Hidden Testing and Selective Disclosure of Evidence*

Claudia Herresthal†

September 14, 2018

Abstract

This paper compares a decision maker's payoff under public and private information acquisition by a biased advisor. Both players agree on the optimal choice under certainty, but differ in how they trade off the loss from errors. The advisor can sequentially acquire informative test outcomes. If acquisition is private he decides in the final period which realizations to verifiably disclose. If players' preferences are sufficiently misaligned, the decision maker is weakly better off under private rather than public information acquisition. The effect on the advisor's payoff depends on the direction of his bias.

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

*I would like to thank Inga Deimen, Matt Elliott, Robert Evans, Pawel Gola, Ole Jann, Meg Meyer, Arina Nikandrova, Anna Sanktjohanser and conference and seminar audiences for valuable comments and feedback.
†Queens' College, University of Cambridge, Silver Street, Cambridge CB3 9ET, UK, cush2@cam.ac.uk.

1 Introduction

Pharmaceutical companies have recently come under scrutiny for selectively reporting outcomes of clinical trials. As a response to these pressures, several of these companies have pledged to register all of their trials in open online databases and to publish their findings (Goldacre et al. (2017)). At first sight, it seems that demanding such transparency would improve regulation, because pharmaceutical companies can no longer omit 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 studies when and why a decision maker (e.g. a regulator) prefers hidden to observable information acquisition by a biased advisor (e.g. a company). The decision maker has to either accept or reject a given hypothesis. Both agree on the optimal action under certainty, but when there is some uncertainty about whether the hypothesis is true or false they may trade off the mistake from inappropriate actions differently. For example, both the regulator and the company may agree that a drug should be approved only if it is safe, but the regulator is relatively more averse to the mistake of approving an unsafe drug than the company. I first consider a setting of *hidden testing*, which corresponds to not having a trial registry. The advisor can in private sequentially acquire binary test outcomes over a fixed number of periods.[1] In the final period, the advisor chooses which outcomes to disclose and then the decision maker takes an action. The advisor can verifiably disclose outcomes, i.e. he can omit but not fake outcomes, but he cannot verifiably disclose the period in which these outcomes were discovered. In the example, this represents the fact that the regulator does not know how many trials have been run in total by the time he receives an application for approval. I contrast this with a setting of *observable testing*, which corresponds to having a trial registry. When testing is observable, the advisor has no private information.[2]

My main focus lies on situations in which the decision maker (DM) cannot commit to a decision rule, e.g. if on the basis of evidence already presented the regulator prefers approval to rejection, then he has limited scope to condition acceptance on additional positive evidence.[3] I first study a two-period model and then show that the main insights are robust to allowing for an arbitrary fixed number of periods. My main result fully characterizes the situations in which the DM is strictly better off and the situations in which she is strictly worse off under hidden rather than observable testing. I identify two distinct economic effects which cause the DM to benefit from hidden testing, and I refer to them as the skepticism effect and the insurance effect.

[1]The finite horizon represents the fact that companies may face time and money constraints in their development. This can be seen as a limiting case of a convex cost of testing, as discussed in Section 6.

[2]For my results, the important aspect of observable testing is that all realizations are disclosed. It is irrelevant whether or not the time at which these outcomes were discovered is disclosed.

[3]Although a regulator can demand additional evidence, he may find it costly to delay approval of a drug, e.g. because he is facing pressure by patient advocacy groups, as for example in the case of the drug Flibanserin (Pollack (2015)) or drugs for HIV treatment (Epstein (1998)).

The skepticism effect describes situations in which the DM can credibly set a tougher standard for acceptance in equilibrium under hidden than under observable testing, which allows her to take a more informed decision. In a two-period model, suppose the DM wants to accept if and only if there are strictly more positive than negative outcomes and the advisor is more inclined to accept than the DM. Under observable testing, the advisor stops testing as soon as he has found one positive outcome and the DM accepts. However, if the advisor were to reveal a single positive outcome under hidden testing, the DM may reject. The reason is that the DM suspects that the advisor may only discovered a positive outcome on the second test and then omitted contradicting evidence. If the DM is sufficiently averse to falsely accepting, the advisor needs to show two positive outcomes for her to accept. In the example, if trials are registered then fewer trials with positive findings may be needed to convince the regulator. Although the evidence taken at face-value is stronger when trials are registered, the regulator also lowers his threshold for how much positive evidence is needed for approval. The skepticism effect shows that this can lead to less safe drugs appearing on the market.

The insurance effect describes situations in which the DM gains better information when testing is hidden, because the advisor has greater incentives to learn whether or not the hypothesis is true. In a two-period model, suppose the DM wants to reject if and only if at least one in two tests is negative, whereas the advisor wants to reject if and only if both are negative.[4] Under observable testing, the advisor runs no test at all and the DM accepts. To see why, suppose the advisor were to find a negative outcome. Then the DM would reject and she would reject irrespective of what a second test yields. However, a single negative outcome is not sufficient for the advisor to prefer rejection. By not testing, the advisor can guarantee that the DM acts in his interest. By contrast, under hidden testing, the advisor can test in private and hide outcomes if he wants to accept, but reveal all outcomes if he wants to reject. Therefore, the DM rejects if and only if the evidence is sufficiently negative to convince the advisor to reject. The insurance effect shows that a trial registry may deter companies from investigating whether or not certain side effects exist. To see why, suppose the regulator would approve the drug in the absence of further information. Without a registry, the company may run investigations in private and withdraw its application if it finds strong evidence that these side effects exist. However, with a registry, it may not conduct any investigations out of fear that it will find some evidence of these side effects, which is too weak for the company to lose interest in selling the drug, but strong enough for the regulator to decline approval.

However, the DM does not necessarily benefit from hidden testing. Although the advisor becomes better informed when testing is hidden, he also has a larger scope for strategic disclosure

[4]In this example, I assume that the test technology is such that a false positive outcome is more likely to arise than a false negative outcome. In a model with more than two periods, the insurance effect can exist even when this assumption does not hold.

because he has acquired more test outcomes. I show that if players' preferences are sufficiently misaligned then the DM weakly benefits from hidden testing. This demonstrates that whether or not a trial registry is optimal may depend on how much the regulator values the drug relative to the pharmaceutical company. For example, the regulator may be strongly averse to approving an unsafe drug relative to the company when a drug is intended mostly for cosmetic use, and then it would be better if the company selectively disclosed trials. Overall, hidden testing allows the DM to achieve her first-best expected payoff for a larger set of parameter combinations than observable testing, where her first-best expected payoff is defined as the one achieved if the DM could acquire evidence herself.

If the DM had the power to commit to her decision rule, she can achieve her first-best expected payoff, whether testing is hidden or observable. Therefore, the DM would ideally like to commit to a decision rule, but if this option is not available to her, then she may compensate for this to some extent by allowing the advisor to test in private. In addition, I show that if the advisor has the power to commit to a testing and a disclosure plan, then the skepticism effect ceases to exist, whereas the insurance effect persists. This provides a rationale for why pharmaceutical companies faced demands to register their trials for any drug, thereby precluding that companies could choose to side-step registries for individual drugs. In addition, I show that delegating decision rights to the advisor never makes the DM strictly better off than hidden testing. This implies that it is not necessary to give up decision rights in order to improve the advisor's incentives for information acquisition provided this information acquisition is private.

Although I use the drug approval process as my leading example, the insights also apply to the judicial process. By law, both the prosecutor and the judge or jury should prefer to convict defendants they know are guilty and acquit defendants they know are innocent. However, if the evidence is inconclusive, the prosecutor is more eager to convict than the judge. The US Supreme Court ruled that suppressing exculpatory exculpatory evidence violates the defendant's right to due process (Brady v. Maryland, 373 U.S. 83 (1963)). Yet, charges for prosecutorial misconduct are rare (Davis (2006)). It may be optimal not to deter the prosecutor from hiding evidence. This is because the prosecutor then has to provide stronger evidence that the defendant is guilty compared to a situation in which prosecutors are deterred from hiding evidence (skepticism effect). In addition, if the prosecutor is deterred from hiding evidence, then he may not pursue certain lines of investigation out of fear that the evidence in favor of the defendant's innocence could be weak, yet lead the judge to acquit (lack of insurance effect).

Another example is 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.[5] However, if scientists have to register their experiment then it is harder for an 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 paper is structured as follows. Section 3 introduces the model. Section 4 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 any arbitrary finite number of periods. Section 5 analyzes this comparison when the DM delegates decision rights to the advisor and when one of the players has commitment power. Finally, Section 6 analyzes the comparison when the advisor has to commit to the number of tests ex ante and when the horizon is infinite and the advisor incurs a constant cost per test. Section 7 concludes. All proofs can be found in the Appendix.

2 Related Literature

My work builds on the extensive literature on optimal persuasion when the sender strategically acquires information and no contracts can be written.[6] In particular, it forms part of the literature that treats the sender's testing technology as given and focuses on manipulation by different means.[7] One strand of this literature assumes that information acquisition is observable and asks how the sender can manipulate the receiver through strategically stopping the flow of information. Brocas and Carrillo (2007) were the first to study this type of manipulation in a discrete time model, while Henry and Ottaviani (2018) study a continuous time version.[8] Another strand of the literature assumes information acquisition is hidden and asks how the sender can manipulate the receiver by disclosing evidence selectively. 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.

My paper compares the receiver's payoff under hidden and observable information acquisition. Matthews and Postlewaite (1985) is an early paper to address a related question.[9] They assume that a seller faces a choice of whether or not to acquire a single costless signal about product quality.

[5] E.g. see the Berkeley Initiative for Transparency in the Social Sciences.

[6] A large literature studies persuasion when the sender 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] Following Kamenica and Gentzkow (2011), recent work has focused on how the sender can design the testing technology to strategically influence the receiver's action.

[8] See also McClellan (2017).

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

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 quality is good, which allows buyers to learn about quality. The reason the seller tests is because buyers interpret a lack of disclosure as bad news. More recently, 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 data base can result in full transparency, but reduce the company's incentive to test.[10] Unlike in these models, the advisor in my paper can acquire multiple signals, which expands his opportunities for manipulation through selective disclosure.

The assumption of multiple signals is also shared by the following closely related papers: Brocas and Carrillo (2007), Henry (2009) and Felgenhauer and Loerke (2017). Henry (2009) and Felgenhauer and Loerke (2017) find that the DM is weakly better off under hidden testing than observable testing, whereas Brocas and Carrillo (2007) find that the DM is equally well off in either regime. By contrast, I find that there are Pareto-undominated equilibria in which the DM is strictly worse off under hidden testing than under observable testing. Furthermore, my paper is the first to identify the insurance effect as a reason for why the DM can benefit from hidden testing. This effect has not been identified previously, because existing work in this area has assumed that players do not always agree on the optimal action under certainty.[11]

Brocas and Carrillo (2007) use a similar framework to mine, but assume that the advisor can verifiably disclose his posterior belief, whereas I assume that he can verifiably disclose test outcomes. As a consequence, in my setting, the advisor has more scope for strategic disclosure under hidden testing, which introduces equilibria in which the DM's payoff differs from that under observable testing.

In contrast to my paper, 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, because he can then report additional favorable realizations and hide unfavorable ones. However, in equilibrium, the DM can perfectly infer the quantity chosen by the advisor and, therefore, unraveling occurs in the vein of Milgrom (1981) and Grossman (1981). This effect is similar to the skepticism effect identified in this paper, but the skepticism effect does not exist for all parameter values in my setting. Since the advisor acquires test outcomes sequentially, the total number of tests run is path-dependent. Consequently, in equilibrium under hidden testing, the DM cannot infer the total

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

[11] Existing work has assumed that 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)), or, in models with continuous action choice, that the advisor's ideal action always differs from the DM's ideal action by a constant amount 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 4.2.

number of tests run.[12] This allows for equilibria in which the advisor manipulates the DM's choice to his advantage by hiding outcomes.

In Felgenhauer and Loerke (2017)'s framework an advisor cannot only decide how many tests to run, but also how informative each test will be. When the advisor discloses an outcome he also discloses the technology of the test by which it was generated. 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. While they study full flexibility in test design, I assume a fixed testing technology. 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.[13]

Di Tillio et al. (2017a) 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 observation and one in which she restricts him to only collect a single observation in public. They show that depending on the distribution, the DM's payoff may either be higher or lower when the advisor can selectively disclose his preferred outcome. I allow the advisor to disclose as many outcomes as desired, but restrict my attention to a binary distribution. Holding preference parameters fixed, my analysis gives conditions on the distribution parameters under which the DM strictly benefits from hidden testing.

My work is related to the literature on the ideal bias of an advisor when the DM can choose how closely aligned the advisor's preferences are with her own, e.g. Che and Kartik (2009), Gerardi and Yariv (2008) and Dur and Swank (2005). Che and Kartik (2009) primarily study a situation in which players' prior beliefs differ. My work relates to the part of their paper where they assume differences in preferences but not in prior beliefs. In their setting, 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, which is bad for the DM. However, since the DM responds less favorably if no signal is reported, a more biased advisor has greater incentives to exert effort. In my setting, a more biased advisor 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. Much of this literature assumes

[12] In addition, since the DM's choice is discrete, she does not necessarily adjust her response when her beliefs change.

[13] In a related paper, Liebgober (2017) studies how a scientist's choice of experiment is affected by whether the experimental "protocol" is observable or not. Through manipulating protocol, the scientist can increase the chances of a positive outcome. Without transparency, the scientist cannot commit to stick to standard protocol and this increases his incentives to invest in a more informative experiment which is less susceptible to manipulation.

that the advisor is exogenously informed, e.g. Li and Suen (2004), who use the same assumption on preferences as this paper, 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. In models with information acquisition, Argenziano et al. (2016) and Deimen and Szalay (2018) conclude that the DM prefers communication to delegation assuming talk is cheap. I show that this finding extends to the case in which disclosure is verifiable.

3 Model

I will first outline the model of hidden testing and then outline the benchmark model of observable testing.

Setting: Time is discrete and there are finitely many periods, $n = 1,...,N$. The players are a Decision Maker (DM) and an advisor (A).

Timing and Information: In period $n = 1$, 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 $a = 1$ or not $a = 0$. If he tests then Nature draws a test outcome $s_n \in \{+,-\}$, with state-dependent accuracy given by

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

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

where $\frac{1}{2} < p_i < 1$ for $i = F,T$ and the realizations of outcomes are independent conditional on the state.[14] If the advisor does not test then Nature does not draw a test outcome, and this event is denoted by $s_n = \emptyset$. The test outcome is privately observed by the advisor. At the end of period N, the advisor sends a message $m \in M$ to the DM, where the message space M is defined below. Then the DM chooses an action $\tau \in \{accept, reject\}$. Finally, payoffs are realized. It is assumed that players do not have commitment power.

