state of the world, players may disagree on what the optimal action is, because players may face a different trade-off between the loss of $\lambda_i$ from accepting if false and the loss of 1 from rejecting if true. The advisor is *more reluctant* to accept if he cares more about the loss from falsely accepting than the DM, i.e. if $\lambda_A \geq \lambda_{DM}$. Conversely, the advisor is *more enthusiastic* about accepting if he cares less about the loss from falsely rejecting than the DM, i.e. if $\lambda_A \leq \lambda_{DM}$. Each player $i$ maximizes his expected payoff $\widehat{u}_i(a) = E_\Omega(u_i(a,\omega))$.[16]

The choice of state-dependent payoffs $u_i(a,\omega)$ implies that the loss $\lambda_i$ from falsely accepting is a sufficient statistic for determining whether acceptance or rejection maximizes player $i$'s expected payoff at a given belief about the state. However, the analysis applies to any payoff function which satisfies i) $u_i(reject, false) > u_i(accept, false)$ and ii) $u_i(accept, true) > u_i(reject, true)$ for $i = A, DM$. Then the sufficient statistic is the ratio of the excess payoff from rejection if the hypothesis is false to the excess payoff from acceptance if the hypothesis is true.[17]

I assume the DM's loss $\lambda_{DM}$ from falsely accepting is sufficiently high such that her expected payoff at the prior belief $q$ is maximized when choosing *reject*, i.e. *reject* is her default choice:

$$\widehat{u}_{DM}(reject) > \widehat{u}_{DM}(accept) \Leftrightarrow \frac{q}{1-q} < \lambda_{DM}.\text{[18]} \tag{3}$$

***Histories:*** Define the *history of test outcomes* at the end of period $n = 0, 1, ..., N$, denoted by $h_n \in H_n$, to be an ordered list of past outcomes collected by the advisor, i.e. $h_0 = \emptyset$ and $h_n = (s_1, s_2, ..., s_n)$ for $n \geq 1$, where $s_n \in \{+, -, \emptyset\}$ and $s_n = \emptyset$ denotes the event that the advisor did not test. Define the *unordered history of test outcomes* at the end of period $N$, denoted by $\widetilde{h} \in \widetilde{H}$, to be the set of past outcomes realizations collected by the advisor, which does not contain information on the order in which these realizations

[16]Since the cost of testing is infinitesimal, it can be ignored in the payoff calculation. Generally, when there is a cost of testing, the advisor may decide not to obtain some information even if it is valuable to him. I assume that the cost of testing is infinitesimal to abstract from such concerns and focus solely on how the strategic information acquisition and transmission are affected by the conflict of interest between players.

[17]Consider a general payoff function for $i = A, DM$:

| $u_i(a,\omega)$ | $false$ | $true$ |
|---|---|---|
| $reject$ | $\alpha_i$ | $\beta_i$ |
| $accept$ | $\gamma_i$ | $\delta_i$ |

where $\alpha_i, \beta_i, \gamma_i, \delta_i \in \mathbb{R}$ and $\alpha_i > \beta_i$ and $\delta_i > \gamma_i$. Then my results apply if and only if $\lambda_i = \frac{\alpha_i - \gamma_i}{\delta_i - \beta_i}$.

[18]If I instead assumed $\frac{q}{1-q} \geq \lambda_{DM}$ such that *accept* is the DM's default choice, the results would be a mirror image of the results presented. Equilibrium play would be as described but the labels of states $\omega \in \{false, true\}$, test outcomes $s_n \in \{+, -\}$ and final decisions $a \in \{accept, reject\}$ would have to be switched. For results on payoff comparisons the likelihood ratio would become the inverse likelihood ratio.

were obtained. For example, if $N = 2$ then both $h_2 = (+, -)$ and $h_2 = (-, +)$ map into $\widetilde{h} = \{+, -\}$, and both $h_2 = (+, \emptyset)$ and $h_2 = (\emptyset, +)$ map into $\widetilde{h} = \{+\}$.

***Message:*** The message space is given by

$$\mathcal{M} \equiv \{\{s_1, ..., s_N\} | s_i \in \{+, -\}, i \in \{1, ..., N\}\}. \tag{4}$$

The advisor's set of feasible messages at unordered history $\widetilde{h}$ is given by $M\left(\widetilde{h}\right) = \mathcal{P}\left(\widetilde{h}\right)$, where $\mathcal{P}\left(\widetilde{h}\right)$ denotes the power set of $\widetilde{h}$, i.e. the set of all subsets of $\widetilde{h}$ including the empty set. This means that the advisor can report any subset of outcomes he has collected, e.g. if $N = 2$ and $\widetilde{h} = \{+, -\}$ then the set of feasible messages is given by $M(\{+, -\}) = \{\emptyset, \{+\}, \{-\}, \{+, -\}\}$. In particular, this implies that the realizations of outcomes are verifiable, i.e. the advisor cannot forge outcomes but he is able to hide outcomes. In addition, it implies that the period in which outcomes were acquired is not verifiable, i.e. if $N = 2$ and the advisor reports $m = \{+\}$ it is not verifiable whether $h_2 = (+, \cdot)$ or $h_2 = (\cdot, +)$.

***Strategies:*** The advisor has a testing and a disclosure strategy. A testing strategy for the advisor is: $\sigma_A^T : H_n \rightarrow \{0, 1\}$. It selects action $\tau \in \{0, 1\}$ conditional on history $h_n \in H_n$ for $n = 0, 1, ..., N - 1$. A reporting strategy for the advisor is $\sigma_A^M : H_N \rightarrow \mathcal{M}$. It selects a message $m \in \mathcal{M}$ conditional on history $h_N \in H_N$. A strategy for the DM is $\sigma_{DM} : \mathcal{M} \rightarrow \{accept, reject\}$. It selects action $a \in \{accept, reject\}$ conditional on message $m \in \mathcal{M}$. I assume that the DM accepts if she is indifferent between accepting and rejecting.[19]

***Equilibrium Concept:*** The solution concept is a sequential equilibrium in pure strategies.[20] A sequential equilibrium of $\Gamma$ consists of both a profile of strategies $\sigma \equiv \left(\sigma_A^T, \sigma_A^M, \sigma_{DM}\right)$ and a system of beliefs $\mu \equiv (\mu_A, \mu_{DM})$, where $\mu_A : \cup_{n=0}^N H_n \rightarrow [0, 1]$ and $\mu_{DM} : \mathcal{M} \rightarrow [0, 1]$. The advisor's beliefs select a probability that the state is $\omega = true$ for each history of outcomes $h_n \in H_n$ where $n = 0, 1, ..., N$ and the DM's beliefs select such a probability for each message $m \in \mathcal{M}$. Let $\Pi(\sigma, \mu; \Gamma) = (\pi_A(\sigma, \mu; \Gamma), \pi_{DM}(\sigma, \mu; \Gamma))$ denote the vector of expected payoffs given strategy profile $\sigma$ and system of beliefs $\mu$. By definition of a sequential equilibrium,

1. the advisor's strategy $\left(\sigma_A^T, \sigma_A^M\right)$ maximizes $\pi_A(\sigma, \mu; \Gamma)$ at any history $h_n$ for $n = 0, 1, ..., N$ given $\sigma_{DM}$ and $\mu_A$, and

[19] Given this assumption, there is a unique equilibrium under observable testing, whereas without it, there are multiple equilibria which only differ in equilibrium play at knife-edge cases. To keep payoff comparisons more tractable, it is helpful to have a unique equilibrium under observable testing.

[20] The concept of sequential equilibrium is adequate, because it imposes restrictions on off-paths beliefs and, therefore, allows me to show that the equilibria studied are not constructed choosing arbitrary off-path beliefs. The restriction to pure strategies is for tractability. The main insights of Theorem 1 (for $N = 2$) and Theorem 2 (for $N > 2$) would apply if mixed strategies are allowed, but the boundaries of regions would change.

2. the DM's strategy $\sigma_{DM}$ maximizes $\pi_{DM}(\sigma, \mu; \Gamma)$ at any message $m$ given $\mu_{DM}$, and

3. there exists a sequence of completely mixed strategies $\{\sigma^k\}_{k=1}^{\infty}$, with $\lim_{k\to\infty} \sigma^k = \sigma$, such that the system of beliefs $\mu = \lim_{k\to\infty} \mu^k$, where $\mu^k$ denotes the beliefs derived from strategy profile $\sigma^k$ using Bayes' rule.

***Observable Testing:*** In the benchmark case of observable testing, the advisor's actions and Nature's draws of test outcomes are observed by both players. This renders the advisor's message superfluous. The DM's strategy is given by $\sigma_{DM}^o : H_N \to A$. It selects action $a \in A$ conditional on history $h_N \in H_N$ (rather than conditional on a message). A sequential equilibrium of $\Gamma^o$ consists of a strategy profile $\sigma^o \equiv (\sigma_A^o, \sigma_{DM}^o)$ and a system of beliefs $\mu^o : \cup_{n=1}^N H_n \to [0, 1]$. The vector of expected payoffs given strategy profile $\sigma^o$ and system of beliefs $\mu^o$ is given by $\Pi(\sigma^o, \mu^o; \Gamma^o) = (\pi_A(\sigma^o, \mu^o; \Gamma^o), \pi_{DM}(\sigma^o, \mu^o; \Gamma^o))$.

***Remarks on Modeling Choices:*** I assume the content of evidence is verifiable because I want to focus on the effect of strategically omitting evidence. As the literature on verifiable disclosure has frequently argued, in many settings it is relatively harder to fabricate than to omit evidence and this model assumes an extreme case in which fabrication is prohibitively costly. In addition, I assume that the time at which evidence was collected cannot be verifiably disclosed and the advisor cannot report before the final period. These assumptions are made to capture a setting in which the DM can only make limited inferences from the calendar time of reports. For example, this could be because both the time at which the advisor first has the possibility to test the hypothesis as well as the total number of tests feasible in any interval of time is private information to the advisor and cannot be verifiably disclosed by him. The finite horizon represents a constraint on the advisor's resources to conduct tests. This setting with an infinitesimal cost of testing and a finite horizon can be understood as a tractable limit of a setting in which the cost of testing increases with each test and the horizon is infinite, as discussed in Section 5.[21]

# 3 Hidden vs. Observable Testing

This section's main focus is to compare the DM's expected payoff under hidden and observable testing. For the case $N = 2$, I fully characterize under which conditions the DM is strictly better or strictly worse off under hidden relative to observable testing. I identify the insurance and skepticism effects as reasons for why the DM's expected payoff improves with

[21] When the cost of testing increases with each test, then the advisor's choice to continue testing depends not only on his beliefs about the state, but also on how many tests he has already run, independent of whether the horizon is finite or not. This important feature is preserved in this model with a finite horizon and an infinitesimal cost of testing.

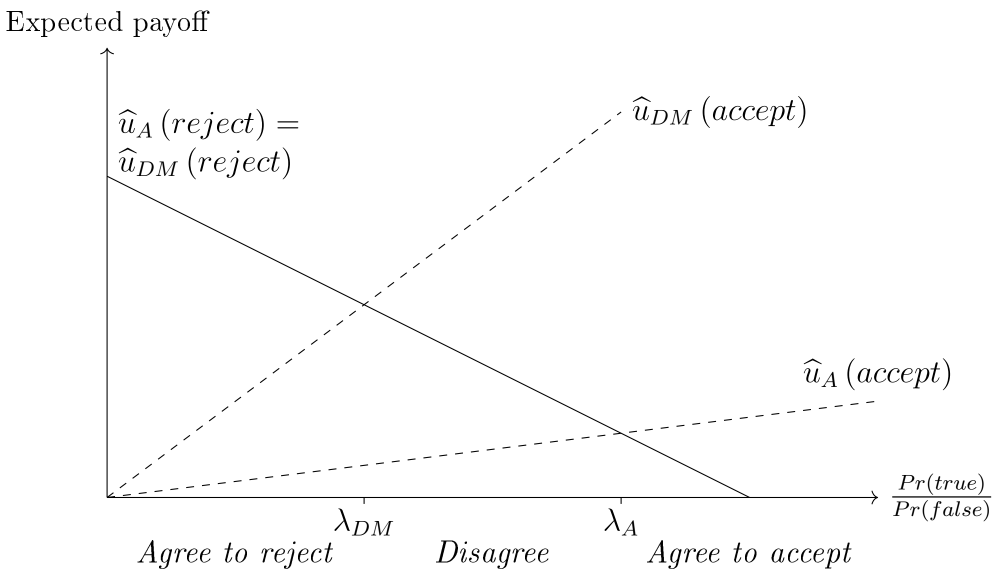


Figure 1: This figure illustrates how the conflict of interest between players depends on the likelihood ratio and their preference parameters $(\lambda_A, \lambda_{DM})$ for the case in which the advisor is more reluctant to accept than the DM.

hidden testing. Then I give sufficient conditions for these effects to exist when $N > 2$. For any $N$, I show that the DM is weakly better off under hidden relative to observable testing when preferences are sufficiently misaligned.

As a first step, it is helpful to note that a player $i$ prefers acceptance over rejection at a given probability that the hypothesis is true if and only if

$$\widehat{u}_i\left(accept\right) = -Pr\left(false\right)\lambda_i \geq -Pr\left(true\right) = \widehat{u}_i\left(reject\right) \Leftrightarrow \lambda_i \leq \frac{Pr\left(true\right)}{Pr\left(false\right)}. \tag{5}$$

This observation helps to understand the nature of conflict between players, as illustrated in Figure 1. Suppose that the advisor is more reluctant to accept than the DM, i.e. $\lambda_A > \lambda_{DM}$, and fix a likelihood that the hypothesis is true. Then if the DM prefers rejection, the advisor must also prefer rejection. In addition, if the advisor prefers acceptance, the DM must also prefer acceptance. The reverse situation arises if the advisor is more enthusiastic about accepting than the DM.

As a next step, I turn to analyze equilibrium payoffs under both hidden and observable testing.

**Lemma 1 (Equilibrium Existence)** *An equilibrium exists under both observable and hidden testing. Under observable testing, the equilibrium is unique. Under hidden testing, multiple equilibria may exist.*

When testing is observable, there is a unique equilibrium $(\sigma^o, \mu^o)$ and, hence, a unique equilibrium payoff vector $\Pi(\sigma^o, \mu^o; \Gamma^o)$. However, when testing is hidden, there are multiple equilibria. As the aim of this paper is a welfare comparison across testing regimes, it seems natural to focus on Pareto-undominated equilibria. Define $E$ to be the set of equilibria which are Pareto-undominated, i.e. $(\sigma, \mu) \in E$ if and only if there exists no equilibrium $(\sigma', \mu')$ of $\Gamma$ such that $\pi_i(\sigma, \mu; \Gamma) \geq \pi_i(\sigma', \mu'; \Gamma)$ and $\pi_j(\sigma, \mu; \Gamma) > \pi_j(\sigma', \mu'; \Gamma)$ for $i, j = A, DM$ and $j \neq i$. and any equilibrium $(\sigma', \mu')$ of $\Gamma$. There exists some Pareto-undominated equilibrium in which the DM's payoff is weakly higher than in any other Pareto-undominated equilibrium, i.e. there exists $(\overline{\sigma}, \overline{\mu}) \in E$ such that $\pi_{DM}(\overline{\sigma}, \overline{\mu}; \Gamma) \geq \pi_{DM}(\sigma, \mu; \Gamma)$ for any $(\sigma, \mu) \in E$. I will refer to $(\overline{\sigma}, \overline{\mu})$ as a *DM-preferred equilibrium.* In addition, there exists some Pareto-undominated equilibrium in which the advisor's payoff is weakly higher than in any other Pareto-undominated equilibrium, i.e. there exists $\left(\underline{\sigma}, \underline{\mu}\right) \in E$ such that $\pi_A\left(\underline{\sigma}, \underline{\mu}; \Gamma\right) \geq \pi_A(\sigma, \mu; \Gamma)$ for any $(\sigma, \mu) \in E$. I will refer to $\left(\underline{\sigma}, \underline{\mu}\right)$ as an *advisor-preferred equilibrium.*

**Lemma 2 (Pareto-undominated Equilibria)** *Suppose testing is hidden.*

1. *Any DM-preferred equilibrium results in a unique payoff vector* $\Pi(\overline{\sigma}, \overline{\mu}; \Gamma)$. *The advisor is strictly worse off in a DM-preferred equilibrium than in any other Pareto-undominated equilibrium, i.e.* $\pi_A(\overline{\sigma}, \overline{\mu}; \Gamma) < \pi_A(\sigma, \mu; \Gamma)$ *for any* $(\sigma, \mu) \in E$.

2. *Any advisor-preferred equilibrium results in a unique payoff vector* $\Pi\left(\underline{\sigma}, \underline{\mu}; \Gamma\right)$. *The DM is strictly worse off in an advisor-preferred equilibrium than in any other Pareto-undominated equilibrium:* $\pi_{DM}\left(\underline{\sigma}, \underline{\mu}; \Gamma\right) < \pi_{DM}(\sigma, \mu; \Gamma)$ *for any* $(\sigma, \mu) \in E$.

## 3.1 Two-period Model

Throughout this subsection, I assume that $N = 2$. When comparing players' expected payoffs between hidden and observable testing, it turns out that there are four preference parameter regions of interest. These region's boundaries correspond to posterior likelihood ratios conditional on some subset of outcome realizations from two tests. The reason boundaries are likelihood ratios has to do with the feature that a player's preferred final decision depends on the comparison between his loss from falsely accepting and the posterior likelihood ratio that the hypothesis is true, as shown by (5).

To define the boundaries, denote the complete set of ordered outcome realizations from two tests by $\Phi$, where

$$\Phi = \{(+,+), (+,-), (-,+), (-,-)\}. \tag{6}$$

Then define $l_\varphi$ to be the posterior likelihood ratio conditional on some subset $\varphi$ of $\Phi$. Given the prior $q$ and the test accuracy $(p_F, p_T)$, the relevant boundaries are given by:

$$l_{(+,+)} \equiv l_{\{(+,+)\}} = \frac{q}{1-q}\frac{p_T^2}{(1-p_F)^2}, \tag{7}$$

$$l_{(+,-)} \equiv l_{\{(+,-)\}} = l_{\{(-,+)\}} = l_{\{(+,-),(-,+)\}} = \frac{q}{1-q}\frac{p_T(1-p_T)}{p_F(1-p_F)}, \tag{8}$$

$$l_{(+,\cdot)} \equiv l_{\{(+,+),(+,-)\}} = \frac{q}{1-q}\frac{p_T}{1-p_F}, \tag{9}$$

$$l_{\neg(-,-)} \equiv l_{\Phi\setminus\{(-,-)\}} = \frac{q}{1-q}\frac{1-(1-p_T)^2}{1-p_F^2}, \tag{10}$$

$$l_\Phi = \frac{q}{1-q}. \tag{11}$$

The four regions of interest are defined as follows and depicted in Figure 2:

$$S \equiv \left\{(\lambda_A, \lambda_{DM}) : \lambda_A < l_{(+,-)}, l_{\neg(-,-)} < \lambda_{DM} \leq l_{(+,\cdot)}\right\}, \tag{12}$$

$$I \equiv \left\{(\lambda_A, \lambda_{DM}) : l_\Phi < \lambda_{DM} \leq l_{(+,-)}, l_{(+,\cdot)} \leq \lambda_A < l_{(+,+)}\right\}, \tag{13}$$

$$W_S \equiv \left\{(\lambda_A, \lambda_{DM}) : \lambda_A < l_{(+,-)}, \max\left\{l_\Phi, l_{(+,-)}\right\} < \lambda_{DM} \leq l_{\neg(-,-)}\right\}, \tag{14}$$

$$W_I \equiv \left\{(\lambda_A, \lambda_{DM}) : l_\Phi < \lambda_{DM} < l_{(+,-)} < \lambda_A < l_{(+,\cdot)}\right\}. \tag{15}$$

I refer to $S$ as the *skepticism region* and to $I$ as the *insurance region*.[22]

**Theorem 1 (DM Payoff Comparison)**

1. *Given $(\lambda_A, \lambda_{DM})$, the DM has a strictly higher payoff in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing if and only if $(\lambda_A, \lambda_{DM})$ lies in the skepticism or the insurance region, i.e. for any $(\lambda_A, \lambda_{DM})$ it holds that $\pi_{DM}(\sigma, \mu; \Gamma) > \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ for all $(\sigma, \mu) \in E$ if and only if $(\lambda_A, \lambda_{DM}) \in S \cup I$.*

2. *Given $(\lambda_A, \lambda_{DM})$, there exists a Pareto-undominated equilibrium under hidden testing which gives the DM a strictly lower payoff than the unique equilibrium under observable testing if and only if $(\lambda_A, \lambda_{DM})$ lies in region $W_S$ or $W_I$, i.e. for any $(\lambda_A, \lambda_{DM})$ there exists $(\sigma, \mu) \in E$ such that $\pi_{DM}(\sigma, \mu; \Gamma) < \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ if and only if $(\lambda_A, \lambda_{DM}) \in W_S \cup W_I$.*

Theorem 1 shows that the DM always strictly benefits from hidden testing for some parameters, while for other parameters it is possible that he strictly suffers from hidden testing. Interestingly, the effect which causes the DM to be better off in the skepticism

[22] The motivation for this terminology will be explained later.

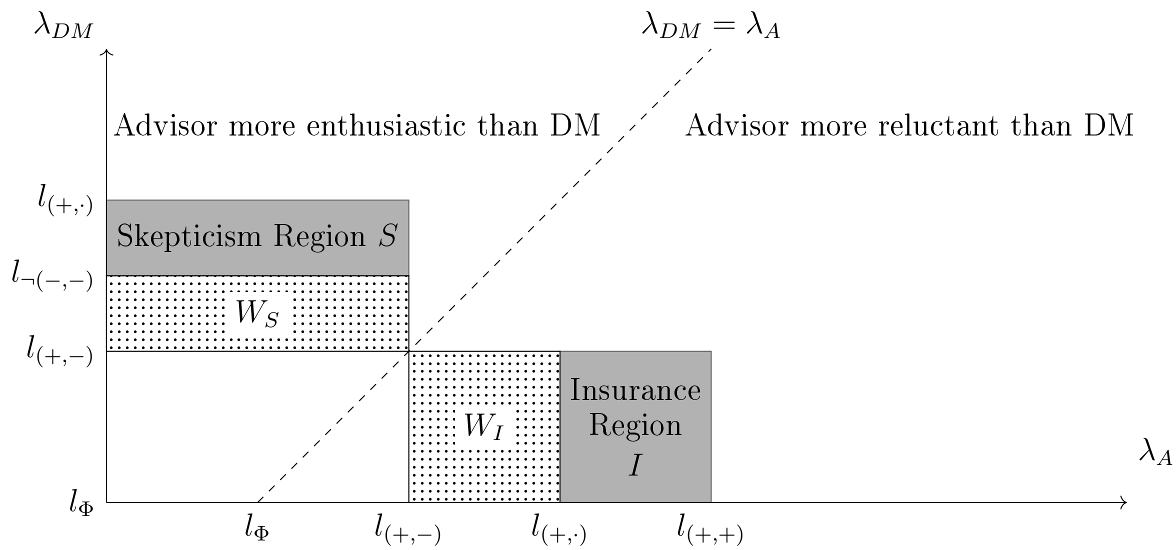


Figure 2: This figure shows the regions in which the DM has a strictly higher payoff in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing, i.e. the skepticism region $S$ and the insurance region $I$, and the regions in which there exists some Pareto-undominated equilibrium under hidden testing which gives the DM a strictly lower payoff than the unique equilibrium under observable testing, labeled $W_S$ and $W_I$, in $(\lambda_A, \lambda_{DM})$ −space for $p_F > p_T$. If $p_T \geq p_F$ then regions $I$ and $W_I$ would be empty because there exists no value of $\lambda_{DM}$ such that $l_\Phi < \lambda_{DM} \leq l_{(+,-)}$.

region, where the advisor is more enthusiastic about accepting than the DM ($\lambda_A < \lambda_{DM}$), turns out to be distinct from the effect at work in the insurance region, where the advisor is more reluctant to accept than the DM ($\lambda_A > \lambda_{DM}$). The effect at work in the insurance region is novel, while the effect at work in the skepticism region is related to findings in the existing literature.

To explain these results, the following Lemmas characterize the associated equilibria for each of the aforementioned regions in turn.[23] The proofs can be found in the appendix in Sections A.4 and A.5.[24] When testing is hidden, I focus on the advisor-preferred equilibrium $(\underline{\sigma}, \underline{\mu})$. This is because the DM's payoff in the advisor-preferred equilibrium is a lower bound on her payoffs in any Pareto-undominated equilibrium under hidden testing by Lemma 2. For any message $m$ that contains two outcomes, the associated belief is derived by Bayes' rule independent of equilibrium play and denoted by $l_m$.[25]

[23]The characterizations require additional likelihood ratios defined based on $l_\varphi$ where $\varphi \in \Phi$: $l_{(-,\cdot)} \equiv l_{\{(-,+),(-,-)\}} = \frac{q}{1-q}\frac{1-p_T}{p_F}$ and $l_{(-,-)} \equiv l_{\{(-,-)\}} = \frac{q}{1-q}\frac{(1-p_T)^2}{p_F^2}$ and $l_{\neg(+,+)} \equiv l_{\Phi\setminus\{(+,+)\}} = \frac{q}{1-q}\frac{1-p_T^2}{1-(1-p_F)^2}$.

[24]Characterizations of equilibria for the remaining preference parameter regions can also be found in the appendix in Sections A.4 and A.5.

[25]This is due to the assumption of verifiable disclosure.

**Lemma 3 (Insurance Region)**
*Suppose* $(\lambda_A, \lambda_{DM}) \in I$.

- *In the unique equilibrium under observable testing: the DM accepts if and only if* $\widetilde{h} \in \{\{+,+\}, \{+,-\}, \{+\}\}$*; the advisor never tests at any history of outcomes.*
- *In an advisor-preferred equilibrium under hidden testing: the DM accepts if and only if* $m \in \{\{+,+\}, \{+,-\}, \{+\}\}$*; the advisor always tests in period 1, whereas he tests in period 2 if and only if the first test is positive, and he discloses only negative outcomes unless he found two positive outcomes, i.e.*

$$\underline{\sigma}_A^M = \begin{cases} \emptyset & \textit{if } \widetilde{h} \in \{\emptyset, \{+\}\} \\ \{-\} & \textit{if } \widetilde{h} \in \{\{-\}, \{+,-\}\} \\ \widetilde{h} & \textit{otherwise;} \end{cases} \tag{16}$$

*The DM's beliefs conditional on messages satisfy*

$$\frac{\underline{\mu}_{DM}(m)}{1-\underline{\mu}_{DM}(m)} = \begin{cases} l_{\neg(+,+)} & \textit{if } m = \{-\} \\ l_{(+,\cdot)} & \textit{if } m = \{+\} \\ l_{\Phi} & \textit{if } m = \emptyset \\ l_m & \textit{otherwise.}^{26} \end{cases} \tag{17}$$

In the insurance region $I$, the DM prefers rejection at the prior belief, but if she knew that two tests were run and outcomes were mixed then this would be sufficient for her to prefer acceptance. Note that for the insurance region to be non-empty, it must be that mixed outcomes are evidence in favor of the hypothesis, i.e. $p_F > p_T$ .[27] The advisor is more reluctant and prefers acceptance if and only if both outcomes are positive.

On the equilibrium path under observable testing, the advisor does not test at all and the DM rejects. If the advisor tested in period 1, he would never test in period 2. This

[26] Note that $l_m$ is a slight abuse of notation since $m$ is an unordered set of outcomes whereas $l_\varphi$ with $\varphi \in \Phi$ is defined based on a set of complete lists of ordered outcomes. However, if $m$ contains two outcomes then the order of these outcomes does not matter. The notation is short for

$$\frac{\underline{\mu}_{DM}(m)}{1-\underline{\mu}_{DM}(m)} = \begin{cases} l_{(+,+)} & \text{if } m = \{+,+\} \\ l_{(-,-)} & \text{if } m = \{-,-\} \\ l_{(+,-)} = l_{(-,+)} & \text{if } m = \{+,-\}. \end{cases} \tag{18}$$

[27] If $N > 2$ an insurance region exists even if $p_F = p_T$ (see Section 3.3).

is because if the first test outcome was positive, the DM would accept regardless of what the second test shows. And if the first test outcome was negative, the advisor would prefer rejection regardless of the second outcome and, if he stopped testing, the DM would reject. But then the advisor is better off if he does not test at all, because a single test can never provide sufficient evidence for him to prefer acceptance.

By contrast, on the equilibrium path under hidden testing, the advisor tests once and again if the first outcome is positive. If he finds two positives, he reveals these and the DM accepts. Otherwise, he reveals a single negative. Then the DM does not know if the advisor either found a negative outcome on the first test and then stopped, in which case the DM would prefer rejection, or if he first found a positive and then a negative outcome, in which case the DM would prefer acceptance. However, on average, the DM infers that a report of a single negative must be bad news about the hypothesis being true and rejects.[28] Therefore, the DM always acts in the advisor's interest, which gives the advisor an incentive to test.

In conclusion, under observable testing, the advisor strategically avoids testing because testing has the downside that the evidence could turn out to support the hypothesis enough for the DM to accept, yet not enough for the advisor to prefer acceptance. By allowing the advisor to test in private, the DM insures the advisor against this downside. As a consequence, the DM at least learns whether or not the evidence is strong enough for both to agree that acceptance is the better choice. In particular, hidden testing can be compared to providing the advisor with limited liability. Under observable testing, running tests is a gamble for the advisor which does not pay off on average. But under hidden testing, he has the upside from testing, which is that acceptance is chosen following two positives, but no downside, because rejection is chosen following mixed outcomes. Interestingly, offering this limited liability also makes the DM better off. Although she forgoes the option to accept following mixed outcomes, she benefits from the fact that the advisor then takes the gamble and tests.

**Lemma 4 (Skepticism Region)**
*Suppose* $(\lambda_A, \lambda_{DM}) \in S$.

- *In the equilibrium under observable testing: the DM accepts if and only if* $\widetilde{h} \in \{\{+,+\}, \{+\}\}$*; the advisor tests once.*[29]
- *In an advisor-preferred equilibrium under hidden testing: the DM accepts if and only if* $m \in \{+,+\}$*; the advisor always tests in period 1, whereas he tests in period 2 if and*

[28] The DM either receives a report of two positives or of a single negative. A report of two positives raises her posterior belief that the hypothesis is true. Since her expected posterior must equal her prior belief, this implies that a report of a single negative must lower her posterior belief.

[29] He always tests in period 1, whereas he tests in period 2 if and only if he did not test in period 1.

*only if the first test is positive, and he discloses only positive outcomes, i.e.*

$$\underline{\sigma}_A^M = \begin{cases} \{+,+\} & \text{if } \widetilde{h} = \{+,+\} \\ \emptyset & \text{if } \widetilde{h} \in \{\emptyset, \{-\}, \{-,-\}\} \\ \{+\} & \text{otherwise.} \end{cases} \tag{19}$$

*The DM's beliefs conditional on messages satisfy*

$$\frac{\underline{\mu}_{DM}(m)}{1-\underline{\mu}_{DM}(m)} = \begin{cases} l_{(+,-)} & \text{if } m = \{+\} \\ l_{(-,\cdot)} & \text{if } m = \emptyset \\ l_{\neg\{+,+\}} & \text{if } m = \{-\} \\ l_m & \text{otherwise.} \end{cases} \tag{20}$$

In the skepticism region $S$, the DM would prefer acceptance if she knew that a single test was run and the outcome was positive. However, she would prefer rejection if she knew that two tests were run and outcomes were mixed.[30] By contrast, the advisor is more enthusiastic and would prefer acceptance even if he knew two tests were run and the outcomes were mixed.[31]

On the equilibrium path under observable testing, the advisor tests only in the first period and the DM accepts if and only if the outcome is positive. Following a first positive outcome, the advisor prefers acceptance irrespective of what a second test shows and the DM accepts. Following a first negative outcome, the DM would reject irrespective of the second outcome. If the DM and the advisor agree that rejection is preferred at the prior belief, the DM also chooses the advisor's preferred action following either test outcome. Hence, by testing, the advisor increases the chances that the DM's action is appropriate given the state. If the DM and the advisor disagree at the prior belief, then the advisor benefits from the fact that the DM acts in his interest if the outcome is positive, while the DM's choice is unaffected if the outcome is negative.

On the equilibrium path under hidden testing, the advisor tests to see if he can find two positive outcomes and discloses only positive outcomes. The DM accepts if and only if two tests are positive. It cannot be part of an equilibrium that the advisor reveals a single positive outcome and the DM accepts, as was the case under observable testing. To see why, suppose the DM were to accept based on the report of a single positive outcome. Then the advisor would stop testing following a positive outcome in period 1, but he would be tempted to keep

[30] Recall that if $p_F > p_T$, then $l_\Phi < l_{(+,-)}$ and a preference for rejection given mixed outcomes implies a preference for rejection at the prior.

[31] In the extreme case, the advisor prefers acceptance regardless of test outcomes, i.e. $\lambda_A = 0$.

testing following a negative outcome in period 1. If he finds a positive outcome in period 2 he can hide the negative outcomes discovered previously and still achieve acceptance. The DM could then not be sure whether the reported positive outcome was found on the first or on the second test. Conditional on the report of a single positive outcome, her beliefs would result in a posterior likelihood ratio equal to $l_{\neg(-,-)}$, i.e. all she can infer is that not both tests were negative. On balance, the DM would find this too weak evidence in favor of the hypothesis and would reject. Therefore, the advisor can only convince the DM to accept by revealing two positive outcomes.

In conclusion, under hidden testing, the DM learns whether a positive outcome is backed up by an additional test, and this allows her to increase her expected payoff. Because the advisor has the temptation to hide negative evidence, the DM can be credibly skeptical towards any report of weak evidence in favor of the hypothesis. As a result, she requires a larger number of positive test outcomes to accept.

While the DM is better off in any Pareto-undominated equilibrium under hidden testing in the skepticism or the insurance region, there are other regions in which the DM can end up strictly worse off when testing is hidden. In these regions, preferences are more aligned than in either the skepticism or the insurance region. Although the advisor conducts more tests when testing is hidden, the DM does not benefit, because the additional tests are only used to expand the advisor's scope for manipulation.

**Lemma 5 (Regions $W_S$ and $W_I$)**

1. *Suppose $(\lambda_A, \lambda_{DM}) \in W_S$.*

   - *The equilibrium under observable testing is as in the skepticism region $S$.*
   - *In an advisor-preferred equilibrium under hidden testing: the DM accepts if and only if $m \in \{\{+,+\},\{+\}\}$; the advisor always tests once, if and only if the first test is negative he tests again, and his disclosure strategy is given by (19). The DM's beliefs conditional on messages satisfy*

$$\frac{\underline{\mu}_{DM}(m)}{1-\underline{\mu}_{DM}(m)} = \begin{cases} l_{\neg(-,-)} & \text{if } m = \{+\} \\ l_{(-,\cdot)} & \text{if } m = \{-\} \\ l_{(-,-)} & \text{if } m = \emptyset \\ l_m & \text{otherwise.} \end{cases} \tag{21}$$

2. *Suppose* $(\lambda_A, \lambda_{DM}) \in W_I$.

   - *In the equilibrium under observable testing: the DM accepts if and only if* $\widetilde{h} \in \{\{+,+\},\{+,-\},\{+\}\}$*; the advisor tests once.*
   - *In an advisor-preferred equilibrium under hidden testing: the DM accepts if and only if* $m = \{\{+\},\{+,-\},\{+,+\}\}$*; the advisor always tests once, if and only if the first test is positive he tests again, and discloses*

$$\underline{\sigma}_A^M = \begin{cases} \{-\} & \text{if } \widetilde{h} \in \{\{-\},\{+,-\}\} \\ \widetilde{h} & \text{otherwise.} \end{cases} \tag{22}$$

   *The DM's beliefs conditional on messages satisfy (17).*

Consider $(\lambda_A, \lambda_{DM}) \in W_S$. The DM has a lower loss $\lambda_{DM}$ from falsely accepting than in the skepticism region and this changes what happens on the equilibrium path under hidden testing. Just as when $(\lambda_A, \lambda_{DM}) \in S$, if a single positive outcome is reported, all the DM can infer is that not both tests were negative. However, in contrast to when $(\lambda_A, \lambda_{DM}) \in S$, she is willing to accept based on this inference, i.e. $\lambda_{DM} < l_{\neg(-,-)}$. Therefore, the DM only learns whether or not both tests were negative under hidden testing, which is worse for her than learning whether or not the first test was positive.

In addition, consider $(\lambda_A, \lambda_{DM}) \in W_I$. The advisor has a lower loss $\lambda_A$ from falsely accepting than in the insurance region $I$ and, as a consequence, he would prefer acceptance if he knew that a single test was run and the outcome was positive. Hence, under observable testing, in contrast to when $(\lambda_A, \lambda_{DM}) \in I$, the advisor tests once and the DM acts in her interest following either outcome. Under hidden testing, just as when $(\lambda_A, \lambda_{DM}) \in I$, the advisor tests to see if he can find two positive outcomes and the DM accepts if and only if both outcomes are positive. Therefore, like in the insurance region, hidden testing can be thought of providing limited liability to the advisor. It guarantees that, if outcomes are mixed, the final decision is in line with the advisor's interests. Therefore, the advisor has an incentive to run a second test following a positive outcome. However, unlike in the insurance region, this is not necessary to encourage him to run the first test. Therefore, the DM does not benefit from providing this limited liability, but instead suffers because rejection is chosen when outcomes are mixed.

The DM can be strictly harmed by hidden testing if and only if $(\lambda_A, \lambda_{DM}) \in W_S \cup W_I$, which implies that the DM is weakly better off under hidden rather observable testing if her preferences are sufficiently misaligned with the advisor's preferences.

**Corollary 1 (Preference Alignment)** *There exists a threshold $d > 0$ such that for any $(\lambda_A, \lambda_{DM})$ where $|\lambda_A - \lambda_{DM}| > d$ the DM has a weakly higher payoff in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing, i.e. for any $(\lambda_A, \lambda_{DM})$, if $|\lambda_A - \lambda_{DM}| > d$ then $\pi_{DM}(\sigma, \mu; \Gamma) \geq \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ for all $(\sigma, \mu) \in E$.*

As a benchmark, consider a scenario in which the DM is in charge both of collecting evidence and of taking the final decision. I will refer to the payoff she could achieve in this scenario as her first-best expected payoff, denoted by $\pi_{DM}^{FB}$.

**Proposition 1 (First-Best Benchmark)**

1. *Given any $(\lambda_A, \lambda_{DM})$, if the DM achieves her first-best expected payoff under observable testing, she also does so in any Pareto-undominated equilibrium under hidden testing, but the converse is not true; i.e. for any $(\lambda_A, \lambda_{DM})$, if $\pi_{DM}(\sigma^o, \mu^o; \Gamma^o) = \pi_{DM}^{FB}$ then $\pi_{DM}(\sigma, \mu; \Gamma) = \pi_{DM}^{FB}$ for all $(\sigma, \mu) \in E$, and if there exists $(\sigma, \mu) \in E$ such that $\pi_{DM}(\sigma, \mu; \Gamma) < \pi_{DM}^{FB}$ then $\pi_{DM}(\sigma^o, \mu^o; \Gamma^o) < \pi_{DM}^{FB}$.*

2. *Given any $(\lambda_A, \lambda_{DM})$, there exists a Pareto-undominated equilibrium under hidden testing in which the DM achieves her first-best expected payoff, i.e. for any $(\lambda_A, \lambda_{DM})$, there exists $(\sigma, \mu) \in E$ such that $\pi_{DM}(\sigma, \mu; \Gamma) = \pi_{DM}^{FB}$ .*

This shows that the parameter region in which the DM achieves her first-best expected payoff in any Pareto-undominated equilibrium under hidden testing is larger than the region in which she achieves her first-best expected payoff under observable testing.

Furthermore, there always exists an equilibrium under hidden testing in which the DM achieves her first-best expected payoff. A characterization of these equilibria can be found in the appendix in Section A.5. These equilibria have in common that, in some situations, the DM believes that the advisor omits unfavorable outcomes when he discloses fewer than two outcomes. The DM learns more, because the advisor then has an incentive to test to prove the DM wrong.[32] The multiplicity of equilibria in discussed further in Section 3.2.

So far I have focused on the changes in the DM's expected payoff. The following proposition characterizes changes in the advisor's expected payoff.

[32] For example, if $(\lambda_A, \lambda_{DM}) \in W_S$, then there exists a DM-preferred equilibrium in which the DM believes that a single positive outcome necessarily indicates that the advisor has both a positive and a negative outcome and is hiding the negative outcome. Consequently, the advisor cannot convince the DM to accept unless he reveals two positive outcomes. Therefore, the advisor tests in period 2 if the first test is positive. If he finds another positive, he discloses both and the DM accepts. If he ends up with mixed outcomes, he reveals only the positive and the DM rejects.

**Proposition 2 (Advisor Payoff Comparison)**

1. *Given any* $(\lambda_A, \lambda_{DM})$*, there exists a Pareto-undominated equilibrium under hidden testing which gives the advisor a weakly lower payoff than the unique equilibrium under observable testing, i.e. for any* $(\lambda_A, \lambda_{DM})$*, there exists* $(\sigma, \mu) \in E$ *such that* $\pi_A(\sigma, \mu; \Gamma) \leq \pi_A(\sigma^o, \mu^o; \Gamma^o)$.

2. *Given* $(\lambda_A, \lambda_{DM})$*, the advisor has a strictly lower payoff in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing if and only if* $(\lambda_A, \lambda_{DM}) \in S$*, i.e. for any* $(\lambda_A, \lambda_{DM})$ *it holds that* $\pi_A(\sigma, \mu; \Gamma) < \pi_A(\sigma^o, \mu^o; \Gamma^o)$ *for all* $(\sigma, \mu) \in E$ *if and only if* $(\lambda_A, \lambda_{DM}) \in S$.

3. *Given* $(\lambda_A, \lambda_{DM})$*, there exists a Pareto-undominated equilibrium under hidden testing which gives the advisor a strictly higher payoff than the unique equilibrium under observable testing if and only if* $(\lambda_A, \lambda_{DM}) \in I \cup W_S \cup W_I$*, i.e. for any* $(\lambda_A, \lambda_{DM})$ *there exists* $(\sigma, \mu) \in E$ *such that* $\pi_A(\sigma, \mu; \Gamma) > \pi_A(\sigma^o, \mu^o; \Gamma^o)$ *if and only if* $(\lambda_A, \lambda_{DM}) \in I \cup W_S \cup W_I$.

In the skepticism region, the advisor is strictly worse off under hidden testing because the DM has a higher acceptance threshold than under observable testing and, therefore, the advisor is less likely to convince the DM to act in his interest. By contrast, in the insurance region $I$ as well as in regions $W_S$ and $W_I$, the advisor can be strictly better off under hidden testing, because he can manipulate the DM to act in his interest by strategically withholding information.

**Corollary 2 (Hidden Testing Pareto-improving)**

1. *Given* $(\lambda_A, \lambda_{DM})$*, there exists a Pareto-undominated equilibrium under hidden testing which gives both the DM and the advisor a strictly higher payoff off than the unique equilibrium under observable testing if and only if* $(\lambda_A, \lambda_{DM}) \in I$*, i.e. for any* $(\lambda_A, \lambda_{DM})$ *there exists* $(\sigma, \mu) \in E$ *such that* $\pi_i(\sigma, \mu) > \pi_i(\sigma^o, \mu^o; \Gamma^o)$ *for* $i = A, DM$ *if and only if* $(\lambda_A, \lambda_{DM}) \in I$.

2. *Given any* $(\lambda_A, \lambda_{DM})$*, the DM and the advisor never both have a strictly higher payoff in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing, i.e. for any* $(\lambda_A, \lambda_{DM})$*, there exists* $(\sigma, \mu) \in E$ *such that* $\pi_A(\sigma, \mu; \Gamma) \leq \pi_A(\sigma^o, \mu^o; \Gamma^o)$ *or* $\pi_{DM}(\sigma, \mu; \Gamma) \leq \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$.

Interestingly, it is only possible that both players are strictly better off under hidden than observable testing if the advisor is more reluctant than the DM. In particular, the skepticism effect benefits the DM, but harms the advisor, whereas the insurance effect benefits both players.

## 3.2 Discussion

Under hidden testing, there is no unique sequential equilibrium, and the DM does not perfectly infer all information acquired by the advisor, i.e. the unraveling result by Milgrom (1981) and Grossman (1981) does not apply. DM-preferred equilibria have the feature that the DM learns all the decision-relevant information that can be acquired from $N$ tests. In particular, the DM believes that some reports with fewer than $N$ outcomes are omitting unfavorable outcomes. As a best response, the advisor tests more and fully discloses what he found such that the DM acts more in line with his interests.[33] By contrast, in an advisor-preferred equilibrium, the DM takes all reports at face value, unless the advisor's best response would render such beliefs inconsistent. In particular, if fewer than $N$ outcomes are disclosed, the DM may believe that the advisor is not withholding evidence, but that he stopped testing early. This in turn means that the advisor has no reason to keep testing if there is a report that leads the DM to act in his interest and additional outcomes are not pivotal to his preferred final decision.

There are two reasons for why there are multiple sequential equilibria. The main reason is that the DM's choice is discrete. This implies that there is a range of posterior beliefs for which the DM's optimal action choice is unchanged. Consequently, even if the DM believes that the advisor is withholding some information, this does not necessarily translate into an action choice that is less desirable for the advisor. A second reason for why unraveling does not occur is that the advisor faces a cost of testing. Therefore, the DM cannot infer how many test outcomes the advisor has acquired in total and, hence, how many outcomes he has withheld.

However, if tests were assumed to be costless, all results on payoff comparisons would remain unchanged. In particular, holding fixed all $N$ draws by Nature, the mapping from the list of Nature's $N$ draws to final decisions is unchanged. Suppose the advisor would run all $N$ tests under hidden testing. In an DM-preferred equilibrium, just as in a framework with unraveling, the DM would believe that any withheld outcome is an unfavorable one. The advisor would then have an incentive to make the strongest possible case for his preferred

[33] These equilibria are sensitive to introducing the possibility of a failed test or the possibility that the advisor can only test once. This is because then there is an alternative sequence of events following which only a single positive outcome is revealed and if the DM could be sure that this alternative sequence of events had taken place she would prefer to accept.

final decision and reveal all favorable ones. By contrast, in an advisor-preferred equilibrium, the advisor would at most reveal a certain number of favorable outcomes, even if he has acquired more favorable outcomes, and the DM accepts if at least this number of favorable outcomes is revealed. This means that the advisor effectively pools some realizations of the $N$ outcomes for which the DM would ideally accept with some realizations for which the DM would ideally reject in such a way that on average the DM prefers to accept. In these situations, the DM acts in the advisor's interest, even though the fact that information is being withheld leaves her uncertain that this is also in her interest.

When testing is costless, an advisor-preferred equilibrium has the undesirable feature that for some outcomes the advisor does not make the strongest case for the final decision he prefers. Instead, he sometimes withholds favorable outcomes when indifferent between disclosing them or not. However, when testing is costly, the advisor does not acquire these additional outcomes in the first place. Therefore, the advisor can make the strongest case for his preferred decision given what he found, yet the DM does not learn about all outcomes, because she cannot infer how many tests were conducted in total.[34] This lack of knowledge about the total number of tests conducted is realistic in many contexts.

## 3.3 N Periods

In this subsection, I show that the key insights of Subsection 3.1 are not specific to a model with $N = 2$ by providing sufficient conditions for their existence for any $N > 2$. For tractability, I assume that the test is symmetric in the sense that false negatives are equally likely as false positives, i.e. $p_T = p_F = p$.

I denote the number of positive outcomes in $h_n$ by history of outcomes at the end of period $n$ by $\nu_n^+$ and the number of negative outcomes by $\nu_n^-$. Given the assumption that the test is symmetric, a sufficient statistic for the posterior belief is how many more positive than negative outcomes were found, independent of the total number of outcomes. I denote the number of excess positive outcomes by $x_n \equiv \nu_n^+ - \nu_n^-$ and use $l\left(x_n\right)$ to denote the posterior likelihood ratio that the hypothesis is true conditional on $x_n$, i.e.

$$l\left(x_n\right) = \frac{q}{1-q}\frac{(p)^{\nu_n^+}\left(1-p\right)^{\nu_n^-}}{\left(1-p\right)^{\nu_n^+}(p)^{\nu_n^-}} = \frac{q}{1-q}\frac{p^{\nu_n^+-\nu_n^-}}{\left(1-p\right)^{\nu_n^+-\nu_n^-}} = \frac{q}{1-q}\frac{p^{x_n}}{\left(1-p\right)^{x_n}}. \tag{23}$$

To describe the equilibrium, it is helpful to denote the largest number of excess positive

[34]Note that in all equilibrium characterizations in this paper the advisor makes the strongest possible case for his preferred decision.

outcomes at which a player prefers rejection by $r_i$ for $i = A, DM$, i.e.

$$r_{DM} \equiv sup \{j | j \in \mathbb{N}_{\geq 0}, l(j) < \lambda_{DM}\}, \tag{24}$$

$$r_A \equiv sup \{j | j \in \mathbb{Z}, l(j) < \lambda_A\}. \tag{25}$$

To compare players' payoffs across regimes, the following lemma describes on-path equilibrium play under observable testing.[35]

**Lemma 6 (Observable Testing for $N > 2$)** *Suppose testing is observable. For any $N > 2$ and $(\lambda_A, \lambda_{DM})$, the following occurs in equilibrium. If $r_{DM} \geq N$ the advisor never starts testing and the DM rejects.*

1. *If the advisor is more reluctant, i.e. $\lambda_A \geq \lambda_{DM}$, and $r_{DM} \leq N - 1$, there exists a critical value $\overline{\lambda}_A \geq \lambda_{DM}$ such that the advisor never starts testing and the DM rejects if and only if $\lambda_A > \overline{\lambda}_A$.*

2. *If the advisor is more enthusiastic, i.e. $\lambda_A \leq \lambda_{DM}$, and $r_{DM} \leq N - 1$, there exists a critical value $\underline{\lambda}_A \leq \lambda_{DM}$ such that if $\lambda_A < \underline{\lambda}_A$ the advisor tests in period $n + 1$ if and only if $r_{DM} + 1 - (N - n) \leq x_n \leq r_{DM}$ for $n \in \{0, ..., N - 1\}$. If the advisor stops at $x_n = r_{DM} + 1$ then the DM accepts, otherwise the DM rejects.*

At each history, the advisor's optimal action depends not just on the current evidence and the associated posterior belief about the state, but also on how many tests he can still run. Suppose the advisor is more enthusiastic than the DM. The advisor keeps testing if the DM would reject given the current evidence, but would accept for some realization of future outcomes. The reason is that if future outcomes are such that the DM accepts, then they also lead the advisor to prefer acceptance. By contrast, the advisor may face a trade-off if the evidence collected up to now leads the DM to accept, but for some realization of future outcomes the DM rejects. The downside to testing is that the additional evidence could lead the DM to reject, yet be insufficient for the advisor to prefer rejection. The upside to testing is that the additional evidence could lead both of them to agree that rejecting is optimal. This upside ceases to exist if the advisor's loss from falsely accepting is sufficiently low. Figure 3 illustrates a situation in which the advisor always stops as soon as the evidence leads the DM to accept.

[35] The proof contains a complete equilibrium characterization.

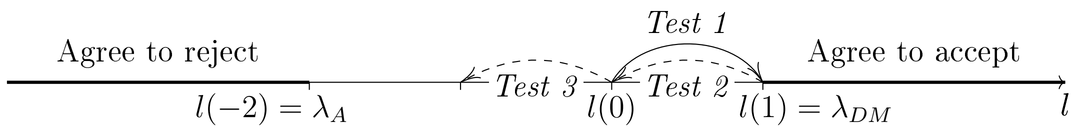


Figure 3: An example of a combination of preference parameters given $N = 3$ for which the advisor optimally stops testing as soon as the likelihood ratio leads the DM to accept: if the first outcome is positive, the advisor stops testing because the DM acts in line with his interest and the remaining two outcomes can never convince him that rejecting is optimal, not even if both turned out negative.

An analogous reasoning applies if the advisor is more reluctant than the DM. I find the smallest critical value $\overline{\lambda}_A$ such that an advisor who has a higher loss from falsely accepting, i.e. $\overline{\lambda}_A < \lambda_A$, does not test at all, although there are realizations of future outcomes for which both players would agree that accepting is optimal. Consider the following example illustrated in Figure 4, where $N = 3$. The DM is just willing to accept if there is one more positive outcome than there are negative outcomes, i.e. $\lambda_{DM} = l\,(1)$, and the advisor is just willing to accept if there are three more positive than negative outcomes, i.e. $\lambda_A = l\,(3)$. If the first two outcomes happened to be positive, at which point the likelihood ratio is $l\,(2)$, the DM accepts regardless of what the final test shows. Therefore, the advisor has no reason to do the final test at $l\,(2)$. However, then the advisor can never reach a likelihood ratio at which he prefers acceptance. Hence, the advisor is better off if he does not test. More generally, consider the strongest possible evidence in favor of the hypothesis at which the DM accepts regardless of what the remaining tests show. If and only if the advisor prefers rejection at this likelihood ratio, then he is sufficiently reluctant to not start testing.

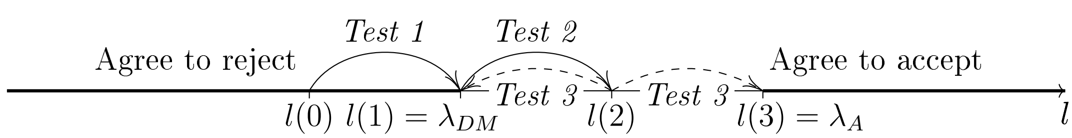


Figure 4: An example of a combination of preference parameters given $N = 3$ for which the advisor optimally does not start testing, although there exists some realization of test outcomes such that both players agree that accepting is optimal.

The following lemma describes equilibrium play under hidden testing, which can generate the expected payoffs attained in the advisor-preferred equilibrium. Let $m = (m^+, m^-)$ denote a message that discloses $m^+$ positive outcomes and $m^-$ negative outcomes.

**Lemma 7 (Hidden Testing for $N > 2$)** *Suppose testing is hidden. For any $N > 2$ and $(\lambda_A, \lambda_{DM})$, there exists an advisor-preferred equilibrium in which the following occurs.*

1. *If the advisor is more reluctant, i.e. $\lambda_A \geq \lambda_{DM}$, there exists a threshold $\overline{\nu}_+$ of positive outcomes and a threshold $\overline{\nu}_-$ of negative outcomes given by*

$$\overline{\nu}_+ \equiv inf\left\{ j | j \in \mathbb{N}, j \geq \frac{r_A + 1 + N}{2} \right\}, \tag{26}$$

$$\overline{\nu}_- \equiv N - \overline{\nu}_+ + 1, \tag{27}$$

   *such that the advisor tests in period $n+1$ given $h_n$ if and only if both $\nu_n^+ < \overline{\nu}_+$ and $\nu_n^- < \overline{\nu}_-$ for $n \in \{0, ..., N-1\}$. In period $N$, either $\nu_n^+ = \overline{\nu}_+$ or $\nu_n^- = \overline{\nu}_-$. If $\nu_n^+ = \overline{\nu}_+$, he reports $m = (\overline{\nu}_+, 0)$ and the DM accepts. If $\nu_n^- = \overline{\nu}_-$, he reports $m = (0, \overline{\nu}_-)$ and the DM rejects.*

2. *If the advisor is more enthusiastic, i.e. $\lambda_A \leq \lambda_{DM}$, there exist thresholds $\widehat{\nu}_+$ and $\widehat{\nu}_-$ given by*

$$\widehat{\nu}_+ \equiv inf\left\{ j | j \in \mathbb{N}_{\geq 0}, \lambda_{DM} \leq \frac{q}{1-q} \frac{\sum_{t=j}^{t=N} \binom{N}{t} p^t (1-p)^{N-t}}{\sum_{t=j}^{t=N} \binom{N}{t} (1-p)^t p^{N-t}} \right\}, \tag{28}$$

$$\widehat{\nu}_- \equiv N - \widehat{\nu}_+ + 1, \tag{29}$$

   *such that if $\widehat{\nu}_+ \leq \overline{\nu}_+$ then the equilibrium path is as in Part 1. If $\widehat{\nu}_+ > \overline{\nu}_+$, the following occurs. The advisor tests in period $n+1$ given $h_n$ if and only if $\nu_n^+ < \widehat{\nu}_+$ and $\nu_n^- < \widehat{\nu}_-$ for $n \in \{0, ..., N-1\}$. In period $N$, either $\nu_n^+ = \widehat{\nu}_+$ or $\nu_n^- = \widehat{\nu}_-$. If $\nu_n^+ = \widehat{\nu}_+$ then the advisor reports $m = (\widehat{\nu}_+, 0)$ and the DM accepts. If $\nu_n^- = \widehat{\nu}_-$ and $\nu_n^+ \leq r_A + \widehat{\nu}_-$ then the advisor reports $m = (0, \widehat{\nu}_-)$ and the DM rejects. If $\nu_n^- = \widehat{\nu}_-$ and $\nu_n^+ > r_A + \widehat{\nu}_-$ then the advisor reports $m = (\nu_n^+, 0)$ and the DM rejects.*

Suppose the advisor is more reluctant than the DM. Lemma 7 shows that the advisor tests until the remaining tests cannot be pivotal to determining which action he prefers the DM to take.[36] Once he has concluded testing, if he prefers to accept, then he reports all positive outcomes and no negative outcomes. Otherwise, he reports all negative outcomes and no positive outcomes. Hence, a report of only positive outcomes indicates to the DM

[36]The threshold $\overline{\nu}_+$ is defined to be the smallest number of positive outcomes such that, if the advisor has observed $\overline{\nu}_+$ outcomes, then he prefers acceptance, irrespective of the realization of the remaining $N - \overline{\nu}_+$ outcomes. Similarly, the threshold $\overline{\nu}_-$ is defined to be the smallest number of negative outcomes such that, if the advisor has observed $\overline{\nu}_-$ outcomes, then he prefers rejection, irrespective of the realization of the remaining $N - \overline{\nu}_-$ outcomes.

that the advisor prefers acceptance. Since the advisor is more reluctant, the DM must then also prefer acceptance. By contrast, a report of only negative outcomes indicates that the advisor prefers rejection. It is not clear whether the DM would also prefer rejection if she could observe all that the advisor has observed. However, the fact that the advisor prefers rejection must lead the DM to revise her beliefs about the hypothesis being true downwards and, therefore, she optimally rejects.[37]

Equilibrium play may differ when the advisor is more enthusiastic than the DM. Note that the advisor prefers acceptance at $h_n$ irrespective of future test outcomes if $\nu_n^+ = \overline{\nu}_+$. Suppose the advisor stops testing as soon as $\nu_n^+ = \overline{\nu}_+$ and reports $\overline{\nu}_+$ positive outcomes and no negative outcomes. Then, in equilibrium, the DM can only infer that at least $\overline{\nu}_+$ outcomes in $N$ tests were positive. The threshold $\widehat{\nu}_+$ is defined to be the smallest number of positive outcomes such that, if the DM believes that at least this number of outcomes in $N$ tests are positive, she optimally accepts. Hence, the DM accepts given a report of $\overline{\nu}_+$ positive outcomes if and only if $\widehat{\nu}_+ \leq \overline{\nu}_+$. Therefore, if $\widehat{\nu}_+ \leq \overline{\nu}_+$ the advisor indeed stops testing as soon as $\nu_n^+ = \overline{\nu}_+$. However, if $\widehat{\nu}_+ > \overline{\nu}_+$ the DM rejects given a report of $\overline{\nu}_+$ positive outcomes. Therefore, if $\widehat{\nu}_+ > \overline{\nu}_+$ the advisor has a reason to keep testing until either $\nu_n^+ = \widehat{\nu}_+$ or $\nu_n^- = \widehat{\nu}_-$. If $\nu_n^+ = \widehat{\nu}_+$ he can report $\widehat{\nu}_+$ positive outcomes and the DM accepts. If $\nu_n^- = \widehat{\nu}_-$, he can never find $\widehat{\nu}_+$ positive outcomes and the DM rejects independent of the report. Therefore, if $\widehat{\nu}_+ > \overline{\nu}_+$, a more enthusiastic advisor faces an additional constraint compared to a more reluctant advisor, which implies that the DM does not always act in the advisor's interest.

The following proposition shows that analogues of regions $S$, $I$, $W_S$ and $W_I$ exist for any number of periods $N > 2$.

**Theorem 2 (DM Payoff Comparison for $N > 2$)**

1. *For any $N > 2$ and any $(\lambda_A, \lambda_{DM})$ where $\lambda_A \geq \lambda_{DM}$,*

   (a) *there exists an insurance region $\mathcal{I} \subset \mathbb{R}^2$ such that if $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$ the DM has a strictly higher payoff in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing, i.e. $\pi_{DM}(\sigma, \mu; \Gamma) > \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ for all $(\sigma, \mu) \in E$, and*

   (b) *there exists a region $W_{\mathcal{I}} \subset \mathbb{R}^2$ such that if $(\lambda_A, \lambda_{DM}) \in W_{\mathcal{I}}$ there exists a Pareto-undominated equilibrium under hidden testing which gives the DM a strictly lower payoff than the unique equilibrium under observable testing, i.e. there exists $(\sigma, \mu) \in E$ such that $\pi_{DM}(\sigma, \mu; \Gamma) < \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$.*

[37]A report of only positives raises her posterior belief that the hypothesis is true. Since her expected posterior must equal her prior belief, this implies that a report of only negatives must lower her posterior belief.

*2. For any $N > 2$ and any $(\lambda_A, \lambda_{DM})$ where $\lambda_A \leq \lambda_{DM}$,*

*(a) there exists a skepticism region $\mathcal{S} \subset \mathbb{R}^2$ such that if $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$ the DM has a strictly higher payoff in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing, i.e. $\pi_{DM}(\sigma, \mu; \Gamma) > \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ for all $(\sigma, \mu) \in E$, and*

*(b) there exists a region $W_{\mathcal{S}} \subset \mathbb{R}^2$ such that if $(\lambda_A, \lambda_{DM}) \in W_{\mathcal{S}}$ there exists a Pareto-undominated equilibrium under hidden testing which gives the DM a strictly lower payoff than the unique equilibrium under observable testing, i.e. there exists $(\sigma, \mu) \in E$ such that $\pi_{DM}(\sigma, \mu; \Gamma) < \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$.*

In regions $\mathcal{S}$ and $W_{\mathcal{S}}$, the advisor must be sufficiently more enthusiastic than the DM such that, under observable testing, the advisor stops testing when he has acquired just enough excess positive outcomes for the DM to accept. In addition, the DM's loss from falsely accepting must be low enough that she would benefit from observing further tests. Furthermore, in region $\mathcal{S}$, the DM's loss from falsely accepting is high enough for her to require a strictly higher number of positive outcomes for acceptance under hidden testing than she required excess positive outcomes under observable testing. This implies that she obtains more decision-relevant information about the hypothesis under hidden testing. However, in the region $W_{\mathcal{S}}$, the DM's loss from falsely accepting is sufficiently low such that she requires a lower number of positive outcomes for acceptance under hidden testing than excess positive outcomes under observable testing. Since the advisor may hide negative outcomes under hidden testing, this implies that she learns less about the hypothesis when testing is hidden.

To construct region $\mathcal{I}$, it is crucial that players agree on the optimal action under certainty. The advisor's loss from falsely accepting needs to be low enough such that, if he were to run all tests, there exist realizations of test outcomes at which both players agree that acceptance is optimal. This ensures that the advisor has an incentive to test under hidden testing. In addition, the advisor's loss from falsely accepting needs to be high enough such that he does not test under observable testing.

Lastly, in region $W_{\mathcal{I}}$, preferences must be sufficiently aligned for the advisor to test even when testing is observable. Yet preferences need to be sufficiently misaligned for the advisor to run more tests under hidden testing. This additional testing is not beneficial for the DM. To the contrary, the DM now rejects following histories of outcomes at which she would have preferred to accept.

The construction of region $\mathcal{I}$ helps to explain why the insurance effect does not arise in a two-period model with a symmetric test, i.e. $p_T = p_F$. Recall that the DM prefers rejection at the prior belief and, hence, she must also prefer rejection when outcomes are mixed. The advisor is more reluctant and, therefore, prefers rejection whenever the DM does. In addition, he needs to prefer acceptance following two positive outcomes, otherwise, he would never prefer acceptance after two tests and would have no incentives to test when testing is hidden. However, this implies that whatever the realization of two test outcomes, the two players always agree on what the preferred final decision is. Hence, the advisor is better off running both tests than not testing at all even when testing is observable.[38]

To summarize the DM's payoff comparison, Corollary 1 continues to hold for $N > 2$.[39]

**Corollary 3 (Preference Alignment for $N > 2$)** *Take any $N > 2$. There exists a threshold $d > 0$ such that for any $(\lambda_A, \lambda_{DM})$ where $|\lambda_A - \lambda_{DM}| > d$ the DM has a weakly higher payoff in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing, i.e. if $|\lambda_A - \lambda_{DM}| > d$ then $\pi_{DM}(\sigma, \mu; \Gamma) \geq \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ for any $(\sigma, \mu) \in E$.*

The following proposition gives sufficient conditions for results in Proposition 1 and 2 and Corollary 2 for any $N > 2$.

**Proposition 3 (Further Payoff Comparisons for $N > 2$)**

1. *DM's first-best benchmark: Given any $N > 2$ and $(\lambda_A, \lambda_{DM})$, if $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$ then the DM achieves her first-best expected payoff in any Pareto-undominated equilibrium under hidden testing, but not in the unique equilibrium under observable testing, i.e. for any $N > 2$ and $(\lambda_A, \lambda_{DM})$, if $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$ then $\pi^{FB}_{DM} = \pi_{DM}(\sigma, \mu; \Gamma) > \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ for all $(\sigma, \mu) \in E$.*

2. *Advisor's payoff comparison: Given any $N > 2$ and $(\lambda_A, \lambda_{DM})$, if $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$ then the advisor achieves a strictly lower payoff in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing, i.e. for any $N > 2$ and $(\lambda_A, \lambda_{DM})$, if $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$ then $\pi_A(\sigma, \mu; \Gamma) < \pi_A(\sigma^o, \mu^o; \Gamma^o)$ for all $(\sigma, \mu) \in E$.*

3. *Hidden testing a Pareto-improvement over observable testing: Given any $N > 2$ and $(\lambda_A, \lambda_{DM})$, if $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$ then there exists a Pareto-undominated equilibrium under hidden testing in which both the DM and the advisor have a strictly higher payoff than*

[38] If $p_F > p_T$ then $x_{(+,-)} > x_\Phi$. Hence, it is possible that the DM prefers rejection at the prior, but acceptance following mixed outcomes. To generate the insurance effect, the advisor must prefer acceptance if and only if two tests were run and both show positive outcomes.