[14] In Section 4.1, I allow for the informativeness of the test to depend on the state of the world, i.e. $p_T \neq p_F$.

Payoffs: A player $i = DM, A$ incurs the following state-dependent loss, denoted by $L_i(\tau,\omega)$:[15]

$L_i(\tau,\omega)$	$false$	$true$
$reject$	0	1
$accept$	λ_i	0

where $\lambda_i \geq 0$. Hence, if the state of the world were known then players would agree on what the optimal action is. Each player would prefer to accept if and only if the hypothesis is true. However, when the state of the world is uncertain then players may disagree on what the optimal action is, because players may face a different trade-off between the loss from accepting if false and the loss from rejecting if true. Without loss of generality, I normalize the loss from rejecting if true to be 1 for each player, but allow for the loss from accepting if false to differ between players. 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 $\lambda_A \leq \lambda_{DM}$.

Player i's expected payoff is given by $\pi_i(\tau) = -E(L_i(\tau))$, and the vector of expected payoffs is given by $\Pi = (\pi_A, \pi_{DM})$. Each player i maximizes their expected payoff π_i arising from the statistical decision problem.[16] Without loss of generality, I assume the DM's loss λ_{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:[17]

$$\pi_{DM}(reject) > \pi_{DM}(accept) \quad \Leftrightarrow \frac{q}{1-q} < \lambda_{DM}.$$

In addition, I assume that whenever the advisor is indifferent between testing or not, he will not test.[18]

[15] This framework is based on DeGroot (1970)'s framework for optimal statistical decisions. Consider the most general loss function $L_i(\tau,\omega)$:

$L_i(\tau,\omega)$	$false$	$true$
$reject$	a	b
$accept$	c	d

where $a,b,c,d \in \mathbb{R}$ and $b > a$ and $c > d$. For the analysis to follow, a sufficient statistic of this loss function is $\frac{c-d}{b-a} \geq 0$, i.e. the ratio of the excess loss from accepting if the hypothesis is true rather than false to the excess loss from rejecting if the hypothesis is true rather than false. For tractability, I choose the parameters such that the sufficient statistic is equal to the loss λ of falsely accepting and the narrative is that a player incurs a loss from type I and type II errors, but no loss from choosing an appropriate action given the state.

[16] I assume tests are costless for tractability, i.e. the advisor stops testing if and only if he can strategically manipulate the DM's choice. All my results continue to hold if the cost of testing is sufficiently small.

[17] If I restricted preferences such that *accept* is the DM's default choice, the results would be a "mirror image" of the results presented here.

[18] This corresponds to making the assumption that testing is costly and considering the limit when this cost becomes vanishingly small.

Histories: Let the *history of test outcomes* at the end of period $n = 1, ..., N$ be denoted by $h_n \in H_n$, which is an ordered list of past draws by Nature, i.e. $h_n = (s_1, s_2, ..., s_n)$ where $s_n \in \{+, -, \emptyset\}$. Let h_0 denote the history of outcomes at the start of the game. Let the *unordered history of test outcomes* at the end of period N be denoted by $\widetilde{h}$, and the set of all $\widetilde{h}$ be denoted by $\widetilde{H}$. The unordered history $\widetilde{h}$ is a set of past draws by Nature, which contains the realization of past draws, but not the order in which they were obtained. For example, if $N = 2$ then both $h_2 = (+, -)$ and $h_2 = (-, +)$ map into $\widetilde{h} = \{+, -\}$.

Message: The advisor's message space is given by $M = \mathscr{P}\left(\widetilde{h}\right) \in \mathscr{P}\left(\widetilde{H}\right)$, where $\mathscr{P}\left(\widetilde{h}\right)$ denotes the power set of $\widetilde{h}$, i.e. the set of all subsets of $\widetilde{h}$ including the empty set. For example, if $N = 2$ and $\widetilde{h} = \{+, -\}$ then the message space is given by $M = \{\{\emptyset\}, \{+\}, \{-\}, \{+, -\}\}$. The advisor can report any subset of test outcomes he has generated, which implies that he cannot make up outcomes but he is able to hide outcomes (verifiable disclosure). In addition, it implies that the advisor cannot verifiably disclose the period in which the outcome was found (not datable), i.e. if $N = 2$ and he reports $m = \{+\}$ he cannot convey 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: $\sigma_A : H_n \to \{0, 1\}$. It selects action $a \in \{0, 1\}$ conditional on history $h_n \in H_n$ for $n = 0, ..., N-1$. A reporting strategy for the advisor is $\sigma_M : H_N \to M$. It selects a message $m \in M$ conditional on history $h_N \in H_N$. A strategy for the DM is $\sigma_{DM} : M \to T$ where $T \in \{accept, reject\}$. It selects action $\tau \in T$ conditional on message $m \in M$. I assume that the DM accepts if indifferent.[19]

Equilibrium Concept: The solution concept is a sequential equilibrium in pure strategies.[20] A sequential equilibrium consists of both a profile of strategies $\sigma \equiv (\sigma_A, \sigma_M, \sigma_{DM})$ and a system of beliefs $\mu : \cup_{n=1}^{N} H_n \cup M \to \triangle\Omega$, where $\Omega = \{true, false\}$. Beliefs select a probability distribution over states $\omega \in \Omega$ for each history of outcomes $h_n \in H_n$ where $n = 1, ..., N$ and for each message $m \in M$. By definition of a sequential equilibrium,

1. the advisor's strategy (σ_A, σ_M) at any history h_n for $n = 0, ..., N-1$ maximizes his expected payoff given the DM's strategy σ_{DM} and given the system of beliefs μ, and

2. the DM's strategy σ_{DM} maximizes her expected payoff at any message m given the system of beliefs μ, and

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

[19]With this assumption, there is a unique sequential equilibrium under observable testing, whereas without it, there are multiple equilibria at knife-edge cases. To keep payoff comparisons more tractable, it is helpful to have uniqueness under observable testing.

[20]The restriction to pure strategies is for tractability. The main insights of Theorem 1 (for $N = 2$) and Proposition 3 (for $N > 2$) would apply if mixed strategies are allowed, but the boundaries of regions would change.

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 report superfluous. The DM's strategy is given by $\overline{\sigma}_{DM} : H_N \to T$. It selects action $\tau \in T$ conditional on history $h_N \in H_N$ (rather than conditional on a message). A sequential equilibrium consists of a strategy profile $\overline{\sigma} \equiv (\overline{\sigma}_A, \overline{\sigma}_{DM})$ and a system of beliefs $\overline{\mu} : \cup_{n=1}^{N} H_n \to \triangle\Omega$, where $\Omega = \{true, false\}$. By definition of a sequential equilibrium:

1. the advisor's strategy $\overline{\sigma}_A$ at any history h_n for $n = 0, ..., N$ maximizes his expected payoff given the DM's strategy $\overline{\sigma}_{DM}$ and given the system of beliefs $\overline{\mu}$, and

2. the DM's strategy $\overline{\sigma}_{DM}$ maximizes her expected payoff at any any history h_N given the system of beliefs $\overline{\mu}$, and

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

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 inference from the calendar time of reports. This is 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. For example, in a research context, it is often unknown how many experiments a scientist has carried out before he presents his findings, because it is unknown at what point the scientist started his inquiry, or because the time an experiment lasts is unknown or random, or because it is unclear whether or not the scientist had to devote time to other projects. The finite horizon represents a constraint on the advisor's resources to conduct tests. This setting with costless tests 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 6.[21]

[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 costless testing.

4 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 two effects for why the DM's expected payoff improves with 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 observable testing when preferences are sufficiently misaligned.

4.1 Two-period Model

Throughout this subsection, I assume that $N = 2$.

Lemma 1 (Equilibrium Existence and Payoff Uniqueness) *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, the vector of expected payoffs arising in equilibrium is unique. However, when testing is hidden, multiple expected payoff vectors can arise in equilibrium. As the aim of this paper is a welfare comparison across testing regimes, it seems natural to focus on the set of Pareto-undominated equilibrium payoff vectors, denoted by S^*. Then $S^* = \left\{\underline{\Pi}, \overline{\Pi}\right\}$, where $\underline{\Pi}$ is the payoff vector which assigns the lowest expected payoff to the DM and the highest expected payoff to the advisor within the set $S^*$of Pareto-undominated equilibrium payoff vectors, i.e.

$$\underline{\Pi} \equiv \underset{\Pi \in S^*}{\text{argmin}}\ \pi_{DM} = \underset{\Pi \in S^*}{\text{argmax}}\ \pi_A$$

and $\overline{\Pi}$ assigns the highest expected payoff to the DM and the lowest expected payoff to the advisor within S^*, i.e.

$$\overline{\Pi} \equiv \underset{\Pi \in S^*}{\text{argmax}}\ \pi_{DM} = \underset{\Pi \in S^*}{\text{argmin}}\ \pi_A.$$

I will refer to any equilibrium that gives rise to payoff vector $\underline{\Pi}$ as an *advisor-preferred equilibrium*, which reflects that this equilibrium assigns the advisor his maximum payoff across all equilibria, and similarly, I will refer to any equilibrium that gives rise to payoff vector $\overline{\Pi}$ as a *DM-preferred equilibrium.*

As a first step, it is worth noting that a player prefers to accept if and only if their loss λ from falsely accepting lies below their posterior likelihood ratio that the hypothesis is true. In particular, player i at some information set I_i prefers to accept if and only if

$$\pi_i(accept|I_i) = -Pr(false|I_i)\lambda_i \geq -Pr(true|I_i) = \pi_i(reject|I_i)$$
$$\Leftrightarrow$$
$$\lambda_i \leq \frac{Pr(true|I_i)}{Pr(false|I_i)}. \tag{3}$$

This observation is helpful to illustrate the nature of the conflict of interest between players. Suppose testing is observable. Then at the end of period 2, $I_i = h_2$ for $i \in \{A, DM\}$, and both players prefer to accept if

$$max\{\lambda_A, \lambda_{DM}\} < \frac{Pr(true|h_2)}{Pr(false|h_2)}$$

and both prefer to reject if

$$min\{\lambda_A, \lambda_{DM}\} > \frac{Pr(true|h_2)}{Pr(false|h_2)}$$

and, otherwise, players disagree on what their preferred action is.

When comparing players' expected payoffs between hidden and observable testing, it turns out that there are four preference parameter regions of interest. I will first define these regions and then state the main result on how the DM's payoffs compare across testing regimes. It is helpful to know that the region's boundaries turn out to correspond to posterior likelihood ratios conditional on some sets of outcome realizations from two tests. Denote the complete set of ordered outcome realizations from two tests by Φ, where

$$\Phi = \{(+,+), (+,-), (-,+), (-,-)\}. \tag{4}$$

Then define $x_{\cup\Phi}$ to be the posterior likelihood ratio conditional on some union of elements of Φ. Given the prior q and the test accuracy (p_F, p_T), the relevant boundaries are given by:

$$x_{(+,+)} = \frac{qp_T^2}{(1-q)(1-p_F)^2} \tag{5}$$

$$x_{(+,-)} = x_{(-,+)} = x_{(+,-)\cup(-,+)} = \frac{qp_T(1-p_T)}{(1-q)p_F(1-p_F)} \tag{6}$$

$$x_{(+,\cdot)} \equiv x_{(+,+)\cup(+,-)} = \frac{qp_T}{(1-q)(1-p_F)} \tag{7}$$

$$x_{\Phi\setminus(-,-)} = \frac{q\left(1-(1-p_T)^2\right)}{(1-q)\left(1-p_F^2\right)} \tag{8}$$

$$x_{\Phi} = \frac{q}{1-q}. \tag{9}$$

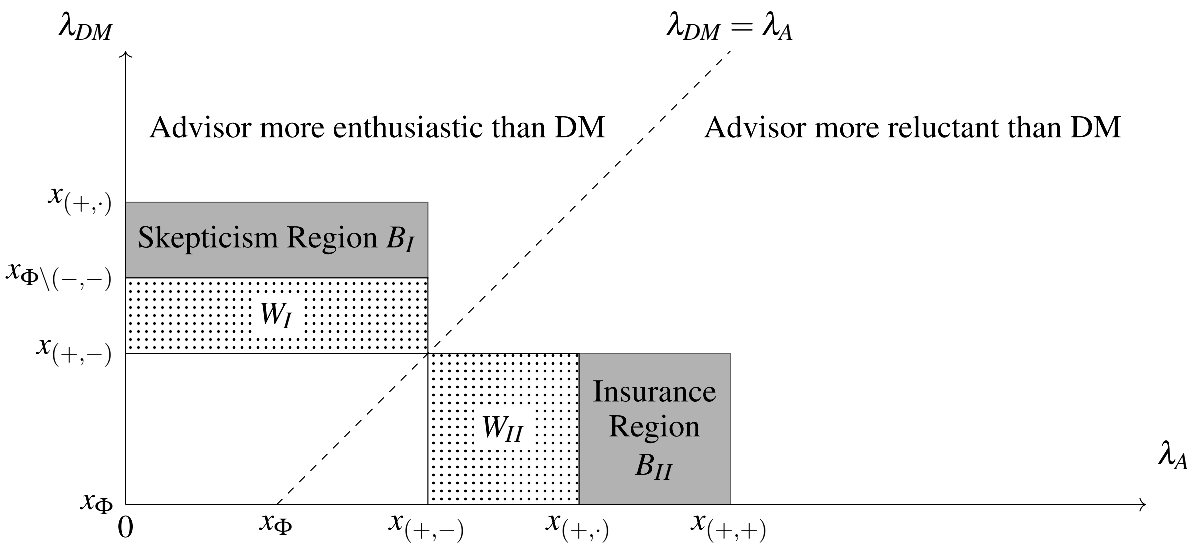

Figure 1: This figure shows the regions in which the DM is strictly better off under hidden rather than observable testing in any Pareto-undominated equilibrium, i.e. the skepticism region B_I and the insurance region B_{II}, and the regions in which the DM is strictly worse off under some Pareto-undominated equilibrium (labeled W_I and W_{II}) in $(\lambda_A, \lambda_{DM})-$space given $p_F > p_T$. If $p_T \geq p_F$ then regions B_{II} and W_{II} would be empty because there exists no value of λ_{DM} such that $x_\Phi < \lambda_{DM} \leq x_{(+,-)}$.