[39] I call the result below a corollary because of its correspondence to Corollary 1, but some additional analysis is required for its proof given Theorem 2.

*in the unique equilibrium under observable testing, i.e. for any* $N > 2$ *and* $(\lambda_A, \lambda_{DM})$, *if* $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$ *then there exists* $(\sigma, \mu) \in E$ *such that* $\pi_i(\sigma, \mu; \Gamma) > \pi_i(\sigma^o, \mu^o; \Gamma^o)$ *for* $i = A, DM$.

# 4 Delegation and Commitment

In this section, I first compare the DM's expected payoffs under hidden and observable testing to her expected payoff when she delegates decision-making authority to the advisor. In addition, I study how the comparison between hidden and observable testing changes when either party has the power to commit to a strategy ex ante. This analysis helps to further illustrate differences between the skepticism and the insurance effect. Throughout I allow for $N > 2$ and assume that the test has accuracy $p$ independent of the state.

When the DM delegates decision-making authority, the game is described by $\Gamma^o$ except that the advisor chooses both $\sigma_A^o$ and $\sigma_{DM}^o$. Denote the game with delegation by $\Gamma^d$ and its equilibrium by $\left(\sigma^d, \mu^d\right)$.

**Proposition 4 (Delegation)** *Suppose the DM could delegate decision-making authority to the advisor.*

1. *For any* $N > 2$ *and* $(\lambda_A, \lambda_{DM})$, *under hidden testing, delegating decision-making authority never makes the DM strictly better off, i.e. for any* $N > 2$ *and* $(\lambda_A, \lambda_{DM})$, $\pi_{DM}\left(\sigma^d, \mu^d; \Gamma^d\right) \leq \pi_{DM}(\sigma, \mu; \Gamma)$ *for all* $(\sigma, \mu) \in E$.

2. *For any* $N > 2$ *and* $(\lambda_A, \lambda_{DM})$, *under observable testing, delegating decision-making authority*

   (a) *makes the DM strictly better off if* $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$,
   *i.e.* $\pi_{DM}\left(\sigma^d, \mu^d; \Gamma^d\right) > \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ *if* $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$, *and*

   (b) *makes the DM strictly worse off if* $(\lambda_A, \lambda_{DM}) \in W_{\mathcal{I}} \cup \mathcal{S}$,
   *i.e.* $\pi_{DM}\left(\sigma^d, \mu^d; \Gamma^d\right) < \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ *if* $(\lambda_A, \lambda_{DM}) \in W_{\mathcal{I}} \cup \mathcal{S}$.

A disadvantage of delegating decision-making authority for the DM is that, for some test outcomes and some preferences of the advisor, the advisor chooses a different action than what the DM finds optimal. However, an advantage of delegating decision-making authority can exist if the advisor has incentives to conduct more tests than he would without delegation.

When testing is hidden, the DM has no upside from delegating decision-making. When the advisor is more reluctant than the DM, then even if the DM is in charge of the final decision, she always follows the advisor's recommendation. Therefore, the advisor's incentives to test are the same as if he was in charge of the final decision. When the advisor is more enthusiastic than the DM, then the advisor has even less incentive to test if he is in charge of the final decision rather than the DM. Therefore, when testing is hidden, the DM weakly prefers communication to delegating the right to take the final decision.

When testing is observable, then in the insurance region $\mathcal{I}$, delegating decision-making authority to the advisor gives him an incentive to test. This is because delegating ensures the advisor that the final decision is made in line with his interest, just as under hidden testing.[40] By contrast, in region $\mathcal{S}$ delegating reduces the advisor's incentives to test and the DM accepts based on weaker evidence. In region $W_{\mathcal{I}}$, delegating increases the advisor's incentive to test, but this works to the DM's disadvantage. This is because these additional tests are pivotal for the advisor's but not for the DM's preferred choice, yet in equilibrium the DM's choice depends on their outcomes.

In what follows, I explore what would change if the DM could commit ex ante to what action she will take for any evidence presented to her. Suppose the DM could commit to $\sigma_{DM}$. Under observable testing, denote the unique equilibrium by $(\sigma_C^o, \mu_C^o)$, and under hidden testing, denote the set of Pareto-undominated equilibria by $E_C$.

**Lemma 8 (DM Commitment)** *For any $N \geq 2$ and $(\lambda_A, \lambda_{DM})$, if the DM has the power to commit ex ante to her actions contingent on (reported) outcomes, then she achieves her first-best expected payoff, whether testing is hidden or observable, i.e. for any $(\sigma_C, \mu_C) \in E_C$:*

$$\pi_{DM}\left(\sigma_C^o, \mu_C^o\right) = \pi_{DM}\left(\sigma_C, \mu_C\right) = \pi_{DM}^{FB}. \tag{30}$$

This shows that, in some circumstances, the DM can use hidden testing to circumvent a commitment problem.

**Proposition 5 (Commitment vs. Hidden Testing)** *Suppose testing is observable and the DM does not have the power to commit to actions contingent on outcomes.*

1. *For any $N \geq 2$, if $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$ then the DM's marginal benefit from allowing testing to be hidden is equally high as her marginal benefit from gaining commitment power,*

[40] With $N > 2$ periods, delegation is not necessarily beneficial only in the insurance region. Without delegation it is possible that the advisor stops testing once the evidence is just strong enough for the DM to accept, even if it is possible that the remaining outcomes could lead both players to prefer rejection. The reason the advisor stops testing is that it is too likely that the remaining outcomes provide only weak evidence against the hypothesis being true, causing the DM to reject when the advisor would prefer to accept. In this situation, delegating decision-making has an upside for the DM because the advisor would then continue testing as he can ensure that rejection is chosen if and only if he prefers rejection given the evidence collected.

*i.e. for any* $(\sigma, \mu) \in E$

$$[\pi_{DM}(\sigma, \mu; \Gamma) - \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)] = [\pi_{DM}(\sigma_C^o, \mu_C^o) - \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)] > 0. \tag{31}$$

2. *For any* $N > 2$*, if* $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$ *then the DM's marginal benefit from allowing testing to be hidden is weakly smaller than her marginal benefit from gaining commitment power: for any* $(\sigma, \mu) \in E$*,*

$$[\pi_{DM}(\sigma, \mu; \Gamma) - \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)] \leq [\pi_{DM}(\sigma_C^o, \mu_C^o) - \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)] > 0. \tag{32}$$

Both hidden testing and commitment can be used to generate the same expected payoffs if $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$. For example, if $N = 2$, the DM accepts if she knows one test was run and the outcome was positive, but ideally she would like to make her decision dependent on whether or not an additional test also shows a positive outcome. To generate the skepticism effect under observable testing, the DM needs to commit to act less in the advisor's interest and reject based on a single positive outcome. By contrast, under hidden testing, this threat is credible even without commitment power, because the DM has reason to suspect that the advisor has omitted some negative evidence.

Commitment can be used to generate the same expected payoffs as hidden testing if $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$, but it allows the DM to do even better. Under observable testing, the advisor tests if the DM commits to act in his interest for any outcome realization. Under hidden testing, such commitment is not necessary, because the advisor has the possibility to omit outcomes and thereby prevent that acceptance is chosen when he prefers rejection. However, the DM can do better and achieve her first-best expected payoff if she commits to accept unless the advisor shows her evidence that leads her to reject, i.e. she commits to act less in the advisor's interest.

Next, suppose the advisor had the power to commit ex ante to a testing and a disclosure strategy. Denote the unique equilibrium under observable testing by $(\sigma_{CA}^o, \mu_{CA}^o)$ and the set of Pareto-undominated equilibria under hidden testing by $E_{CA}$.

**Proposition 6 (Advisor Commitment)** *Suppose the advisor has the power to commit to a testing and a disclosure strategy ex ante.*

1. *For any* $N \geq 2$*, if preferences lie in the skepticism region, then there exists a Pareto-undominated equilibrium under hidden testing in which the DM has a weakly lower expected payoff than in the unique equilibrium under observable testing, i.e. if* $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$*, there exists some* $(\sigma_{CA}, \mu_{CA}) \in E_{CA}$ *such that* $\pi_{DM}(\sigma_{CA}, \mu_{CA}) \leq \pi_{DM}(\sigma_{CA}^o, \mu_{CA}^o)$*.*

2. *For any $N > 2$, if preferences lie in the insurance region, the DM has a strictly higher payoff in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing, i.e. for any $(\sigma_{CA}, \mu_{CA}) \in E_{CA}$, if $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$, then $\pi_{DM}(\sigma_{CA}, \mu_{CA}) > \pi_{DM}(\sigma^o_{CA}, \mu^o_{CA})$.*

If the advisor has a choice to commit to disclosing all outcomes, the skepticism effect no longer exists. If $N = 2$ and the advisor has no commitment power, the DM benefits from hidden testing for preferences in the skepticism region $S$, because she can credibly reject based on the report of a single positive outcome. By contrast, if the advisor has commitment power then he would optimally commit to reporting a single positive outcome if and only if this outcome was obtained on the first test. As a consequence, it is no longer possible for the DM to credibly reject a single positive outcome, just as in the case of observable testing. By contrast, in the insurance region, the comparison between hidden and observable testing is unaffected by whether the advisor has commitment power or not.

# 5 Robustness Checks

This section revisits the key comparison between expected payoffs under hidden and observable testing when i) the advisor has to commit to the number of tests in advance and ii) the horizon is infinite and the advisor incurs a non-infinitesimal cost per test. I again assume the test has accuracy $p$ independent of the state. Proofs can be found in the supplementary appendix Sections B.4 and B.5.

In some circumstances, the advisor may have to decide how much evidence to acquire before he learns about any of the outcomes. If $N = 2$, the results in Theorem 1 still apply. However, if $N > 2$ then there is not always a region of preference parameters for which the skepticism effect exists.

**Proposition 7 (Simultaneous Testing)** *Suppose $N > 2$ and the advisor has to commit to the number of tests ex ante.*

1. *The skepticism effect does not exist for all $N$, i.e. for some $N$ there exists no $(\lambda_A, \lambda_{DM})$ such that $\pi_{DM}(\sigma, \mu; \Gamma) > \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ for all $(\sigma, \mu) \in E$.*

2. *The insurance effect exists for all $N$, i.e. for any $N$, there exists some set $\mathcal{I}' \subset \mathbb{R}^2$ such that if $(\lambda_A, \lambda_{DM}) \in \mathcal{I}'$ then $\pi_{DM}(\sigma, \mu; \Gamma) > \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ for any $(\sigma, \mu) \in E$.*

Suppose the advisor is more enthusiastic and testing is observable. When the advisor tests sequentially, he can choose to continue testing or not conditional on the history of realized outcomes. Therefore, it is possible that the advisor strategically stops testing when the existing evidence leads the DM to accept. By contrast, suppose the advisor has to choose the number of tests in advance. When the advisor chooses whether to run $n$ or $n+1$ tests it is unclear whether the DM were to accept or reject after $n$ tests. In expectation, it is more likely that further test outcomes lead the DM to accept when he would have otherwise rejected rather than lead the DM to reject when he would have otherwise accepted. Therefore, it is possible that the advisor runs all $N$ tests for any configuration of preference parameters under observable testing and, hence, the DM cannot be strictly better off under hidden testing. This reasoning does not apply if the advisor is more reluctant. When the advisor tests sequentially, it is possible that the advisor does not start testing. This is because the chances that the test outcomes lead the two players to disagree are too high. These same concerns are present when the advisor has to choose the number of tests in advance.

In some circumstances, the advisor may not face resource constraints, i.e. the horizon may be infinite.[41]

**Proposition 8 (Infinite Horizon)** *Suppose $N \to \infty$ and the advisor incurs a constant non-infinitesimal cost per test.*

1. *There exist parameter combinations for which the skepticism effect exists.*

2. *There exists no parameter combination for which the insurance effect exists.*

The reason for why the skepticism effect exists is similar to the one discussed in the preceding analysis with a finite horizon and an infinitesimal cost of testing. Suppose $\lambda_A = 0$, i.e. the advisor weakly prefers the DM to accept independent of the state. Then under observable testing, the advisor stops testing for one of two reasons. Either he has found enough evidence to convince the DM to accept or he gives up because he expects that it would be too costly to convince the DM to accept given his evidence. As the DM's threshold of acceptance rises, it becomes relatively more likely that the advisor gives up when the hypothesis is false relative to when it is true and, therefore, his report becomes more informative.[42]

However, an insurance effect does not exist in this alternative model. The reason is that the advisor's optimal testing strategy is stationary and depends only on the current posterior belief. A necessary condition for the insurance effect to exist is that, under hidden testing, the advisor must find it worthwhile to keep testing at beliefs which lie between the

[41] With an infinitesimal cost of testing, analyzing the conflict of interest between the advisor and the DM would cease to be interesting, since they agree on the optimal action under certainty.

[42] Felgenhauer and Schulte (2014) study an infinite-horizon model of hidden testing with a constant cost of testing and an advisor with state-independent preferences. They show that an increase in the DM's threshold for acceptance improves how informative it is that the advisor meets the threshold.

prior belief and the belief at which he just prefers acceptance. Otherwise, the advisor would never find evidence that leads him to prefer acceptance. However, if the advisor optimally keeps testing in this range of beliefs under hidden testing, then he must also optimally keep testing under observable testing. This is because under observable testing stopping is an even less attractive option for the advisor, as it can result in the DM acting against his interest. Therefore, the insurance that the DM acts in the advisor's interest under hidden testing does not lead to additional information acquisition.

For the insurance effect to exist, the cost per test must be increasing in the number of tests, e.g. a finite horizon can be interpreted as an extreme case of a convex cost of testing. The reason is that when costs are convex, the advisor's optimal testing strategy depends not only on his posterior belief, but also on how much time has passed. Therefore, it might be that the posterior belief is such that the DM accepts when the advisor would prefer to reject, yet the advisor stops testing because the cost has become too high (or he has run out of time). To avoid finding himself in such a situation, the advisor may instead stop testing sooner at a posterior belief at which both players agree that rejecting is optimal. By contrast, under hidden testing, since the DM always acts in the advisor's interest, the consequences of stopping at certain histories of outcomes are not as negative as under observable testing and, hence, the advisor may acquire more information than under observable testing.

In many situations, it is reasonable to assume that the cost of testing is convex, e.g. a pharmaceutical company may find it increasingly difficult to recruit subjects for their trials the more trials they run (Kolata (2017)), or that overall resources are limited, e.g. the budget for development of a drug is limited. Similar consequences would arise if with some probability the final decision has to be taken in a given period and this probability increases in the number of periods.

# 6 Conclusion

In many situations, a decision maker has to rely on an advisor to become better informed about certain aspects on which her optimal choice depends. Since the decision maker and the advisor's interests are not always aligned, a natural question is whether or not the decision maker should monitor more closely how the advisor goes about collecting information. This paper shows that making the collection of information transparent can have adverse consequences for decision making. On the one hand, transparency ensures that an advisor cannot strategically omit findings. On the other hand, transparency can discourage information acquisition, even if the advisor and the decision maker agree on the optimal action under certainty.

The paper has distinguished between two effects which cause the DM to be strictly better off when testing is hidden: the skepticism effect, which can arise when the advisor is more enthusiastic about accepting than the DM, and the insurance effect, which can arise when the advisor is more reluctant about accepting than the DM. Both effects cause the advisor to become better informed when testing is hidden rather than observable. The insurance effect is novel to this paper. It shows that, when testing is hidden, the advisor does not face the risk that additional evidence leads the DM to act against his interest. Therefore, the advisor has an incentive to explore whether or not additional evidence leads both to agree on a different action choice. By contrast, the skepticism effect shows that, when testing is hidden, the DM can credibly raise her threshold for acceptance, because she has reason to suspect that the advisor is hiding contradicting evidence.

These effects are only present when preferences are sufficiently misaligned. When preferences are not sufficiently misaligned, the DM may be worse off under hidden than observable testing. Although the advisor runs more tests when testing is hidden, these additional tests yield no valuable information for the DM and increase the advisor's scope for manipulation through selective disclosure. From this analysis, I conclude that the DM is weakly better off under hidden than observable testing if preferences are sufficiently misaligned.

The insurance and skepticism effect have different implications for the advisor's payoff. While the insurance effect also causes the advisor's expected payoff to increase, the skepticism effect causes the advisor's expected payoff to decrease. This implies that only the insurance effect can cause hidden testing to be a Pareto-improvement over observable testing.

While the current paper analyzes the interaction between an individual DM and advisor, there are other interesting consequences of transparency when there are several decision makers using the evidence as a basis for their choice or several advisors supplying evidence. I leave this for future research.

# A Appendix

## Contents



### A.1 Preliminaries

Let $\Phi$ denote the set of ordered lists of Nature's $N$ draws, i.e.

$$\Phi \equiv \{(s_1, ..., s_N)\,|s_i \in \{-,+\}\,, i \in \{1, ..., N\}\}\,. \tag{33}$$

**Definition 1 (Equilibrium Acceptance Set)** *Given an equilibrium* $(\sigma, \mu)$ *of* $\Gamma$*, there exists a function* $f_{(\sigma,\mu;\Gamma)} : \Phi \to \{accept, reject\}$ *such that if Nature's draws were fixed to be* $\phi$ *ex ante then the DM chooses* $f_{(\sigma,\mu;\Gamma)}(\phi)$ *in equilibrium. Then the equilibrium acceptance set is*

$$\Phi_{(\sigma,\mu;\Gamma)} \equiv \left\{\phi \in \Phi | f_{(\sigma,\mu;\Gamma)}(\phi) = accept\right\}. \tag{34}$$

*Similarly, given an equilibrium* $(\sigma^o, \mu^o)$ *of* $\Gamma^o$*, the equilibrium acceptance set is*

$$\Phi_{(\sigma^o,\mu^o;\Gamma^o)} \equiv \left\{\phi \in \Phi | f_{(\sigma^o,\mu^o;\Gamma^o)}(\phi) = accept\right\}. \tag{35}$$

**Remark 1** $\Phi_{(\sigma,\mu;\Gamma)}$ *is sufficient to characterize the expected payoff vector* $\Pi(\sigma, \mu; \Gamma)$ *in equilibrium* $(\sigma, \mu)$ *of* $\Gamma$*, and* $\Phi_{(\sigma^o,\mu^o;\Gamma^o)}$ *is sufficient to characterize the expected payoff vector* $\Pi(\sigma^o, \mu^o; \Gamma^o)$ *in equilibrium* $(\sigma^o, \mu^o)$ *of* $\Gamma^o$*.*

Under hidden testing, given $f_{(\sigma,\mu;\Gamma)}(\phi)$, the expected payoff of player $i = A, DM$ in equilibrium $(\sigma, \mu)$ of $\Gamma$ is

$$\begin{aligned}\pi_i(\sigma, \mu; \Gamma) =& \sum_{\omega\in\Omega} Pr(\omega) \sum_{\phi\in\Phi} u_i\left(f_{(\sigma,\mu;\Gamma)}(\phi), \omega\right) Pr(\phi|\omega) \\ =& -\lambda_i Pr(false) \sum_{\phi\in\Phi} 1_{f_{(\sigma,\mu;\Gamma)}(\phi)=accept} Pr(\phi|false) \\ & - Pr(true) \sum_{\phi\in\Phi} 1_{f_{(\sigma,\mu;\Gamma)}(\phi)=reject} Pr(\phi|true)\end{aligned} \tag{36}$$

and similarly for observable testing.

**Definition 2 (First-best Acceptance Set)** *For* $i = A, DM$*, there exists a mapping* $g_i : \Phi \to \{accept, reject\}$ *such that* $g_i(\phi) = accept$ *if and only if*

$$\sum_{\omega\in\Omega} u_i(accept, \omega) Pr(\omega|\phi) \geq \sum_{\omega\in\Omega} u_i(reject, \omega) Pr(\omega|\phi). \tag{37}$$

*Then the first-best acceptance set is*

$$\Phi_i^{FB} \equiv \{\phi \in \Phi | g_i(\phi) = accept\}. \tag{38}$$

*Player i's first-best expected payoff is*

$$\pi_i^{FB} \equiv \sum_{\omega\in\Omega} Pr(\omega) \sum_{\phi\in\Phi} u_i(g_i(\phi), \omega) Pr(\phi|\omega). \tag{39}$$

**Remark 2** *If the advisor is more enthusiastic, i.e.* $\lambda_A \leq \lambda_{DM}$*, then* $\Phi_{DM}^{FB} \subseteq \Phi_A^{FB}$*. If the advisor is more reluctant, i.e.* $\lambda_A \geq \lambda_{DM}$*, then* $\Phi_{DM}^{FB} \supseteq \Phi_A^{FB}$*.*

By (5), $\phi \in \Phi_i^{FB}$ if and only if $\lambda_i \leq \frac{Pr(true|\phi)}{Pr(false|\phi)}$ for any $\phi \in \Phi$. If $\lambda_A \leq \lambda_{DM}$ and $\lambda_{DM} \leq \frac{Pr(true|\phi)}{Pr(false|\phi)}$ then $\lambda_A \leq \frac{Pr(true|\phi)}{Pr(false|\phi)}$ for any $\phi \in \Phi$. If $\lambda_A \geq \lambda_{DM}$ and $\lambda_A \leq \frac{Pr(true|\phi)}{Pr(false|\phi)}$ then $\lambda_{DM} \leq \frac{Pr(true|\phi)}{Pr(false|\phi)}$ for any $\phi \in \Phi$.

**Lemma 9 (Acceptance Sets and Payoffs)** *Consider any two equilibria characterized by acceptance sets $\Phi', \Phi''$ where $\Phi' \subset \Phi''$. If $\Phi' \subset \Phi'' \subset \Phi_i^{FB}$ then player $i$ is worse off in an equilibrium characterized by $\Phi'$ than in an equilibrium characterized by $\Phi''$. If $\Phi_i^{FB} \subset \Phi' \subset \Phi''$ then player $i$ is better off in an equilibrium characterized by $\Phi'$ than in an equilibrium characterized by $\Phi''$.*

Proof: Consider any $\hat{\phi}$ such that $\hat{\phi} \in \Phi''$ and $\hat{\phi} \notin \Phi'$. Suppose $\Phi' \subset \Phi'' \subset \Phi_i^{FB}$. Since $\Phi'' \subset \Phi_i^{FB}$, $\hat{\phi} \in \Phi_i^{FB}$ and, hence, player $i$'s payoff from acceptance is higher than from rejection at any $\hat{\phi}$ by Definition 2. Suppose $\Phi_i^{FB} \subset \Phi' \subset \Phi''$. Since $\Phi_i^{FB} \subset \Phi'$, $\hat{\phi} \notin \Phi_i^{FB}$ and, hence, player $i$'s payoff from rejection is higher than from acceptance at any $\hat{\phi}$ by Definition 2.

## A.2 Lemma 1 [Equilibrium Existence]

Under observable testing, the system of beliefs $\mu^o$ is consistent if and only if it is derived using Bayes' rule given the testing technology. Hence, $\mu^o$ is independent of strategy profile $\sigma^o$. Given $\mu^o$, there must exist $\sigma_{DM}^o$ which selects $a \in \{accept, reject\}$ at any $h_N$ such that $\pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ is maximized, independent of $\sigma_A^o$. In addition, given $\mu^o$, for any $\sigma_{DM}^o$, there must exist $\sigma_A^o$ which selects $\tau = 0$ or $\tau = 1$ at any $h_n$ for $n = 0, 1, ...N-1$ such that $\pi_A(\sigma^o, \mu^o; \Gamma^o)$ is maximized. For the case of hidden testing, see Lemma 11 for $N = 2$ and Lemma 7 for $N > 2$.

## A.3 Lemma 2 [Pareto-undominated Equilibria]

Suppose testing is hidden. Let $l_\phi \equiv \frac{Pr(true|\phi)}{Pr(false|\phi)}$ denote the posterior likelihood ratio conditional on list $\phi$ of Nature's outcome draws and let $M(\phi)$ denote the set of feasible messages if $h_N = \phi$, where $\phi \in \Phi$ and $\Phi$ is defined by (33).

**Claim 1:** *Assume the advisor runs all $N$ tests, i.e. $h_N = \phi$. Consider $\phi', \phi'' \in \Phi$, where $\phi'$ contains weakly fewer positive outcomes than $\phi''$. If there exists a message $m' \in M(\phi')$ such that $\sigma_{DM}(m') = accept$, there also exists a message $m'' \in M(\phi'')$ such that $\sigma_{DM}(m'') = accept$.*

If $m'' \in M(\phi')$ then choose $m'' = m'$. If $m' \notin M(\phi'')$, choose $m''$ to contain all outcomes in $\phi''$. Due to verifiability, $m''$ can only be sent following any $\phi$ with as many positive outcomes as $\phi''$. In addition, since $m' \notin M(\phi'')$ it must be that $m'$ contains strictly more negative outcomes than $\phi''$, i.e. it cannot be sent following any $\phi$ with weakly more positive outcomes than $\phi''$. In addition, for any $\phi_1, \phi_2 \in \Phi$ where $\phi_1$ contains weakly fewer positive outcomes than $\phi_2$, $l_{\phi_1} \leq l_{\phi_2}$ given the testing technology. Hence, in equilibrium, consistent beliefs must satisfy $\mu_{DM}(m'') \geq \mu_{DM}(m')$. By (5), the DM accepts if and only if $\mu_{DM} \geq \frac{\lambda_{DM}}{1-\lambda_{DM}}$. Hence, if $\mu_{DM}\left(m^{'}\right) \geq \frac{\lambda_{DM}}{1-\lambda_{DM}}$ then also $\mu_{DM}\left(m^{''}\right) \geq \frac{\lambda_{DM}}{1-\lambda_{DM}}$.

**Claim 2:** *Consider the function $f_{(\sigma,\mu;\Gamma)}$ defined in Definition 1. In any equilibrium $(\sigma, \mu)$ of $\Gamma$, the function $f_{(\sigma,\mu;\Gamma)}$ satisfies the following single-crossing property. Consider $\phi', \phi'' \in \Phi$, where $\phi'$ contains weakly fewer positive outcomes than $\phi''$. If $f(\phi') = accept$, then $f(\phi'') = accept$.*

Note that $l_{\phi'} \leq l_{\phi''}$. If a player $i = A, DM$ prefers acceptance at $\phi'$, he also does at $\phi'$, i.e. by (5), if $\lambda_i \leq l_{\phi'}$ then $\lambda_i \leq l_{\phi''}$. Suppose $\lambda_A \leq \lambda_{DM}$. If the advisor prefers rejection, the DM also does, i.e. if $\lambda_A > l_\phi$ then $\lambda_{DM} > l_\phi$ for any $\phi \in \Phi$. Assume the advisor runs all $N$ tests, i.e. $h_N = \phi$. If $f(\phi') = accept$, then it must be that $\lambda_A \leq l_{\phi'}$. This is because, if $\lambda_A > l_{\phi'}$ and $f(\phi') = accept$, then the advisor can deviate and disclose all outcomes and the DM rejects, since $\lambda_{DM} > l_{\phi'}$. If $f(\phi') = accept$ and $\lambda_A \leq l_{\phi'}$, then $f(\phi'') = accept$ by Claim 1.

Next, suppose $\lambda_A > \lambda_{DM}$. Note that if the advisor prefers acceptance, the DM also does, i.e. if $\lambda_A \leq l_\phi$ then $\lambda_{DM} \leq l_\phi$ for any $\phi \in \Phi$. Assume the advisor runs all $N$ tests, i.e. $h_N = \phi$. If $f(\phi') = accept$ and $\lambda_A \leq l_{\phi'}$, then $f(\phi'') = accept$ since $\lambda_{DM} < \lambda_A \leq l_\phi$. If $f(\phi') = accept$ and $\lambda_A > l_{\phi'}$, then it must be that there exists no $m' \in M(\phi')$ such that $\sigma_{DM}(m') = reject$. By the same reasoning as used in the proof of Claim 1, if there exists no $m' \in M(\phi')$ such that $\sigma_{DM}(m') = reject$ then there also exists no $m'' \in M(\phi'')$ such that $\sigma_{DM}(m'') = reject$. Hence, if $f(\phi') = accept$ and $\lambda_A > l_{\phi'}$, then $f(\phi'') = accept$.

Finally, the argument also holds when $\sigma_A^T$ is chosen optimally. The advisor stops if and only if he is indifferent between running all tests and stopping. The advisor's payoff depends on the state and the DM's action. Therefore, there are only two situations in which the advisor stops. First, he stops if the remaining outcomes cannot be pivotal to the DM's action given $\sigma_A^M$. This is because, holding the DM's choice fixed, the advisor's expected payoff is linear in the belief that the hypothesis is true and the expected posterior belief is equal to the prior belief. In addition, the advisor stops in period $n$ if the remaining outcomes cannot be pivotal to his preferred final decision and the DM acts in his interest given $\sigma_A^M$. This is because $\sigma_{DM}\left(\sigma_A^M(h_N)\right)$ is constant for all possible $h_N$ given $h_n$ and $\sigma_{DM}\left(\sigma_A^M(h_N)\right) = \sigma_{DM}\left(\sigma_A^M(h_n)\right)$.

**Claim 3:** *Lemma 2 holds.*

By Remark 1, any DM-preferred equilibrium $(\overline{\sigma}, \overline{\mu})$ must be characterized by a unique equilibrium acceptance set $\Phi_{(\overline{\sigma},\overline{\mu};\Gamma)}$ (see Definition 1). In addition, given that $\Phi_{(\overline{\sigma},\overline{\mu})}$ is unique,

$\pi_A(\overline{\sigma},\overline{\mu};\Gamma)$ must be unique. By the same reasoning, any advisor-preferred equilibrium $\left(\underline{\sigma},\underline{\mu}\right)$ must have a unique $\pi_{DM}\left(\underline{\sigma},\underline{\mu};\Gamma\right)$.

By Claim 2, for any two equilibria characterized by $\Phi'$ and $\Phi''$, either $\Phi' \subseteq \Phi''$ or $\Phi' \supseteq \Phi''$. Consider the first-best acceptance sets $\Phi_A^{FB}$ and $\Phi_{DM}^{FB}$ (see Definition 2). If $\lambda_A \leq \lambda_{DM}$ then $\Phi_{DM}^{FB} \subseteq \Phi_A^{FB}$ by Remark 2. The following statements are implied by Lemma 9. Suppose there exists an equilibrium characterized by $\Phi'$ where $\Phi' \subset \Phi_{DM}^{FB} \subseteq \Phi_A^{FB}$. Then either $\Phi' = \Phi_{(\overline{\sigma},\overline{\mu};\Gamma)}$ or $\Phi'$ characterizes an equilibrium which is Pareto-dominated by the DM-preferred equilibrium. Suppose there exists an equilibrium characterized by $\Phi''$ where $\Phi_{DM}^{FB} \subseteq \Phi_A^{FB} \subset \Phi''$. Then either $\Phi'' = \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)}$ or $\Phi''$ characterizes an equilibrium which is Pareto-dominated by the advisor-preferred equilibrium. Furthermore, if $\Phi_{DM}^{FB} \subseteq \Phi_{(\overline{\sigma},\overline{\mu};\Gamma)}$, there cannot be any equilibrium characterized by $\Phi_1$ such that $\Phi_{DM}^{FB} \subseteq \Phi_1 \subset \Phi_{(\overline{\sigma},\overline{\mu};\Gamma)}$, because otherwise $\Phi_1 = \Phi_{(\overline{\sigma},\overline{\mu};\Gamma)}$. Similarly, if $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} \subseteq \Phi_A^{FB}$, there cannot be any equilibrium characterized by $\Phi_2$ such that $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} \subset \Phi_2 \subseteq \Phi_A^{FB}$, because otherwise $\Phi_2 = \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)}$. Therefore, any Pareto-undominated equilibrium $(\sigma,\mu)$ must satisfy $\Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} \subseteq \Phi_{(\sigma,\mu;\Gamma)} \subseteq \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)}$. Claim 3 follows by Lemma 9. An analogous argument applies for the case of $\lambda_A > \lambda_{DM}$.

## A.4 Equilibrium Characterization under Observable Testing

Lemma 10 below characterizes equilibrium under observable testing and contains the proof of Part 1 of Lemmas 3 - 5. It also states the equilibrium acceptance set $\Phi_{(\sigma^o,\mu^o;\Gamma^o)}$ (see Definition 1). Posterior likelihood ratios are denoted by $l_\varphi$ with $\varphi \in \Phi$ as defined in equations (7)-(11) and footnote 23. The advisor's strategy is denoted by

$$\sigma_A^o \equiv \left(\sigma_A^o\left(h_0\right), \sigma_A^o\left(h_1 = (+)\right), \sigma_A^o\left(h_1 = (-)\right), \sigma_A^o\left(h_1 = (\emptyset)\right)\right) \in [0,1]^4 . \tag{40}$$

It is helpful to divide the range of $\lambda_{DM} \in \left(\frac{q}{1-q}, \infty\right)$ into four regions such that $\sigma_{DM}^o$ is unchanged within each region, as indicated along the vertical axis in Figure 5.[43]

**Lemma 10 (Equilibrium Characterization Observable Testing)** *Consider $\Gamma^o$.*
*In **Region 1**, i.e. if*

$$l_\Phi < \lambda_{DM} \leq l_{(+,-)}, \tag{41}$$

*the DM accepts if and only if $\widetilde{h} \in \{\{+,+\},\{+\},\{+,-\}\}$. The advisor's strategy is*

$$\sigma_A^o = \begin{cases} (1,0,1,1) & \text{if } \lambda_A < l_{(+,-)} \\ (1,0,0,1) & \text{if } l_{(+,-)} \leq \lambda_A < l_{(+,\cdot)} \\ (0,0,0,0) & \text{if } l_{(+,\cdot)} \leq \lambda_A . \end{cases} \tag{42}$$

[43] Region 1 is non-empty if and only if $p_F > p_T$.

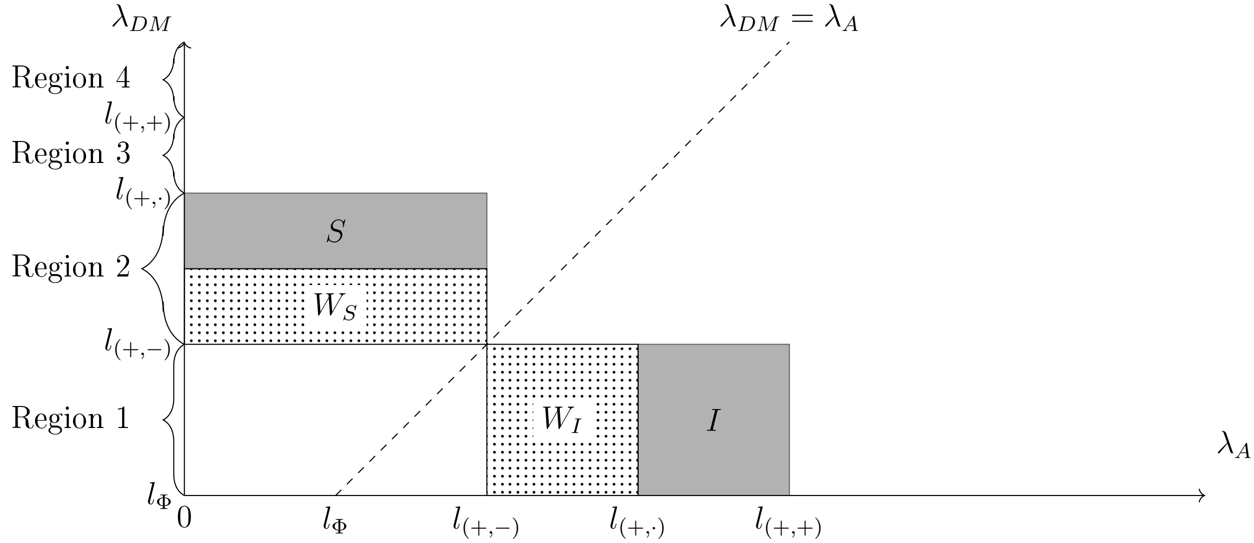


Figure 5: This Figure indicates along the vertical axis the regions in which the DM's optimal decision rule under observable testing is unchanged, in $(\lambda_A, \lambda_{DM})$ −space given $p_T < p_F$.

*The equilibrium acceptance set is given by*

$$\Phi_{(\sigma^o,\mu^o;\Gamma^o)} = \begin{cases} \{(-,+),(+,-),(+,+)\} & \text{if } \lambda_A < l_{(+,-)} \\ \{(+,+),(+,-)\} & \text{if } l_{(+,-)} \leq \lambda_A < l_{(+,\cdot)} \\ \emptyset & \text{if } l_{(+,\cdot)} \leq \lambda_A. \end{cases} \tag{43}$$

*In **Region 2**, i.e. if*

$$max\left\{l_\Phi, l_{(+,-)}\right\} < \lambda_{DM} \leq l_{(+,\cdot)}, \tag{44}$$

*the DM accepts if and only if* $\widetilde{h} \in \{\{+,+\},\{+\}\}$. *The advisor's strategy is*

$$\sigma_A^o = \begin{cases} (1,0,0,1) & \text{if } \lambda_A \leq l_{(+,-)} \\ (1,1,0,1) & \text{if } l_{(+,-)} < \lambda_A < l_{(+,\cdot)} \\ (1,1,0,0) & \text{if } l_{(+,\cdot)} \leq \lambda_A < l_{(+,+)} \\ (0,1,0,0) & \text{if } l_{(+,+)} \leq \lambda_A. \end{cases} \tag{45}$$

*The equilibrium acceptance set is given by*

$$\Phi_{(\sigma^o,\mu^o;\Gamma^o)} = \begin{cases} \{(+,-),(+,+)\} & \text{if } \lambda_A < l_{(+,-)} \\ \{(+,+)\} & \text{if } l_{(+,-)} \leq \lambda_A < l_{(+,+)} \\ \emptyset & \text{if } l_{(+,+)} \leq \lambda_A. \end{cases} \tag{46}$$

*In **Region 3**, i.e. if*

$$l_{(+,\cdot)} < \lambda_{DM} \leq l_{(+,+)}, \tag{47}$$

*the DM accepts if and only if* $\widetilde{h}\{+,+\}$. *The advisor's strategy is*

$$\sigma_A^o = \begin{cases} (1,1,0,0) & \text{if } \lambda_A < l_{(+,+)} \\ (0,0,0,0) & \text{if } l_{(+,+)} \leq \lambda_A. \end{cases} \tag{48}$$

*The equilibrium acceptance set is given by*

$$\Phi_{(\sigma^o,\mu^o;\Gamma^o)} = \begin{cases} \{(+,+)\} & \text{if } \lambda_A < l_{(+,+)} \\ \emptyset & \text{if } l_{(+,+)} \leq \lambda_A. \end{cases} \tag{49}$$

*In **Region 4**, i.e. if*

$$\lambda_{DM} > l_{(+,+)}, \tag{50}$$

*then the DM never accepts. The advisor's strategy is* $\sigma_A^o = (0,0,0,0)$. *The equilibrium acceptance set is given by* $\Phi_{(\sigma^o,\mu^o;\Gamma^o)} = \emptyset$.
*In any region, beliefs satisfy* $\frac{\mu^o(h_0)}{1-\mu^o(h_0)} = l_\Phi$, $\frac{\mu^o(h_2)}{1-\mu^o(h_2)} = l_{h_2}$ *and*

$$\frac{\mu^o(h_1)}{1-\mu^o(h_1)} = \begin{cases} l_{(+,\cdot)} & \text{if } h_1 = (+) \\ l_{(-,\cdot)} & \text{if } h_1 = (-) \\ l_\Phi & \text{if } h_1 = (\emptyset). \end{cases} \tag{51}$$

**Proof:** All equilibria derived below are sequential equilibria. I can propose a completely mixed strategy for the advisor as a function of some $\varepsilon \in (0,1)$ which allows each history of outcomes to be reached with positive probability and which converges to $\sigma_A^o$ as $\varepsilon \to 0$. Then as $\varepsilon \to 0$ the ratio of beliefs $\frac{\mu^o(h_n)}{1-\mu^o(h_n)}$ at each history $h_n$ for $n = 0,1,2$ are equal to the posterior likelihood ratios stated above, independent of $\sigma_A^o$.

**Region 1.** Suppose $p_F > p_T$ and $l_\Phi < \lambda_{DM} \leq l_{(+,-)}$. In period 2, given $h_1 = (+)$, the DM accepts independent of whether or not another test is run. Hence, the advisor stops testing.[44] Given $h_1 = (-)$, if the advisor does not test, the DM rejects. If the advisor tests, the DM accepts if and only if $h_2 = (-,+)$. The advisor tests if and only if

$$\begin{aligned} \mu^o(h_1 = (-)) &> \lambda_A [1 - \mu^o(h_1 = (-))](1-p_F) + \mu^o(h_1 = (-))(1-p_T) \\ \lambda_A &< \frac{\mu^o(h_1 = (-))\, p_T}{[1-\mu^o(h_1 = (-))](1-p_F)} = \frac{q p_T (1-p_T)}{(1-q)(1-p_F)\, p_F} = l_{(+,-)}. \end{aligned} \tag{52}$$

[44] Recall the assumption that the advisor prefers not to test whenever he is indifferent between testing or not and the DM accepts whenever she is indifferent between accepting and rejecting.

Given $h_1 = (\emptyset)$, if the advisor does not test, the DM rejects. If the advisor tests, the DM accepts if and only if $h_2 = (\emptyset, +)$. The advisor tests if and only if

$$\lambda_A < \frac{\mu^o(h_1 = (\emptyset))\, p_T}{[1 - \mu^o(h_1 = (\emptyset))](1 - p_F)} = \frac{q p_T}{(1 - q)(1 - p_F)} = l_{(+,\cdot)}. \tag{53}$$

In period 1, if the advisor does not test, the DM rejects. For $\lambda_A < l_{(+,-)}$, if he tests, he tests in period 2 if and only if $h_1 = (-)$ and the DM accepts if and only if $h_2 \in \{(+, \emptyset), (-, +)\}$. The advisor tests in period 1 since $l_{(+,-)} < l_{\neg(-,-)}$ and

$$\mu^o(h_0) > \lambda_A [1 - \mu^o(h_0)] \left(1 - p_F^2\right) + \mu^o(h_0)(1 - p_T)^2$$
$$\lambda_A < \frac{q\left(1 - (1 - p_T)^2\right)}{(1 - q)\left(1 - p_F^2\right)} = l_{\neg(-,-)}. \tag{54}$$

For $\lambda_A \geq l_{(+,-)}$, if he tests, then he stops testing in period 2 and the DM accepts if and only if $h_2 = (+, \emptyset)$. The advisor tests if and only if (53) holds.

**Region 2.** In period 2, given $h_1 = (-)$, the DM rejects independent of whether or not another test is run. Hence, the advisor stops testing. Given $h_1 = (+)$, if the advisor does not test, the DM accepts. If the advisor tests, the DM accepts if and only if $h_2 = (+, +)$. The advisor tests if and only if

$$\lambda_A [1 - \mu^o(h_1 = (+))] > \lambda_A [1 - \mu^o(h_1 = (+))](1 - p_F) + \mu^o(h_1 = (+))(1 - p_T)$$
$$\lambda_A > \frac{q p_T (1 - p_T)}{(1 - q)(1 - p_F) p_F} = l_{(+,-)}. \tag{55}$$

Given $h_1 = (\emptyset)$, if the advisor does not test, the DM rejects. If the advisor tests, the DM accepts if and only if $h_2 = (\emptyset, +)$. The advisor tests if and only if (53) holds. In period 1, if the advisor does not test, the DM rejects. For $\lambda_A \leq l_{(+,-)}$, if the advisor tests, he stops testing in period 2 and the DM accepts if and only if $h_2 = (+, \emptyset)$. The advisor tests since $\lambda_A \leq l_{(+,-)}$ implies that (53) holds. For $\lambda_A > l_{(+,-)}$, if the advisor tests, he tests in period 2 if and only if $h_1 = (+)$ and the DM accepts if and only if $h_2 = (+, +)$. The advisor tests if and only if

$$\mu^o(h_0) > \lambda_A (1 - \mu^o(h_0))(1 - p_F)^2 + \mu^o(h_0)\left(1 - p_T^2\right)$$
$$\lambda_A < \frac{q p_T^2}{(1 - q)(1 - p_F)^2} = l_{(+,+)}. \tag{56}$$

**Region 3.** Suppose $l_{(+,\cdot)} < \lambda_{DM} \leq l_{(+,+)}$. In period 2, given $h_1 = (-)$, the DM rejects independent of whether or not another test is run. Hence, the advisor stops testing. The

same is true given $h_1 = (\emptyset)$. Given $h_1 = (+)$, if the advisor does not test, the DM rejects. If the advisor tests, the DM accepts if and only if $h_2 = (+,+)$. The advisor tests if and only if (56) holds. In period 1, if the advisor does not test, the DM rejects. For $\lambda_A < l_{(+,+)}$, since the advisor tests at $h_1 = (+)$ he tests in period 1. For $\lambda_A \geq l_{(+,+)}$, if he tests, he stops testing in period 2 and the DM rejects and, therefore, the advisor does not test in period 1.
**Region 4.** Suppose $\lambda_{DM} > l_{(+,+)}$. Given that the DM never accepts at any $h_2$, the advisor never tests.

## A.5 Equilibrium Characterization under Hidden Testing

Lemma 11 below characterizes an advisor-preferred equilibrium under hidden testing and contains the proof of Part 2 of Lemmas 3 - 5. Lemma 12 below characterizes an DM-preferred equilibrium under hidden testing.
Let $\Phi_{(\sigma,\mu;\Gamma)}$ denote the equilibrium acceptance set of equilibrium $(\sigma,\mu)$ of $\Gamma$ (see Definition 1). Posterior likelihood ratios are denoted by $l_\varphi$ with $\varphi \in \Phi$ as defined in equations (7)-(11) and footnote 23. By (5), the DM's first-best acceptance set (see Definition 2) is given by

$$\Phi_{DM}^{FB} = \begin{cases} \{(+,+),(+,-),(-,+)\} & \text{if } l_\Phi < \lambda_{DM} \leq l_{(+,-)} \\ \{(+,+)\} & \text{if } max\left\{l_{(+,-)}, l_\Phi\right\} \leq \lambda_{DM} \leq l_{(+,+)} \\ \emptyset & \text{if } \lambda_{DM} \geq l_{(+,+)}. \end{cases} \tag{57}$$

Similarly, the advisor's first-best acceptance set is by

$$\Phi_{A}^{FB} = \begin{cases} \Phi & \text{if } \lambda_A \leq l_{(-,-)} \\ \{(+,+),(+,-),(-,+)\} & \text{if } l_{(-,-)} \leq \lambda_A \leq l_{(+,-)} \\ \{(+,+)\} & \text{if } l_{(+,-)} \leq \lambda_A \leq l_{(+,+)} \\ \emptyset & \text{if } \lambda_A \geq l_{(+,+)}. \end{cases} \tag{58}$$

### A.5.1 Advisor-preferred Equilibria

**Lemma 11 (Advisor-preferred Equilibrium Characterization Hidden Testing)**
*Consider $\Gamma$. $(\underline{\sigma},\underline{\mu})$ specified below is an advisor-preferred equilibrium. In all regions, $\underline{\mu}_A : \cup_{n=0}^N H_n \to [0,1]$ is equal to $\mu^o : \cup_{n=0}^N H_n \to [0,1]$ in Lemma 10.*
*In **Region 1**, i.e. $l_\Phi < \lambda_{DM} \leq l_{(+,-)}$.*

*(a) If $\lambda_A < l_{(+,-)}$, $\underline{\sigma}_{DM}(m) =$ accept if and only if $m = \{\{+\},\{+,-\},\{+,+\}\}$, $\underline{\sigma}_A$ and $\underline{\mu}_{DM}$ are as in $(\lambda_A,\lambda_{DM}) \in W_S$ (Region 2a). The equilibrium acceptance set is given*

by $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \{(-,+),(+,-),(+,+)\} = \Phi^{FB}_{DM}$.

(b) *If $l_{(+,-)} \leq \lambda_A < l_{(+,\cdot)}$ then $(\lambda_A, \lambda_{DM}) \in W_I$. Equilibrium is characterized in Part 2 of Lemma 5 and $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \{(+,+)\} = \Phi^{FB}_{A}$.*

(c) *If $l_{(+,\cdot)} \leq \lambda_A < l_{(+,+)}$ then $(\lambda_A, \lambda_{DM}) \in I$. Equilibrium is characterized in Lemma 3 and $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \{(+,+)\} = \Phi^{FB}_{A}$.*

*In **Region 2**, i.e. $\max\left\{l_\Phi, l_{(+,-)}\right\} < \lambda_{DM} \leq l_{(+,\cdot)}$.*

(a) *If $\max\left\{l_\Phi, l_{(+,-)}\right\} < \lambda_{DM} \leq l_{\neg(-,-)}$ and $\lambda_A < l_{(+,-)}$ then $(\lambda_A, \lambda_{DM}) \in W_S$. Equilibrium is characterized in Part 1 of Lemma 5 and $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \{(-,+),(+,-),(+,+)\}$.*

(b) *If $l_{\neg(-,-)} < \lambda_{DM} \leq l_{(+,\cdot)}$ and $\lambda_A < l_{(+,-)}$ then $(\lambda_A, \lambda_{DM}) \in S$. Equilibrium is characterized in Lemma 4 and $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \{(+,+)\} = \Phi^{FB}_{DM}$.*

(c) *If $l_{(+,-)} \leq \lambda_A < l_{(+,\cdot)}$, $\underline{\sigma}_{DM}(m) = accept$ if and only if $m \in \{\{+\},\{+,+\}\}$. $\underline{\sigma}_A$ and $\underline{\mu}_{DM}$ are as in when $(\lambda_A, \lambda_{DM}) \in W_I$ (Region 1b). The equilibrium acceptance set is given by $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \{(+,+)\} = \Phi^{FB}_{DM} = \Phi^{FB}_{A}$.*

(d) *If $l_{(+,\cdot)} \leq \lambda_A < l_{(+,+)}$, strategies are: $\underline{\sigma}_{DM}(m) = accept$ if and only if $m \in \{\{+\},\{+,+\}\}$. $\underline{\sigma}_A$ and $\underline{\mu}_{DM}$ are as in when $(\lambda_A, \lambda_{DM}) \in I$ (Region 1c). The equilibrium acceptance set is given by $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \{(+,+)\} = \Phi^{FB}_{DM} = \Phi^{FB}_{A}$.*

*In **Region 3**, i.e. suppose $l_{(+,\cdot)} < \lambda_{DM} \leq l_{(+,+)}$. $\underline{\sigma}_{DM}(m) = accept$ if and only if $m = \{+,+\}$. If $\lambda_A < l_{(+,+)}$, then $\underline{\sigma}^T_A = (1,1,0,0)$. If $\lambda_A < l_{(+,-)}$, $\underline{\sigma}^M_A$ and $\underline{\mu}_{DM}$ are as in Region 2b. If $l_{(+,-)} < \lambda_A < l_{(+,\cdot)}$, $\underline{\sigma}^M_A$ and $\underline{\mu}_{DM}$ are as in Region 2c. If $l_{(+,\cdot)} < \lambda_A < l_{(+,+)}$, $\underline{\sigma}^M_A$ and $\underline{\mu}_{DM}$ are as in Region 2d. The equilibrium acceptance set is given by $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \{(+,+)\} = \Phi^{FB}_{DM}$.*

*In **Region 4**, i.e. suppose $\lambda_{DM} > l_{(+,+)}$. Then strategies are: $\underline{\sigma}_{DM}(m) = reject$ for all $m$, $\underline{\sigma}^T_A = (0,0,0,0)$, $\underline{\sigma}^M_A(h_2) = \widetilde{h}$ for all $h_2$. Beliefs conditional on messages satisfy*

$$\frac{\underline{\mu}_{DM}(m)}{1-\underline{\mu}_{DM}(m)} = \begin{cases} l_{(+,\cdot)} & \text{if } m = \{+\} \\ l_\Phi & \text{if } m = \emptyset \\ l_{(-,\cdot)} & \text{if } m = \{-\} \\ l_m & \text{otherwise.} \end{cases} \tag{59}$$

*The equilibrium acceptance set is given by $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \emptyset = \Phi^{FB}_{DM}$.*

*In all regions, if $\lambda_A < l_{(+,+)}$, then $\underline{\sigma}^T_A = (0,0,0,0)$ and $\underline{\sigma}^M_A(h_2) = \emptyset$ for all $h_2$. $\underline{\mu}_{DM}$ satisfies (59). The equilibrium acceptance set is given by $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \emptyset = \Phi^{FB}_{A}$.*

**Proof:** I show that $\left(\underline{\sigma}, \underline{\mu}\right)$ constitutes an advisor-preferred PBE. For a proof that this is a sequential equilibrium, see Section B.1. In any region, $\underline{\sigma}_{DM}(m)$ is optimal by (5), since $\underline{\sigma}_{DM}(m) = accept$ if and only if $\frac{\underline{\mu}(m)}{1-\underline{\mu}(m)} \geq \lambda_{DM}$.

**Region 1a and Region 2a)**, i.e. $\lambda_A < l_{(+,-)}$ and $\lambda_{DM} < l_{\neg(-,-)}$. $\underline{\sigma}_A^M$ is optimal: if $\widetilde{h} \in \{\{-,-\},\{-\},\emptyset\}$, $\underline{\sigma}_{DM}(m) = reject$ for any feasible $m \in M\left(\widetilde{h}\right)$. Otherwise, $\underline{\sigma}_{DM}\left(\underline{\sigma}_A^M(h_2)\right) = accept$ and $\pi_A(accept) > \pi_A(reject)$ since $\lambda_A < l_{(+,-)}$. $\underline{\sigma}_A^T$ is optimal since (52), (53) and (54) hold. On-path beliefs are consistent: $m = \emptyset$ occurs if and only if $h_2 = \phi = (-,-)$, and $m = \{+\}$ occurs if and only if $h_2 \in \{(-,+),(+,\emptyset)\}$, i.e. if and only if $\phi \in \{(+,+),(+,-),(-,+)\}$. The equilibrium is an advisor-preferred equilibrium. If $l_{(-,-)} \leq \lambda_A < l_{(+,-)}$ then $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_A^{FB}$ and if $\lambda_A \leq l_{(-,-)}$, then $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} \cup (-,-) = \Phi_A^{FB} = \Phi$, but it cannot be part off an equilibrium that the DM always accepts since $l_\Phi < \lambda_{DM}$.

**Region 1b).** $\underline{\sigma}_A^M$ is optimal: if $h_2$ is such that $\widetilde{h} \in \{\{+,+\},\{+\}\}$, $\underline{\sigma}_{DM}\left(\underline{\sigma}_A^M(h_2)\right) = accept$ and $\pi_A(accept) > \pi_A(reject)$ since $\lambda_A < l_{(+,\cdot)}$. Otherwise, $\underline{\sigma}_{DM}\left(\underline{\sigma}_A^M(h_2)\right) = reject$ and $\pi_A(reject) > \pi_A(accept)$ since $\lambda_A > l_{(+,-)}$. $\underline{\sigma}_A^T$ is optimal since (53), (55) and (56) hold. On-path beliefs are consistent: $m = \{-\}$ occurs if and only if $h_2 \in \{(-,\emptyset),(+,-)\}$, i.e. if and only if $\phi \in \{(-,-),(+,-),(-,+)\}$ and $m = \{+,+\}$ occurs if and only if $h_2 = \phi = (+,+)$. The equilibrium is an advisor-preferred equilibrium since $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_A^{FB}$.

**Region 1c).** $\underline{\sigma}_A^M$ is optimal: If $h_2 = (+,+)$, $\underline{\sigma}_{DM}\left(\underline{\sigma}_A^M(h_2)\right) = accept$ and $\pi_A(accept) > \pi_A(reject)$ since $\lambda_A < l_{(+,+)}$. Otherwise, $\underline{\sigma}_{DM}\left(\underline{\sigma}_A^M(h_2)\right) = reject$ and $\pi_A(reject) > \pi_A(accept)$ since $\lambda_A \geq l_{(+,\cdot)}$. $\underline{\sigma}_A^T$ is optimal since (55) and (56) hold, but (53) does not hold. On-path beliefs are consistent by the same reason as in Region 1b). The equilibrium is advisor-preferred since $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_A^{FB}$.

**Region 2b).** $\underline{\sigma}_A^M$ is optimal: If $h_2 = (+,+)$, $\underline{\sigma}_{DM}\left(\underline{\sigma}_A^M(h_2)\right) = accept$ and $\pi_A(accept) > \pi_A(reject)$ since $\lambda_A < l_{(+,+)}$. Otherwise, $\underline{\sigma}_{DM}(m) = reject$ for any feasible message $m \in M\left(\widetilde{h}\right)$. $\underline{\sigma}_A^T$ is optimal since a single test is not pivotal to the DM's choice and (56) holds. On-path beliefs are consistent: $m = \{+\}$ occurs if and only if $h_2 = \phi = (+,-)$ and $m = \emptyset$ occurs if and only if $h_2 = (-,\emptyset)$, i.e. if and only if $\phi \in \{(-,-),(-,+)\}$. The equilibrium is an advisor-preferred equilibrium. Given $\Phi_A^{FB} = \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} \cup (+,-) \cup (-,+)$. By Lemma 9, in an equilibrium which makes the advisor strictly better off than $\left(\underline{\sigma}, \underline{\mu}\right)$ at least $\phi = (+,-)$ or $\phi = (-,+)$ must result in acceptance. But if a message that leads the DM to accept is feasible given $h_2 = (+,-)$ then it would also be feasible given $h_2 = (-,+)$ and it cannot be part of an equilibrium that the DM accepts if $\phi \in \{(+,-),(-,+),(+,+)\}$ since $l_{\neg(-,-)} < \lambda_{DM}$.

**Region 2c).** $\underline{\sigma}_A^M$ and $\underline{\sigma}_A^T$ are optimal and on-path beliefs are consistent by the same reasoning as in Region 1b. The equilibrium is an advisor-preferred equilibrium since $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_A^{FB}$.

**Region 2d).** $\underline{\sigma}_A = \left(\underline{\sigma}_A^T, \underline{\sigma}_A^M\right)$ is optimal and on-path beliefs are consistent by the same reasoning as in Region 1c. The equilibrium is advisor-preferred since $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_A^{FB}$.

**Region 3.** $\underline{\sigma}_A^M$ is optimal: If $h_2 = (+,+)$, $\underline{\sigma}_{DM}\left(\underline{\sigma}_A^M(h_2)\right) = accept$ and $\pi_A(accept) > \pi_A(reject)$ since $\lambda_A < l_{(+,+)}$. Otherwise, $\underline{\sigma}_{DM}(m) = reject$ for any feasible message $m \in M\left(\widetilde{h}\right)$. $\underline{\sigma}_A^T$ is optimal since (56) holds, but (53) does not hold. On-path beliefs are consistent by the same reasoning as in Region 2. The equilibrium is an advisor-preferred equilibrium. If $l_{(+,-)} \leq \lambda_A < l_{(+,+)}$, then $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_A^{FB}$. If $\lambda_A < l_{(+,-)}$ then $\Phi_A^{FB} \supseteq \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} \cup (+,-) \cup (-,+)$ and the equilibrium is advisor-preferred by the same reasoning as in Region 2b.
**Region 4.** $\underline{\sigma}_A^M$ and $\underline{\sigma}_A^T$ are optimal since $\underline{\sigma}_{DM}(m) = reject$ for any $m \in \mathcal{M}$. On-path beliefs are consistent: $m = \emptyset$ occurs if and only if $h_2 = (\emptyset,\emptyset)$, i.e. for any $\phi \in \Phi$. The equilibrium is an advisor-preferred equilibrium. For any $\lambda_A < l_{(+,+)}$, $\Phi_A^{FB} \supset \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \emptyset$. By Lemma 9, in an equilibrium which makes the advisor strictly better off than $\left(\underline{\sigma},\underline{\mu}\right)$ at least $\phi = (+,+)$ must be result in acceptance. But it can never be part of an equilibrium that the DM accepts for any $\phi$ since $\lambda_{DM} > l_{(+,+)}$.
Finally, if $\lambda_A < l_{(+,+)}$ then $\underline{\sigma}_A^M$ is optimal since, for any $h_2$, $\underline{\sigma}_{DM}\left(\underline{\sigma}_A^M(h_2)\right) = reject$ and $\pi_A(reject) > \pi_A(accept)$. On-path beliefs are consistent by the same reasoning as in Region 4. The equilibrium is an advisor-preferred equilibrium since $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_A^{FB}$.

### A.5.2 DM-preferred Equilibria

**Lemma 12 (DM-preferred Equilibrium Characterization Hidden Testing)**
*Consider $\Gamma$. $(\overline{\sigma},\overline{\mu})$ is a DM-preferred equilibrium. In all regions, $\underline{\mu}_A : \cup_{n=0}^N H_n \to [0,1]$ is equal to $\mu^o : \cup_{n=0}^N H_n \to [0,1]$ in Lemma 10. For any $m \in \{\{+,+\},\{+,-\},\{-,-\}\}$, $\overline{\mu}_{DM} : \mathcal{M} \to [0,1]$ is given by (18).*
*In **Region 1**, i.e. $l_\Phi < \lambda_{DM} \leq l_{(+,-)}$. Then $\overline{\sigma}_{DM}(m) = reject$ if and only if $m = \{-,-\}$ and $\overline{\sigma}_A^T = (1,0,1,0)$.*

1. *If $l_{(+,-)} \leq \lambda_A < l_{(+,\cdot)}$, then $(\lambda_A,\lambda_{DM}) \in W_I$ and*

$$\overline{\sigma}_A^M = \begin{cases} \{-\} & \text{if } \widetilde{h} \in \{\{-\},\{+,-\}\} \\ \widetilde{h} & \text{otherwise,} \end{cases} \tag{60}$$

*and beliefs satisfy*

$$\frac{\overline{\mu}_{DM}(m)}{1-\overline{\mu}_{DM}(m)} = \begin{cases} l_{(+,-)} & \text{if } m \in \{\{-\},\emptyset\} \\ l_{(+,\cdot)} & \text{if } m = \{+\}. \end{cases} \tag{61}$$

2. *If* $l_{(+,\cdot)} \leq \lambda_A < l_{(+,+)}$, *then* $(\lambda_A, \lambda_{DM}) \in I$ *and*

$$\overline{\sigma}_A^M = \begin{cases} \emptyset & \text{if } \widetilde{h} \in \{\emptyset, \{+\}\} \\ \{-\} & \text{if } \widetilde{h} \in \{\{-\}, \{+,-\}\} \\ \widetilde{h} & \text{otherwise,} \end{cases} \tag{62}$$

*and beliefs satisfy*

$$\frac{\overline{\mu}_{DM}(m)}{1-\overline{\mu}_{DM}(m)} = \begin{cases} l_{(+,-)} & \text{if } m = \{-\} \\ l_{(+,\cdot)} & \text{if } m = \emptyset \\ l_{\neg(-,-)} & \text{if } m = \{+\}. \end{cases} \tag{63}$$

3. *If* $\lambda_A \geq l_{(+,+)}$,

$$\overline{\sigma}_A^M = \begin{cases} \emptyset & \text{if } h \in \{\emptyset, \{+\}, \{+,+\}\} \\ \{-\} & \text{if } h \in \{\{-\}, \{+,-\}\} \\ \widetilde{h} & \text{otherwise,} \end{cases} \tag{64}$$

*and beliefs satisfy (63).*

4. *In* ***Region 2****, i.e. max* $\left\{l_\Phi, l_{(+,-)}\right\} < \lambda_{DM} \leq l_{\neg(-,-)}$. *If* $\lambda_A \leq l_{(+,-)}$, *then* $(\lambda_A, \lambda_{DM}) \in W_S$. $\overline{\sigma}_{DM}(m) = accept$ *if and only if* $m = \{+,+\}$, $\overline{\sigma}_A^T = (1,1,0,0)$ *and*

$$\overline{\sigma}_A^M = \begin{cases} \{+\} & \text{if } \widetilde{h} \in \{\{+\}, \{+,-\}\} \\ \widetilde{h} & \text{otherwise,} \end{cases} \tag{65}$$

*and beliefs satisfy*

$$\frac{\overline{\mu}_{DM}(m)}{1-\overline{\mu}_{DM}(m)} = \begin{cases} l_{(+,-)} & \text{if } m = \{+\} \\ l_{(-,\cdot)} & \text{if } m = \{-\} \\ l_\Phi & \text{if } m = \emptyset. \end{cases} \tag{66}$$

5. *In* ***Region 2 and 3****, i.e. max* $\left\{l_\Phi, l_{(+,-)}\right\} < \lambda_{DM} \leq l_{(+,+)}$. *If* $\lambda_A > l_{(+,+)}$ *then* $\overline{\sigma}_{DM}(m) = accept$ *if and only if* $m \in \{\emptyset, \{+\}, \{+,+\}\}$, $\overline{\sigma}_A^T = (1,1,0,1)$ *and*

$$\overline{\sigma}_A^M = \begin{cases} \{-\} & \text{if } \widetilde{h} \in \{\{-\}, \{+,-\}, \{-,-\}\} \\ \emptyset & \text{otherwise,} \end{cases} \tag{67}$$

*and beliefs satisfy*

$$\frac{\overline{\mu}_{DM}(m)}{1-\overline{\mu}_{DM}(m)} = \begin{cases} l_{(+,+)} & \text{if } m \in \{\emptyset, \{+\}\} \\ l_{\neg(+,+)} & \text{if } m = \{-\}. \end{cases} \tag{68}$$

6. *Otherwise, the DM-preferred and the advisor-preferred equilibrium coincide and the equilibrium is characterized by Lemma 11.*

7. *In any DM-preferred equilibrium,* $\Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \Phi^{FB}_{DM}$.

**Proof:** For Parts 1-5, I show that $(\overline{\sigma}, \overline{\mu})$ constitutes a DM-preferred PBE. For a proof that this is a sequential equilibrium, see Section B.2. For Parts 1-5, $\overline{\sigma}_{DM}(m)$ is optimal by (5), since $\overline{\sigma}_{DM}(m) = accept$ if and only if $\frac{\overline{\mu}_{DM}(m)}{1-\overline{\mu}_{DM}(m)} \geq \lambda_{DM}$.