The four regions of interest are defined as follows:

$$B_I \equiv \left\{(\lambda_{DM}, \lambda_A) : \lambda_A < x_{(+,-)}, x_{\Phi\backslash(-,-)} < \lambda_{DM} \leq x_{(+,\cdot)}\right\},$$
$$B_{II} \equiv \left\{(\lambda_{DM}, \lambda_A) : x_\Phi < \lambda_{DM} \leq x_{(+,-)}, x_{(+,\cdot)} \leq \lambda_A < x_{(+,+)}\right\},$$
$$W_I \equiv \left\{(\lambda_{DM}, \lambda_A) : \lambda_A < x_{(+,-)}, \max\left\{x_\Phi, x_{(+,-)}\right\} < \lambda_{DM} \leq x_{\Phi\backslash(-,-)}\right\},$$
$$W_{II} \equiv \left\{(\lambda_{DM}, \lambda_A) : x_\Phi < \lambda_{DM} < x_{(+,-)} \leq \lambda_A < x_{(+,\cdot)}\right\}.$$

I will refer to B_I as the *skepticism region* and to B_{II} as the *insurance region*.[22]

Theorem 1 (DM Payoff Comparison)

1. *In any Pareto-undominated equilibrium under hidden testing the DM is strictly better off than in the unique equilibrium under observable testing if and only if*

   (a) *preferences lie in the skepticism region, i.e.* $(\lambda_{DM}, \lambda_A) \in B_I$*, or*

   (b) *preferences lie in the insurance region, i.e.* $(\lambda_{DM}, \lambda_A) \in B_{II}$.

2. *There exist a Pareto-undominated equilibrium under hidden testing in which the DM is strictly worse than in the unique equilibrium under observable testing if and only if* $(\lambda_{DM}, \lambda_A) \in$

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

$W_I \cup W_{II}$.

The comparison of the DM's expected payoff is illustrated in Figure 1. Theorem 1 shows that the DM strictly benefits from hidden testing for some parameters, while for other parameters it is possible that he strictly suffers from hidden testing, considering only Pareto-undominated equilibria. Interestingly, the economic effect which causes the DM to be better off in the skepticism region, where the advisor is more enthusiastic about accepting than the DM ($\lambda_A < \lambda_{DM}$), turns out to be distinct from the economic 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 skepticism region is similar to effects identified in the existing literature, while the insurance effect is novel.

To explain these results, I will consider each of the aforementioned regions in turn.[23] For the case of observable testing, I will characterize the unique equilibrium. For the case of hidden testing, I will characterize an advisor-preferred equilibrium.[24] Then I will compare the DM's expected payoffs across both testing regimes and explain the effects at work. I focus on advisor-preferred equilibria because if and only if the DM is strictly better off under hidden testing in an advisor-preferred equilibrium then the DM is strictly better off under hidden testing in any Pareto-undominated equilibrium.

Lemma 2 (Equilibria Insurance Region)
Suppose $(\lambda_{DM}, \lambda_A) \in B_{II}$.

- *Equilibrium under observable testing: the DM accepts if and only if* $\widetilde{h} \in \{\{+,+\},\{+,-\},\{+\}\}$*; the advisor never tests at any history of outcomes.*
- *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*

$$\sigma_M = \begin{cases} \{\emptyset\} & \text{if } \widetilde{h} \in \{\{\emptyset\},\{+\}\} \\ \{-\} & \text{if } \widetilde{h} \in \{\{-\},\{+,-\}\} \\ \widetilde{h} & \text{otherwise}; \end{cases}$$

[23] I characterize equilibria for the preference parameter regions in which expected payoffs remain unchanged in Sections A.1 and A.2 in the Appendix.

[24] The characterization requires additional likelihood ratios defined based on $x_{\cup\Phi}$: $x_{(-,\cdot)} \equiv x_{(-,+)\cup(-,-)} = \frac{q(1-p_T)}{(1-q)p_F}$ and $x_{(-,-)} = \frac{q(1-p_T)^2}{(1-q)p_F^2}$ and $x_{\Phi\setminus(+,+)} = \frac{q\left(1-p_T^2\right)}{(1-q)\left(1-(1-p_F)^2\right)}$.

The DM's beliefs conditional on messages satisfy

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{\Phi\backslash(+,+)} & \textit{if } m = \{-\} \\ x_{(+,\cdot)} & \textit{if } m \in \{\emptyset, +\} \\ x_m & \textit{otherwise.}^{25} \end{cases}$$

In the insurance region B_{II}, without any further information the DM prefers to reject, 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 a positive and a negative outcome together are evidence in favor of the hypothesis, i.e. $p_F > p_T$.[26] The advisor is more reluctant and prefers to accept 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 had tested in period 1 then the advisor would never test in period 2. This is because if the first test outcome was positive, the DM would accept regardless of whether the second test turned out positive or negative, so there is no reason for the advisor to test in period 2. And if the first test outcome was negative, the advisor would prefer rejection regardless of the second outcome and without a second test the DM would act in his interest and reject. But the advisor is better off if he does not test than if he runs a single test. The reason is that, whatever the outcome of a single test, the advisor would always prefer to reject. However, the DM accepts following a single positive outcome, but rejects if no test is run. By not testing, the advisor ensures that the DM acts in his interest.

By contrast, on the equilibrium path under hidden testing, the advisor tests to find out if both outcomes are positive. If he indeed discovers two positive outcomes, then he reveals these outcomes and the DM acts in his interest and accepts. Otherwise, he reveals only a single negative outcome and the DM again acts in his interest and rejects. When the advisor reveals a single negative, the DM does not know if the advisor either found a negative outcome on the first test and then stopped or if he first found a positive and then a negative outcome and is hiding the positive outcome. However, since a report of two positive outcomes is good news about the hypothesis being true, the DM infers that a report of a single negative outcome must be bad news.[27] Therefore, the

[25] Note that x_m is a slight abuse of notation since m is an unordered set of outcomes whereas $x_{\cup\Phi}$ is defined based on a set of complete list of ordered outcomes. However, if m contains two outcomes then the order of these outcomes does not matter. The notation is short for

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{(+,+)} & \text{if } m = \{+,+\} \\ x_{(-,-)} & \text{if } m = \{-,-\} \\ x_{(+,-)} = x_{(-,+)} & \text{if } m = \{+,-\}. \end{cases}$$

[26] If $N > 2$ an insurance region exists even if $p_F = p_T$ (see Section 4.2).

[27] The DM either receives a report of two positives or of a single negative. A report of two positives raises her

DM optimally rejects if the advisor reports one negative outcome.

In conclusion, under observable testing, the advisor strategically avoids testing because testing has the disadvantage that the evidence could turn out to be strong enough to convince the DM to accept, but too weak to convince the advisor that acceptance is optimal. By allowing the advisor to test in private, the DM insures the advisor against this disadvantage. As a consequence, the DM at least learns whether or not the evidence is strong enough for both to agree to accept. 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 benefits the DM. Although she forgoes the option to accept following mixed outcomes, the fact that the advisor then takes the gamble and tests results in an overall gain for her.

Lemma 3 (Equilibria Skepticism Region)
Suppose $(\lambda_{DM}, \lambda_A) \in B_I$.

- *Equilibrium under observable testing: the DM accepts if and only if* $\widetilde{h} \in \{\{+,+\},\{+\}\}$*; the advisor tests once.*[28]

- *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*

$$\sigma_M = \begin{cases} \{+,+\} & \text{if } \widetilde{h} = \{+,+\} \\ \{+\} & \text{if } \widetilde{h} \in \{\{+\},\{+,-\}\} \\ \{\emptyset\} & \text{if } \widetilde{h} \in \{\{-\},\{-,-\}\}. \end{cases} \tag{10}$$

The DM's beliefs conditional on messages satisfy

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{(+,-)} & \text{if } m = \{+\} \\ x_{(-,\cdot)} & \text{if } m = \{\emptyset\} \\ x_{\Phi \setminus \{+,+\}} & \text{if } m = \{-\} \\ x_m & \text{otherwise.} \end{cases}$$

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.

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

In the skepticism region B_I, the DM would prefer to accept if she knew that a single test was run and the outcome was positive. However, she would prefer to reject if she knew that two tests were run and outcomes were mixed.[29] By contrast, the advisor is more enthusiastic and would prefer to accept even if two tests were run and the outcomes were mixed.[30]

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, there is no reason to test in the second period. The advisor prefers acceptance irrespective of what a second test shows and if he stops testing the DM acts in his interest and accepts. Following a first negative outcome, there is no reason to test because the DM would reject irrespective of the second outcome. But it is optimal to run a single test. 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, the advisor tests because gathering additional information increases the chances that the DM chooses acceptance if the hypothesis is true or rejection if it is false. If the DM and the advisor disagreed on their preferred action at the prior belief, then the advisor also benefits from the fact that the DM chooses his preferred action 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, just as under observable testing. However, unlike under observable testing, the advisor would be tempted to keep 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. However, 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 $x_{\Phi\setminus(-,-)}$, 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. However, by revealing two positive outcomes, the advisor could convince the DM to accept.

In conclusion, under hidden testing the DM learns whether a positive outcome is backed up by an additional test, and this allows her to reduce her expected loss. 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

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

[30] The advisor may even want to accept irrespective of what the test outcomes are.

outcomes to be convinced.

While the DM is better off in any Pareto-undominated equilibrium under hidden testing in the skepticism region, there are other regions in which the DM can end up strictly worse off when testing is hidden.

Lemma 4 (Equilibria Regions W_I and W_{II})

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

   - *Equilibrium under observable testing is as in* $(\lambda_{DM}, \lambda_A) \in B_I$.
   - *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 (10). The DM's beliefs conditional on messages satisfy*

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{\Phi\setminus(-,-)} & \text{if } m = \{+\} \\ x_{(-,\cdot)} & \text{if } m = \{-\} \\ x_{\Phi} & \text{if } m = \{\emptyset\} \\ x_m & \text{otherwise.} \end{cases}$$

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

   - *Equilibrium under observable testing: the DM accepts if and only if* $\widetilde{h} \in \{\{+,+\}, \{+,-\}, \{+\}\}$*; the advisor tests once.*
   - *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*

$$\sigma_M = \begin{cases} \{-\} & \text{if } \widetilde{h} \in \{\{-\}, \{+,-\}\} \\ \widetilde{h} & \text{otherwise.} \end{cases} \tag{11}$$

   The DM's beliefs conditional on messages satisfy

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{\Phi\setminus(+,+)} & \text{if } m = \{-\} \\ x_{(+,\cdot)} & \text{if } m = \{+\} \\ x_{\Phi} & \text{if } m = \{\emptyset\} \\ x_m & \text{otherwise.} \end{cases}$$

Consider $(\lambda_{DM}, \lambda_A) \in W_I$. Under observable testing, the advisor tests once and the DM accepts if and only if the outcome is positive, as is the case in the skepticism region B_I. However, the DM has a lower loss from falsely accepting than in the skepticism region B_I and this changes what happens on the equilibrium path under hidden testing. Just as when $(\lambda_{DM}, \lambda_A) \in B_I$, if the DM were to accept based on a single positive outcome, then the advisor would report a single positive outcome if either he discovered it in period 1 or if he discovered it in period 2 following a negative outcome. When a single positive outcome is reported, all the DM can infer is that not both tests were negative, but in contrast to when $(\lambda_{DM}, \lambda_A) \in B_I$, she is willing to accept based on this inference, i.e. $\lambda_{DM} < x_{\Phi \setminus (-,-)}$. 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 $(\lambda_{DM}, \lambda_A) \in W_{II}$. The advisor has a lower loss from falsely accepting than in the insurance region B_{II} and, as a consequence, he would prefer to accept if he knew that a single test was run and the outcome was positive. Under observable testing, the advisor tests once and the DM accepts if and only if the outcome is positive. The advisor has an incentive to run a single test, because following either outcome the DM acts in her interest. However, he has no reason to run a second test. This is because, following a positive outcome, the DM accepts regardless of the outcome of the second test. In addition, following a negative outcome, the advisor prefers to reject regardless of the second test and the DM rejects without a second test. Under hidden testing, the advisor tests to find see if he can find two positive outcomes and the DM accepts if and only if both outcomes are positive. Therefore, as in the insurance region, hidden testing can be thought of providing limited liability to the advisor. It guarantees that whenever outcomes are mixed the advisor's preferred action is chosen instead of the DM's preferred action. However, unlike in the insurance region, the DM does not benefit from this. The limited liability encourages the advisor to run a second test following a positive outcome, but, unlike in the insurance region, this is not necessary to encourage him to run the first test. Therefore, the DM has no advantage from encouraging the advisor to run a second test, only the disadvantage that rejection is chosen when outcomes are mixed.

Theorem 1 has shown that the DM can be strictly harmed by hidden testing, even when she acts perfectly rational. However, the DM is weakly better off in any Pareto-undominated equilibrium under hidden rather than in the unique equilibrium under 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 the DM is weakly better off in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing if*

$$|\lambda_A - \lambda_{DM}| > d.$$

It is also interesting to compare the DM's expected payoff under hidden testing and observable testing to the expected payoff the DM could achieve if she collected the evidence herself. I will refer to the expected payoff in this benchmark case as the DM's *first-best expected payoff*. Let $Z_{OT} \in \mathbb{R}^2$ denote the set of all preference parameters $(\lambda_{DM}, \lambda_A)$ for which the DM achieves her first-best expected payoff under observable testing, and let $Z_{HT} \in \mathbb{R}^2$ denote the set of all preference parameters $(\lambda_{DM}, \lambda_A)$ for which the DM achieves her first-best expected payoff in any Pareto-undominated equilibrium under hidden testing.

Proposition 1 (First-Best Benchmark)

1. *The DM achieves her first-best expected payoff for a larger set of parameter combinations in any Pareto-undominated equilibrium under hidden rather than in the unique equilibrium under observable testing, i.e.* $Z_{OT} \subset Z_{HT}$.

2. *There always exist a Pareto-undominated equilibrium under hidden testing in which the DM achieves her first-best expected payoff.*

The DM never achieves her first-best expected payoff under observable testing in any of the regions B_I, B_{II}, W_I or W_{II} because the advisor strategically avoids conducting some tests which would be pivotal to the DM's choice. However, the DM achieves her first-best expected payoff in any Pareto-undominated equilibrium under hidden testing in the skepticism region B_I.

Furthermore, it is worth noting that 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.2. These equilibria have the common feature that the DM acts against the advisor's interest unless the evidence convinces her to do otherwise, similar to the unraveling logic in Henry (2009). For example, if $(\lambda_{DM}, \lambda_A) \in W_I$, there exists an 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, he will continue testing in period 2 if the first test was positive. If he ends up with one positive and one negative outcome, the advisor is indifferent between all feasible messages, but if he chooses to reveal only the positive, then the DM's beliefs are correct. [31]

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.