**Part 1-3:** $\sigma^M_A$ is optimal: if $h_2 = (-,-)$, $\overline{\sigma}_{DM}\left(\overline{\sigma}^M_A(h_2)\right) = reject$ and $\pi_A(reject) > \pi_A(accept)$ since $\lambda_A > l_{(-,-)}$. Otherwise, $\overline{\sigma}_{DM}\left(\overline{\sigma}^M_A(h_2)\right) = accept$ for any $m \in M\left(\widetilde{h}\right)$. $\overline{\sigma}^T_A$ is optimal: In period 2, following $h_1 \in \{(+),(\emptyset)\}$, the advisor stops testing because irrespective of the second outcome the DM accepts. Following $h_1 = (-)$, the advisor tests. If he stops, the DM accepts. If he tests, the DM rejects if and only if $h_2 = (-,-)$. The advisor tests since

$$\lambda_A [1-\overline{\mu}_A(h_1=(-))] > \lambda_A [1-\overline{\mu}_A(h_1=(-))](1-p_F) + \overline{\mu}_A(h_1=(-))(1-p_T) \Leftrightarrow$$

$$\lambda_A > \frac{q}{1-q}\frac{\overline{\mu}_A(h_1=(-))(1-p_T)}{[1-\overline{\mu}_A(h_1=(-))]p_F} = \frac{q(1-p_T)^2}{(1-q)p_F^2} = l_{(-,-)}.$$

This also implies that he tests in period 1. For Part 1, on-path beliefs are consistent: $m = \{-\}$ occurs if and only if $h_2 = \phi = (-,+)$, and $m = \{+\}$ occurs if and only if $h_2 = (+,\emptyset)$, i.e. $\phi \in \{(+,+),(+,-)\}$, and $m = \{-,-\}$ occurs if and only if $h_2 = \phi = (-,-)$. For Part 2-3, on-path beliefs are consistent: $m = \{-\}$ occurs if and only if $h_2 = \phi = (-,+)$, and $m = \emptyset$ occurs if and only if $h_2 = (+,\emptyset)$, i.e. $\phi \in \{(+,+),(+,-)\}$, and $m = \{-,-\}$ occurs if and only if $h_2 = \phi = (-,-)$. The equilibrium acceptance set is

$$\Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \{(+,+),(+,-),(-,+)\} = \Phi^{FB}_{DM}. \tag{69}$$

**Part 4:** $\overline{\sigma}^M_A$ is optimal: if $h_2 = (+,+)$, $\overline{\sigma}_{DM}\left(\overline{\sigma}^M_A(h_2)\right) = accept$ and $\pi_A(accept) > \pi_A(reject)$ since $\lambda_A < l_{(+,+)}$. Otherwise, $\overline{\sigma}_{DM}(m) = reject$ for all $m \in M\left(\widetilde{h}\right)$. $\overline{\sigma}^T_A$ is optimal: In period 2, following $h_1 \in \{(-),(\emptyset)\}$, the advisor stops testing because irrespective of the second outcome the DM rejects. Following $h_1 = (+)$, he tests since (56) holds. This implies that he tests in period 1. On-path beliefs are consistent: $m = \{-\}$ occurs if and only if $h_2 = (-, \emptyset)$, i.e. $\phi \in \{(-,-), (-,+)\}$, and $m = \{+\}$ occurs if and only if $h_2 = \phi = (+,-)$ and $m = \{+,+\}$ occurs if and only if $h_2 = \phi = (+,+)$. The equilibrium acceptance is given by $\Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \{(+,+)\}$.

**Part 5.** $\overline{\sigma}_A^M$ is optimal: if $h_2$ is such that $\widetilde{h} \in \{\{-\}, \{+,-\}, \{-,-\}\}$, $\overline{\sigma}_{DM}\left(\overline{\sigma}_A^M(h_2)\right) = reject$ and $\pi_A(reject) > \pi_A(accept)$ since $\lambda_A > l_{(+,-)}$. Otherwise, $\overline{\sigma}_{DM}(m) = accept$ for all $m \in M\left(\widetilde{h}\right)$. $\overline{\sigma}_A^T$ is optimal: In period 2, following $h_1 = (-)$, because the advisor prefers to reject irrespective of the second outcome and if he stops the DM rejects. Following $h_1 = (+)$, if he stops the DM accepts, but if he tests, the DM accepts if and only if $h_2 = (+,+)$ and (52) does not hold. Following $h_1 = (\emptyset)$, if he stops, the DM accepts, but if he tests, the DM accepts if and only if $h_2 = (\emptyset, +)$. The advisor tests since

$$\lambda_A [1 - \overline{\mu}_A(h_0)] > \overline{\mu}_A(h_0)(1 - p_T) + \lambda_A [1 - \overline{\mu}_A(h_0)](1 - p_F) \Leftrightarrow$$
$$\lambda_A > \frac{q(1-p_T)}{(1-q)p_F} = l_{(-,\cdot)}. \tag{70}$$

This implies that he tests in period 1. On-path beliefs are consistent: $m = \{-\}$ occurs if and only if $h_2 \in \{(-,\emptyset), (+,-)\}$, i.e. $\phi \in \{(-,-), (-,+), (+,-)\}$, and $m = \emptyset$ occurs if and only if $h_2 = \phi = (+,+)$. The equilibrium acceptance set is given by $\Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \{(+,+)\}$.

**Part 6:** By Lemma 2, if $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_{DM}^{FB}$ then $\Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \Phi_{DM}^{FB}$. Hence, by Lemma 11, for any $(\lambda_A, \lambda_{DM})$ not covered in Parts 1-5 it holds that $\Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \Phi_{DM}^{FB}$ .

**Part 7:** $\Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \Phi_{DM}^{FB}$ by the proofs of Parts 1-5 and by Part 6.

## A.6 Theorem 1 [DM Payoff Comparison], Proposition 2 [Advisor Payoff Comparison]

The proof uses Definitions 1 and 2. The equilibrium acceptance set under observable testing, $\Phi_{(\sigma^o,\mu^o;\Gamma^o)}$, is given by Lemma 10 in Section A.4, and the equilibrium acceptance set for any advisor-preferred equilibrium under hidden testing, $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)}$ is given by Lemma 11 and for any DM-preferred equilibrium under hidden testing, $\Phi_{(\overline{\sigma},\overline{\mu};\Gamma)}$ , is given by Lemma 12 in Section A.5. The DM's first-best acceptance set, $\Phi_{DM}^{FB}$, is given by (57) and the advisor's first-best acceptance set, $\Phi_A^{FB}$, is given by (58).

**Region 1**: Suppose $l_\Phi < \lambda_{DM} \leq l_{(+,-)}$.

(a) If $\lambda_A < l_{(+,-)}$ then

$$\Phi_{(\sigma^o,\mu^o;\Gamma^o)} = \Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \Phi_{DM}^{FB} = \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} \subseteq \Phi_A^{FB}. \tag{71}$$

By Lemma 9, $\pi_i(\overline{\sigma}, \overline{\mu}; \Gamma) = \pi_i\left(\underline{\sigma}, \underline{\mu}; \Gamma\right) = \pi_i(\sigma^o, \mu^o; \Gamma^o)$ for $i = A, DM$.

(b) If $l_{(+,-)} \leq \lambda_A < l_{(+,\cdot)}$ then $(\lambda_A, \lambda_{DM}) \in W_I$ and

$$\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_A^{FB} \subset \Phi_{(\sigma^o,\mu^o;\Gamma^o)} \subset \Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \Phi_{DM}^{FB}. \tag{72}$$

By Lemma 9, $\pi_{DM}\left(\underline{\sigma},\underline{\mu};\Gamma\right) < \pi_{DM}\left(\sigma^o,\mu^o;\Gamma^o\right) < \pi_{DM}\left(\overline{\sigma},\overline{\mu};\Gamma\right)$ and $\pi_A\left(\underline{\sigma},\underline{\mu};\Gamma\right) > \pi_A\left(\sigma^o,\mu^o;\Gamma^o\right) > \pi_A\left(\overline{\sigma},\overline{\mu};\Gamma\right)$.

(c) If $l_{(+,\cdot)} \leq \lambda_A < l_{(+,+)}$ then $(\lambda_A, \lambda_{DM}) \in I$ and

$$\Phi_{(\sigma^o,\mu^o;\Gamma^o)} \subset \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_A^{FB} \subset \Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \Phi_{DM}^{FB}. \tag{73}$$

By Lemma 9, $\pi_{DM}\left(\sigma^o,\mu^o;\Gamma^o\right) < \pi_{DM}\left(\underline{\sigma},\underline{\mu};\Gamma\right) < \pi_{DM}\left(\overline{\sigma},\overline{\mu};\Gamma\right)$ and $\pi_A\left(\underline{\sigma},\underline{\mu};\Gamma\right) > \pi_A\left(\sigma^o,\mu^o;\Gamma^o\right) > \pi_A\left(\overline{\sigma},\overline{\mu};\Gamma\right)$, where $\pi_A\left(\sigma^o,\mu^o;\Gamma^o\right) > \pi_A\left(\overline{\sigma},\overline{\mu};\Gamma\right)$ since

$$\begin{aligned} Pr\,(true) < Pr\,(true)\,(1-p_T)^2 + \lambda_A Pr\,(false)\left(1-p_F^2\right) &\Leftrightarrow \\ q < q\,(1-p_T)^2 + \lambda_A\,(1-q)\left(1-p_F^2\right) &\Leftrightarrow \\ l_{\neg(-,-)} = \frac{q\left(1-(1-p_T)^2\right)}{(1-q)\left(1-p_F^2\right)} < \lambda_A. \end{aligned} \tag{74}$$

(d) If $\lambda_A \geq l_{(+,+)}$ then

$$\Phi_{(\sigma^o,\mu^o;\Gamma^o)} = \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_A^{FB} \subseteq \Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \Phi_{DM}^{FB}. \tag{75}$$

By Lemma 9, $\pi_{DM}\left(\sigma^o,\mu^o;\Gamma^o\right) = \pi_{DM}\left(\underline{\sigma},\underline{\mu};\Gamma\right) < \pi_{DM}\left(\overline{\sigma},\overline{\mu};\Gamma\right)$ and $\pi_A\left(\sigma^o,\mu^o;\Gamma^o\right) = \pi_A\left(\underline{\sigma},\underline{\mu};\Gamma\right) > \pi_A\left(\overline{\sigma},\overline{\mu};\Gamma\right)$.

**Region 2:** Suppose $max\left\{l_\Phi, l_{(+,-)}\right\} < \lambda_{DM} \leq l_{(+,\cdot)}$.

(a) If $max\left\{l_\Phi, l_{(+,-)}\right\} < \lambda_{DM} \leq l_{\neg(-,-)}$ and $\lambda_A < l_{(+,-)}$ then $(\lambda_A, \lambda_{DM}) \in W_S$ and

$$\Phi_{DM}^{FB} = \Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} \subset \Phi_{(\sigma^o,\mu^o;\Gamma^o)} \subset \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} \subseteq \Phi_A^{FB}. \tag{76}$$

By Lemma 9, $\pi_{DM}\left(\underline{\sigma},\underline{\mu};\Gamma\right) < \pi_{DM}\left(\sigma^o,\mu^o;\Gamma^o\right) < \pi_{DM}\left(\overline{\sigma},\overline{\mu};\Gamma\right)$ and $\pi_A\left(\underline{\sigma},\underline{\mu};\Gamma\right) > \pi_A\left(\sigma^o,\mu^o;\Gamma^o\right) > \pi_A\left(\overline{\sigma},\overline{\mu};\Gamma\right)$.

(b) If $l_{\neg(-,-)} < \lambda_{DM} \leq l_{(+,\cdot)}$ and $\lambda_A < l_{(+,-)}$ then $(\lambda_A, \lambda_{DM}) \in S$ and

$$\Phi_{DM}^{FB} = \Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} \subset \Phi_{(\sigma^o,\mu^o;\Gamma^o)} \subset \Phi_A^{FB}. \tag{77}$$

By Lemma 9, $\pi_{DM}(\sigma^o,\mu^o;\Gamma^o) < \pi_{DM}\left(\underline{\sigma},\underline{\mu};\Gamma\right) = \pi_{DM}(\overline{\sigma},\overline{\mu};\Gamma)$ and $\pi_A(\sigma^o,\mu^o;\Gamma^o) > \pi_A\left(\underline{\sigma},\underline{\mu};\Gamma\right) = \pi_A(\overline{\sigma},\overline{\mu};\Gamma)$.

(c) If $l_{(+,-)} \leq \lambda_A < l_{(+,+)}$ then (71) holds.

(d) If $\lambda_A \geq l_{(+,+)}$ then (75) holds.

**Region 3:** Suppose $l_{(+,\cdot)} < \lambda_{DM} \leq l_{(+,+)}$. If $\lambda_A < l_{(+,+)}$ then (71) holds. If $\lambda_A \geq l_{(+,+)}$ then (75) holds.

**Region 4:** Suppose $\lambda_{DM} > l_{(+,+)}$. If $\lambda_A < l_{(+,+)}$ then (71) holds. If $\lambda_A \geq l_{(+,+)}$ then

$$\Phi_A^{FB} \subseteq \Phi_{(\sigma^o,\mu^o;\Gamma^o)} = \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \Phi_{(\overline{\sigma},\overline{\mu};\Gamma)} = \Phi_{DM}^{FB}. \tag{78}$$

**Theorem 1, Part 1:** $\pi_{DM}\left(\underline{\sigma},\underline{\mu};\Gamma\right) > \pi_{DM}(\sigma^o,\mu^o;\Gamma^o)$ if and only if $(\lambda_A,\lambda_{DM}) \in S \cup I$. By Lemma 2, $\pi_{DM}(\sigma,\mu;\Gamma) > \pi_{DM}(\sigma^o,\mu^o;\Gamma^o)$ for all $(\sigma,\mu) \in E$ if and only if $(\lambda_A,\lambda_{DM}) \in S \cup I$.

**Theorem 1, Part 2:** $\pi_{DM}\left(\underline{\sigma},\underline{\mu};\Gamma\right) < \pi_{DM}(\sigma^o,\mu^o;\Gamma^o)$ if and only if $(\lambda_A,\lambda_{DM}) \in W_S \cup W_I$. By Lemma 2, there exists $(\sigma,\mu) \in E$ such that $\pi_{DM}(\sigma,\mu;\Gamma) < \pi_{DM}(\sigma^o,\mu^o;\Gamma^o)$ if and only if $(\lambda_A,\lambda_{DM}) \in W_S \cup W_I$.

**Proposition 2, Part 1:** $\pi_A(\overline{\sigma},\overline{\mu};\Gamma) \leq \pi_A(\sigma^o,\mu^o;\Gamma^o)$ for all $(\lambda_A,\lambda_{DM})$.

**Proposition 2, Part 2:** $\pi_A\left(\underline{\sigma},\underline{\mu};\Gamma\right) < \pi_A(\sigma^o,\mu^o;\Gamma^o)$ if and only if $(\lambda_A,\lambda_{DM}) \in S$. By definition of advisor-preferred equilibrium, $\pi_A(\sigma,\mu;\Gamma) < \pi_A(\sigma^o,\mu^o;\Gamma^o)$ for all $(\sigma,\mu) \in E$ if and only if $(\lambda_A,\lambda_{DM}) \in S$.

**Proposition 2, Part 3:** $\pi_A\left(\underline{\sigma},\underline{\mu};\Gamma\right) > \pi_A(\sigma^o,\mu^o;\Gamma^o)$ if and only if $(\lambda_A,\lambda_{DM}) \in I \cup W_S \cup W_I$. By Lemma 2, there exists $(\sigma,\mu) \in E$ such that $\pi_A(\sigma,\mu;\Gamma) > \pi_A(\sigma^o,\mu^o;\Gamma^o)$ if and only if $(\lambda_A,\lambda_{DM}) \in I \cup W_S \cup W_I$.

## A.7 Corollary 1 [Preference Alignment]

By Part 2 of Theorem 1, $\pi_{DM}(\sigma,\mu;\Gamma) \geq \pi_{DM}(\sigma^o,\mu^o;\Gamma^o)$ for all $(\sigma,\mu) \in E$ if and only if both $(\lambda_A,\lambda_{DM}) \notin W_S$ and $(\lambda_A,\lambda_{DM}) \notin W_I$. Using simple geometry in Figure 2, one can find $d_S = l_{\neg(-,-)}$ to be the smallest value such that if $\lambda_{DM} - \lambda_A > d_S$ then $(\lambda_A,\lambda_{DM}) \notin W_S$. Similarly, one can find $d_I = l_{(+,\cdot)} - l_\Phi$ to be the smallest value such that if $\lambda_A - \lambda_{DM} > d_I$ then $(\lambda_A,\lambda_{DM}) \notin W_I$. Choose $d = max\left\{d_I, d_S\right\}$.

## A.8 Proposition 1 [First-Best Benchmark]

**Part 1:** Define $Z$ to be the set of $(\lambda_A,\lambda_{DM})$ such that $\pi_{DM}(\sigma^o,\mu^o;\Gamma^o) = \pi_{DM}^{FB}$ if and only if $(\lambda_A,\lambda_{DM}) \in Z$. Denote its complement by $Z^c$. Define $Y$ to be the set of $(\lambda_A,\lambda_{DM})$ such that

$\pi_{DM}(\sigma,\mu;\Gamma)=\pi_{DM}^{FB}$ for any $(\sigma,\mu)\in E$ if and only if $(\lambda_A,\lambda_{DM})\in Y$. Denote its complement by $Y^c$. By the proof of Theorem 1 and Lemma 2,

$$Z^c=W_S\cup W_I\cup S\cup I\cup\left\{(\lambda_{DM},\lambda_A):l_\Phi<\lambda_{DM}<l_{(+,+)},\lambda_A\geq l_{(+,+)}\right\}, \tag{79}$$

$$Y^c=W_S\cup W_I\cup I\cup\left\{(\lambda_{DM},\lambda_A):l_\Phi<\lambda_{DM}<l_{(+,+)},\lambda_A\geq l_{(+,+)}\right\}. \tag{80}$$

Hence, $Z^c=Y^c\cup S$ and, hence, $Z\subset Y$.

**Part 2:** This is shown by Part 7 of Lemma 12 in Section A.5.

## A.9 Corollary 2 [Hidden Testing Pareto-improving]

**Part 1:** By Part 1 of Theorem 1, $\pi_{DM}(\sigma,\mu;\Gamma)>\pi_{DM}(\sigma^o,\mu^o;\Gamma^o)$ for all $(\sigma,\mu)\in E$ if and only if $(\lambda_A,\lambda_{DM})\in S\cup I$. By Part 3 of Proposition 2, $\pi_A\left(\underline{\sigma},\underline{\mu};\Gamma\right)>\pi_A(\sigma^o,\mu^o;\Gamma^o)$ if and only if $(\lambda_A,\lambda_{DM})\in I\cup W_S\cup W_I$. By Lemma 2, if $\pi_A\left(\underline{\sigma},\underline{\mu};\Gamma\right)\leq\pi_A(\sigma^o,\mu^o;\Gamma^o)$ then $\pi_A(\sigma,\mu;\Gamma)\leq\pi_A(\sigma^o,\mu^o;\Gamma^o)$ for all $(\sigma,\mu)\in E$.

**Part 2:** This follows by Part 1 of Proposition 2.

## A.10 Lemma 6 [Observable Testing Equilibrium for $N>2$]

Suppose testing is observable. Denote the likelihood ratio at the end of period $n$ by $l_n$ for any $n\in\{0,...,N\}$. In equilibrium, beliefs are consistent if and only if $\frac{\mu^o(h_n)}{1-\mu^o(h_n)}=l(x_n)$ for any $n\in\{0,...,N\}$, where $x_n$ is the number of excess positive outcomes in $h_n$. Note that the advisor's expected payoff is independent of future test outcomes if these outcomes do not affect the DM's choice. This is because, holding the DM's choice fixed, the advisor's expected payoff is linear in the belief that the hypothesis is true and the expected posterior belief is equal to the prior belief.

If $r_{DM}\geq N$ then the DM optimally rejects following any $h_N$, because the highest likelihood ratio that can be realized at the end of period $N$ is $l_N=l(N)$ and at this likelihood ratio the DM prefers rejection since $l(N)\leq l(r_{DM})<\lambda_{DM}$.

**Part 1.** Choose $\overline{\lambda}_A=l(k)$, where

$$k\equiv inf\left\{j|j\in\mathbb{Z},j\geq\frac{N+r_{DM}+1}{2}\right\}. \tag{81}$$

Note that if $\lambda_A>\overline{\lambda}_A$ then the advisor is more reluctant than the DM since $r_A\geq k\geq r_{DM}+1$.[45] $\sigma_{DM}^o(h_N)=accept$ if and only if $x_N\geq r_{DM}+1$ by (24). Claim 1a-1e shows that if $\lambda_A>\overline{\lambda}_A=l(k)$ then $\sigma_A^o(h_n)=1$ if $r_{DM}+1\leq x_n\leq k-1$ and, otherwise, $\sigma_A^o(h_n)=0$ for

[45]Note that $k$ rises less than one-for-one with $r_{DM}$ and even if $r_{DM}=N-1$ then $k=N\geq N=r_{DM}+1$.

$n \in \{0, ..., N-1\}$. Claim 1f shows that $\overline{\lambda}_A = l(k)$ is the smallest value of $\overline{\lambda}_A$ such that if $\lambda_A > \overline{\lambda}_A$ then$\sigma_A^o(h_n) = 0$ if $x_n = 0$ for $n \in \{0, ..., N-1\}$.

**Claim 1a** $\sigma_A^o(h_n) = 0$ if $x_n = k$ *for* $n \in \{k, ..., N-1\}$.

First, note that given likelihood ratio $l(k)$ the DM prefers to accept since $l(k) > l(r_{DM}+1)$. Second, if $l_n = l(k)$ then no realization of the remaining test outcomes can lead the DM to prefer rejection. If all remaining $N-n$ outcomes are negative then $l_N = l(k-(N-n))$. This is the lowest possible value of $l_N$ given $l_n = l(k)$. Independent of $n$, the DM prefers acceptance at the lowest possible value of $l_N$. This is because $\lambda_{DM} \leq l(r_{DM}+1) \leq l(k-(N-n))$ since $n \geq k$ for $l_n = l(k)$ to be feasible. Therefore, there is no reason to test.

**Claim 1b** $\sigma_A^o(h_n) = 0$ if $x_n \geq k+1$ *for* $n \in \{k+1, ..., N-1\}$.

Whatever the advisor's action in any period $n+1$ for $n \in \{k+1, ..., N-1\}$, if $l_n > l(k)$, then by Claim 1a it must be that in any period $n' > n$, $l_{n'} \geq l(k)$. This implies that the DM always accepts since $\lambda_{DM} \leq l(r_{DM}+1) < l(k)$. Therefore, there is no reason to test.

**Claim 1c** $\sigma_A^o(h_n) = 0$ if $x_n \leq r_{DM} - (N-n)$ *for* $n \in \{0, ..., N-1\}$.

First, note that if the advisor does not test in any period $n+1$ for $n \in \{0, ..., N-1\}$, then $l_N \leq l(r_{DM}-(N-n))$ and the DM rejects since $l(r_{DM}-(N-n)) < l(r_{DM}) < \lambda_{DM}$. Second, no realization of the remaining test outcomes can lead the DM to accept. If $l_n = l(r_{DM}-(N-n))$ and all remaining $N-n$ outcomes are positive then $l_N = l(r_{DM})$. Therefore, if $l_n \leq l(r_{DM}-(N-n))$ then it must be that $l_N \leq l(r_{DM})$ and the DM rejects by (24). Therefore, there is no reason to test.

**Claim 1d** $\sigma_A^o(h_n) = 0$ if $r_{DM} - (N-n) \leq x_n \leq r_{DM}$ *for* $n \in \{0, ..., N-1\}$.

First, I show that if the advisor tests in any period $n+1$ if $l(r_{DM}-(N-n)) \leq l_n \leq l(r_{DM})$ then with a strictly positive probability the DM accepts when the advisor prefers rejection. If the advisor tests in any period $n+1$ if $l(r_{DM}-(N-n)) \leq l_n \leq l(r_{DM})$ then eventually either Claim 1c applies and the DM rejects or at some future period $n' > n$ it holds that $l_{n'} \geq l(r_{DM}+1)$. I have not yet specified what the advisor will do in period $n+1$ if $l(r_{DM}+1) \leq l_n \leq l(k-1)$. Note that there is no reason for the advisor to stop and then restart. If he were to stop at some period $n+1$ if $l(r_{DM}+1) \leq l_n \leq l(k-1)$ then $l(r_{DM}+1) \leq l_N \leq l(k-1)$ and the DM accepts but the advisor prefers rejection since $\lambda_{DM} \leq l_N < \overline{\lambda}_A$. If he were to test at any $n+1$ if $l(r_{DM}+1) \leq l_n \leq l(k-1)$ then eventually either Claim 1a applies and the DM accepts or Claim 1c applies and the DM rejects. If Claim 1a applies then the advisor prefers rejection since $l(k) \leq \overline{\lambda}_A < \lambda_A$. Second, if the advisor does not test in in any period $n+1$ if $l(r_{DM}-(N-n)) \leq l_n \leq l(r_{DM})$ then $l_N \leq l(r_{DM})$ and the DM rejects since $l(r_{DM}) < \lambda_{DM}$. Lastly, note that if the DM rejects then the advisor prefers rejection given that the advisor is more reluctant than the DM. This implies that testing has no upside but a strict downside.

**Claim 1e** $\sigma_A^o(h_n) = 1$ if $r_{DM}+1 \leq x_n \leq k-1$ *for* $n \in \{r_{DM}+1, ..., N-1\}$.

For any likelihood ratio $l(r_{DM}+1) \leq l_n \leq l(k-1)$, the advisor prefers rejection since $l(k-1) < l(k) = \overline{\lambda}_A$. If he does not test then $l(r_{DM}+1) \leq l_N \leq l(k-1)$ and the DM accepts. If he tests in any period $n+1$ then eventually he stops either because Claim 1a applies, or Claim 1d applies or he runs out of tests. If Claim 1a applies or if he runs out of tests then the DM accepts, but the advisor prefers rejection since $l_N \leq l(k) < \lambda_A$. If Claim 1d applies, the DM rejects and the advisor prefers rejection. Therefore, testing has a strict upside but no downside.

***Claim 1f*** *$k$ is chosen to be the smallest value such that $\sigma_A^o(h_n) = 0$ if $x_n = 0$ for $n \in \{0, ..., N-1\}$.*

The argument above rests on Claim 1a being true, i.e. the likelihood ratio at which the DM accepts irrespective of the remaining outcomes in any period $n$, given by $l(k)$, must lead the advisor to prefer rejection, i.e. $\lambda_A < l(k)$. Claim 1a holds if and only if $\lambda_{DM} \leq l(r_{DM}+1) \leq l(k-(N-n))$ holds at any $n$, where $n \geq k$. Hence, it is sufficient that $k$ is the smallest value to satisfy $k-(N-k) \geq r_{DM}+1$, which is given by (81).

**Part 2.** Consider $r_{DM} < N$. Choose $\underline{\lambda}_A = l(2(r_{DM}+1)-N)$. Suppose $\lambda_A < \underline{\lambda}_A$, i.e. $r_A + 1 \leq 2(r_{DM}+1) - N$. Note that this implies that the advisor is more enthusiastic than the DM, since $r_A \leq 2(r_{DM}+1) - N - 1 \leq r_{DM}$.[46] $\sigma_{DM}^o(h_N) = accept$ if and only if $x_N \geq r_{DM}+1$ by (24). I show that $\sigma_A^o(h_n) = 1$ if $r_{DM}+1-(N-n) \leq x_n \leq r_{DM}$ and, otherwise, $\sigma_A^o(h_n) = 0$ for $n \in \{0, ..., N-1\}$.

**Claim 2a** $\sigma_A^o(h_n) = 0$ *if* $x_n = r_{DM}+1$ *for* $n \in \{r_{DM}+1, ..., N-1\}$.

First, note that at $x_n = r_{DM}+1$, then $l_n = l(r_{DM}+1)$ and the advisor prefers to accept since $\lambda_A \leq \lambda_{DM} \leq l(r_{DM}+1)$. Second, if $x_n = r_{DM}+1$ then no realization of the remaining test outcomes can lead the advisor to prefer rejection. If $l_n = l(r_{DM}+1)$ and all remaining $N-n$ outcomes are negative then $l_N = l(r_{DM}+1-(N-n))$. This is the lowest possible value of $l_N$ given $l_n = l(r_{DM}+1)$. Independent of $n$, the advisor prefers acceptance at the lowest possible value of $l_N$ since $\lambda_A < \underline{\lambda}_A = l(2(r_{DM}+1)-N) \leq l(r_{DM}+1-(N-n))$ and it must be that $n \geq r_{DM}+1$ for $l_n = l(r_{DM}+1)$ to be feasible. Lastly, if $l_n = l(r_{DM}+1)$ and the advisor does not test in any period $n+1$ for $n \in \{r_{DM}+1, ..., N-1\}$, then $l_N = l(r_{DM}+1)$ and the DM accepts. Therefore, there is no reason to test.

**Claim 2b** $\sigma_A^o(h_n) = 0$ *if* $x_n \geq r_{DM}+2$ *for* $n \in \{r_{DM}+2, ..., N-1\}$.

Whatever the advisor's action in any period $n+1$ for $n \in \{r_{DM}+2, ..., N-1\}$, if $x_n \geq r_{DM}+2$ , then by Claim 2a it must be that in any period $n' > n$, $x_{n'} \geq r_{DM}+1$ . This implies that the DM always accepts since $\lambda_{DM} \leq l(r_{DM}+1) \leq l_{n'}$. Therefore, there is no reason to test.

**Claim 2c** $\sigma_A^o(h_n) = 0$ *if* $x_n \leq r_{DM} - (N-n)$ *for* $n \in \{0, ..., N-1\}$.

[46] $2(r_{DM}+1) - N - 1$ rises more than one-for-one with $r_{DM}$ and even if $r_{DM} = N-1$ then $2(r_{DM}+1) - N - 1 = N - 1 \leq r_{DM}$.

There is no reason to test by the same argument as in Claim 1c.

**Claim 2d** $\sigma_A^o(h_n) = 1$ *if* $r_{DM} + 1 - (N-n) \leq x_n \leq r_{DM}$ *for* $n \in \{0, ..., N-1\}$.

For any $x_n$ where $r_{DM} + 1 - (N-n) \leq x_n \leq r_{DM}$, there are realizations of the remaining $N-n$ outcomes such that at some future period $n' > n$ it holds that $l_{n'} = l(r_{DM}+1)$. Hence, if the advisor tests in any period $1+n$ if $l(r_{DM}+1-(N-n)) \leq l_n \leq l(r_{DM})$, then eventually the advisor stops either because Claim 2a applies or because Claim 2c applies or because he runs out of tests. If Claim 2a applies, the DM accepts. If Claim 2c applies or if the advisor runs out of tests, the DM rejects. There is no reason for the advisor to stop and then restart. If the advisor stopped in some period $n+1$ for $n \in \{0, ..., N-1\}$ if $l(r_{DM}+1-(N-n)) \leq l_n \leq l(r_{DM})$, then the DM rejects. Therefore, testing is pivotal to the DM's choice if and only if $l(r_{DM}+1)$ is reached. The advisor prefers acceptance at $l(r_{DM}+1)$ since $\lambda_A \leq l(r_A+1) \leq l(r_{DM}+1)$. Hence, there is a strict upside but no downside to testing.

## A.11 Lemma 7 [Hidden Testing for $N > 2$]

Suppose testing is hidden. Define $r_A$ and $r_{DM}$ by (24) and (25). Denote the likelihood ratio at the end of period $n$ by $l_n$ for any $n \in \{0, ..., N\}$.

**Part 1:** The advisor is more reluctant, i.e. $r_{DM} + 1 \leq r_A$. The following is an advisor-preferred sequential equilibrium. The DM's beliefs are given by

$$\frac{\underline{\mu}_{DM}(m)}{1-\underline{\mu}_{DM}(m)} = \begin{cases} \frac{q\sum_{s=m^+}^{s=N-m^-}\binom{N}{s}p^s(1-p)^{N-s}}{(1-q)\sum_{s=m^+}^{s=N-m^-}\binom{N}{s}(1-p)^s p^{N-s}} & \text{if } m^+ \geq \overline{\nu}_+ \text{ and } m^- \geq 0, \\ \frac{q\sum_{s=max\{m^-,\overline{\nu}_-\}}^{s=N-m^+}\binom{N}{s}p^{N-s}(1-p)^s}{(1-q)\sum_{s=max\{m^-,\overline{\nu}_-\}}^{s=N-m^+}\binom{N}{s}(1-p)^{N-s}p^s} & \text{otherwise.} \end{cases} \tag{82}$$

which means that if the advisor reports $m^+ \geq \overline{\nu}_+$ and $m^- \geq 0$, the DM believes that at least $m^+$ and at most $N - m^-$ in $N$ tests were positive, otherwise, the DM believes that at least $max\{m^-, \overline{\nu}_-\}$ and at most $N - m^+$ outcomes in $N$ tests were negative.

The DM's strategy is$\underline{\sigma}_{DM}(m) = accept$ if and only if $m^+ \geq min\left\{\overline{\nu}_+, \widehat{Y}(m^-)\right\}$, where

$$\widehat{Y}\left(m^-\right) \equiv inf\left\{j | j \in \mathbb{N}, \frac{q\sum_{s=max\{m^-,\overline{\nu}_-\}}^{s=N-j}\binom{N}{s}p^{N-s}(1-p)^s}{(1-q)\sum_{s=max\{m^-,\overline{\nu}_-\}}^{s=N-j}\binom{N}{s}(1-p)^{N-s}p^s} \geq \lambda_{DM}\right\}. \tag{83}$$

The advisor's strategy is

$$\underline{\sigma}_A^T(h_n) = \begin{cases} 0 & \text{if } \nu_n^+ \geq \overline{\nu}_+ \text{ or } \nu_n^- \geq \overline{\nu}_-, \\ 1 & \text{otherwise,} \end{cases} \tag{84}$$

and

$$\underline{\sigma}_A^M(h_N) = \begin{cases} \left(0, \nu_N^-\right) & \text{if } x_N \leq r_A, \\ \left(\nu_N^+, 0\right) & \text{otherwise,} \end{cases} \tag{85}$$

and the advisor's beliefs satisfy $\frac{\underline{\mu}_A(h_n)}{1-\underline{\mu}_A(h_n)} = l(h_n)$ for any $n \in \{0, ..., N-1\}$.