[31] 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.

Proposition 2 (Advisor Payoff Comparison)

1. *There always exists a Pareto-undominated equilibrium under hidden testing in which the advisor is weakly worse off than in the unique equilibrium under observable testing.*

2. *In any Pareto-undominated equilibrium under hidden testing, the advisor is strictly worse off than in the unique equilibrium under observable testing if and only if preferences lie in the skepticism region, i.e.* $(\lambda_{DM}, \lambda_A) \in B_I$.

3. *There exists a Pareto-undominated equilibrium under hidden testing in which the advisor is strictly better off than in the unique equilibrium under observable testing if and only if* $(\lambda_{DM}, \lambda_A) \in B_{II} \cup W_I \cup W_{II}$.

In the skepticism region, the advisor is strictly worse off under hidden testing because the DM has a higher acceptance threshold of acceptance 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 B_{II} as well as in regions W_I and W_{II}, 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. This illustrates a further point of contrast between the skepticism and the insurance region: in the insurance region, not just the DM but also the advisor can strictly benefit from testing being hidden rather than observable.

4.2 N Periods

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

To build up to comparing the DM's expected payoff under hidden and observable testing, I will first show that under observable testing, it is part of the equilibrium that the advisor stops testing strategically at a history of outcomes at which the two parties agree on the optimal action, provided preferences are sufficiently misaligned. Second, I will show that in the advisor-preferred equilibrium under hidden testing, the DM will learn only whether or not the number of positive or negative outcomes found reaches some threshold.

I will denote a history of outcomes at the end of period n by $h_n = (y, z)$, where $y \in \mathbb{N}_{\geq 0}$ denotes the number of positive outcomes and $z \in \mathbb{N}_{\geq 0}$ denotes the number of negative outcomes. 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 will use x_j to denote the posterior likelihood ratio that the hypothesis is true conditional on observing j excess positive outcomes, i.e. $h_n = (y,z)$ where $y - z = j$ and $j \in \{-N, ..., N\}$. Then x_j is given by

$$x_j = \frac{qp^y (1-p)^z}{(1-q)(1-p)^y p^z} = \frac{qp^{y-z}}{(1-q)(1-p)^{y-z}} = \frac{qp^j}{(1-q)(1-p)^j}.$$

To compare players' payoffs across regimes, the following lemma make statements about on-path equilibrium play for any possible realization of Nature's draws of outcomes. A complete equilibrium characterization can be found in the proofs in the Appendix.

Lemma 5 (Observable Testing) *Suppose testing is observable, $x_l < \lambda_{DM} \leq x_{l+1}$ where $l \in \{0, ..., N-1\}$, which implies that the DM prefers to accept for some realizations of the N test outcomes. The following occurs on the equilibrium path.*

1. *(Advisor more reluctant.) For any $N > 2$ and any λ_{DM}, there exists a critical value $\overline{\lambda}_A \geq \lambda_{DM}$, such that the advisor never starts testing and the DM rejects.*

2. *(Advisor more enthusiastic.) For any $N > 2$ and any λ_{DM}, there exists a critical value $\underline{\lambda}_A \leq \lambda_{DM}$, such that if $\lambda_A < \underline{\lambda}_A$ an equilibrium exists in which the advisor tests in period $n+1$ given history $h_n = (y,z)$ if and only if $l - (N-n) < y - z < l + 1$ for $n \in \{0, ..., N-1\}$. If the advisor stops at $y - z = l + 1$ then the DM accepts, otherwise the DM rejects.*

Under observable testing, a sufficiently enthusiastic advisor finds it optimal to stop testing if he has discovered just enough evidence for the DM to accept, while a sufficiently reluctant advisor finds it optimal to never start testing, which guarantees that the DM rejects.

In particular, suppose the advisor is more enthusiastic than the DM. The advisor keeps testing if the evidence collected up to now leads the DM to reject but for some realization of future outcomes the DM will accept. The reason is that if future outcomes turn out to be such that the DM accepts, then they also lead the advisor to prefer acceptance. Hence, if outcomes turn out to be pivotal to the DM's choice, then they cause the DM to act in the advisor's interest. 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 will reject. 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 that no realization of future outcomes leads him to prefer rejection. A situation in which the advisor always stops as soon as the evidence leads the DM to accept is illustrated in Figure 2.

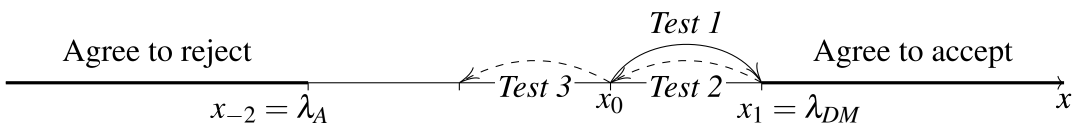

Figure 2: An example of a combination of preference parameters given $N = 3$ for which the advisor optimally stops testing as soon as the likelihood ratio leads the DM to accept: if the first outcome is positive, the advisor stops testing because the DM acts in line with his interest and the remaining 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. In particular, the advisor faces a trade-off if the evidence collected up to now leads the DM to reject, but for some realization of future outcomes the DM will accept. I find a critical value $\overline{\lambda}_A$ such that an advisor who has a higher loss from falsely accepting, i.e. $\overline{\lambda}_A < \lambda_A$, will not test in this situation, although there are realizations of future outcomes for which both players would agree that accepting is optimal. Consider the following example illustrated in Figure 3, where $N = 3$, the DM is just willing to accept if there is at least one more positive outcome than there are negative outcomes, i.e. $\lambda_{DM} = x_1$, and the advisor is just willing to accept if there are at least three more positive than negative outcomes, i.e. $\lambda_A = x_3$. If the first two outcomes happened to be positive, at which point the likelihood ratio is x_2, the DM will accept regardless of what the final test will show and, therefore, the advisor has no reason to do the final test. However, at likelihood ratio x_2, the advisor prefers to reject. If the advisor tests and stops at likelihood ratio x_2 the outcomes will never lead the advisor to prefer acceptance, but they could lead the DM to accept. 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 will accept regardless of what the remaining tests show. In the example, this corresponds to the likelihood ratio x_2. If the advisor prefers rejection at this likelihood ratio, then he is sufficiently reluctant to not start testing.

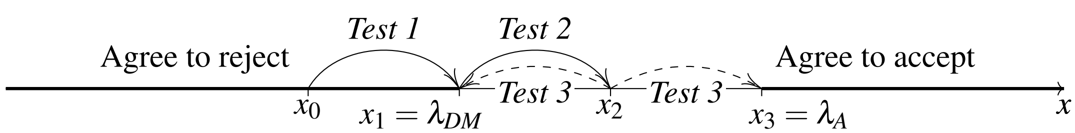

Figure 3: An example of a combination of preference parameters given $N = 3$ for which the advisor optimally does not start testing, although there exists some realization of test outcomes such that both players agree that accepting is optimal.

Under hidden testing, the following lemma gives an example of equilibrium play which can generate the expected payoffs attained in the advisor-preferred equilibrium. Let $m = (\widetilde{y}, \widetilde{z})$ denote a message that contains $\widetilde{y}$ positive outcomes and $\widetilde{z}$ negative outcomes.

Lemma 6 (Hidden Testing) *Suppose testing is hidden and $x_l < \lambda_{DM} \leq x_{l+1}$ where $l \in \{0,...,N-1\}$ and $x_r < \lambda_A \leq x_{r+1}$ where $r \in \mathbb{Z}$. There exist a threshold $\bar{y}$ of positive outcomes and a threshold $\bar{z}$ of negative outcomes given by*

$$\bar{y} \equiv inf\left\{j \in \mathbb{N}, j \geq \frac{r+1+N}{2}\right\},$$
$$\bar{z} \equiv inf\left\{j \in \mathbb{N}, j \geq \frac{N-r}{2}\right\} = N - \bar{y} + 1,$$

such that the following occurs on the equilibrium path in an advisor-preferred equilibrium.

1. *(Advisor more reluctant.) For any $N > 2$ and any $(\lambda_A, \lambda_{DM})$ where $\lambda_A > \lambda_{DM}$, the advisor tests in period $n+1$ given $h_n = (y,z)$ if and only if $y < \bar{y}$ and $z < \bar{z}$ for $n \in \{0,...,N-1\}$. In period N, either $y = \bar{y}$ or $z = \bar{z}$. If $y = \bar{y}$, he reports $m = (\bar{y}, 0)$ and the DM accepts. If $z = \bar{z}$, he reports $m = (0, \bar{z})$ and the DM rejects.*

2. *(Advisor more enthusiastic.) There exist further thresholds Y and Z given by*

$$Y \equiv inf\left\{j | j \in \mathbb{N}_{\geq 0}, \lambda_{DM} \leq \frac{q\sum_{s=j}^{s=N}\binom{N}{s}p^s(1-p)^{N-s}}{(1-q)\sum_{s=j}^{s=N}\binom{N}{s}(1-p)^s p^{N-s}}\right\}, \tag{12}$$
$$Z \equiv N - Y + 1,$$

such that for any $N > 2$ and any $(\lambda_A, \lambda_{DM})$ where $\lambda_A < \lambda_{DM}$, if $Y \leq \bar{y}$ then the equilibrium path is as in Part 1. Otherwise, the advisor tests in period $n+1$ given $h_n = (y,z)$ if and only if $y < Y$ and $z < Z$ for $n \in \{0,...,N-1\}$. In period N, either $y = Y$ or $z = Z$. If $y = Y$ then the advisor reports $m = (Y, 0)$ and the DM accepts. If $z = Z$ and $y \leq r + Z$ then the advisor reports $m = (0, Z)$ and the DM rejects. If $z = Z$ and $y > r + Z$ then the advisor reports $m = (y, 0)$ and the DM rejects.

First, consider an advisor who is more reluctant than the DM. The threshold $\bar{y}$ is defined to be the smallest number of positive outcomes such that if the advisor has observed $\bar{y}$ outcomes, then he prefers to accept, irrespective of the realization of the remaining $N - \bar{y}$ outcomes. Similarly, the threshold $\bar{z}$ is defined to be the smallest number of negative outcomes such that if the advisor has observed $\bar{z}$ outcomes, then he prefers to reject, irrespective of the realization of the remaining $N - \bar{z}$ outcomes. Lemma 6 shows that the advisor tests until the remaining tests cannot be pivotal to determining which action he prefers. 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 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 acceptance 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.

Second, consider an advisor who is more enthusiastic than the DM. Additional tests cannot be pivotal to the advisor's preferred action if $y \geq \bar{y}$. Suppose the advisor stopped testing as soon as $y = \bar{y}$ and reported $\bar{y}$ positive outcomes and no negative outcomes. Then the DM can only infer that at least $\bar{y}$ outcomes in N tests were positive. The threshold Y is defined to be the smallest number of positive outcomes such that if the DM believes that at least Y outcomes in N tests are positive, she optimally accepts. Hence, if $Y \leq \bar{y}$ then the DM would optimally accept if the advisor reported $\bar{y}$ positive outcomes. Therefore, if $Y \leq \bar{y}$ the advisor indeed stops testing as soon as $y = \bar{y}$. However, if $Y > \bar{y}$ the DM would optimally reject if the advisor reported $\bar{y}$ positive outcomes. Therefore, if $Y > \bar{y}$ the advisor has reason to keep testing until either $y = Y$ or $z = N - Y + 1$. If $y = Y$ he can report all positive outcomes and the DM accepts. If $z = N - Y + 1$ then he can never find Y positive outcomes and the DM rejects independent of the report. Therefore, if $Y > \bar{y}$, a more enthusiastic advisor faces an additional constraint compared to a more reluctant advisor and due to this additional constraint the DM will not always act in his interest.

The following proposition shows that analogues of regions B_I, B_{II}, W_I and W_{II} exist for any number of periods $N > 2$. In particular, it shows that the skepticism and insurance effects arise for any number of periods $N > 2$.

Proposition 3 (DM Payoff Comparison)

1. *Suppose the advisor is more enthusiastic than the DM, i.e. $\lambda_A < \lambda_{DM}$. Then for any $N \geq 2$ there exists*
   (a) *a skepticism region $B_1 \subset \mathbb{R}^2$ such that if $(\lambda_{DM}, \lambda_A) \in B_1$ the DM is strictly better off in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing, and*
   (b) *a region $W_1 \subset \mathbb{R}^2$ such that if $(\lambda_{DM}, \lambda_A) \in W_1$ there exists a Pareto-undominated equilibrium under hidden testing in which the DM is strictly worse than in the unique equilibrium under observable testing.*
2. *Suppose the advisor is more reluctant than the DM, i.e. $\lambda_A > \lambda_{DM}$. Then for any $N > 2$, there exists*
   (a) *an insurance region $B_2 \subset \mathbb{R}^2$ such that if $(\lambda_{DM}, \lambda_A) \in B_2$ the DM is strictly better off in any Pareto-undominated equilibrium under hidden than in the unique equilibrium under observable testing, and*

(b) a region $W_2 \subset \mathbb{R}^2$ such that if $(\lambda_{DM}, \lambda_A) \in W_2$ there exists a Pareto-undominated equilibrium under hidden testing in which the DM is strictly worse than in the unique equilibrium under observable testing.

The regions B_1 and W_1 are chosen in a way that the DM does not achieve her first-best expected payoff under observable testing. The advisor is 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 (see Lemma 5). In addition, the DM's loss from falsely accepting (λ_{DM}) is low enough that it is possible she would benefit from observing further tests.

By contrast, under hidden testing, the advisor has the possibility to test more and hide negative outcomes. In region B_1 the DM's loss from falsely accepting is high enough for her to require a higher number of positive outcomes for acceptance under hidden testing than she required excess positive outcomes under observable testing.[32] As a consequence, she obtains additional decision-relevant information about the hypothesis under hidden than observable testing. However, in the region W_1, the DM's loss from falsely accepting is so low 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 also hide negative outcomes when testing is hidden, she learns less about the hypothesis than when testing is observable.

To construct the region B_2, it is crucial that the players agree on the optimal action under certainty. In particular, the advisor needs to have a sufficiently high loss from falsely accepting such that, under observable testing, he does not test and both agree that rejection is optimal at the prior (see Lemma 5). In addition, the advisor's loss from falsely accepting needs to be low enough such that there exists a history of outcomes at which he agrees with the DM that acceptance is optimal. This ensures that the advisor tests when testing is hidden. This is because the DM acts in line with his interest whatever he finds and testing increases the chances that the DM's action choice is appropriate given the state (see Lemma 6). Note that no additional condition is needed to ensure that the DM acts in the advisor's interest under hidden testing. The reason is that a report of only positives occurs when the advisor prefers acceptance, which implies that the DM prefers acceptance, given the advisor is more reluctant than the DM. This implies that a report of only positives must raise the DM's posterior belief that the hypothesis is true. Since her expected posterior must equal her prior belief, this implies that a report of only negative outcomes must lower her posterior belief that the hypothesis is true.