**Claim 1a:** *$\underline{\sigma}_A^M$ is optimal given $\underline{\sigma}_{DM}$ and $\underline{\mu}_A$.*

For any $h_N$, reporting $m = (m^+, m^-) = \left(0, \nu_N^-\right)$ always leads to rejection. Therefore, if the advisor prefers rejection at $h_N$, i.e. if $l_N = l(x_N) \leq l(r_A) < \lambda_A$, then it is optimal to report $m = \left(0, \nu_N^-\right)$. For any $h_N$, if reporting $m = \left(\nu_N^+, 0\right)$ does not lead to acceptance then no feasible message $m \in M\left(\widetilde{h}\right)$ leads to acceptance. Therefore, if the advisor prefers acceptance it is optimal to report $m = \left(\nu_N^+, 0\right)$.

**Claim 1b:** *$\underline{\sigma}_A^T$ is optimal given $\underline{\sigma}_A^M$ and $\underline{\sigma}_{DM}$ and $\underline{\mu}_A$.*

First, if $\nu_n^- \geq \overline{\nu}_-$, the advisor prefers rejection irrespective of the realization of the remaining outcomes. At $h_n$, $l_n = l\left((n - \nu_n^-) - \nu_n^-\right)$ and if all remaining $N-n$ outcomes are positive then $l_N = l\left(n - 2\nu_n^- + (N-n)\right) = l\left(N - 2\nu_n^-\right)$. This is the highest possible likelihood ratio at the end of period $N$. If $\nu_n^- \geq \overline{\nu}_-$ then the advisor prefers rejection even at the highest possible likelihood ratio since $\lambda_A > l(r_A) \geq l\left(N - 2\nu_n^-\right) \geq l\left(N - 2\overline{\nu}_-\right)$ by (26) and (27). Given $\underline{\sigma}_A^M$, if he stops testing he reports $m = \left(\nu_n^-, 0\right)$ and the DM rejects. Hence, there is no reason to test. Second, if $\nu_n^+ \geq \overline{\nu}_+$, the advisor prefers acceptance irrespective of the realization of the remaining outcomes. At $h_n$, $l_n = l\left(\nu_n^+ - (n - \nu_n^+)\right)$ and if all remaining $N-n$ outcomes are negative then $l_N = l\left(2\nu_n^+ - n - (N-n)\right) = l\left(2\nu_n^+ - N\right)$. This is the lowest possible likelihood ratio at the end of period $N$. If $\nu_n^+ \geq \overline{\nu}_+$ then the advisor prefers acceptance even at the lowest possible likelihood ratio since $\lambda_A \leq l(r_A + 1) \leq l\left(2\overline{\nu}_+ - N\right) \leq l\left(2\nu_n^+ - N\right)$ by (26). Given $\underline{\sigma}_A^M$, if he stops testing he reports $m = \left(\nu_n^+, 0\right)$ and the DM accepts. Hence, there is no reason to test. Otherwise, the advisor benefits from testing because the remaining outcomes may be pivotal to his preferred action. At some period $n \leq N$, either $\nu_n^+ \geq \overline{\nu}_+$ or $\nu_n^- \geq \overline{\nu}_-$ and in each case the DM acts in his interest.

**Claim 1c:** *$\underline{\sigma}_{DM}$ is optimal given $\underline{\mu}_{DM}$.*

Given $m^+ \geq \overline{\nu}_+$, the DM always accepts because, as shown in Claim 1b, the advisor prefers to accept irrespective of the remaining outcomes and since the DM is more enthusiastic than the advisor, the DM must optimally accept. Otherwise, the DM optimally accepts if and only if $\nu_n^+ \geq \widehat{Y}(m^-)$ given (83).

**Claim 1d:** *There exists a sequence of completely mixed strategies $\left\{\sigma^k\right\}_{k=1}^{\infty}$, with $\lim_{k\to\infty} \sigma^k = \underline{\sigma}$, such that the system of beliefs $\underline{\mu} = \lim_{k\to\infty} \mu^k$, where $\mu^k$ denotes the beliefs derived from strategy profile $\sigma^k$ using Bayes' rule.*

See Section B.3 in the supplementary appendix.

**Claim 1e:** *This is an advisor-preferred equilibrium.*

By period $N$, either $\nu_n^+ = \overline{\nu}_+$ and the DM accepts or $\nu_n^- = \overline{\nu}_-$ and the DM rejects. By Claim 1b, this shows that whatever the realization of the complete list of Nature's $N$ draws, the advisor achieves the same payoff as if he could himself choose to accept or reject, i.e. his first-best expected payoff. Hence, $\Phi_{(\underline{\sigma},\underline{\mu};\Gamma)} = \Phi_A^{FB}$ using Definitions 1 and 2.

**Part 2:** The advisor is more enthusiastic, i.e. $r_A + 1 \leq r_{DM}$. Define

$$\widehat{Z}\left(\nu^+\right) \equiv sup\left\{j | j \in \mathbb{N}_{\geq 0}, \lambda_{DM} \leq \frac{q\sum_{s=\nu^+}^{s=N-j}\binom{N}{s}p^s\left(1-p\right)^{N-s}}{\left(1-q\right)\sum_{s=\nu^+}^{s=N-j}\binom{N}{s}\left(1-p\right)^s p^{N-s}}\right\}, \tag{86}$$

i.e. $\widehat{Z}\left(\nu^+\right)$ is the largest number such that if at least $\nu^+$ but at most $N - \widehat{Z}\left(\nu^+\right)$ outcomes in $N$ tests were positive the DM accepts.

First, if $max\left\{\widehat{\nu}_+, \overline{\nu}_+\right\} = \overline{\nu}_+$, which implies $min\left\{\overline{\nu}_-, \widehat{\nu}_-\right\} = \overline{\nu}_-$ then the following is an advisor-preferred sequential equilibrium: the DM's beliefs are given by (82). $\underline{\sigma}_{DM}\left(m\right) =$ *accept* if and only if $m^+ \geq \widehat{\nu}_+$ and $m^- \leq \widehat{Z}\left(m^+\right)$, and $\underline{\sigma}_A$ is given by (84) and (85) and $\frac{\underline{\mu}_A(h_n)}{1-\underline{\mu}_A(h_n)} = l\left(h_n\right)$ for any $n \in \{0, ..., N-1\}$.

Second, if $max\left\{\widehat{\nu}_+, \overline{\nu}_+\right\} = \widehat{\nu}_+$, which implies $min\left\{\overline{\nu}_-, \widehat{\nu}_-\right\} = \widehat{\nu}_-$ then the following is a sequential equilibrium: the DM's beliefs are given by

$$\frac{\underline{\mu}_{DM}\left(m\right)}{1-\underline{\mu}_{DM}\left(m\right)} = \begin{cases} \frac{q\sum_{s=m^+}^{s=N-m^-}\binom{N}{s}p^s(1-p)^{N-s}}{(1-q)\sum_{s=m^+}^{s=N-m^-}\binom{N}{s}(1-p)^s p^{N-s}} & \text{if Condition 1 is satisfied,} \\ \frac{q\sum_{s=m^-}^{s=N-m^+}\binom{N}{s}p^{N-s}(1-p)^s}{(1-q)\sum_{s=m^-}^{s=N-m^+}\binom{N}{s}(1-p)^{N-s}p^s} & \text{if Condition 2 is satisfied,} \\ \frac{q\sum_{s=N-\widehat{\nu}_++1}^{s=N-m^+}\binom{N}{s}p^{N-s}(1-p)^s}{(1-q)\sum_{s=N-\widehat{\nu}_++1}^{s=N-m^+}\binom{N}{s}(1-p)^{N-s}p^s} & \text{if } m^- < \widehat{\nu}_- \text{ and } m^+ < \overline{\nu}_+, \end{cases} \tag{87}$$

where Condition 1 is satisfied if either i) $m^+ \geq \overline{\nu}_+$ and $m^- \geq 0$ or ii) $\overline{\nu}_- > m^- \geq \widehat{\nu}_-$ and $r_A + m^- + 1 \leq m^+ < \overline{\nu}_+$, Condition 2 is satisfied if either i) $m^- \geq \overline{\nu}_-$ and $m^+ < \overline{\nu}_+$ or ii) $\overline{\nu}_- > m^- \geq \widehat{\nu}_-$ and $m^+ < min\left\{\overline{\nu}_+, r_A + m^- + 1\right\}$.

$\underline{\sigma}_{DM}\left(m\right) =$ *accept* if and only if $m^+ \geq \widehat{\nu}_+$ and $m^- \leq \widehat{Z}\left(m^+\right)$. The advisor's strategy is

$$\underline{\sigma}_A^T\left(h_n\right) = \begin{cases} 0 & \text{if } \nu_n^+ \geq \widehat{\nu}_+ \text{ or if } \nu_n^- \geq \widehat{\nu}_-, \\ 1 & \text{otherwise,} \end{cases} \tag{88}$$

and $\underline{\sigma}_A^M$ is given by (85) and the advisor's beliefs satisfy $\frac{\underline{\mu}_A(h_n)}{1-\underline{\mu}_A(h_n)} = l\left(h_n\right)$ for any $n \in \{0, ..., N-1\}$.

**Claim 2a:** *$\underline{\sigma}_A$ is optimal given $\underline{\sigma}_{DM}$ and $\underline{\mu}_A$.*

$\underline{\sigma}_A^M$ is optimal by Claim 1a above. If $\nu_n^- \geq \widehat{\nu}_-$ then by period $N$ it cannot be that $\nu_N^+ \geq \widehat{\nu}_+$. Therefore, there cannot be a feasible message $m \in M\left(\widetilde{h}\right)$ for which the DM accepts and, hence, there is no reason to test. In addition, by the argument in the proof of Claim 1a above, it follows that the advisor's preferred action is not affected by future outcomes if $\nu_n^- \geq \overline{\nu}_-$ or $\nu_n^+ \geq \overline{\nu}_+$. If $\nu_n^- \geq \overline{\nu}_-$, the advisor can always induce the DM to reject and, hence, there is no reason to test. If $\nu_n^+ \geq \overline{\nu}_+$ the advisor can induce the DM to accept if and only if $\nu_n^+ \geq \widehat{\nu}_+$. Hence, there is no reason to test if $\nu_n^+ \geq max\{\widehat{\nu}_+, \overline{\nu}_+\}$. Otherwise, the advisor benefits from testing either because future outcomes can affect his preferred action, or because future outcomes can affect feasible messages and, therefore, give rise to the possibility that the DM chooses the preferred action.

**Claim 2b:** *$\underline{\sigma}_{DM}$ is optimal given $\underline{\mu}_{DM}$.*

By the definition of $\widehat{\nu}_+$ in (28), the DM optimally accepts if $m = (m^+, 0)$ where $m^+ \geq \widehat{\nu}_+$. Given $m^+ \geq \widehat{\nu}_+$, the DM optimally accepts if and only if $m^- \leq \widehat{Z}(m^+)$ by definition of $\widehat{Z}(m^+)$ in (86). Suppose $max\{\widehat{\nu}_+, \overline{\nu}_+\} = \overline{\nu}_+$. Then given $m^+ < \widehat{\nu}_+$, unless $\overline{\nu}_- > m^- \geq \widehat{\nu}_-$ and $r_A + m^- + 1 \leq m^+ < \overline{\nu}_+$(i.e. unless Condition 1 applies), the DM believes that at least $\overline{\nu}_-$ outcomes in $N$ tests are negative. If exactly $\overline{\nu}_-$ outcomes in $N$ tests are negative then the advisor prefers rejection by Claim 2a, and since the advisor is more enthusiastic, it must be that the DM would optimally reject. Given $m^+ < \widehat{\nu}_+$ and Condition 1 applies, the DM must reject given the definition of $\widehat{\nu}_+$ in (28). Suppose $max\{\widehat{\nu}_+, \overline{\nu}_+\} = \widehat{\nu}_+$. Then given $m^+ < \widehat{\nu}_+$, the DM believes that at least $\nu_n^-$ outcomes in $N$ tests are negative, where $\nu_n^- > \widehat{\nu}_-$. Knowing that at least $\widehat{\nu}_+$ outcomes in $N$ tests are positive raises her posterior belief above her prior by the definition of $\widehat{\nu}_+$. Since her expected posterior must equal her prior, the complementary event that at least $\widehat{\nu}_-$ outcomes in $N$ tests are negative must lower her posterior belief below her prior and make it optimal to reject. Hence, it must also optimal to reject if at least $\nu_n^-$ outcomes in $N$ tests are negative where $\nu_n^- > \widehat{\nu}_-$.

**Claim 2c:** *There exists a sequence of completely mixed strategies $\{\sigma^k\}_{k=1}^{\infty}$, with $\lim_{k\to\infty} \sigma^k = \underline{\sigma}$, such that the system of beliefs $\underline{\mu} = \lim_{k\to\infty} \mu^k$, where $\mu^k$ denotes the beliefs derived from strategy profile $\sigma^k$ using Bayes' rule.*

See Section B.3.

**Claim 2d:** *This is an advisor-preferred equilibrium.*

By period $N$, either $\nu_n^+ = max\{\widehat{\nu}_+, \overline{\nu}_+\}$ and the DM accepts or $\nu_n^- = min\{\widehat{\nu}_-, \overline{\nu}_-\}$ and the DM rejects. If $max\{\widehat{\nu}_+, \overline{\nu}_+\} = \overline{\nu}_+$ then the advisor achieves his first-best payoff by the same reasoning as in Claim 1e. If $max\{\widehat{\nu}_+, \overline{\nu}_+\} = \widehat{\nu}_+$ , then given $\nu_n^+ = \widehat{\nu}_+ \geq \overline{\nu}_+$ the advisor prefers acceptance by Claim 2a. However, if $\nu_n^- = \widehat{\nu}_-$ the advisor may prefer acceptance. To increase the advisor's payoff the DM would have to accept if $m^+ \geq \widehat{\nu}$ for some $\widehat{\nu} < \widehat{\nu}_+$. But this cannot be part of an equilibrium because if the DM were to accept for $m^+ \geq \widehat{\nu}$ then the advisor would stop testing as soon as $\nu_n^+ = \widehat{\nu}$ and send $m = (\widehat{\nu}, 0)$. Given $m = (\widehat{\nu}, 0)$, in equilibrium the DM would infer that at least $\widehat{\nu}$ outcomes in $N$ tests were positive. But then it is optimal for the DM to reject given the definition of $\widehat{\nu}_+$ in (28).

## A.12 Theorem 2 [DM Payoff Comparison for $N > 2$]

The proof uses the definition of an equilibrium acceptance set and a first-best acceptance set (see Definitions 1 and 2).

**Part 1a.** Suppose $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$, where

$$\mathcal{I} \equiv \{(\lambda_A, \lambda_{DM}) : l(k) < \lambda_A \leq l(N), l(0) < \lambda_{DM} \leq l(N-2)\}, \tag{89}$$

and $k$ is a function of $\lambda_{DM}$ and given by (81) in the proof of Lemma 6. $k$ depends on $r_{DM}$ defined by (24). Since $\lambda_{DM} \leq l(N-2)$ implies $r_{DM} \leq N-3$, $k \leq N-1$. By Part 1 of Lemma 6, since $l(k) < \lambda_A$ in the unique equilibrium $(\sigma^o, \mu^o)$ of $\Gamma^o$, the advisor does not test and the DM always rejects. Hence, $\Phi_{(\sigma^o,\mu^o;\Gamma^o)} = \emptyset$.

For an advisor-preferred equilibrium $(\underline{\sigma}, \underline{\mu})$ of $\Gamma$, apply Part 1 of Lemma 7. Given $l(k) < \lambda_A \leq l(N)$, then $k \leq r_A \leq N-1$ where $r_A$ is defined by (25). Hence, $N \geq \overline{\nu}_+ \geq \frac{k+1+N}{2}$, where $\overline{\nu}_+$ is given by (26). In period $N$, either $\nu_N^+ = \overline{\nu}_+$ and the advisor reports $m = (\overline{\nu}_+, 0)$ and the DM accepts, or $\nu_N^- = N - \overline{\nu}_+ + 1$ and the advisor reports $m = (0, N - \overline{\nu}_+ + 1)$ and the DM rejects. As shown in the proof of Part 1 of Lemma 7, $\Phi_{(\underline{\sigma},\underline{\mu};\Gamma)} = \Phi_A^{FB}$. Since $\lambda_A \leq l(N)$, the advisor prefers acceptance if all $N$ outcomes are positive and, hence, $\Phi_{(\sigma^o,\mu^o;\Gamma^o)} \subset \Phi_{(\underline{\sigma},\underline{\mu};\Gamma)}$. Since $\lambda_A > l(k) > \lambda_{DM}$, $\Phi_A^{FB} \subset \Phi_{DM}^{FB}$. Hence, $\Phi_{(\sigma^o,\mu^o;\Gamma^o)} \subset \Phi_{(\underline{\sigma},\underline{\mu};\Gamma)} \subset \Phi_{DM}^{FB}$. By Lemma 9, $\pi_{DM}(\sigma^o, \mu^o; \Gamma^o) < \pi_{DM}(\underline{\sigma}, \underline{\mu}; \Gamma) < \pi_{DM}^{FB}$. The result follows by Lemma 2.

**Part 1b.** Suppose $(\lambda_A, \lambda_{DM}) \in W_{\mathcal{I}}$, where

$$W_{\mathcal{I}} \equiv \{(\lambda_A, \lambda_{DM}) : l(N-2) < \lambda_A \leq l(N-1), l(N-3) < \lambda_{DM} \leq l(N-2)\}. \tag{90}$$

I show that in the unique equilibrium $(\sigma^o, \mu^o)$ of $\Gamma^o$, $\sigma_{DM}^o(h_N) = accept$ if and only if $x_N \geq N-2$, $\sigma_A^o(h_n) = 1$ if $n-1 \leq x_n \leq N-2$ and, otherwise, $\sigma_A^o(h_n) = 0$, and $\frac{\mu_A^o(h_n)}{1-\mu_A^o(h_n)} = l_n$ for $n \in \{0, ..., N-1\}$. First, given $l_n = l(n-2)$, the advisor prefers rejection irrespective of the remaining outcomes, because even if the remaining $N-n$ outcomes are positive, $l_N = l(n-2+(N-n)) = l(N-2) < \lambda_A$. For any $n \leq N-1$, if the advisor stops at $l_n = l(n-2)$, the DM rejects since $l_N \leq l(N-3) < \lambda_{DM}$. Therefore, if $l_n \leq l(n-2)$, there is no reason to test. Second, given $l_n = l(N-1)$, it must be that $n \geq N-1$, and the DM accepts regardless of the remaining test outcome, because even if $n = N-1$ and the remaining test outcome is negative then $\lambda_{DM} \leq l(N-2)$. Therefore, there is no reason to test if $l_n \geq l(N-1)$. Finally, given $l(N-1) \leq l_n \leq l(N-2)$, there is a reason to test. If the advisor continues testing, the remaining outcomes either lead to $l_N = l(N-1)$ and the DM accepts, or they lead to $l_N = l(N-3)$ and the DM rejects. In both cases, the DM's action is in line with the advisor's interest. Hence, it is optimal to test. On the equilibrium path, the advisor stops if and only if either he has found a single negative outcome or he has run $N-1$ tests. If he finds $N-1$ positive outcomes, the DM accepts, otherwise, the DM rejects. Hence, $\Phi_{(\sigma^o,\mu^o;\Gamma^o)} = \{\phi \in \Phi | s_1 = ... = s_{N-1} = (+)\}$.

For an advisor-preferred equilibrium $\left(\underline{\sigma}, \underline{\mu}\right)$ of $\Gamma$, apply Part 1 of Lemma 7. Given (25), $r_A = N-1$ and, hence, $\overline{\nu}_+ = N$ and $\overline{\nu}_- = 1$. In period $N$, either $\nu_n^+ = N$ and the advisor reports $m = (N, 0)$ and the DM accepts, or $\nu_n^- = 1$ and the advisor reports $m = (0, 1)$ and the DM rejects. Hence, $\Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} = \{(+, ..., +)\}$.

Therefore, $\Phi_{(\sigma^o,\mu^o;\Gamma^o)} = \Phi_{\left(\underline{\sigma},\underline{\mu};\Gamma\right)} \cup \{(+, ..., +, -)\}$. Given $\phi = (+, ..., +, -)$, the DM prefers acceptance since $\lambda_{DM} \leq l\,(N-2)$. Hence, $\pi_{DM}\left(\underline{\sigma}, \underline{\mu}; \Gamma\right) < \pi_{DM}\left(\sigma^o, \mu^o; \Gamma^o\right)$.

**Part 2a.** Suppose $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$, where

$$\mathcal{S} \equiv \left\{(\lambda_A, \lambda_{DM}) : \lambda_A < l\,(N-2)\,, \bar{l} < \lambda_{DM} \leq l\,(N-1)\right\}, \tag{91}$$

and

$$\bar{l} \equiv \frac{q \sum_{s=N-1}^{s=N} \binom{N}{s} p^s (1-p)^{N-s}}{(1-q) \sum_{s=N-1}^{s=N} \binom{N}{s} (1-p)^s p^{N-s}}, \tag{92}$$

i.e. $\bar{l}$ is the likelihood ratio conditional on at least $N-1$ outcomes in $N$ tests being positive. Note that if $N = 2$ then $\mathcal{S} = S$.

For the unique equilibrium $(\sigma^o, \mu^o)$ of $\Gamma^o$, apply Part 2 of the proof of Lemma 6 where $r_{DM}$ is defined by (24) and, hence, $r_{DM} = N-2$. Then $\underline{\lambda}_A = l\,(2\,(r_{DM}+1) - N) = l\,(N-2)$. Since $\lambda_A < l\,(N-2) = \underline{\lambda}_A$, the advisor stops as soon as the likelihood ratio satisfies $l_n = l\,(r_{DM}+1) = l\,(N-1)$ or $l_n = l\,(r_{DM} - (N-n)) = l\,(n-2)$. If $l_N = l\,(N-1)$ then the DM accepts since $\lambda_{DM} \leq l\,(N-1)$, otherwise the DM rejects.

For an advisor-preferred equilibrium $\left(\underline{\sigma}, \underline{\mu}\right)$ of $\Gamma$, apply Part 2 of the proof of Lemma 7.

$$\widehat{\nu}_+ \equiv inf\left\{j | j \in \mathbb{N}, \lambda_{DM} \leq \frac{q \sum_{s=j}^{s=N} \binom{N}{s} p^s (1-p)^{N-s}}{(1-q) \sum_{s=j}^{s=N} \binom{N}{s} (1-p)^s p^{N-s}}\right\} = N, \tag{93}$$

since

$$l\,(N) = \frac{qp^N}{(1-q)(1-p)^N} > \lambda_{DM} > \frac{q \sum_{s=N-1}^{s=N} \binom{N}{s} p^s (1-p)^{N-s}}{(1-q) \sum_{s=N-1}^{s=N} \binom{N}{s} (1-p)^s p^{N-s}} \equiv \bar{l}. \tag{94}$$

Hence, $\widehat{\nu}_- \equiv N - \widehat{\nu}_+ + 1 = 1$. Given $\lambda_A < l\,(N-2)$, then $r_A$ defined by (25) satisfies $r_A \leq N-3$ and hence $\overline{\nu}_+ \leq N-1$. Therefore, $max\left\{\overline{\nu}_+, \widehat{\nu}_+\right\} = \widehat{\nu}_+$. Hence, the advisor tests until either $N$ outcomes are positive or one outcome is negative. If $N$ outcomes are positive, he reports $m = (N, 0)$ and the DM accepts. If one outcome is negative, the advisor reports $m = (0, 1)$ if $l\,(-1) < \lambda_A$ or $m = (0, 0)$ if $\lambda_A \leq l\,(-1)$ and in either case the DM rejects.

Consider the set $\Phi$ of the lists of Nature's $N$ draws given by (33). The DM optimally accepts if all $N$ outcomes are positive since $\lambda_{DM} < l(N)$, but not if $N-1$ outcomes in $N$ tests are positive since $l((N-1)-1) = l(N-2) < \lambda_{DM}$. Hence, the DM's first-best acceptance set is $\Phi^{FB}_{DM} = \{(+,...,+)\}$. In any $(\underline{\sigma},\underline{\mu})$ of $\Gamma$, the DM accepts if and only if all $N$ test outcomes are positive. Hence, $\Phi_{(\underline{\sigma},\underline{\mu};\Gamma)} = \Phi^{FB}_{DM}$. For the unique equilibrium $(\sigma^o,\mu^o)$ of $\Gamma^o$, the DM accepts if and only if the first $N-1$ in $N$ test outcomes are positive. Hence, $\Phi_{(\sigma^o,\mu^o;\Gamma^o)} = \{\phi \in \Phi | s_1 = ... = s_{N-1} = (+)\}$. Hence, $\Phi_{(\sigma^o,\mu^o;\Gamma^o)} \subset \Phi_{(\underline{\sigma},\underline{\mu};\Gamma)} = \Phi^{FB}_{DM}$. Therefore, $\pi_{DM}(\sigma^o,\mu^o;\Gamma^o) < \pi_{DM}(\underline{\sigma},\underline{\mu};\Gamma) = \pi^{FB}_{DM}$. The result follows by Lemma 2.

**Part 2b.** Suppose $(\lambda_A,\lambda_{DM}) \in W_{\mathcal{S}}$, where

$$W_{\mathcal{S}} \equiv \left\{(\lambda_A,\lambda_{DM}) : \lambda_A < l(N-4), \bar{\bar{l}} < \lambda_{DM} \leq \bar{l}\right\}, \tag{95}$$

and $\bar{l}$ is defined by (92) and

$$\bar{\bar{l}} \equiv \frac{q\sum_{s=N-2}^{s=N}\binom{N}{s}p^s(1-p)^{N-s}}{(1-q)\sum_{s=N-2}^{s=N}\binom{N}{s}(1-p)^s p^{N-s}}, \tag{96}$$

i.e. $\bar{\bar{l}}$ is the likelihood ratio conditional on at least $N-2$ outcomes in $N$ tests being positive. $r_{DM}$ is defined by (24) and satisfies $r_{DM} = N-2$ since $l(N-2) < \bar{\bar{l}}$ and $\bar{l} < l(N-1)$. For the unique equilibrium $(\sigma^o,\mu^o)$ of $\Gamma^o$, the same argument as in Part 2a applies. For an advisor-preferred equilibrium $(\underline{\sigma},\underline{\mu})$ of $\Gamma$, apply Part 2 of the proof of Lemma 7.

$$\widehat{\nu}_+ \equiv inf\left\{j | j \in \mathbb{N}, \lambda_{DM} \leq \frac{q\sum_{s=j}^{s=N}\binom{N}{s}p^s(1-p)^{N-s}}{(1-q)\sum_{s=j}^{s=N}\binom{N}{s}(1-p)^s p^{N-s}}\right\} = N-1, \tag{97}$$

since

$$\bar{l} \equiv \frac{q\sum_{s=N-1}^{s=N}\binom{N}{s}p^s(1-p)^{N-s}}{(1-q)\sum_{s=N-1}^{s=N}\binom{N}{s}(1-p)^s p^{N-s}} \geq \lambda_{DM} > \frac{q\sum_{s=N-2}^{s=N}\binom{N}{s}p^s(1-p)^{N-s}}{(1-q)\sum_{s=N-2}^{s=N}\binom{N}{s}(1-p)^s p^{N-s}} \equiv \bar{\bar{l}}. \tag{98}$$

Hence, $\widehat{\nu}_- \equiv N - \widehat{\nu}_+ + 1 = 2$. $r_A$ is defined by (25) and satisfies $r_A \leq N-5$ since $\lambda_A \leq l(N-4)$. Hence, $\overline{\nu}_+ \leq N-2$. Therefore, $max\{\overline{\nu}_+,\widehat{\nu}_+\} = \widehat{\nu}_+$. Hence, the advisor tests until either $N-1$ outcomes are positive or two outcomes are negative. If $N-1$ outcomes are positive, he reports $m = (N-1,0)$ and the DM accepts. If two outcomes are negative, the advisor reports $m = (0,2)$ if $l(-2) < \lambda_A$ or $m = (0,0)$ if $\lambda_A \leq l(-2)$ and in either case the DM rejects.

Consider the set $\Phi$ of the lists of Nature's $N$ draws given by (33). Let $\Phi'$ denote the set of lists that contain exactly one negative outcome and in which this negative outcomes is drawn before the final period, i.e.

$$\Phi' \equiv \{\phi \in \Phi | s_i = (-), s_{j \neq i} = (+), i \in \{1, ..., N-1\}, j \in \{1, ..., N\}\}. \tag{99}$$

If $\phi \in \Phi'$, the DM accepts in $(\underline{\sigma}, \underline{\mu})$ but rejects in $(\sigma^o, \mu^o)$. Otherwise, the DM chooses the same action in $(\underline{\sigma}, \underline{\mu})$ and $(\sigma^o, \mu^o)$. Therefore, $\Phi_{(\sigma^o, \mu^o; \Gamma^o)} \cup \Phi' = \Phi_{(\underline{\sigma}, \underline{\mu}; \Gamma)}$. If $\phi \in \Phi'$, the DM prefers rejection since $l((N-1)-1) = l(N-2) < \lambda_{DM}$. Hence, $\pi_{DM}(\underline{\sigma}, \underline{\mu}; \Gamma) < \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$.

## A.13 Corollary 3 [Preference Alignment for $N > 2$]

**Claim 1:** *Take any* $(\lambda_A, \lambda_{DM})$. *For any* $N > 2$ *and* $(\sigma, \mu) \in E$, *if* $\lambda_{DM} - \lambda_A > \bar{l}$ *then* $\pi_{DM}(\sigma, \mu; \Gamma) \geq \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$, *where* $\bar{l}$ *is defined by (92).*

At $(\lambda_A, \lambda_{DM}) = (0, \bar{l})$, $\lambda_{DM} - \lambda_A = \bar{l}$. I consider all $(\lambda_A, \lambda_{DM})$ for which $\lambda_{DM} \geq \bar{l} + \lambda_A$. First, if $\lambda_{DM} \geq \bar{l} + \lambda_A$ where $\bar{l} < \lambda_{DM} \leq l(N-1)$ and $\lambda_A < l(N-1-\bar{l})$, then $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$ and $\pi_{DM}(\sigma, \mu; \Gamma) > \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ for any $(\sigma, \mu) \in E$ by the proof of Part 2a of Theorem 2. Second, if $\lambda_{DM} \geq \bar{l} + \lambda_A$ where $l(N-1) < \lambda_{DM} \leq l(N)$ and $\lambda_A \leq l(N-\bar{l})$, then $\Phi_{(\sigma^o, \mu^o; \Gamma^o)} = \Phi_{(\underline{\sigma}, \underline{\mu}; \Gamma)} = \{(+, ..., +)\}$ for any $(\sigma, \mu) \in E$. The reason is that in $\Gamma^o$, independent of $\sigma_A^o$, the DM's optimal strategy is $\sigma_{DM}^o(h_N) =$ *accept* if and only if $h_N = (+, ..., +)$. In $\Gamma$, independent of $\sigma_A$, her optimal strategy is $\sigma_{DM}(m) =$ *accept* if and only if $m = (N, 0)$. Since the advisor prefers acceptance given $h_N = (+, ..., +)$ as $\lambda_A \leq l(N-\bar{l}) < l(N)$, he runs all $N$ tests and, if testing is hidden, discloses all $N$ outcomes if $h_N = (+, ..., +)$. Hence, $\Phi_{(\sigma^o, \mu^o; \Gamma^o)} = \Phi_{(\underline{\sigma}, \underline{\mu}; \Gamma)}$. Third, if $\lambda_{DM} \geq \bar{l} + \lambda_A$ where $\lambda_{DM} > l(N)$, then in $\Gamma^o$, independent of $\sigma_A^o$, the DM's optimal strategy is $\sigma_{DM}^o(h_N) =$ *reject* for any $h_N$ and, in $\Gamma$, independent of $\sigma_A$, her optimal strategy is $\sigma_{DM}(m) =$ *reject* for any $m = (0, N)$. Hence, $\Phi_{(\sigma^o, \mu^o; \Gamma^o)} = \Phi_{(\underline{\sigma}, \underline{\mu}; \Gamma)}$. Claim 1 follows by Lemma 2.

**Claim 2:** *Take any* $(\lambda_A, \lambda_{DM})$. *For any* $N > 2$ *and* $(\sigma, \mu) \in E$, *if* $\lambda_A - \lambda_{DM} > l(\underline{k}) - l(0)$ *where* $\underline{k} \equiv inf\left\{j | j \in \mathbb{N}, j \geq \frac{N+1}{2}\right\}$ *then* $\pi_{DM}(\sigma, \mu; \Gamma) \geq \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$.