Lastly, region W_2 is chosen such that preferences are sufficiently aligned for the advisor to test even when testing is observable. This implies there must exist a history of outcomes for which

[32] By Lemma 6, in equilibrium under hidden testing, acceptance is chosen given a belief that y or more outcomes in N tests are positive, where the value of y depends on the players' preferences. To ensure that the DM learns more under hidden testing, it is sufficient that this value of y exceeds the number of excess positive outcomes at which the DM is just willing to accept under observable testing.

the advisor prefers acceptance and, therefore, the equilibrium under hidden testing is unchanged from B_2. Under observable testing, the advisor stops either if he prefers rejection irrespective of the realization of the remaining outcomes or if the DM prefers acceptance irrespective of the realization of the remaining test outcomes. Under hidden testing, the advisor stops testing if his own preferred choice is independent of the realization of the remaining outcomes. Therefore, the additional testing does not benefit the DM. To the contrary, the DM now rejects at histories at which she would have preferred to accept.

The description of the construction of region B_2 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 testing 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 action is and, hence, the advisor is better off running both tests than not testing at all even when testing is observable.[33]

To summarize the payoff comparison for the DM, Proposition 1 continues to hold for $N > 2$.

Proposition 4 (Preference Alignment) *Given any $N > 2$, there exists a threshold $d > 0$ such that the DM is weakly better off in any Pareto-undominated equilibrium under hidden rather than in the unique equilibrium under observable testing if*

$$|\lambda_A - \lambda_{DM}| > d.$$

The following continues to hold regarding the DM's first-best benchmark in the case of $N \geq 2$.

Proposition 5 (First-Best Benchmark) *Given any $N \geq 2$, if $(\lambda_{DM}, \lambda_A) \in B_1$ 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.*

Finally, the following statement about the advisor's payoff comparison applies when $N \geq 2$.

Proposition 6 (Advisor Payoff Comparison) *Given any $N \geq 2$, if $(\lambda_{DM}, \lambda_A) \in B_1$ then the advisor is strictly worse off in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing.*

[33]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. Even if the advisor prefers acceptance if both outcomes are positive, players may disagree on the optimal action following mixed outcomes. To generate the insurance effect, the advisor must prefer rejection following a single positive outcome. Under observable testing, the advisor stops after a single positive outcome because the DM accepts regardless of the second outcome, but this implies that the advisor can never reach a history at which he prefers acceptance. Hence, he is better off not testing and ensuring that the DM rejects.

5 Delegation and Commitment

This section first compares expected payoffs in the baseline model to those achieved when the DM delegates decision-making authority to the advisor. It then studies 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 $N \geq 2$ and assume that the test has accuracy p independent of the state.

The following proposition shows that the DM weakly prefers to retain decision-making authority rather than delegating it to the advisor if testing is hidden, regardless of how aligned preferences are. However, the same is not true if testing is observable.

Proposition 7 (Delegation) *Suppose the DM could delegate decision-making authority to the advisor.*

1. *The DM is never strictly better off under delegation than hidden testing.*
2. *For any $N > 2$, if $(\lambda_{DM}, \lambda_A) \in B_2$ then the DM is strictly better off under delegation than observable testing.*
3. *For $N > 2$, if $(\lambda_{DM}, \lambda_A) \in W_2$ then the DM is strictly worse off under delegation than observable testing.*

A disadvantage of delegation 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 delegation can exist if the advisor has incentives to conduct more tests than he would without delegation.

When testing is hidden, there is no upside to delegation for the DM. When the advisor is more reluctant than the DM, then the advisor tests until additional outcomes cannot be pivotal to whether or not he prefers acceptance and the DM follows his recommendation (see Part 1 of Lemma 6). Therefore, the DM's payoff is the same as if she delegated decision-making authority to the advisor. When the advisor is more enthusiastic than the DM, then delegation reduces the advisor's incentives to test. Under both hidden testing and delegation, if the advisor prefers rejection irrespective of what additional test outcomes may show, he stops testing and the DM rejects. However, under hidden testing, even if the advisor prefers acceptance irrespective of what additional test outcomes may show, he may continue testing. This is because he needs additional positive outcomes to convince the DM to accept (see Part 2 of Lemma 6). Therefore, the DM weakly prefers hidden testing with communication over delegation.

When testing is observable, then delegating decision-making authority can increase incentives to test for a more reluctant advisor. Suppose the advisor would not have tested without decision-making authority. Then delegation provides an incentive to test by eliminating the possibility that outcomes lead the DM to act against the advisor's interest, just like hidden testing provides this incentive in the insurance region.[34] This clearly benefits the DM. However, if the advisor would have tested even without having decision-making authority, the incentive to test more works to the DM's disadvantage. This is because these additional tests are only pivotal for the advisor's preferred choice and not for the DM's, yet in equilibrium the DM's action depends on their realization.

So far I have assumed that the DM has no commitment power. In what follows, I will explore what would change if the DM could commit ex ante to which action she will take for any evidence presented to her.

Lemma 7 (DM Commitment) *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.*

If the advisor is more enthusiastic about accepting than the DM, the DM can achieve her first-best expected payoff by committing to reject if and only if the advisor reports fewer than N outcomes or the N outcomes lead her to prefer rejection. Similarly, if the advisor is more reluctant to accept than the DM, the DM can achieve her first-best expected payoff by committing to accept if and only if the advisor reveals fewer than N outcomes or the N outcomes lead her to prefer acceptance. To compare the DM's welfare across different settings, denote her expected payoff with commitment power by $\pi_{DM}(C)$ and denote her expected payoff without commitment by $\pi_{DM}(NC,i)$ where $i \in \{OT, HT\}$ indicates whether testing is observable (OT) or hidden (HT).

Corollary 2 (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 $(\lambda_{DM}, \lambda_A) \in B_1$ then in any Pareto-undominated equilibrium the DM's marginal benefit from allowing testing to be hidden is equally high as her marginal benefit from gaining commitment power:*

$$\pi_{DM}(NC, OT) < \pi_{DM}(NC, HT) = \pi_{DM}(C).$$

[34] 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.

2. *For any $N > 2$, if $(\lambda_{DM}, \lambda_A) \in B_2$ then in any Pareto-undominated equilibrium the DM's marginal benefit from allowing testing to be hidden is weakly smaller than her marginal benefit from gaining commitment power:*

$$\pi_{DM}(NC, OT) < \pi_{DM}(NC, HT) \leq \pi_{DM}(C).$$

This proposition shows that, in some circumstances, the DM can use hidden testing to circumvent a commitment problem. In particular, both hidden testing and commitment can be used to generate the skepticism effect if $(\lambda_{DM}, \lambda_A) \in B_1$. 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. In particular, she needs to commit to reject based on a single positive outcome. By contrast, under hidden testing, the threat to reject a single positive outcome 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 insurance effect, but it allows the DM to do even better. Suppose preferences lie in the insurance region B_2. Under observable testing, the DM can induce the advisor to test by committing to act in his interest for any realization of outcomes. Under hidden testing, such commitment is not necessary because the advisor has the possibility to omit outcomes and thereby prevent the DM from accepting if he would like to reject. However, the DM could do even better 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. With this commitment the DM can achieve her first-best expected payoff.

It is also insightful to see what would happen if the advisor had the power to commit ex ante to a testing and a disclosure strategy.

Proposition 8 (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 the advisor is more enthusiastic about accepting than the DM, then the DM is never strictly better off under hidden rather than observable testing.*

2. *For any $N > 2$, if the advisor is more reluctant to accept than the DM, then the DM is strictly better off under hidden testing than observable testing if preferences lie in the insurance region, i.e. $(\lambda_{DM}, \lambda_A) \in B_2$.*

Suppose the advisor is more enthusiastic about accepting than the DM. If $N = 2$ and the advisor has no commitment power, the DM benefits from hidden testing for preferences in the skepticism region B_I, 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. Put differently, when the advisor has commitment power he can do at least as well as under observable testing and therefore the skepticism effect ceases to exist. If the advisor is more reluctant to accept than the DM then he achieves the same expected payoff under hidden testing as when the decision-making authority was delegated to him. Therefore, he does not strictly benefit from commitment. In this situation, the DM's payoff comparison between hidden and observable testing is the same as when the advisor has no commitment power.

6 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, i.e. testing is simultaneous rather than sequential, and ii) the horizon is infinite and the advisor incurs a constant cost per test. I will again assume the test has accuracy p independent of the state.

In the preceding analysis, I assumed that the advisor acquires test outcomes sequentially. In some circumstances, the advisor may have to decide in advance 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 9 (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 does not exist an open set $B^{'} \subset \mathbb{R}^2$ such that the DM is strictly better off under hidden rather than observable testing for all $(\lambda_{DM}, \lambda_A) \in B^{'}$.*

2. *The insurance effect exists for all N, i.e. for any N, there exists some open set $\overline{B} \subset \mathbb{R}^2$ such that the DM is strictly better off under hidden rather than observable testing for all $(\lambda_{DM}, \lambda_A) \in \overline{B}$.*

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, he anticipates which action the DM will take if he were to stop, and that there are preference parameter configurations for which the advisor stops when the existing evidence leads the DM to accept (as shown in Lemma 5). 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, there are preference parameter configurations for which the advisor does not start testing (as shown in Lemma 5). He is concerned that even though he has the ability to act conditional on the realized outcomes up to this point, the chances that the test outcomes lead the two players to disagree are too high. When the advisor has to choose the number of tests in advance, he has even less scope to prevent the test outcomes from leading the DM to act against his interest, and hence, even less incentive to start testing.

Furthermore, the preceding analysis was built on the assumption that the horizon is finite and tests are costless. This finite horizon was used to represent the fact that the advisor is constrained in the total resources he can devote to testing. In some circumstances, the advisor may not face such constraints, i.e. the horizon may be infinite. Given an infinite horizon and costless tests, 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. To avoid this, I introduce a constant cost per test and use the solution concept of Perfect Bayesian Equilibrium.

Proposition 10 (Infinite Horizon) *Suppose $N \to \infty$ and the advisor incurs a constant 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 very similar to the one discussed in the preceding analysis with a finite horizon and costless 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. Therefore, on the equilibrium path, the DM accepts if and only if the evidence is just strong enough to lead him to accept. As in the preceding analysis, there exist parameter combinations for which the DM requires a larger number of positive outcomes to accept when testing is hidden than she requires excess positive outcomes when testing is observable. 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. This benefits the DM.[35]

However, an insurance effect does not exist in this alternative model. To understand why, the first step is to realize that the advisor's optimal testing strategy is stationary and depends only on the current posterior belief. Under observable testing, this follows from the fact the DM's optimal action choice depends only on the posterior belief. Under hidden testing, this follows because in equilibrium the more reluctant advisor's reporting strategy leads the DM to act in his interest at any posterior belief, as was the case in the preceding analysis. 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 prior belief and the belief at which he just prefers acceptance. Otherwise, the advisor will 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 than under hidden testing, 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, 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.

[35]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.

7 Conclusion

When taking decisions under uncertainty we often have to rely on others to carry out research. Since their preferences are not always aligned with ours, a natural question is whether or not we should monitor more closely how they go 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. One 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 information acquisition is private: 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. The skepticism effect shows that when information acquisition is private the DM can credibly raise her threshold for acceptance because she has reason to suspect that the advisor is hiding contradicting evidence. By contrast, the insurance effect shows that when information acquisition is private the advisor does not face the risk that additional evidence leads the DM to act against his interest. Therefore, he can freely explore whether or not additional evidence is sufficiently strong for both to agree on a different action choice. Due to the presence of these effects, there are combination of preference parameters for which the DM achieves her first-best expected payoff under private, but not under public information acquisition. While the insurance effect also causes the advisor's expected payoff to increase, the skepticism effect causes the advisor's expected payoff to decrease.

In addition, this paper has shown that there are combination of preferences for which the DM is strictly worse off when information acquisition is private. A sufficient condition for the DM to be weakly better off under private information acquisition is for preference to be sufficiently misaligned, i.e. for differences in the relative loss from inappropriate actions to be large between players.

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

A.1 Equilibrium Characterization under Observable Testing

Lemma 8 below will characterize the unique equilibrium under observable testing. Therefore, it contains the proof of Part 1 of Lemmas 2 - 4. It will be also used to prove statements about payoff comparisons.

To build a basis for payoff comparisons, Lemma 8 below will define the set of lists of Nature's draws which are mapped into acceptance on the equilibrium path. Note that for any fixed list of draws by Nature, the DM's action choice is deterministic.[36] Define the set of complete lists of Nature's draws by Φ, with representative element ϕ, as in (4). Then define a subset, denoted by

$$\Phi^* \subseteq \Phi = \{(+,+),(+,-),(-,+),(-,-)\},$$

such that if Nature's draws were fixed to be ϕ and $\phi \in \Phi^*$ then acceptance is chosen on the equilibrium path.[37] I will use Φ^*_{OT} to denote the set $\Phi^*$for the unique equilibrium under observable testing.

As shown by (3), the DM's optimal strategy $\overline{\sigma}_{DM}$ is to accept if and only if the likelihood ratio she assigns to history h_2 exceeds λ_{DM}. It is helpful to divide the range of $\lambda_{DM} \in \left(\frac{q}{1-q},\infty\right)$ into four regions such that the DM's optimal strategy is unchanged within each region, as indicated along the vertical axis in Figure 4.[38] I use the notation of posterior likelihood ratios $x_{\cup\Phi}$ as defined in equations (5)-(9) as well as

$$x_{(-,-)} = \frac{q\,(1-p_T)^2}{(1-q)\,p_F^2} \tag{13}$$

$$x_{(-,\cdot)} \equiv x_{(-,-)\cup(-,+)} = \frac{q\,(1-p_T)}{(1-q)\,p_F} \tag{14}$$

$$x_{\Phi\setminus(+,+)} = \frac{q\left(1-p_T^2\right)}{(1-q)\left(1-(1-p_F)^2\right)} \tag{15}$$

I denote the advisor's strategy by

$$\overline{\sigma}_A \equiv \left(\sigma_A(h_0),\sigma_A(h_1=(+)),\sigma_A(h_1=(-)),\sigma_A(h_1=(\emptyset))\right) \in [0,1]^4. \tag{16}$$

Lemma 8 (Equilibrium Characterization Observable Testing)

[36]Recall that I focus on pure-strategy equilibria.

[37]If the advisor runs a test in period 1 (2), he observes the first (second) element of $\phi \in \Phi^*$.

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

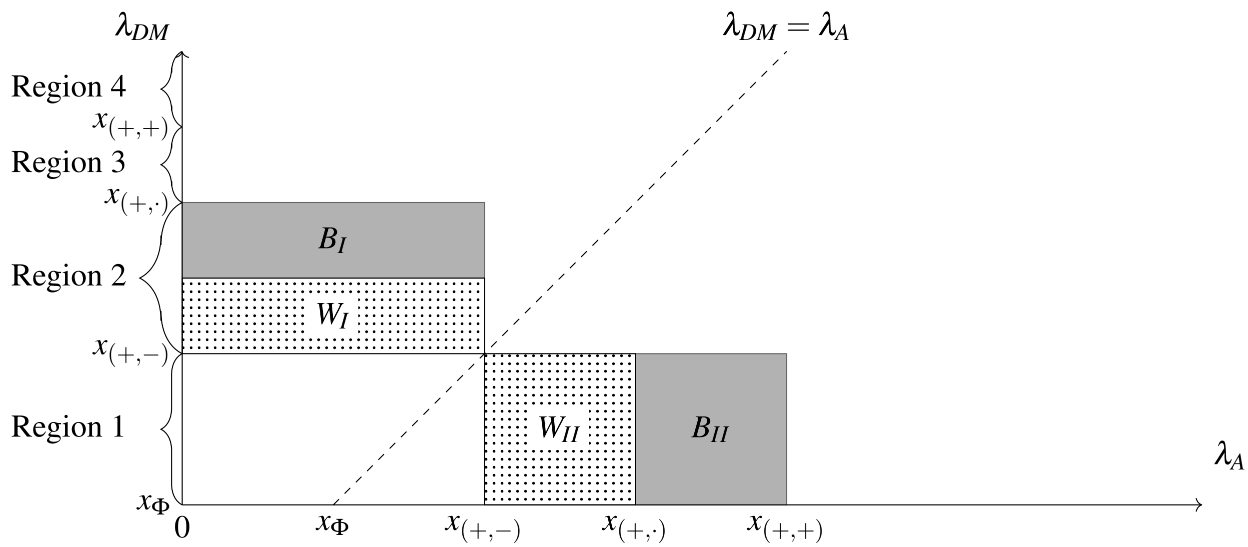

Figure 4: This Figure indicates along the vertical axis the regions in which the DM's optimal decision rule under observable testing is unchanged, in $(\lambda_A, \lambda_{DM})-$space given $p_T < p_F$.

*In **Region 1**, i.e. if*

$$x_\Phi < \lambda_{DM} \leq x_{(+,-)},$$

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

$$\overline{\sigma}_A = \begin{cases} (1,0,1,1) & \textit{if } \lambda_A < x_{(+,-)} \\ (1,0,0,1) & \textit{if } x_{(+,-)} \leq \lambda_A < x_{(+,\cdot)} \\ (0,0,0,0) & \textit{if } x_{(+,\cdot)} \leq \lambda_A. \end{cases}$$

The set of complete lists of outcome draws leading to acceptance is given by

$$\Phi^*_{OT} = \begin{cases} \{(-,+),(+,-),(+,+)\} & \textit{if } \lambda_A < x_{(+,-)} \\ \{(+,+),(+,-)\} & \textit{if } x_{(+,-)} \leq \lambda_A < x_{(+,\cdot)} \\ \emptyset & \textit{if } x_{(+,\cdot)} \leq \lambda_A. \end{cases}$$

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

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

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

$$\overline{\sigma}_A = \begin{cases} (1,0,0,1) & \text{if } \lambda_A \leq x_{(+,-)} \\ (1,1,0,1) & \text{if } x_{(+,-)} < \lambda_A < x_{(+,\cdot)} \\ (1,1,0,0) & \text{if } x_{(+,\cdot)} \leq \lambda_A < x_{(+,+)} \\ (0,1,0,0) & \text{if } x_{(+,+)} \leq \lambda_A. \end{cases}$$

The set of complete lists of outcome draws leading to acceptance is given by

$$\Phi^*_{OT} = \begin{cases} \{(+,-),(+,+)\} & \text{if } \lambda_A < x_{(+,-)} \\ \{(+,+)\} & \text{if } x_{(+,-)} \leq \lambda_A < x_{(+,+)} \\ \emptyset & \text{if } x_{(+,+)} \leq \lambda_A. \end{cases}$$

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

$$x_{(+,\cdot)} < \lambda_{DM} \leq x_{(+,+)},$$

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

$$\overline{\sigma}_A = \begin{cases} (1,1,0,1) & \text{if } \lambda_A < x_{(+,+)} \\ (0,0,0,0) & \text{if } x_{(+,+)} \leq \lambda_A. \end{cases}$$

The set of complete lists of outcome draws leading to acceptance is given by

$$\Phi^*_{OT} = \begin{cases} \{(+,+)\} & \text{if } \lambda_A < x_{(+,+)} \\ \emptyset & \text{if } x_{(+,+)} \leq \lambda_A. \end{cases}$$

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

$$\lambda_{DM} > x_{(+,+)},$$

then the DM never accepts. The advisor's strategy is $\overline{\sigma}_A = (0,0,0,0)$. *The set of complete lists of outcome draws leading to acceptance is given by* $\Phi^*_{OT} = \emptyset$.

In any region, posterior beliefs are formed according to Bayes' rule for any history of outcomes, whether on- or off-path. Hence, the system of beliefs satisfies

$$\frac{Pr\left(true|\widetilde{h}\right)}{Pr\left(false|\widetilde{h}\right)} = \begin{cases} x_{(+,\cdot)} & \text{if } \widetilde{h} = \{+\} \\ x_{(-,\cdot)} & \text{if } \widetilde{h} = \{-\} \\ x_{\Phi} & \text{if } \widetilde{h} = \{\emptyset\} \\ x_{\widetilde{h}} & \text{otherwise}. \end{cases} \tag{17}$$

Proof: For each region, I will use backward induction to show that the advisor's strategy maximizes his expected payoff given the DM's strategy and given the system of posterior beliefs. The strategy profile $(\overline{\sigma}_{DM}, \overline{\sigma}_A)$ and the system of posterior beliefs must constitute a sequential equilibrium for the following reason: 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 will converge to $\overline{\sigma}_A$ as $\varepsilon \to 0$. Since posterior beliefs are formed according to Bayes' rule conditional on any history of outcomes, the posterior beliefs derived based on the completely mixed strategy also satisfy (17).

Region 1. Suppose $p_F > p_T$ and $x_\Phi < \lambda_{DM} \leq x_{(+,-)}$. At the start of period 2, given $h_1 = (+)$, the DM accepts independent of whether or not another test is run. Hence, the advisor stops testing.[39] Given $h_1 = (-)$, if the advisor does not test the DM rejects. If the advisor tests then the DM accepts if and only if $h_2 = (-,+)$. The advisor tests if and only if

$$\begin{aligned} Pr(true|h_1 = (-)) &> \lambda_A Pr(false|h_1 = (-))(1-p_F) + Pr(true|h_1 = (-))(1-p_T) \\ \lambda_A &< \frac{Pr(true|h_1 = (-))\,p_T}{Pr(false|h_1 = (-))(1-p_F)} = \frac{qp_T(1-p_T)}{(1-q)(1-p_F)\,p_F} = x_{(+,-)}. \end{aligned} \tag{18}$$

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

$$\begin{aligned} Pr(true) &\leq \lambda_A Pr(false)(1-p_F) + Pr(true)(1-p_T) \\ \lambda_A &\geq \frac{Pr(true)\,p_T}{Pr(false)(1-p_F)} = \frac{qp_T}{(1-q)(1-p_F)} = x_{(+,\cdot)}. \end{aligned} \tag{19}$$

At the start of period 1, if the advisor does not test the DM rejects. For $\lambda_A < x_{(+,-)}$, if he tests then he will run another test 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 $x_{(+,-)} < x_{\Phi\backslash(-,-)}$ and

$$\begin{aligned} Pr(true) &> \lambda_A Pr(false)\left(1-p_F^2\right) + Pr(true)(1-p_T)^2 \\ \lambda_A &< \frac{q\left(1-(1-p_T)^2\right)}{(1-q)\left(1-p_F^2\right)} = x_{\Phi\backslash(-,-)}. \end{aligned} \tag{20}$$

For $\lambda_A \geq x_{(+,-)}$, if he tests then he will stop testing in period 2 and the DM will accept if and only if $h_2 = (+,\emptyset)$. The advisor tests if and only if (19) holds.

Region 2. At the start of 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

[39] 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.

the DM accepts. If the advisor tests then the DM accepts if and only if $h_2 = (+,+)$. The advisor stops testing if and only if

$$\lambda_A Pr(false|h_1 = (+)) \leq \lambda_A Pr(false|h_1 = (+))(1-p_F) + Pr(true|h_1 = (+))(1-p_T)$$

$$\lambda_A \leq \frac{Pr(true|h_1 = (+))(1-p_T)}{Pr(false|h_1 = (+))p_F} = \frac{qp_T(1-p_T)}{(1-q)(1-p_F)p_F} \equiv x_{(+,-)} \tag{21}$$

Given $h_1 = (\emptyset)$, if the advisor does not test the DM rejects. If the advisor tests then the DM accepts if and only if $h_2 = (\emptyset,+)$. The advisor tests if and only if equation (19) holds. At the start of period 1, if the advisor does not test the DM rejects. For $\lambda_A \leq x_{(+,-)}$, if the advisor tests then he will stop testing in period 2 and the DM will accept if and only if $h_2 = (+,\emptyset)$. The advisor tests since $\lambda_A \leq x_{(+,-)}$ implies that equation (19) holds. For $\lambda_A > x_{(+,-)}$, if the advisor tests then he will continue if and only if $h_1 = (+)$ and the DM will accept if and only if $h_2 = (+,+)$. The advisor tests if and only if

$$Pr(true) > \lambda_A Pr(false)(1-p_F)^2 + Pr(true)\left(1-p_T^2\right)$$

$$\lambda_A < \frac{qp_T^2}{(1-q)(1-p_F)^2} \equiv x_{(+,+)}. \tag{22}$$

Region 3. Suppose $x_{(+,\cdot)} < \lambda_{DM} \leq x_{(+,+)}$. At the start of 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 then the DM accepts if and only if $h_2 = (+,+)$. The advisor stops testing if and only if

$$Pr(true|h_1 = (+)) \leq \lambda_A Pr(false|h_1 = (+))(1-p_F) + Pr(true|h_1 = (+))(1-p_T)$$

$$\lambda_A \geq \frac{Pr(true|h_1 = (+))p_T}{Pr(false|h_1 = (+))(1-p_F)} = \frac{qp_T^2}{(1-q)(1-p_F)^2} \equiv x_{(+,+)}, \tag{23}$$

At the start of period 1, if the advisor does not test the DM rejects. For $\lambda_A < x_{(+,+)}$, since the advisor will test at $h_1 = (+)$ he tests in period 1. For $\lambda_A \geq x_{(+,+)}$, if he tests then he will stop testing in period 2 and the DM will reject and, therefore, the advisor does not test in period 1.

Region 4. Suppose $\lambda_{DM} > x_{(+,+)}$. Given that the DM will never accept for any h_2, the advisor never tests.

A.2 Equilibrium Characterization under Hidden Testing

Lemma 9 will characterize the advisor-preferred equilibrium under hidden testing and contain the proof of the Part 2 of Lemmas 2 - 4. Lemma 10 below will characterize an DM-preferred equilibrium under hidden testing. Both Lemmas will be used to prove statements about payoff comparisons.

I use the notation of posterior likelihood ratios $x_{\cup\Phi}$ as defined in equations (5)-(9) and (13)-(15). The advisor's testing is given in the format of equation (16). Just as in Section A.1, I will use Φ^* to denote the set of complete lists of Nature's draws which are mapped into acceptance along the equilibrium path. I will use $\underline{\Phi}^*_{HT}$ to denote the set Φ^* for any advisor-preferred equilibrium under hidden testing and $\overline{\Phi}^*_{HT}$ to denote the set Φ^* for any DM-preferred equilibrium under hidden testing.

The following two definitions will be used in the proofs of Lemmas 9 and 10. The *DM's first-best acceptance set,* denoted by Φ^{FB}_{DM}, is a set of complete lists of Nature's draws which contains all $\phi \in \Phi$ such that the DM prefers to accept if and only if she observed $h_2 = \phi$. If an equilibrium satisfies $\Phi^* = \Phi^{FB}_{DM}$ then the DM obtains her first-best expected payoff in equilibrium. Given that the DM optimally accepts if and only if the likelihood ratio conditional on ϕ exceeds λ_{DM}, as shown in (3), the DM's first-best acceptance set is given by

$$\Phi^{FB}_{DM} = \begin{cases} \{(+,+),(+,-),(-,+)\} & \text{if } x_\Phi < \lambda_{DM} \leq x_{(+,-)} \\ \{(+,+)\} & \text{if } max\left\{x_{(+,-)}, x_\Phi\right\} \leq \lambda_{DM} \leq x_{(+,+)} \\ \emptyset & \text{if } \lambda_{DM} \geq x_{(+,+)}. \end{cases} \tag{24}$$

Similarly, the *advisor's first-best acceptance set* is denoted by Φ^{FB}_A and given by

$$\Phi^{FB}_A = \begin{cases} \Phi & \text{if } \lambda_A \leq x_{(-,-)} \\ \{(+,+),(+,-),(-,+)\} & \text{if } x_{(-,-)} \leq \lambda_A \leq x_{(+,-)} \\ \{(+,+)\} & \text{if } x_{(+,-)} \leq \lambda_A \leq x_{(+,+)} \\ \emptyset & \text{if } \lambda_A \geq x_{(+,+)}. \end{cases} \tag{25}$$

A.2.1 Advisor-preferred Equilibria

Lemma 9 (Advisor-preferred Equilibrium Characterization Hidden Testing) *Suppose $\lambda_A < x_{(+,+)}$. In **Region 1**, i.e. $x_\Phi < \lambda_{DM} \leq x_{(+,-)}$.*

(a) If $\lambda_A < x_{(+,-)}$, the DM accepts if and only if $m = \{\{+\},\{+,-\},\{+,+\}\}$ and the advisor's strategy is

$$\sigma_A = (1,0,1,1)$$

$$\sigma_M = \begin{cases} \{+,+\} & \text{if } \widetilde{h} = \{+,+\} \\ \{+\} & \text{if } \widetilde{h} \in \{\{+,-\},\{+\}\} \\ \{\emptyset\} & \text{if } \widetilde{h} \in \{\{-,-\},\{-\},\{\emptyset\}\}. \end{cases} \tag{26}$$

and beliefs conditional on messages satisfy

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{\Phi\setminus(-,-)} & \text{if } m = \{+\} \\ x_{(-,-)} & \text{if } m = \{\emptyset\} \\ x_{(-,\cdot)} & \text{if } m = \{-\} \\ x_m & \text{otherwise.} \end{cases}$$

The set of complete lists of Nature's draws leading to acceptance is given by

$$\underline{\Phi}^*_{HT} = \{(-,+),(+,-),(+,+)\}.$$

(b) *If* $x_{(+,-)} \leq \lambda_A < x_{(+,\cdot)}$ *then* $(\lambda_A, \lambda_{DM}) \in W_{II}$ *and the equilibrium is as in Part 1 of Lemma 4. The set of complete lists of Nature's draws leading to acceptance is given by* $\underline{\Phi}^*_{HT} = \{(+,+)\}$.