At $(\lambda_A, \lambda_{DM}) = (l(\underline{k}), l(0))$, $\lambda_A - \lambda_{DM} = l(\underline{k}) - l(0)$, where $\underline{k}$ is chosen to be $k$ defined by (81) evaluated at $r_{DM} = 0$, which is the minimum value of $k$ for any $r_{DM}$. I consider all $(\lambda_A, \lambda_{DM})$ for which $l(0) \leq \lambda_{DM} \leq l(0) - l(\underline{k}) + \lambda_A$. First, if $l(0) \leq \lambda_{DM} \leq l(0) - l(\underline{k}) + \lambda_A$ where $l(\underline{k}) < \lambda_A \leq l(N)$ and $l(0) \leq \lambda_{DM} \leq l(0) - l(\underline{k}) + l(N) < l(N-2)$ then $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$ and, by the proof of of Part 1a of Theorem 2, $\pi_{DM}(\sigma, \mu; \Gamma) > \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ for any $(\sigma, \mu) \in E$. Second, if $l(0) \leq \lambda_{DM} \leq l(0) - l(k) + \lambda_A$ where $\lambda_{DM} > l(N)$ then by the same reasoning as in Claim 1 above, $\Phi_{(\sigma^o, \mu^o; \Gamma^o)} = \Phi_{(\underline{\sigma}, \underline{\mu}; \Gamma)}$. Claim 2 follows by Lemma 2.

To prove Corollary 3 choose $d = \max\left\{\bar{l}, l(\underline{k}) - l(0)\right\}$.

## A.14 Proposition 3 [Further Payoff Comparisons for $N > 2$]

**(First-Best Benchmark):** Suppose $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$ where $\mathcal{S}$ is defined in (91). By the proof of Part 2a of Theorem 2, if $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$, then $\pi_{DM}\left(\underline{\sigma}, \underline{\mu}; \Gamma\right) = \pi_{DM}^{FB}$. By Lemma 2, then $\pi_{DM}(\sigma, \mu; \Gamma) = \pi_{DM}^{FB}$ for any $(\sigma, \mu) \in E$.

**(Advisor Payoff Comparison):** Suppose $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$ where $\mathcal{S}$ is defined in (91). By the proof of Part 2a of Theorem 2, $\Phi_{(\sigma^o, \mu^o; \Gamma^o)} = \Phi_{\left(\underline{\sigma}, \underline{\mu}; \Gamma\right)} \cup \{(+, ..., +, -)\}$. The advisor prefers acceptance if exactly $N-1$ in $N$ outcomes are positive since $\lambda_A < l\left((N-1)-1\right)$. Hence, $\pi_A(\sigma^o, \mu^o; \Gamma^o) > \pi_A\left(\underline{\sigma}, \underline{\mu}; \Gamma\right)$. By definition of advisor-preferred equilibrium, then $\pi_A(\sigma^o, \mu^o; \Gamma^o) > \pi_A(\sigma, \mu; \Gamma)$ for any $(\sigma, \mu) \in E$.

**(Hidden Testing Pareto-improving):** Suppose $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$ where $\mathcal{I}$ is defined in (89). By the proof of Part 1a of Theorem 2, $\pi_{DM}(\sigma, \mu; \Gamma) > \pi_{DM}(\sigma^o, \mu^o; \Gamma^o)$ for any $(\sigma, \mu) \in E$ and $\Phi_{\left(\underline{\sigma}, \underline{\mu}; \Gamma\right)} = \Phi_A^{FB}$ and $\Phi_{(\sigma^o, \mu^o; \Gamma^o)} = \emptyset$. By Lemmas 9 and 2, then $\pi_A(\sigma, \mu; \Gamma) = \pi_A^{FB} > \pi_A(\sigma^o, \mu^o; \Gamma^o)$ for any $(\sigma, \mu) \in E$.

## A.15 Proposition 4 [Delegation]

Denote player $i$'s first-best acceptance set by $\Phi_i^{FB}$ (see Definition 2). When the advisor has decision-making authority, the equilibrium acceptance set must be given by $\Phi_{\left(\sigma^d, \mu^d; \Gamma^d\right)} = \Phi_A^{FB}$, because the advisor can run all $N$ tests and then accept if and only if $\widehat{u}_A(\mathit{accept}) > \widehat{u}_{DM}(\mathit{reject})$.

**Part 1.** Consider $\Gamma$ and assume the advisor runs all tests. Suppose $\lambda_A \leq \lambda_{DM}$, which implies $\Phi_A^{FB} \supseteq \Phi_{DM}^{FB}$ by Remark 2. By Lemma 2, for any $(\sigma, \mu) \in E$, $\Phi_{\left(\underline{\sigma}, \underline{\mu}; \Gamma\right)} \supseteq \Phi_{(\sigma, \mu; \Gamma)} \subseteq \Phi_{(\overline{\mu}, \overline{\sigma}; \Gamma)}$ and, hence, $\Phi_{\left(\sigma^d, \mu^d; \Gamma^d\right)} \supseteq \Phi_{(\sigma, \mu; \Gamma)} \subseteq \Phi_{(\overline{\mu}, \overline{\sigma}; \Gamma)}$. For any $(\sigma, \mu) \in E$, either $\Phi_{(\sigma, \mu; \Gamma)} \subseteq \Phi_{DM}^{FB}$ or $\Phi_{DM}^{FB} \subseteq \Phi_{(\sigma, \mu; \Gamma)} = \Phi_{(\overline{\mu}, \overline{\sigma}; \Gamma)}$. The statement follows by Lemma 9. Suppose $\lambda_A \geq \lambda_{DM}$. By the proof of Part 1 of Lemma 7, $\Phi_{\left(\underline{\sigma}, \underline{\mu}; \Gamma\right)} = \Phi_A^{FB}$. Hence, $\Phi_{\left(\sigma^d, \mu^d; \Gamma^d\right)} = \Phi_{\left(\underline{\sigma}, \underline{\mu}; \Gamma\right)}$. The statement follows by Lemma 9. If the advisor choosing $\sigma_A^d$ optimally rather than running all tests, the results are unchanged by the same reasoning as in Lemma 2.

**Part 2.** Consider $\Gamma^o$. Suppose $(\lambda_A, \lambda_{DM}) \in \mathcal{I} \cup W_{\mathcal{I}}$, as defined in (89) and (90). By the proof of Part 1 of Lemma 7, $\Phi_{\left(\underline{\sigma}, \underline{\mu}; \Gamma\right)} = \Phi_A^{FB}$. Hence, $\Phi_{\left(\sigma^d, \mu^d; \Gamma^d\right)} = \Phi_{\left(\underline{\sigma}, \underline{\mu}; \Gamma\right)}$. Then Part 2a follows by Part 1a of Theorem 2. In addition, Part 2b conditional on $(\lambda_A, \lambda_{DM}) \in W_{\mathcal{I}}$ follows by Part 1b of Theorem 2. Suppose $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$, as defined in (91). Then by Part 2a of Theorem 2, $\Phi_{(\sigma^o, \mu^o; \Gamma^o)} = \{\phi \in \Phi | s_1 = ... = s_{N-1} = (+)\}$. The advisor prefers acceptance if exactly $N-1$ in $N$ outcomes are positive since $\lambda_A < l(N-2)$ and, hence, $\Phi_A^{FB} = \Phi_{\left(\sigma^d, \mu^d; \Gamma^d\right)} \supset \Phi_{(\sigma^o, \mu^o; \Gamma^o)}$. The DM prefers rejection if exactly $N-1$ in $N$ outcomes are positive since $\lambda_{DM} > l(N-2)$ and, hence, $\Phi_{\left(\sigma^d, \mu^d; \Gamma^d\right)} \supset \Phi_{(\sigma^o, \mu^o; \Gamma^o)} \supset \Phi_{DM}^{FB}$. By Lemma 9, Part 2b conditional on $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$ follows.

## A.16 Lemma 8 [DM Commitment]

Suppose $\lambda_A \leq \lambda_{DM}$. Under observable testing, $\sigma^o_{C,DM}(h_N) = \textit{accept}$ if and only if $\lambda_{DM} \leq l_N$. The advisor never strictly prefers $\sigma^o_{C,A}(h_n) = 0$ at $n < N$. This is because if additional tests lead the DM to accept then the advisor also prefers acceptance given $\lambda_A < \lambda_{DM} \leq l_N$. Hence, the DM's strategy is optimal.

Under hidden testing, $\sigma_{C,DM}(m) = \textit{accept}$ if and only if $m = (m^+, m^-)$ where $m^+ + m^- = N$ and $\lambda_{DM} \leq l(m^+ - m^-)$. Given $h_N$ where $\nu^+_N + \nu^-_N < N$, there exists no feasible message $m \in M\left(\widetilde{h}\right)$ such that $\sigma_{C,DM}(m) = \textit{accept}$. Hence, the advisor is indifferent between all $m \in M\left(\widetilde{h}\right)$. Given $h_N$ where $\nu^+_N + \nu^-_N = N$ the advisor weakly prefers $\sigma^M_{C,A}(h_N) = \left(\nu^+_N, \nu^-_N\right)$. This is because if disclosing additional outcomes leads the DM to accept, then the advisor also prefers acceptance given $\lambda_A < \lambda_{DM} \leq l_N$. Therefore, the advisor also never strictly prefers $\sigma^T_{C,A}(h_n) = 0$ at $n < N$. Hence, the DM's strategy is optimal. The argument can be adopted for the case when $\lambda_A \geq \lambda_{DM}$.

## A.17 Proposition 5 [Commitment vs. Hidden Testing]

By Proposition 8, under observable testing, if the DM has the power to commit to $\sigma^o_{C,DM}$ then the DM's payoff is $\pi^{FB}_{DM}$.

**Part 1.** Suppose $(\lambda_A, \lambda_{DM}) \in \mathcal{S}$, where $\mathcal{S}$ is defined in (91). By the proof of Part 2a of Theorem 2, $\pi_{DM}(\sigma, \mu; \Gamma) = \pi^{FB}_{DM}$ for any $(\sigma, \mu) \in E$ and $\pi_{DM}(\sigma^o, \mu^o; \Gamma^o) < \pi^{FB}_{DM}$.

**Part 2.** Suppose $(\lambda_A, \lambda_{DM}) \in \mathcal{I}$, where $\mathcal{I}$ is defined in (89). By the proof of Part 1a of Theorem 2, $\pi_{DM}(\sigma^o, \mu^o; \Gamma^o) < \pi_{DM}(\sigma, \mu; \Gamma) < \pi^{FB}_{DM}$ for any $(\sigma, \mu) \in E$.

## A.18 Proposition 6 [Advisor Commitment]

Under observable testing, the equilibrium is unaffected by whether or not the advisor has commitment power. The advisor moves first and has no private information.

**Part 1.** Consider hidden testing and suppose $\lambda_{DM} \geq \lambda_A$. If the advisor commits to full disclosure, i.e. $\sigma^M_{CA,A}(h_N) = \left(\nu^+_N, \nu^-_N\right)$ for any $h_N$, then in the advisor-preferred equilibrium under hidden testing with commitment, expected payoffs are as under observable testing.

**Part 2.** Consider hidden testing and suppose $\lambda_{DM} \leq \lambda_A$. By the proof of Lemma 7, $\Phi_{\left(\underline{\sigma}, \underline{\mu}; \Gamma\right)} = \Phi^{FB}_A$. Therefore, the advisor optimally commits to the same strategy he would choose in a game without commitment. The statement follows by Part 1a of Theorem 2.

# References

R. Argenziano, S. Severinov, and F. Squintani. Strategic information acquisition and transmission. *American Economic Journal: Microeconomics*, 8(3):119–55, August 2016.

A. Berenson. "Eli Lilly Said to Play Down Risk of Top Pill". *New York Times*, December 2006.

I. Brocas and J. Carrillo. Influence through Ignorance. *The RAND Journal of Economics*, 38(4):931–947, 2007.

Y.-K. Che and N. Kartik. Opinions as Incentives. *Journal of Political Economy*, 117(5): 815–860, 2009.

M. Dahm, P. González, and N. Porteiro. Trials, tricks and transparency: How disclosure rules affect clinical knowledge. *Journal of Health Economics*, 28(6):1141 – 1153, 2009. ISSN 0167-6296.

M. DeGroot. New York: McGraw-Hill, 1970.

I. Deimen and D. Szalay. Information, Authority, and Smooth Communication in Organizations. *American Economic Review, forthcoming*, (10969), Dec 2018.

A. Di Tillio, M. Ottaviani, and P. N. Sørensen. Persuasion Bias in Science: Can Economics Help? *The Economic Journal*, 127(605):F266–F304, 2017.

A. Di Tillio, M. Ottaviani, and P. N. Sorensen. Strategic Sample Selection. CEPR Discussion Papers 12202, C.E.P.R. Discussion Papers, Aug 2018.

R. Dur and O. Swank. Producing and Manipulating Information. *The Economic Journal*, 115(500):185–199, 2005.

W. Dziuda. Strategic argumentation. *Journal of Economic Theory*, 146(4):1362–1397, 2011.

J. Farrell and J. Sobel. Voluntary Revelation of Product Information. *Unfinished Mimeo*, 1983.

M. Felgenhauer and P. Loerke. Bayesian Persuasion with Private Experimentation. *International Economic Review*, 58:829–856, 2017.

M. Felgenhauer and E. Schulte. Strategic Private Experimentation. *American Economic Journal: Microeconomics*, 6(4):74–105, 2014.

D. Gerardi and L. Yariv. Costly Expertise. *American Economic Review*, 98(2):187–93, May 2008.

B. Goldacre, S. Lane, K. Mahtani, C. Heneghan, I. Onakpoya, I. Bushfield, and L. Smeeth. Pharmaceutical companies' policies on access to trial data, results, and methods: audit study. *BMJ*, 358, 2017.

S. Grossman. The Informational Role of Warranties and Private Disclosure about Product Quality. *The Journal of Law & Economics*, 24(3):461–483, 1981.

S. Hart, I. Kremer, and M. Perry. Evidence Games: Truth and Commitment. *American Economic Review*, 107(3):690–713, March 2017.

E. Henry. Strategic Disclosure of Research Results: The Cost of Proving Your Honesty. *The Economic Journal*, 119(539):1036–1064, 2009.

E. Henry and M. Ottaviani. Research and the Approval Process: The Organization of Persuasion. *American Economic Review, forthcoming*, 2018.

T. Jefferson, MA. Jones, P Doshi, CB Del Mar, R Hama, MJ Thompson, EA Spencer, IJ Onakpoya, KR Mahtani, D Nunan, J Howick, and CJ Heneghan. Neuraminidase inhibitors for preventing and treating influenza in adults and children. *Cochrane Database of Systematic Reviews*, (4), 2014. ISSN 1465-1858.

E. Kamenica and M. Gentzkow. Bayesian Persuasion. *American Economic Review*, 101(6): 2590–2615, October 2011.

G. Kolata. "A Cancer Conundrum: Too Many Drug Trials, Too Few Patients". *New York Times*, August 2017.

H. Li and W. Suen. Delegating Decisions to Experts. *Journal of Political Economy*, 112(S1): S311–S335, 2004.

J. Liebgober. False Positives and Transparency in Scientific Research. *Unpublished Manuscript*, 2017.

B. Mackowiak and M. Wiederholt. Information Processing and Limited Liability. *American Economic Review*, 102(3):30–34, May 2012.

S. Matthews and A. Postlewaite. Quality Testing and Disclosure. *The RAND Journal of Economics*, 16(3):328–340, 1985.

A. McClellan. Experimentation and Approval Mechanisms. *Unpublished working paper*, 2017.

P. Milgrom. Good News and Bad News: Representation Theorems and Applications. *The Bell Journal of Economics*, 12(2):380–391, 1981.

P. Milgrom and J. Roberts. Price and Advertising Signals of Product Quality. *Journal of Political Economy*, 94(4):796–821, 1986.

R. C. Rabin. "Suicide Data Incorrectly Reported in Drug Trials, Suit Claimed". *New York Times*, September 2017.

S. Shavell. Acquisition and Disclosure of Information Prior to Sale. *The RAND Journal of Economics*, 25(1):20–36, 1994.

# B Supplementary Appendix

## Contents



### B.1 Lemma 11: Advisor-preferred Equilibria

To show that the equilibria are sequential equilibria, I consider a completely mixed strategy for the advisor as a function of some $\varepsilon \in (0,1)$, denoted by $\left(\underline{\sigma}_A^{T\varepsilon}, \underline{\sigma}_A^{M\varepsilon}\right)$, which is constructed as follows:

$$\underline{\sigma}_A^{T\varepsilon}(h_0) = \begin{cases} 1-\varepsilon & \text{if } \underline{\sigma}_A^{T}(h_0) = 1 \\ \varepsilon & \text{if } \underline{\sigma}_A^{T}(h_0) = 0, \end{cases} \tag{100}$$

$$\underline{\sigma}_A^{T\varepsilon}(h_1) = \begin{cases} 1-\varepsilon^2 & \text{if } \underline{\sigma}_A^{T}(h_1) = 1 \\ \varepsilon^2 & \text{if } \underline{\sigma}_A^{T}(h_1) = 0, \end{cases} \tag{101}$$

and for any $\widetilde{h}$ and associated set of feasible messages $M\left(\widetilde{h}\right)$, $\underline{\sigma}_A^M(m)$ is sent with probability $1-\left(|M\left(\widetilde{h}\right)|-1\right)\varepsilon^4$ and any other message $m \in M\left(\widetilde{h}\right)$ is sent with probability $\varepsilon^4$, where $|M\left(\widetilde{h}\right)|$ denotes the total number of feasible messages given $\widetilde{h}$. Clearly, as $\varepsilon \to 0$, $\left(\underline{\sigma}_A^{T\varepsilon}, \underline{\sigma}_A^{M\varepsilon}\right) \to \left(\underline{\sigma}_A^{T}, \underline{\sigma}_A^{M}\right)$. In addition, I show that the beliefs derived from the advisor's mixed strategy profile using Bayes' rule, denoted by $\underline{\mu}^\varepsilon(m)$, satisfy

$$\frac{\underline{\mu}^\varepsilon(m)}{1-\underline{\mu}^\varepsilon(m)} \to \frac{\underline{\mu}(m)}{1-\underline{\mu}(m)}. \tag{102}$$

Note that this must trivially be the case for any message $m \in \{\{+,+\},\{+,-\},\{-,-\}\}$ due to verifiable disclose.

**Region 1a and Region 2a)**, i.e. $\lambda_A < l_{(+,-)}$ and $\lambda_{DM} < l_{\neg(-,-)}$. In the limit as $\varepsilon \to 0$, $m = \{-\}$ is mostly likely observed following $h_2 \in \{(-,-),(-,+)\}$. $m = \{-\}$ always requires a deviation from $\underline{\sigma}_A^M$ . In addition, each $h_2 \in \{(-,-),(-,+)\}$ occurs on path, whereas any other history for which $m = \{-\}$ is feasible is reached only after a deviation from $\underline{\sigma}_A^T$.

**Region 1b).** In the limit as $\varepsilon \to 0$, $m = \{+\}$ is most likely observed following $h_2 = (\emptyset,+)$. $h_2 = (\emptyset,+)$ occurs with a probability of order $\varepsilon$, since it results in $m = \{+\}$ given $\underline{\sigma}_A^M$

and arises after a single deviation from $\underline{\sigma}_A^T$ in period 1. Any other history that results in $m = \{+\}$ occurs with a probability of at least order $\varepsilon^2$. In the limit as $\varepsilon \to 0$, $m = \emptyset$ is mostly likely observed following $h_2 = (\emptyset, \emptyset)$. This history occurs with a probability of order $\varepsilon^3$, since it results in $m = \emptyset$ given $\underline{\sigma}_A^M$ and arises after two deviations from $\underline{\sigma}_A^T$. Any other history resulting in $m = \emptyset$ occurs with a probability of at least order $\varepsilon^4$, because a deviation from $\underline{\sigma}_A^M$ is necessary for it to arise.

**Region 1c).** In the limit as $\varepsilon \to 0$, $m = \{+\}$ is most likely observed following $h_2 \in \{(+,+), (+,-)\}$, since each of these histories occurs given $\underline{\sigma}_A^T$ and results in $m = \{+\}$ if and only if there is a single deviation from $\underline{\sigma}_A^M$, which occurs with a probability of order $\varepsilon^4$. For any other history to result in $m = \{+\}$ both a deviation from $\underline{\sigma}_A^M$ and a deviation from $\underline{\sigma}_A^T$ is necessary, which occurs with a probability of at least order $\varepsilon^5$. In the limit as $\varepsilon \to 0$, $m = \emptyset$ is most likely observed following $h_2 = (\emptyset, \emptyset)$, since this history results in $m = \emptyset$ given $\underline{\sigma}_A^M$ and requires a single deviation from $\underline{\sigma}_A^T$ in the first period, which occurs with a probability of order $\varepsilon$. Any other history which results in $m = \emptyset$ requires at least two deviations from $\underline{\sigma}_A^T$, which occur with a probability of order $\varepsilon^3$.

**Region 2b).** In the limit as $\varepsilon \to 0$, $m = \{-\}$ is most likely observed following $h_2 \in \{(-, \emptyset), (+,-)\}$. Each of these histories occurs on path and results in $m = \{-\}$ with a probability of order $\varepsilon^4$, since a deviation from $\underline{\sigma}_A^M$ is required. For any other history to results in $m = \{-\}$, a deviation from both $\underline{\sigma}_A^T$ and $\underline{\sigma}_A^M$ is required.

**Region 2c).** Apply the same reasoning as in in Region 1b.

**Region 2d).** Apply the same reasoning as in in Region 1c.

**Region 3.** Apply the same reasoning as in Region 2.

**Region 4.** In the limit as $\varepsilon \to 0$, $m = \{-\}$ ($m = \{+\}$) is most likely observed following $h_2 = (-, \emptyset)$ ($h_2 = (+, \emptyset)$). Since this history arises after a single deviation from $\underline{\sigma}_A^T$ in period 1 and no deviation from $\underline{\sigma}_A^M$, it occurs with a probability of order $\varepsilon$. For any other history to result in $m$ at least two deviations from $\underline{\sigma}_A^T$ or a deviation from $\underline{\sigma}_A^M$ are required, which means it occurs with a probability of at least order $\varepsilon^3$.

Finally, if $\lambda_A < l_{(+,+)}$, the equilibrium is a sequential equilibrium by the same reasoning as in Region 4.

## B.2 Lemma 12: DM-preferred Equilibria

**Proof: Part 1:** To show that $(\overline{\sigma}, \overline{\mu})$ is a sequential equilibrium, consider a completely mixed strategy for the advisor as a function of some $\varepsilon \in (0,1)$, denoted by $\left(\overline{\sigma}_A^{T\varepsilon}, \overline{\sigma}_A^{M\varepsilon}\right)$, which is constructed as follows: for any history of outcomes $h$,

$$\overline{\sigma}_A^{T\varepsilon}(h) = \begin{cases} 1-\varepsilon^3 & \text{if } \overline{\sigma}_A^T(h) = 1 \\ \varepsilon^3 & \text{if } \overline{\sigma}_A^T(h) = 0, \end{cases} \tag{103}$$

and if $\widetilde{h} = \{+,-\}$ then $m = \emptyset$ is sent with probability $\varepsilon$, $m \in \{\{+\},\{+,-\}\}$ is sent with probability $\varepsilon^3$ and $m = \{-\}$ is sent with probability $1-\varepsilon-2\varepsilon^3$. Otherwise, for any $\widetilde{h}$ and set of feasible messages $M\left(\widetilde{h}\right)$, the message specified by $\overline{\sigma}_A^M$ is sent with probability $1-\left(|M\left(\widetilde{h}\right)|-1\right)\varepsilon^3$ and any other feasible message $m \in M\left(\widetilde{h}\right)$ is sent with probability $\varepsilon^3$, where $|M\left(\widetilde{h}\right)|$ denotes the total number of feasible messages given $\widetilde{h}$. As $\varepsilon \to 0$, clearly $\left(\overline{\sigma}_A^{T\varepsilon}, \overline{\sigma}_A^{M\varepsilon}\right) \to \left(\overline{\sigma}_A^T, \overline{\sigma}_A^M\right)$. In addition, the beliefs derived from the advisor's mixed strategy profile using Bayes' rule, denoted by $\overline{\mu}^\varepsilon(m)$, satisfy (102). In the limit as $\varepsilon \to 0$, $m = \emptyset$ is most likely observed following $h_2 = (-,+)$. This history occurs on path and results in $m = \emptyset$ after a deviation from $\overline{\sigma}_A^M$. Hence, this history occurs and results in $m = \emptyset$ with a probability of order $\varepsilon$. All other histories occur and result in $m = \emptyset$ with a probability of at least order $\varepsilon^3$.

**Part 2.** Consider a completely mixed strategy for the advisor as a function of some $\varepsilon \in (0,1)$, denoted by $\left(\overline{\sigma}_A^{T\varepsilon}, \overline{\sigma}_A^{M\varepsilon}\right)$, which is constructed as follows: for any history of outcomes $h$,

$$\overline{\sigma}_A^{T\varepsilon}(h) = \begin{cases} 1-\varepsilon & \text{if } \overline{\sigma}_A^T(h) = 1 \\ \varepsilon & \text{if } \overline{\sigma}_A^T(h) = 0, \end{cases} \tag{104}$$

and for any $\widetilde{h}$ and associated set of feasible messages $M\left(\widetilde{h}\right)$, the message specified by $\overline{\sigma}_A^M$ is sent with probability $1-\left(|M\left(\widetilde{h}\right)|-1\right)\varepsilon^3$ and any other feasible message $m \in M\left(\widetilde{h}\right)$ is sent with probability $\varepsilon^3$, where $|M\left(\widetilde{h}\right)|$ denotes the total number of feasible messages given $\widetilde{h}$. As $\varepsilon \to 0$, clearly $\left(\overline{\sigma}_A^{T\varepsilon}, \overline{\sigma}_A^{M\varepsilon}\right) \to \left(\overline{\sigma}_A^T, \overline{\sigma}_A^M\right)$. Furthermore, for any $m$, as $\varepsilon \to 0$ (102) holds for the following reasons: In the limit as $\varepsilon \to 0$, $m = \{+\}$ is mostly likely observed following $h_2 \in \{(+),(-,+)\}$. Each of these histories occurs and result in $m = \{+\}$ with a probability of order $\varepsilon^3$, since they occur on path and $m = \{+\}$ is sent due to a deviation from $\overline{\sigma}_A^M$. All other histories occur and result in $m = \{+\}$ with a probability of at least order $\varepsilon^4$ because they require both a deviation from $\overline{\sigma}_A^T$ and $\overline{\sigma}_A^M$.

**Part 3:** The equilibrium is a sequential equilibrium by the same reasoning as for Part 2.

**Part 4:** Consider a completely mixed strategy constructed as in Part 2. As $\varepsilon \to 0$, $\left(\overline{\sigma}_A^{T\varepsilon}, \overline{\sigma}_A^{M\varepsilon}\right) \to \left(\overline{\sigma}_A^T, \overline{\sigma}_A^M\right)$. Furthermore, for any $m$, as $\varepsilon \to 0$ (102) holds for the following reasons: In the limit as $\varepsilon \to 0$, $m = \emptyset$ is most likely observed following $h_2 \in \{(\emptyset,\emptyset)\}$.

This history occurs and is followed by $m = \emptyset$ with a probability of order $\varepsilon^2$ since it requires two deviations from $\overline{\sigma}_A^T$. All other histories occur and are followed by $m = \emptyset$ with probability of at least order $\varepsilon^3$ because they require a deviation from $\overline{\sigma}_A^M$.

**Part 5.** Consider a completely mixed strategy constructed as in Part 2. As $\varepsilon \to 0$, $\left(\overline{\sigma}_A^{T\varepsilon}, \overline{\sigma}_A^{M\varepsilon}\right) \to \left(\overline{\sigma}_A^T, \overline{\sigma}_A^M\right)$. Furthermore, for any $m$, as $\varepsilon \to 0$ (102) holds for the following reasons: In the limit as $\varepsilon \to 0$, $m = \{+\}$ is most likely observed following $h_2 \in \{(+,+), (+,-)\}$, since each of these histories occurs and results in $m = \{+\}$ with a probability of order $\varepsilon^3$ (due to a deviation from $\overline{\sigma}_A^M$) whereas all other histories occur and result in $m = \{+\}$ with probability of at least order $\varepsilon^4$, because they require both a deviation from $\overline{\sigma}_A^M$ and $\overline{\sigma}_A^T$.

## B.3 Lemma 7 [Hidden Testing for $N > 2$]

**Proof Claim 1d:** Consider the following completely mixed strategy for the advisor with $\varepsilon \in (0,1)$:

$$\underline{\sigma}_A^{\varepsilon T}(h_n) = \begin{cases} \varepsilon^N & \text{if } \nu_n^+ \geq \overline{\nu}_+ \text{ or if } \nu_n^- \geq \overline{\nu}_-, \\ 1 - \varepsilon^N & \text{otherwise,} \end{cases} \tag{105}$$

and $\underline{\sigma}_A^{\varepsilon M}$ is as follows. If $x_N \leq r$, fails to report each negative outcome in $\underline{\sigma}_A^M$ independently with probability $\varepsilon$ and reports any strictly positive number of positive outcomes with probability $\varepsilon$. If $x_N \geq r$ the advisor fails to report each positive outcome in $\underline{\sigma}_A^M$ independently with probability $\varepsilon^N$ and reports any strictly positive number of negative outcome with probability $\varepsilon$.

Clearly, $\left(\underline{\sigma}_A^{T\varepsilon}, \underline{\sigma}_A^{M\varepsilon}\right) \to \left(\underline{\sigma}_A^T, \underline{\sigma}_A^M\right)$ as $\varepsilon \to 0$. I show that the DM's beliefs also converge. If $\varepsilon = 0$, $m = (\overline{\nu}_+, 0)$ occurs following a history in which the advisor stops testing as soon as $\nu_n^+ = \overline{\nu}_+$ and reports all positive outcomes. Therefore, if $\varepsilon = 0$, $m = (\overline{\nu}_+, 0)$ arises if and only if at least $\overline{\nu}_+$ outcomes in $N$ tests are positive. If $\varepsilon = 0$ then $m = (0, \overline{\nu}_-)$ occurs following a history in which when the advisor stops testing as soon as $\nu_n^- = \overline{\nu}_-$ and reports all negative outcomes. Therefore, if $\varepsilon = 0$, $m = (0, \overline{\nu}_-)$ arises if and only if at least $\overline{\nu}_-$ outcomes in $N$ tests are negative.

In the limit as $\varepsilon \to 0$, any message where $m^+ \geq \overline{\nu}_+$ and $m^- \geq 0$, follows a history at which the advisor stops as soon as $\nu_n^+ = m^+$, reports all positive outcomes and a subset of negative outcomes. Since it must be that the advisor prefers acceptance given $\nu_n^+ \geq \overline{\nu}_+$ by Claim 1b, reporting all positive outcomes is in line with $\underline{\sigma}_A^M$. Note that the advisor must deviate from $\underline{\sigma}_A^T$ to achieve $m^+ > \overline{\nu}_+$. The lowest number of deviations to give lead to $m^+ > \overline{\nu}_+$ arises if the advisor stops as soon as $\nu_n^+ = m^+$. Any deviation from $\underline{\sigma}_A^M$ to report some number of negative outcomes is equally likely. Hence, as $\varepsilon \to 0$, a message where $m^+ \geq \overline{\nu}_+$ and $m^- \geq 0$ arises if and only if at least $m^+$ and at most $N - m^-$ outcomes in $N$ tests are positive. By the same reasoning, in the limit as $\varepsilon \to 0$, any message where $m^+ \geq 0$ and $m^- \geq \overline{\nu}_-$ is most likely to arise following a history in which the advisor stops testing as soon as $\nu_n^- = m^-$, reports all negative outcomes and some subset of positive outcomes. Hence, as $\varepsilon \to 0$, any message where $m^+ \geq 0$ and $m^- \geq \overline{\nu}_-$ arises if and only if at least $m^-$ and at most $N - m^+$ outcomes in $N$ tests are negative.

In the limit as $\varepsilon \to 0$, any message where $m^+ < \overline{\nu}_+$ and $m^- < \overline{\nu}_-$, is most likely to arise following a history in which the advisor stops testing as soon as $\nu_n^- = \overline{\nu}_-$, fails to report $\overline{\nu}_- - m^-$ negative outcomes and reports a subset of positive outcomes. To see why, suppose the advisor follows $\underline{\sigma}_A^T$. Then by period $N$, either $\nu_n^+ = \overline{\nu}_+$ or $\nu_n^- = \overline{\nu}_-$. If $\nu_n^+ = \overline{\nu}_+$, then to send any message where $m^+ < \overline{\nu}_+$ and $m^- < \overline{\nu}_-$ the advisor must omit $\overline{\nu}_+ - m^+$ positive outcomes, which occurs with a probability of $\varepsilon^{N\left(\overline{\nu}_+ - m^+\right)}$, and if $m^- > 0$ then he must report a subset of negative outcomes, which occurs with probability $\varepsilon$. If $\nu_n^- = \overline{\nu}_-$, then to send any message where $m^+ < \overline{\nu}_+$ and $m^- < \overline{\nu}_-$ the advisor must omit $\overline{\nu}_- - m^-$ negative outcomes, which occurs with a probability of $\varepsilon^{\left(\overline{\nu}_- - m^-\right)}$, and if $m^+ > 0$ then he must report a subset of positive outcomes, which occurs with probability $\varepsilon$. Since $N > \overline{\nu}_- - m^-$, it must be that $N\left(\overline{\nu}_+ - m^+\right) > \overline{\nu}_- - m^-$. Therefore, even if $m^- = 0$ and $0 < m^+ < \overline{\nu}_+$, it is more likely that $\nu_n^- = \overline{\nu}_-$ than $\nu_n^+ = \overline{\nu}_+$ since $\varepsilon^{(\overline{\nu}_- + 1)} > \varepsilon^{N\left(\overline{\nu}_+ - m^+\right)}$. In addition, a single deviation from $\underline{\sigma}_A^T$ occurs with probability $\varepsilon^N$, and therefore, is less likely than the event that the advisor followed $\underline{\sigma}_A^T$ and deviated from $\underline{\sigma}_A^M$. Hence, as $\varepsilon \to 0$ any message where $m^+ < \overline{\nu}_+$ and $m^- < \overline{\nu}_-$ arises if and only if at least $\overline{\nu}_-$ and at most $N - m^+$ outcomes in $N$ tests are negative.