(c) *If* $x_{(+,\cdot)} \leq \lambda_A < x_{(+,+)}$ *then* $(\lambda_A, \lambda_{DM}) \in B_{II}$ *and the equilibrium is as in Lemma 2. The set of complete lists of Nature's draws leading to acceptance is given by* $\underline{\Phi}^*_{HT} = \{(+,+)\}$.

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

(a) *If* $\max\{x_\Phi, x_{(+,-)}\} < \lambda_{DM} \leq x_{\Phi\setminus(-,-)}$ *and* $\lambda_A < x_{(+,-)}$ *then* $(\lambda_A, \lambda_{DM}) \in W_I$ *and the equilibrium is as in Part 2 of Lemma 4. The set of complete lists of Nature's draws leading to acceptance is given by* $\underline{\Phi}^*_{HT} = \{(-,+),(+,-),(+,+)\}$.

(b) *If* $x_{\Phi\setminus(-,-)} < \lambda_{DM} \leq x_{(+,\cdot)}$ *and* $\lambda_A < x_{(+,-)}$ *then* $(\lambda_A, \lambda_{DM}) \in B_I$ *and the equilibrium is as in Lemma 3. The set of complete lists of Nature's draws leading to acceptance is given by* $\underline{\Phi}^*_{HT} = \{(+,+)\}$.

(c) *If* $x_{(+,-)} \leq \lambda_A < x_{(+,\cdot)}$ *the DM accepts if and only if* $m \in \{\{+\},\{+,+\}\}$. *The advisor's strategy is*

$$\sigma_A = (1,1,0,1)$$

and his disclosure strategy is

$$\sigma_M = \begin{cases} \{\emptyset\} & \text{if } \widetilde{h} = \{\emptyset\} \\ \{-\} & \text{if } \widetilde{h} \in \{\{-\},\{+,-\}\} \\ \widetilde{h} & \text{otherwise.} \end{cases}$$

and beliefs conditional on messages satisfy

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{\Phi\setminus(+,+)} & \text{if } m = \{-\} \\ x_{\Phi} & \text{if } m = \{\emptyset\} \\ x_{(+,\cdot)} & \text{if } m = \{+\} \\ x_m & \text{otherwise.} \end{cases} \tag{27}$$

The set of complete lists of Nature's draws leading to acceptance is given by $\underline{\Phi}^*_{HT} = \{(+,+)\}$.

(d) *If* $x_{(+,\cdot)} \leq \lambda_A < x_{(+,+)}$ *the DM accepts if and only if* $m \in \{\{+\},\{+,+\}\}$. *The advisor's strategy is*

$$\sigma_A = (1,1,0,0)$$

$$\sigma_M = \begin{cases} \{\emptyset\} & \text{if } \widetilde{h} \in \{\{\emptyset\},\{+\}\} \\ \{-\} & \text{if } \widetilde{h} \in \{\{-\},\{+,-\}\} \\ \widetilde{h} & \text{otherwise.} \end{cases}$$

and beliefs conditional on messages satisfy (27). The set of complete lists of Nature's draws leading to acceptance is given by $\underline{\Phi}^*_{HT} = \{(+,+)\}$.

*In **Region 3**, i.e. suppose* $x_{(+,\cdot)} < \lambda_{DM} \leq x_{(+,+)}$. *The DM accepts if and only if* $m = \{+,+\}$. *If* $\lambda_A < x_{(+,+)}$ *the advisor's strategy is*

$$\sigma_A = (1,1,0,0)$$

and his disclosure strategy is $m = \widetilde{h}$ *for all* h_2 *and beliefs conditional on messages satisfy*

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{(+,\cdot)} & \text{if } m = \{+\} \\ x_{\Phi} & \text{if } m = \{\emptyset\} \\ x_{(-,\cdot)} & \text{if } m = \{-\} \\ x_m & \text{otherwise.} \end{cases} \tag{28}$$

The set of complete lists of Nature's draws leading to acceptance is given by $\underline{\Phi}^*_{HT} = \{(+,+)\}$.

*In **Region 4**, i.e. suppose* $\lambda_{DM} > x_{(+,+)}$. *Then the DM never accepts. The advisor's testing strategy is*

$$\sigma_A = (0,0,0,0)$$

and his disclosure strategy is $m = \widetilde{h}$ *for all* h_2 *and beliefs conditional on messages satisfy (28). The* *set of complete lists of Nature's draws leading to acceptance is given by* $\underline{\Phi}^*_{HT} = \emptyset$*.*

In all regions, if $\lambda_A < x_{(+,+)}$*, then for DM's strategy is to accept given* $m \in \{\{+,+\},\{+,-\}\}$ *if* $\lambda_{DM} < x_m$ *and ii)* $m = \{+\}$ *if* $\lambda_{DM} < x_{(+,\cdot)}$ *and to reject otherwise. The advisor's testing strategy is*

$$\sigma_A = (0,0,0,0)$$

and his disclosure strategy is $m = \emptyset$ *for all* h_2 *and beliefs conditional on messages satisfy (28). The set of complete lists of Nature's draws leading to acceptance is given by* $\underline{\Phi}^*_{HT} = \emptyset$*.*

In all regions, belief are formed according to Bayes' rule based on any observable history of outcomes and hence satisfy equation (17).

Proof: First, I will show that each player's strategy is optimal given the other player's strategy and the system of beliefs. Given that the DM optimally accepts if and only if the likelihood ratio conditional on m exceeds λ_{DM}, as shown by (3), it is straightforward to verify that the DM's strategy is sequentially rational given beliefs for all regions.
Next, I will show that the equilibrium is a sequential equilibrium. Consider a completely mixed strategy for the advisor as a function of some $\varepsilon \in (0,1)$, denoted by $\left(\sigma_A^\varepsilon, \sigma_M^\varepsilon\right)$, which is constructed as follows: for any history of outcomes h:

$$\sigma_A^\varepsilon(h) = \begin{cases} 1-\varepsilon & \text{if } \sigma_A(h) = 1 \\ \varepsilon & \text{if } \sigma_A(h) = 0 \end{cases}$$

and for any h_2 and associated message space M, the message specified by σ_M is sent with probability $1-(|M|-1)\varepsilon^3$ and any other feasible message $m \in M$ is sent with probability ε^3, where $|M|$ denotes the total number of feasible messages given h_2. Clearly, as $\varepsilon \to 0$ $\left(\sigma_A^\varepsilon, \sigma_M^\varepsilon\right) \to (\sigma_A, \sigma_M)$. In addition, I will show that the system of beliefs derived from the advisor's mixed strategy profile using Bayes' rule, denoted by $Pr^\varepsilon(\omega|m)$, converges to the system of beliefs proposed in Lemma 9, i.e.

$$\frac{Pr^\varepsilon(true|m)}{Pr^\varepsilon(false|m)} \to \frac{Pr(true|m)}{Pr(false|m)}. \tag{29}$$

Note that this must trivially be the case for any message $m \in \{\{+,+\},\{+,-\},\{-,-\}\}$ due to verifiable disclose.
Region 1a). σ_M is optimal: following $\widetilde{h} \in \{\{-,-\},\{-\},\{\emptyset\}\}$ there is no message available for which the DM would accept, and otherwise, the advisor prefers to accept since $\lambda_A < x_{(+,-)}$ and the messages assigned lead the DM to accept. σ_A is optimal: In period 2, following $h_1 = (+)$, the advisor stops testing because he prefers to accept irrespective of the second outcome and if he

stops the DM will accept. Otherwise, the advisor tests in period 2, because if he stopped the DM would reject, but if and only if the test in period 2 is positive the DM accepts and (18) holds and (19) does not hold. By (20), it follows that he tests in period 1.

Consider mixed strategies $\left(\sigma_A^\varepsilon, \sigma_M^\varepsilon\right)$ constructed as outlined above. As $\varepsilon \to 0$, (29) holds: If $\varepsilon = 0$ then $m = \{\emptyset\}$ occurs on path following $h_2 = (-,-)$ and $m = \{+\}$ occurs on path following $h_2 \in \{(-,+),(+,\emptyset)\}$. In the limit as $\varepsilon \to 0$, $m = \{-\}$ is mostly likely observed following $h_2 \in \{(-,-),(-,+)\}$ since each of these histories occurs with a probability of order ε^3 (due to a reporting mistake), whereas all other histories occur with a probability of at least order ε^4 because they require both a reporting and a testing mistake.

The equilibrium is an advisor-preferred equilibrium. If $x_{(-,-)} < \lambda_A < x_{(+,-)}$ then $\underline{\Phi}^*_{HT} = \Phi^{FB}_A$ and if $\lambda_A < x_{(-,-)}$, then $\underline{\Phi}^*_{HT} \cup (-,-) = \Phi^{FB}_A = \Phi$, but it cannot be that the DM accepts irrespective of Nature's draws because the DM prefer to reject at the prior belief.

Region 1b). σ_M is optimal: Following $\widetilde{h} \in \{\{+,+\},\{+\}\}$ the advisor strictly prefers to accept since $\lambda_A < x_{(+,\cdot)}$ and the messages assigned lead the DM to accept, and otherwise, the advisor weakly prefers to reject since $\lambda_A \geq x_{(+,-)}$ and the messages assigned lead the DM to reject. σ_A is optimal: In period 2, following $h_1 = (-)$, the advisor stops testing because he prefers to reject irrespective of the second outcome and if he stops the DM will reject. Otherwise, the advisor tests in period 2. Given $h_1 = (+)$, if he stopped the DM would accept but if he continues the DM accepts if and only if $h_2 = (+,+)$ and (21) does not hold. Given $h_1 = (\emptyset)$, if he stopped the DM would reject but if he continues the DM accepts if and only if $h_2 = (\emptyset,+)$ and (19) does not hold. By (22), it follows that he tests in period 1.

Consider mixed strategies $\left(\sigma_A^\varepsilon, \sigma_M^\varepsilon\right)$ constructed as outlined above. As $\varepsilon \to 0$, (29) holds: If $\varepsilon = 0$ then $m = \{-\}$ occurs on path following $h_2 \in \{(-,\emptyset),(+,-)\}$. In the limit as $\varepsilon \to 0$, $m = \{+\}$ is most likely observed following $h_2 \in \{(\emptyset,+),(+,\emptyset)\}$ since each of these histories occurs with a probability of order ε (due to a single testing mistake), whereas all other histories occur with a probability of at least order ε^3 because a single testing mistake is insufficient for $m = \{+\}$ to be sent. In the limit as $\varepsilon \to 0$, $m = \{\emptyset\}$ is mostly likely observed following $h_2 = (\emptyset,\emptyset)$, since this histories occurs with a probability of order ε^2 (two testing mistakes) whereas all other histories occur with a probability of at least order ε^3 because they require at least a reporting mistake.

The equilibrium is an advisor-preferred equilibrium. Given $\Phi^* = \{(-,+),(+,-),(+,+)\}$ and $\lambda_A \geq x_{(+,-)}$, for each complete list of outcome draws $\phi \in \Phi$ the DM acts in the advisor's interest.

Region 1c). σ_M is optimal: Following $\widetilde{h} = \{+,+\}$ the advisor strictly prefers to accept since $\lambda_A < x_{(+,+)}$ and the message assigned leads the DM to accept, otherwise, the advisor weakly prefers to reject since $\lambda_A \geq x_{(+,\cdot)}$ and the messages assigned lead the DM to reject. σ_A is optimal: In period 2, following $h_1 = (+)$, the advisor tests. If he stopped the DM would accept but if he continues the DM accepts if and only if $h_2 =$ (+,+) and (21) does not hold. Otherwise, the advisor stops in period 2 because he weakly prefers to reject irrespective of the second outcome and if he stops the DM rejects. By (22), it follows that he tests in period 1.

Consider mixed strategies $\left(\sigma_A^{\varepsilon},\sigma_M^{\varepsilon}\right)$ constructed as outlined above. As $\varepsilon\to 0$, (29) holds: If $\varepsilon=0$ then $m=\{-\}$ occurs on path following $h_2\in\{(-,\emptyset),(+,-)\}$. In the limit as $\varepsilon\to 0$, $m=\{+\}$ and $m=\{\emptyset\}$ each are most likely observed following $h_2\in\{(\emptyset,+),(+,\emptyset)\}$ since each of these histories occurs with a probability of order ε (due to a single testing mistake), whereas all other histories occur with a probability of at least order ε^3 because a single testing mistake is insufficient for $m\in\{\{+\},\{\emptyset\}\}$ to be sent.

The equilibrium is advisor-preferred since $\underline{\Phi}_{HT}^{*}=\Phi_A^{FB}$.

Region 2a). σ_M is optimal: following $\widetilde{h}\in\{\{-,-\},\{-\},\{\emptyset\}\}$ there is no message available for which the DM would accept, and otherwise, the advisor prefers to accept since $\lambda_A<x_{(+,-)}$ and the messages assigned lead the DM to accept. σ_A is optimal: In period 2, following $h_1=(+)$, the advisor stops because he prefers to accept irrespective of a second outcome and if he stops the DM accepts. Otherwise, the advisor tests. Given $h_1=(-)$, if he stopped the DM would reject, but if he continues the DM accepts if and only if $h_2=(-,+)$ and (18) holds. Given $h_1=(\emptyset)$, if he stopped the DM would reject, but if he continues the DM accepts if and only if $h_2=(\emptyset,+)$ and (19) does not hold. By (20), it follows that he tests in period 1.