**Proof Claim 2c:** First, suppose that $max\left\{\widehat{\nu}_+, \overline{\nu}_+\right\} = \overline{\nu}_+$, which implies $min\left\{\overline{\nu}_-, N - \widehat{\nu}_+ + 1\right\} = \overline{\nu}_-$. The equilibrium is sequential by the same reasoning as in Claim 1d above.

Second, suppose that $max\left\{\widehat{\nu}_+, \overline{\nu}_+\right\} = \widehat{\nu}_+$, which implies $min\left\{\overline{\nu}_-, N - \widehat{\nu}_+ + 1\right\} = N - \widehat{\nu}_+ + 1$. Consider the following completely mixed strategy for the advisor with $\varepsilon \in (0,1)$:

$$\underline{\sigma}_A^{T\varepsilon}\left(h_n\right) = \begin{cases} \varepsilon^N & \text{if } \nu_n^+ \geq \widehat{\nu}_+ \text{ or if } \nu_n^- \geq N - \widehat{\nu}_+ + 1, \\ 1 - \varepsilon^N & \text{otherwise,} \end{cases} \tag{106}$$

and $\underline{\sigma}_A^{M\varepsilon}$ is as follows: If $x_N \geq r + 1$ the advisor fails to report each positive outcome in $\underline{\sigma}_A^{T\varepsilon}$ independently with probability $\varepsilon^N$ and reports any strictly positive number of negative outcomes with probability $\varepsilon$. If $x_N \leq r$ the advisor fails to report each negative outcome in $\underline{\sigma}_A^{T\varepsilon}$ independently with probability $\varepsilon$ and reports any strictly positive number of positive outcomes with probability $\varepsilon$.

Clearly, $\left(\underline{\sigma}_A^{T\varepsilon}, \underline{\sigma}_A^{M\varepsilon}\right) \to \left(\underline{\sigma}_A^T, \underline{\sigma}_A^M\right)$ as $\varepsilon \to 0$. I show that the DM's beliefs also converge. In the limit as $\varepsilon \to 0$, any message where $m^+ \geq \widehat{\nu}_+$ and $m^- \geq 0$ arises following a history at which the advisor stopped as soon as $\nu_n^+ = m^+$, reported all positive outcomes and a subset of negative outcomes. Since it must be that the advisor prefers acceptance given $\nu_n^+ \geq \widehat{\nu}_+ \geq \overline{\nu}_+$ by Claim 1a, reporting all positive outcomes is in line with $\underline{\sigma}_A^M$. Note that the advisor must deviated from $\underline{\sigma}_A^T$ to achieve $m^+ > \widehat{\nu}_+$. The lowest number of deviations to lead to $m^+ \geq \widehat{\nu}_+$ arises if the advisor stopped as soon as $\nu_n^+ = m^+$. Any deviation from $\underline{\sigma}_A^M$ to report some number of negative outcomes is equally likely. Hence, as $\varepsilon \to 0$, any message where $m^+ \geq \widehat{\nu}_+$ and $m^- \geq 0$ arises if and only if at least $m^+$ and at most $N - m^-$ outcomes in $N$ tests are positive. In addition, in the limit as $\varepsilon \to 0$, any message where $\widehat{\nu}_+ \geq m^+ \geq \overline{\nu}_+$ and $m^- \geq 0$ arises following a history at which the advisor stopped as soon as $\nu_n^- = N - \widehat{\nu}_+ + 1$, reported all positive outcomes and a subset of negative outcomes. No deviation from $\underline{\sigma}_A^T$ is necessary. Since it must be that the advisor prefers acceptance given $\nu_n^+ \geq \overline{\nu}_+$ by Claim 1a, reporting all positive outcomes is in line with $\underline{\sigma}_A^M$. Therefore, the only deviation is to report some number of negative outcomes if $m^- > 0$ and this kind of deviation is always necessary to generate a message with $m^+ > 0$ and $m^- > 0$. Hence, as $\varepsilon \to 0$, any message where $\widehat{\nu}_+ \geq m^+ \geq \overline{\nu}_+$ and $m^- \geq 0$ arises if and only if at least $m^+$ and at most $N - m^-$ outcomes in $N$ tests are positive.

In the limit as $\varepsilon \to 0$, any message where $\overline{\nu}_- > m^- \geq N - \widehat{\nu}_+ + 1$ and $r + m^- + 1 \leq m^+ < \overline{\nu}_+$, follows a history in which the advisor stops testing as soon as $\nu_n^- = m^-$, reports all positive outcomes and some subset of negative outcomes. Note that the advisor must deviate from $\underline{\sigma}_A^T$ to achieve $m^- > N - \widehat{\nu}_+ + 1$. The lowest number of deviations to allow for $m^- > N - \widehat{\nu}_+ + 1$ arises if the advisor stopped as soon as $\nu_n^- = m^-$. Suppose the advisor preferred acceptance, i.e. $\nu_n^+ - \nu_n^- \geq r + 1$. Then the only necessary deviation from $\underline{\sigma}_A^M$ is that a subset of negative outcomes is reported. If the advisor had preferred rejection then he would need to deviate from $\underline{\sigma}_A^M$ by failing to report some of the positive outcomes as well as reporting some subset of negative outcomes. Hence, as $\varepsilon \to 0$, this message arises if and only if at least $m^-$ and at most $N - m^+$ outcomes in $N$ tests are positive. In addition, in the limit as $\varepsilon \to 0$, any message where $m^- \geq \overline{\nu}_-$ (which implies $m^+ < \overline{\nu}_+$) or where $\overline{\nu}_- > m^- \geq N - \widehat{\nu}_+ + 1$ and $m^+ < min\left\{\overline{\nu_n^+}, r_A + m^- + 1\right\}$, follows a history in which the advisor stops testing as soon as $\nu_n^- = m^-$, reports all negative outcomes and some subset of positive outcomes. Again, the lowest number of deviations to allow for $m^- > N - \widehat{\nu}_+ + 1$ arises if the advisor stopped as soon as $\nu_n^- = m^-$. If $m^- \geq \overline{\nu}_-$ or if $\overline{\nu}_- > m^- \geq N - \widehat{\nu}_+ + 1$ and $m^+ - m^- \leq r$, the only necessary deviation from $\underline{\sigma}_A^M$ is that a subset of positive outcomes is reported and this kind of deviation is always necessary to generate a message with $m^+ > 0$ and $m^- > 0$. Hence, as $\varepsilon \to 0$ this message arises if and only if at least $m^-$ and at most $N - m^+$ outcomes in $N$ tests are negative.

Any message where $m^+ < \overline{\nu}_+$ and $N - \widehat{\nu}_+ + 1 > m^-$, in the limit as $\varepsilon \to 0$, follows a history in which the advisor stops testing as soon as $\nu_n^- = N - \widehat{\nu}_+ + 1$, fails to report $N - \widehat{\nu}_+ + 1 - m^-$

negative outcomes and reports some subset of positive outcomes. Suppose the advisor follows $\underline{\sigma}_A^T$. Then by period $N$ either $\nu_n^- = N - \widehat{\nu}_+ + 1$ or $\nu_n^+ = \widehat{\nu}_+$. If $\nu_n^+ = \widehat{\nu}_+$ then the advisor prefers acceptance and failed to report $\widehat{\nu}_+ - m^+$ positive outcomes, which happens with probability $\varepsilon^{N\left(\widehat{\nu}_+ - m^+\right)}$. If $\nu_n^- = N - \widehat{\nu}_+ + 1$ then if he preferred acceptance he failed to report $\nu_n^+ - m^+$ positive outcomes where $\nu_n^+ < \widehat{\nu}_+$ , which happens with probability $\varepsilon^{N\left(\nu_n^+ - m^+\right)}$, or if he preferred rejection and failed to report $N - \widehat{\nu}_+ + 1 - m^-$ negative outcomes, which happens with probability $\varepsilon^{\left(N - \widehat{\nu}_+ + 1 - m^-\right)}$. Since $N - \widehat{\nu}_+ + 1 - m^- > N\left(\nu_n^+ - m^+\right) > N\left(\widehat{\nu}_+ - m^+\right)$, it is most likely that $\nu_n^- = N - \widehat{\nu}_+ + 1$ and the advisor fails to report $N - \widehat{\nu}_+ + 1 - m^-$ negative outcomes. In addition, a single deviation from $\underline{\sigma}_A^T$ occurs with probability $\varepsilon^N$, and therefore, is less likely than the event that the advisor followed $\underline{\sigma}_A^T$ and deviated from $\underline{\sigma}_A^M$. Hence, as $\varepsilon \to 0$ this message arises if and only if at least $N - \widehat{\nu}_+ + 1$ and at most $N - m^+$ outcomes in $N$ tests are negative.

## B.4 Proposition 7 [Simultaneous Testing]

Consider the following modification to games $\Gamma$ and $\Gamma^o$. The advisor chooses the number of tests ex ante, i.e. his testing strategy is $\sigma_A^T : H_0 \to \{0, ..., N\}$ in $\Gamma$ and $\sigma_A^o : H_0 \to \{0, ..., N\}$ in $\Gamma^o$. The advisor chooses the number of tests at $h_0$ to maximize his expected payoff given his beliefs $\mu_A$ and the DM's strategy. In $\Gamma^o$, $\frac{\mu(h_n)}{1-\mu(h_n)} = l\left(x_n\right)$ and, in $\Gamma$, $\frac{\mu_A(h_n)}{1-\mu_A(h_n)} = l\left(x_n\right)$ for $n \in \{0, N\}$ where $x_n$ is the number of excess positive outcomes in $h_n$. Denote the likelihood ratio at the end of period $n$ by $l_n$.

**Part 1.** For the skepticism effect to exist, it must be that the DM prefers acceptance when all $N$ tests are positive, i.e. $\lambda_{DM} \leq l\left(N\right)$, otherwise, $\sigma_{DM}^o = reject$ for any $h_N$ and $\sigma_{DM} = reject$ for any $m$. Suppose testing is observable, and $r_A$ and $r_{DM}$ are defined by

$$r_{DM} \equiv sup\left\{j | j \in \{0, ..., N-1\}, l\left(j\right) < \lambda_{DM}\right\}, \tag{107}$$

$$r_A \equiv sup\left\{j | j \in \mathbb{Z}, l\left(j\right) < \lambda_A\right\}, \tag{108}$$

and $r_A + 1 \leq r_{DM}$. I show that if $r_{DM}$ is even (odd) the advisor chooses the largest odd (even) number of tests $n$ that satisfies $n \leq N$. Hence, given any even (odd) $r_{DM}$, if $N$ is odd (even) the advisor chooses $n = N$ and, therefore, the DM can never be strictly better off under hidden testing.

First, I show that if $r_{DM}$ is even (odd) then the advisor never chooses an even (odd) number of tests. Suppose the advisor has already committed to run $n$ tests, where $n \leq N - 1$. An additional test only affects the advisor's expected payoff if it affects the DM's choice. There are two situations in which the additional test affects the DM's choice. The first situation is that after $n$ tests the DM rejects, but if the additional test is positive, she accepts. This situation arises if and only if $l_n = l\left(r_{DM}\right)$. The second situation is that after $n$ tests the DM accepts, but if the additional test is negative, she rejects. This situation arises if and only if $l_n = l\left(r_{DM}+1\right)$. Since the advisor prefers accept if at $l_n = l\left(r_{DM}\right)$, he is strictly better off with an additional test in the first situation, but strictly worse off in the second situation. If $r_{DM}$ is even (odd), then the first situation can only arise if $n$ is even (odd) and the second only if $n$ is odd (even). Hence, there is only an upside to an additional test if $n$ is even (odd) and only a downside if $n$ is odd (even). Therefore, the advisor strictly prefers to run an additional test after an even (odd) number of tests.

Consider $r_{DM}$ is even (odd) and $n$ is odd (even). Next, I show that if the advisor benefits from running an additional two tests at some $n$ then he benefits from running the largest odd (even) number of tests such that $n \leq N$. There are two situations in which the additional two tests could affect the DM's choice. The first situations is that after $n$ tests the DM rejects, but if both additional tests are positive, she accepts. This situation arises if and only if $l_n = l\left(r_{DM}-1\right)$, i.e. at a total of $\nu_n^+ = \frac{n+(r_{DM}-1)}{2}$ positive outcomes. The second situation is that after $n$ tests the DM accepts, but if both additional tests are negative, she rejects. This situation arises if and only if $l_n = l\left(r_{DM}+1\right)$, i.e. at a total of $\nu_n^+ = \frac{n+r_{DM}+1}{2}$ positive outcomes. The advisor is strictly better off by conducting these test if and only if

$$\begin{aligned}\lambda_A Pr\left(false\right)\Bigg[&Pr\left(\nu_n^+ = \frac{n+r_{DM}+1}{2}|false\right)p^2 - \\ &Pr\left(\nu_n^+ = \frac{n+(r_{DM}-1)}{2}|false\right)\left(1-p\right)^2\Bigg] \\ +Pr\left(true\right)\Bigg[&Pr\left(\nu_n^+ = \frac{n+(r_{DM}-1)}{2}|true\right)p^2 \\ &-Pr\left(\nu_n^+ = \frac{n+r_{DM}+1}{2}|true\right)\left(1-p\right)^2\Bigg] > 0 \qquad (109)\end{aligned}$$

where

$$Pr\left(\nu_n^+ = j|true\right) = \binom{n}{j}\left(1-p\right)^{n-j}p^j, \qquad (110)$$

$$Pr\left(\nu_n^+ = j|false\right) = \binom{n}{j}\left(1-p\right)^{j}p^{n-j}. \qquad (111)$$

Substituting into (109):

$$\frac{n!\lambda_A\left(1-q\right)\left(1-p\right)^{\frac{n+r_{DM}+1}{2}}p^{\frac{n-\left(r_{DM}-1\right)}{2}}2}{\left(\frac{n+r_{DM}+1}{2}-1\right)!\left(\frac{n-r_{DM}-1}{2}\right)!}\left[\frac{p}{\left(n+r_{DM}+1\right)}-\frac{\left(1-p\right)}{\left(n-r_{DM}+1\right)}\right]$$
$$+\frac{n!qp^{\frac{n+r_{DM}+1}{2}}\left(1-p\right)^{\frac{n-r_{DM}+1}{2}}2}{\left(\frac{n+r_{DM}+1}{2}-1\right)!\left(\frac{n-r_{DM}+1}{2}\right)!}\left[\frac{p}{n-r_{DM}+1}-\frac{1-p}{n+r_{DM}+1}\right]>0\Leftrightarrow$$
$$\lambda_A\left(1-q\right)\left(1-p\right)^{r_{DM}}\left[\left(n-r_{DM}+1\right)p-\left(n+r_{DM}+1\right)\left(1-p\right)\right]$$
$$+qp^{r_{DM}}\left[\left(n+r_{DM}+1\right)p-\left(n-r_{DM}+1\right)\left(1-p\right)\right]>0. \qquad (112)$$

The LHS increases with $n$ since $p>1-p$. Hence, if (112) holds at some $n$ then it must hold at any larger $n$.

What is left to show is that the advisor strictly prefers to choose an odd (even) number of tests large enough that it is possible that the outcomes lead the DM to accept. The advisor is indifferent between running $r_{DM}-1$ or fewer tests, because even if all $r_{DM}-1$ tests were positive, the DM would still reject. However, he must strictly benefit from running $r_{DM}+1$ tests, because if the DM accepts than the advisor also prefers to accept. By the argument above, if the advisor benefits from an additional two tests at $n=r_{DM}-1$ then he must benefit from running two additional tests at any larger $n$. Hence, the advisor chooses the largest odd (even) number of tests feasible.

**Part 2.** Suppose $l\left(0\right)<\lambda_{DM}\leq l\left(1\right)$ and $l\left(N-1\right)<\lambda_A\leq l\left(N\right)$. I show that, under observable testing, the advisor does not to conduct any tests. First, I show that if $r_{DM}$ is even (odd), the advisor never chooses an odd (even) number of tests. There are the two situations in which the additional test affects the DM's choice, as described in Part 1. The advisor is strictly better off with an additional test if $l_n=l\left(1\right)$, but strictly worse off if $l_n=l\left(0\right)$. Given $r_{DM}$ even (odd), $l_n=l\left(1\right)$ can only arise if $n$ is odd (even) and $l_n=l\left(0\right)$ only if $n$ is even (odd). Hence, there is only an upside to an additional test if $n$ is odd (even) and only a downside if $n$ is even (odd). Therefore, the advisor strictly prefers to run an additional test after an odd number of tests. Second, I show that if the advisor benefits from running an additional two tests at some $n$ then he benefits from running the largest even number of tests such that $n\leq N$. Consider only even $n$. There are two situations in which the additional two tests could affect the DM's choice, as described in Part 1. The first situation arises if and only if $l_n=l\left(0\right)$. The second situation arises if and only if $l_n=l\left(2\right)$.

The advisor is strictly better off by conducting two additional tests if and only if

$$\lambda_A Pr(false)\left[Pr\left(\nu_n^+ = \frac{n+2}{2}|false\right)p^2 - Pr\left(\nu_n^+ = \frac{n}{2}|false\right)(1-p)^2\right]$$
$$+Pr(true)\left[Pr\left(\nu_n^+ = \frac{n}{2}|true\right)p^2 - Pr\left(\frac{n+2}{2}|true\right)(1-p)^2\right] > 0 \Leftrightarrow$$
$$\frac{n!\lambda_A(1-q)(1-p)^{\frac{n+2}{2}}p^{\frac{n}{2}}2}{\left(\frac{n+2}{2}-1\right)!\left(\frac{n-2}{2}\right)!}\left[\frac{p}{(n+2)} - \frac{1-p}{n}\right]$$
$$+\frac{n!qp^{\frac{n+2}{2}}(1-p)^{\frac{n}{2}}2}{\left(\frac{n+2}{2}-1\right)!\left(\frac{n}{2}\right)!}\left[\frac{p}{n} - \frac{1-p}{n+2}\right] > 0 \Leftrightarrow$$
$$\lambda_A(1-q)(1-p)[np-(n+2)(1-p)]$$
$$+qp[(n+2)p-(n)(1-p)] > 0. \qquad (113)$$

The LHS increases with $n$ since $p > 1-p$. Hence, if (113) holds at some $n$ then it must hold at any larger $n$. Therefore, if the advisor runs any tests at all then he runs all $N$ tests. Lastly, the advisor prefers no test to $N$ tests if and only if

$$Pr(true) < Pr(false)\lambda_A Pr\left(\nu_n^+ > \frac{N}{2}|false\right) + Pr(true)Pr\left(\nu_n^+ \leq \frac{N}{2}|true\right) \Leftrightarrow$$
$$\lambda_A > \frac{qPr\left(\nu_n^+ > \frac{N}{2}|true\right)}{(1-q)Pr\left(\nu_n^+ > \frac{N}{2}|false\right)}. \qquad (114)$$

This must hold since for $N > 2$,

$$\frac{qPr\left(\nu_n^+ > \frac{N}{2}|true\right)}{(1-q)Pr\left(\nu_n^+ > \frac{N}{2}|false\right)} < l(N-1) \equiv \frac{qp^{N-1}}{(1-q)(1-p)^{N-1}} < \lambda_A. \qquad (115)$$

Finally, suppose testing is hidden and the advisor runs all $N$ tests. By the same argument as in Lemma 7, there is an advisor-preferred equilibrium in which the advisor discloses all positive outcomes if and only if $l_N = l(N)$ since $l(N-1) < \lambda_A \leq l(N)$, and otherwise, discloses only negative outcomes. The DM acts in line with the advisor's interest for any on-path message, by the same reasoning as in Lemma 7. Therefore, the advisor indeed runs all $N$ tests. Hence, the DM must be strictly better off under hidden testing.

## B.5 Proposition 8 [Infinite Horizon]

Consider the following modification to games $\Gamma$ and $\Gamma^o$. There are infinitely many discrete periods, $n = 0, 1, ....$ The advisor incurs a cost $c > 0$ for each test. In $\Gamma^o$, in each period $n$, the advisor chooses to test or not, but once he has stopped testing he cannot test again. When the advisor has stopped, the DM chooses an action, i.e. $\sigma^o_{DM} : H_n \longrightarrow \{accept, reject\}$. In $\Gamma$, in each period $n$, the advisor privately chooses to test or not and when he stops he sends a message $m \in \mathcal{M}$, where $\mathcal{M}_n$ as in (4) and $\mathcal{M} = \cup_{n=0}^{\infty} \mathcal{M}_n$. Given the unordered history of outcomes $\widetilde{h}_n$ at the end of period $n$, the set of feasible messages is $M\left(\widetilde{h}_n\right) = \mathcal{P}\left(\widetilde{h}_n\right)$. Then the DM chooses $a \in \{accept, reject\}$. The DM does not know in which period the message was sent. A reporting strategy for the advisor is $\sigma^M_A : H_n \longrightarrow \mathcal{M}_n$ for $n = 0, 1, ....$ A strategy for the DM is $\sigma_{DM} : \mathcal{M} \longrightarrow \{accept, reject\}$. In each scenario, the solution concept is a PBE.

**Insurance Effect:** Suppose $\lambda_A > \lambda_{DM}$. Under hidden testing, once the advisor has stopped testing at some $h_n$, the DM acts in his interest given $h_n$, by the same reasoning as in the proof of Lemma 7. Therefore, the advisor's optimal testing strategy is the one chosen if the advisor was in charge of taking the final decision. Hence, the advisor's optimal testing strategy is time-independent. In order for the DM to be better off under hidden than observable testing, it is necessary that with some probability the advisor finds $x$ excess positive outcomes such that he prefers acceptance, i.e. $l(x) \geq \lambda_A$. This is because then the DM prefers acceptance, since $l(x) \geq \lambda_A > \lambda_{DM}$, and acceptance is chosen in equilibrium. To be able to find such outcomes, the advisor needs to test at any history where the number $x$ of excess positive outcomes satisfies $l(0) < l(x) \leq \lambda_A$, i.e. for any $x$ satisfying $l(0) < l(x) \leq \lambda_A$ the advisor's continuation value of testing must exceed his continuation value of stopping.

Under observable testing, the DM's optimal strategy is to reject if and only if $l(x_n) < \lambda_{DM}$, where $x_n$ denotes the number of excess positive outcomes when the advisor stops testing. Therefore, the advisor's optimal testing strategy is again time-independent. His value of stopping at $h_n$ is the same as under hidden testing if $l(x_n) < \lambda_{DM}$ or $l(x_n) \geq \lambda_A$, because then the DM and the advisor agree on the optimal final decision. However, it is lower if $\lambda_{DM} \leq l(x_n) < \lambda_A$. Therefore, if the advisor does not test at any history $h_n$ such that $l(0) < \lambda_{DM} < l(x_n) \leq \lambda_A$ when testing is observable, then he also does not test at any such $h_n$ when testing is hidden. Hence, the insurance effect cannot exist.

**Skepticism Effect:** Let $\widetilde{h}_n = (\nu^+_n, \nu^-_n)$ denote the history at the end of period $n$ with $\nu^+_n$ positive and $\nu^-_n$ negative outcomes. Define two thresholds of $c$:

$$\overline{c} = \min\left\{\frac{q(1-p)p}{q(1-p)+(1-q)p}, \frac{q(3p^2-3p^3+p^4)}{2-p^3-q+3pq-3p^2q+2p^3q}\right\}, \tag{116}$$

$$\underline{c} = \max\left\{\frac{q(1-p)}{2(q(1-p)+(1-q)p)}, \frac{q(1-p)^2}{q(1-p)^2+(1-q)p^2}\right\}. \tag{117}$$

The skepticism effect exists if $c$ satisfies $\overline{c} > c > \underline{c}$, $\lambda_A = 0$ and

$$\frac{q\left(1-p^2\right)}{\left(1-q\right)\left(1-\left(1-p\right)^2\right)} < \lambda_{DM} < \frac{qp\left(p+\left(1-p\right)^2+p\left(1-p\right)^2\right)}{\left(1-q\right)\left(1-p\right)\left(1-p+p^2+p^2\left(1-p\right)\right)}. \tag{118}$$

I show that for any $q$, there exists a $p \in \left(\frac{1}{2}, 1\right)$ such that $\overline{c} > \underline{c}$. At $p = 1$, $\overline{c} = \underline{c} = 0$. As $p \to 1$, $\underline{c} = \frac{q(1-p)p^2}{2q(1-p)+2(1-q)p}$ and $\overline{c} = \frac{q(1-p)p}{q(1-p)+(1-q)p}$. As $p \to 1$, $\frac{\partial \overline{c}}{\partial p} < 0$ and $\frac{\partial \underline{c}}{\partial p} < 0$ but $-\frac{\partial \overline{c}}{\partial p} > -\frac{\partial \underline{c}}{\partial p}$. Hence, $\overline{c} > \underline{c}$ for $p$ sufficiently large.

Under observable testing, $\underline{\sigma}_{DM}\left(h_n\right) = accept$ if and only if $l_n \geq l\left(1\right)$, since

$$l\left(0\right) < \lambda_{DM} < \frac{qp\left(p+\left(1-p\right)^2+p\left(1-p\right)^2\right)}{\left(1-q\right)\left(1-p\right)\left(1-p+p^2+p^2\left(1-p\right)\right)} < l\left(1\right). \tag{119}$$

The advisor's optimal strategy is time-independent. $\underline{\sigma}_A\left(h_n\right) = 0$ if $l_n \geq l\left(1\right)$, since he prefers acceptance independent of the state and the DM accepts. In addition, $\underline{\sigma}_A\left(h_n\right) = 0$ if $l_n \leq l\left(-1\right)$ since the expected cost of testing outweighs the expected benefit. In particular, he would stop at $l_n = l\left(-1\right)$ even if the next two outcomes were known to be positive and, hence, would lead the DM to accept since $c > \underline{c}$ and

$$-2c < -Pr\left(true|\widetilde{h}_n = (0,1)\right) = -\frac{q\left(1-p\right)}{q\left(1-p\right)+\left(1-q\right)p}. \tag{120}$$

Given that $\underline{\sigma}_A\left(h_n\right) = 0$ if either $l_n \geq l\left(1\right)$ or $l_n \leq l\left(-1\right)$, $\underline{\sigma}_A\left(h_n\right) = 1$ if $l_n = l\left(0\right)$ since $c < \overline{c}$ and $-q\left(1-p\right) - c > -q \Leftrightarrow qp > c$.

Next, suppose testing is hidden. I show that it cannot be part of an equilibrium that $\underline{\sigma}_{DM}\left(m\right) = accept$ for $m = \left(m^+, 0\right)$ where $m^+ \geq 1$. To see why, suppose this was the case. Then for any $\widetilde{h}_n$ where $\nu_n^+ \geq 1$, the advisor reports only positives, i.e. $\underline{\sigma}_A^M\left(\widetilde{h}_N\right) = \left(\nu_N^+, 0\right)$. Hence, $\underline{\sigma}_A^T\left(\widetilde{h}_n\right) = 0$ if $\nu_n^+ \geq 1$. In addition, $\underline{\sigma}_A^T\left(\widetilde{h}_n\right) = 0$ if $\widetilde{h}_n = (0,2)$, because even if the next test were known to be positive, the advisor is not willing to pay since $c > \underline{c}$ and

$$-c < -Pr\left(true|\widetilde{h}_n = (0,2)\right) = -\frac{q\left(1-p\right)^2}{q\left(1-p\right)^2+\left(1-q\right)p^2}. \tag{121}$$

But the advisor optimally continues testing at $\widetilde{h}_n = (0,1)$, since $c < \overline{c}$ and

$$-Pr\left(true|\widetilde{h}_n = (0,1)\right)\left(1-p\right) - c > -Pr\left(true|\widetilde{h}_n = (0,1)\right) \Leftrightarrow$$
$$\frac{q\left(1-p\right)p}{q\left(1-p\right)+\left(1-q\right)p} > c. \tag{122}$$

Given (122), the advisor also tests at the prior since his belief that the state is true is then even higher. The advisor's strategy means that the DM receives $m = (m^+, 0)$ where $m^+ \geq 1$ if either $\widetilde{h}_1 = (1, 0)$ or $\widetilde{h}_2 = (1, 1)$. In equilibrium, the DM must reject such a report since $\frac{q(1-p^2)}{(1-q)(1-(1-p)^2)} < \lambda_{DM}$.

Next, I show that there exists an equilibrium in which $\underline{\sigma}_{DM}(m) = \textit{accept}$ if and only if $m = (m^+, 0)$ where $m^+ \geq 2$. Suppose the DM followed this strategy. Then for any $\widetilde{h}_n$ where $\nu_n^+ \geq 2$, the advisor reports only positives, i.e. $\underline{\sigma}_A^M\left(\widetilde{h}_n\right) = (\nu_n^+, 0)$. Hence, $\underline{\sigma}_A^T\left(\widetilde{h}_n\right) = 0$ if $\widetilde{h}_n$ where $\nu_n^+ \geq 2$. In addition, $\underline{\sigma}_A^T\left(\widetilde{h}_n\right) = 0$ if $\widetilde{h}_n = (0, 1)$, because even if the next two tests were known to be positive, the advisor is not willing to test since $c > \underline{c}$ and (120). In addition, $\underline{\sigma}_A^T\left(\widetilde{h}_n\right) = 0$ if $\widetilde{h}_n = (1, 3)$, because even if the next test were known to be positive, he is not willing to test since $c > \underline{c}$ and $Pr\left(true|\widetilde{h}_n = (0, 2)\right) = Pr\left(true|\widetilde{h}_n = (1, 3)\right)$ and (121). However, $\underline{\sigma}_A^T\left(\widetilde{h}_n\right) = 1$ if $\widetilde{h}_n = (1, 2)$ since $Pr\left(true|\widetilde{h}_n = (1, 2)\right) = Pr\left(true|\widetilde{h}_n = (0, 1)\right)$ and (122). Suppose $\underline{\sigma}_A^T\left(\widetilde{h}_n\right) = 1$ if $\widetilde{h}_n = (1, 1)$ or if $\widetilde{h}_n = (1, 0)$. Then, $\underline{\sigma}_A^T(h_n) = 1$ if $\widetilde{h}_n = (0, 0)$ since $\overline{c} > c$ and

$$\begin{aligned} -q\left(1-p+p\left(1-p\right)^3\right) - \left(q\left(1-p\right)+\left(1-q\right)p\right)c \\ -\left(qp^2+\left(1-q\right)\left(1-p\right)^2\right)2c - \left(q\left(1-p\right)p^2+\left(1-q\right)\left(1-p\right)^2p\right)3c \\ -\left(q\left(1-p\right)^2p+\left(1-q\right)\left(1-p\right)p^2\right)4c > -q \\ \Leftrightarrow \\ \frac{q\left(3p^2-3p^3+p^4\right)}{2-p^3-q+3pq-3p^2q+2p^3q} > c \end{aligned} \tag{123}$$

Given (123), it must also be that $\underline{\sigma}_A^T\left(\widetilde{h}_n\right) = 1$ if $\widetilde{h}_n = (1, 1)$ or if $\widetilde{h}_n = (1, 0)$ since he only needs one additional positive outcome for acceptance and his posterior belief of the state being true is at least as high as when $\widetilde{h}_n = (0, 0)$. This implies that $m = (m^+, 0)$ where $m^+ \geq 2$ is sent if and only if $h_2 = (+, +)$ or $h_3 = (+, -, +)$ or $h_4 = (+, -, -, +)$. In equilibrium, the DM accepts conditional on $m = (2, 0)$ since

$$\lambda_{DM} \leq \frac{\underline{\mu}_{DM}(m = (2, 0))}{1 - \underline{\mu}_{DM}(m = (2, 0))} = \frac{qp\left(p + (1-p)^2 + p(1-p)^2\right)}{(1-q)(1-p)(1-p+p^2+p^2(1-p))}. \tag{124}$$

Recall that the advisor stops testing in period 4 or earlier on the equilibrium path. Denote the set of lists of Nature's first four draws of outcomes by

$$\Phi_4 \equiv \{(s_1, ..., s_4) \,|s_i \in \{-, +\}, i \in \{1, ..., 4\}\}. \tag{125}$$

Denote the subset of $\Phi_4$ for which acceptance is chosen under hidden and observable testing by $\Phi^o$ and $\Phi^h$ respectively, where

$$\Phi^o = \{(s_1, ..., s_4)\,|s_1 = (+)\}, \tag{126}$$

$$\Phi^h = \{(s_1, ..., s_4)\,|s_1 = (+), s_2 = (+) \vee s_3 = (+) \vee s_4 = (+)\}. \tag{127}$$

Therefore, the only list $\phi$ which satisfies $\phi \in \Phi^o$ and $\phi \notin \Phi^h$ is $\phi = (+, -, -, -)$ and, given this list, the DM prefers rejection since $l(-2) < \lambda_{DM}$. Hence, the DM is better off under hidden testing.