Consider mixed strategies $\left(\sigma_A^{\varepsilon},\sigma_M^{\varepsilon}\right)$ constructed as outlined above. As $\varepsilon\to 0$, (29) holds: if $\varepsilon=0$ then $m=\{+\}$ occurs on path following $h_2\in\{(+,\emptyset),(-,+)\}$. In the limit as $\varepsilon\to 0$, $m=\{-\}$ is most likely observed following $h_2\in\{(\emptyset,-),(-,\emptyset)\}$ since each of these histories occurs with a probability of order ε (due to a single testing mistake), whereas all other histories occur with a probability of at least order ε^3 because a single testing mistake is insufficient for $m=\{-\}$ to be sent. In the limit as $\varepsilon\to 0$, $m=\{\emptyset\}$ is most likely observed following $h_2=(\emptyset,\emptyset)$ since this history occurs with a probability of order ε^2 (two testing mistakes) whereas all other histories occur with a probability of at least order ε^3 because they require at least a reporting mistake.

The equilibrium is advisor-preferred by the reasoning given for Region 1a).

Region 2b). σ_M is optimal: following $\widetilde{h}\neq\{+,+\}$ there is no message available for which the DM would accept, and otherwise, the advisor prefers to accept since $\lambda_A<x_{(+,-)}<x_{(+,+)}$ and the message assigned leads the DM to accept. σ_A is optimal: In period 2, following $h_1=(+)$, the advisor tests. If he stopped the DM would reject, but if he continues the DM accepts if and only if $h_2=(+,+)$ and (23) does not hold. Otherwise, the advisor stops testing because the DM rejects irrespective of the second outcome. By (22), it follows that he tests in period 1.

Consider mixed strategies $\left(\sigma_A^{\varepsilon},\sigma_M^{\varepsilon}\right)$ constructed as outlined above. As $\varepsilon\to 0$, (29) holds: if $\varepsilon=0$ $m=\{+\}$ occurs following $h_2=(+,-)$ and $m=\{\emptyset\}$ occurs following $h_2=(-,\emptyset)$. In the limit as $\varepsilon\to 0$, $m=\{-\}$ is most likely observed following $h_2\in\{(-,\emptyset),(+,-)\}$ since this history occurs with a probability of order ε^3 (reporting mistake) whereas all other histories occur with a probability of at least order $\varepsilon^3(1-\varepsilon)$ because they require at least a reporting and a testing mistake.

The equilibrium is an advisor-preferred equilibrium. Given $\Phi_A^{FB} = \underline{\Phi}_{HT}^* \cup (+,-) \cup (-,+)$. An equilibrium in which the advisor is strictly better off must also map at least $\phi = (+,-)$ or $\phi = (-,+)$ into acceptance. But if a message that leads the DM to accept is available given $h_2 = (+,-)$ then it would also be available given $h_2 = (-,+)$ and it cannot be optimal for the DM to accept if $\phi \in \{(+,-),(-,+),(+,+)\}$ since $x_{\Phi\setminus\widetilde{(-,-)}} < \lambda_{DM}$.

Region 2c). σ_M is optimal: Following $\widetilde{h} \in \{\{+,+\},\{+\}\}$ the advisor strictly prefers to accept since $\lambda_A < x_{(+,\cdot)}$ and the messages assigned lead the DM to accept, and otherwise, the advisor weakly prefers to reject since $\lambda_A \geq x_{(+,-)}$ and the messages assigned lead the DM to reject. σ_A is optimal: In period 2, following $h_1 = (-)$, the advisor stops testing because he weakly prefers to reject irrespective of a second outcome and if he stops the DM rejects. Otherwise, he continues testing. Given $h_1 = (+)$, if he stopped the DM would accept but if he continues the DM accepts if and only if $h_2 = (+,+)$ and (21) does not hold. Given $h_1 = (\emptyset)$, if he stopped the DM would reject but if he continues the DM accepts if and only if $h_2 = (\emptyset,+)$ and (19) does not hold. By (22), it follows that he tests in period 1.

Consider mixed strategies $\left(\sigma_A^\varepsilon, \sigma_M^\varepsilon\right)$ constructed as outlined above. As $\varepsilon \to 0$, (29) holds: if $\varepsilon = 0$ $m = \{-\}$ occurs on path following $h_2 \in \{(-,\emptyset),(+,-)\}$. In the limit as $\varepsilon \to 0$, $m = \{+\}$ is most likely observed following $h_2 \in \{(+,\emptyset),(\emptyset,+)\}$ since each of these histories occurs with a probability of order ε (due to a single testing mistake), whereas all other histories occur with a probability of at least order ε^3 because a single testing mistake is insufficient for $m = \{+\}$ to be sent. In the limit as $\varepsilon \to 0$, $m = \{\emptyset\}$ is most likely observed following $h_2 = (\emptyset,\emptyset)$ since this history occurs with a probability of order ε^2 (two testing mistakes) whereas all other histories occur with a probability of at least order ε^3 because they require at least a reporting mistake.

The equilibrium is an advisor-preferred equilibrium since $\underline{\Phi}_{HT}^* = \Phi_A^{FB}$.

Region 2d). σ_M is optimal: Following $\widetilde{h} = \{+,+\}$ the advisor strictly prefers to accept since $\lambda_A < x_{(+,+)}$ and the message assigned leads the DM to accept, and otherwise, the advisor weakly prefers to reject since $\lambda_A \geq x_{(+,\cdot)}$ and the messages assigned lead the DM to reject. σ_A is optimal: In period 2, following $h_1 = (+)$, the advisor tests. If he stopped the DM would accept but if he continues the DM accepts if and only if $h_2 = (+,+)$ and (21) does not hold. Otherwise, the advisor stops because he weakly prefers to reject irrespective of the outcome and the DM rejects if he stops. By (22), it follows that he tests in period 1.

Consider mixed strategies $\left(\sigma_A^\varepsilon, \sigma_M^\varepsilon\right)$ constructed as outlined above. As $\varepsilon \to 0$, (29) holds: if $\varepsilon = 0$ $m = \{-\}$ occurs on path if $\varepsilon = 0$ following $h_2 \in \{(-,\emptyset),(+,-)\}$. In the limit as $\varepsilon \to 0$, $m = \{+\}$ is most likely observed following $h_2 \in \{(+,+),(+,-)\}$ since each of these histories occurs with a probability of order ε^3 (due to a reporting mistake), whereas all other histories occur with a probability of at least order ε^4 because they involve both a reporting and a testing mistake. In the limit as $\varepsilon \to 0$, $m = \{\emptyset\}$ is most likely observed following $h_2 = (\emptyset,\emptyset)$ since this history occurs with a probability of order ε (a single testing mistake) whereas all other histories occur with a

probability of at least order ε^2 because they require at least two testing mistakes.
The equilibrium is advisor-preferred since $\underline{\Phi}^*_{HT} = \Phi^{FB}_A$.
Region 3. σ_M is optimal: following $\widetilde{h} \neq \{+,+\}$ there is no message available for which the DM would accept, and otherwise, the advisor prefers to accept since $\lambda_A < x_{(+,+)}$ and the message assigned leads the DM to accept. σ_A is optimal: In period 2, following $h_1 = (+)$, the advisor tests. If he stopped the DM would accept but if he continues the DM accepts if and only if $h_2 = $(+,+) and (21) does not hold. Otherwise, the advisor does not test because the DM will reject irrrespective of the outcome. It follows that he tests in period 1. By (22), it follows that he tests in period 1.
Consider mixed strategies $\left(\sigma^\varepsilon_A, \sigma^\varepsilon_M\right)$ constructed as outlined above. As $\varepsilon \to 0$, (29) holds: if $\varepsilon = 0$ $m = \{-\}$ occurs on path following $h_2 = \{(-,\emptyset)\}$. In the limit as $\varepsilon \to 0$, $m = \{+\}$ is most likely observed following $h_2 = (+,\emptyset)$ since this history occurs with a probability of order ε (due to a single testing mistake), whereas all other histories occur with a probability of at least order ε^2 because they involve two testing mistakes or a reporting mistake. In the limit as $\varepsilon \to 0$, $m = \{\emptyset\}$ is most likely observed following $h_2 = (\emptyset,\emptyset)$ since this history occurs with a probability of order ε (due to a single testing mistake), whereas all other histories occur with a probability of at least order ε^2 because they involve two testing mistakes or a reporting mistake.
The equilibrium is an advisor-preferred equilibrium. If $x_{(+,-)} \leq \lambda_A < x_{(+,+)}$, then $\underline{\Phi}^*_{HT} = \Phi^{FB}_A$. If $\lambda_A < x_{(+,-)}$ then $\Phi^{FB}_A = \underline{\Phi}^*_{HT} \cup (+,-) \cup (-,+)$. But there cannot be an equilibrium in which the DM accepts if $\phi \in \{(+,-),(-,+),(+,+)\}$ since $x_{\Phi\backslash(-,-)} < \lambda_{DM}$.
Region 4. σ_M is optimal since there is no feasible message for any $\widetilde{h}$ which leads the DM to accept. The advisor never tests because the DM rejects irrespective of the outcome.
Consider mixed strategies $\left(\sigma^\varepsilon_A, \sigma^\varepsilon_M\right)$ constructed as outlined above. As $\varepsilon \to 0$, (29) holds: if $\varepsilon = 0$ $m = \{\emptyset\}$ occurs on path following $h_2 = (\emptyset,\emptyset)$. In the limit as $\varepsilon \to 0$, $m = \{-\}$ ($m = \{+\}$) is most likely observed following $h_2 = (-,\emptyset)$ ($h_2 = (+,\emptyset)$) since this history occurs with a probability of order ε (due to a single testing mistake), whereas all other histories occur with a probability of at least order ε^2 because they involve two testing mistakes or a reporting mistake.
The equilibrium is an advisor-preferred equilibrium. For any $\lambda_A < x_{(+,+)}$, $(+,+) \in \Phi^{FB}_A \supset \underline{\Phi}^*_{HT} = \emptyset$. An equilibrium in which the advisor is strictly better off must map at least $\phi = (+,+)$ into acceptance. But it is never optimal for the DM to accept if $\phi = (+,+)$ since $\lambda_{DM} > x_{(+,+)}$.
Finally, if $\lambda_A < x_{(+,+)}$ then σ_M is optimal since for any h_2 the advisor prefers rejection and the DM rejects. The advisor never tests as the DM would reject irrespective of the test. This is a sequential equilibrium by the same reasoning as in Region 4. The equilibrium is an advisor-preferred equilibrium since $\underline{\Phi}^*_{HT} = \Phi^{FB}_A$.

A.2.2 DM-preferred Equilibria

Lemma 10 (DM-preferred Equilibrium Characterization Hidden Testing) *Take Region 1, i.e.* $x_\Phi < \lambda_{DM} \leq x_{(+,-)}$.

1. *If* $x_{(+,-)} \leq \lambda_A < x_{(+,\cdot)}$ *then the DM rejects if and only if* $m = \{-,-\}$. *The advisor's testing strategy is*

$$\sigma_A = (1,0,1,0)$$

and his disclosure strategy is

$$\sigma_M = \begin{cases} \{-\} & \text{if } \widetilde{h} \in \{\{-\},\{+,-\}\} \\ \widetilde{h} & \text{otherwise.} \end{cases}$$

and beliefs conditional on messages satisfy

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{(+,-)} & \text{if } m \in \{\{-\},\{\emptyset\}\} \\ x_{(+,\cdot)} & \text{if } m = \{+\} \\ x_m & \text{otherwise.} \end{cases} \tag{30}$$

2. *If* $x_{(+,\cdot)} \leq \lambda_A < x_{(+,+)}$ *then DM rejects if and only if* $m = \{-,-\}$. *The advisor's testing strategy is*

$$\sigma_A = (1,0,1,0)$$

and his disclosure strategy is

$$\sigma_M = \begin{cases} \{\emptyset\} & \text{if } h \in \{\{\emptyset\},\{+\}\} \\ \{-\} & \text{if } h \in \{\{-\},\{+,-\}\} \\ \widetilde{h} & \text{otherwise,} \end{cases}$$

and beliefs conditional on messages satisfy

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{(+,-)} & \text{if } m = \{-\} \\ x_{(+,\cdot)} & \text{if } m = \{\emptyset\} \\ x_{\Phi\setminus(-,-)} & \text{if } m = \{+\} \\ x_m & \text{otherwise.} \end{cases} \tag{31}$$

3. *If* $\lambda_A \geq x_{(+,+)}$ *then DM rejects if and only if* $m = \{-,-\}$. *The advisor's testing strategy is*

$$\sigma_A = (1,0,1,0)$$

and his disclosure strategy is

$$\sigma_M = \begin{cases} \{\emptyset\} & \text{if } h \in \{\{\emptyset\}, \{+\}, \{+,+\}\} \\ \{-\} & \text{if } h \in \{\{-\}, \{+,-\}\} \\ \widetilde{h} & \text{otherwise,} \end{cases}$$

and beliefs conditional on messages satisfy (31).

4. *Take Region 2. If* $\max\{x_\Phi, x_{(+,-)}\} < \lambda_{DM} \le x_{\Phi\setminus(-,-)}$ *and* $\lambda_A < x_{(+,-)}$ *then the DM accepts if and only if* $m \in \{+,+\}$*; the advisor's testing strategy is*

$$\sigma_A = (1,1,0,0)$$

and his disclosure strategy is

$$\sigma_M = \begin{cases} \{+\} & \text{if } \widetilde{h} \in \{\{+\}, \{+,-\}\} \\ \widetilde{h} & \text{otherwise.} \end{cases} \tag{32}$$

The DM's beliefs conditional on messages satisfy

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{(+,-)} & \text{if } m = \{+\} \\ x_{(-,\cdot)} & \text{if } m = \{-\} \\ x_\Phi & \text{if } m = \{\emptyset\} \\ x_m & \text{otherwise.} \end{cases}$$

5. *Take Region 2 and 3, i.e.* $\max\{x_\Phi, x_{(+,-)}\} < \lambda_{DM} \le x_{(+,+)}$ *and* $\lambda_A \ge x_{(+,+)}$ *then the DM rejects if and only if* $m \in \{\{-\}, \{-,-\}\}$*; the advisor's testing strategy is*

$$\sigma_A = (1,1,0,1)$$

and his disclosure strategy is

$$\sigma_M = \begin{cases} \{-\} & \text{if } \widetilde{h} \in \{\{-\}, \{+,-\}, \{-,-\}\} \\ \emptyset & \text{otherwise.} \end{cases} \tag{33}$$

The DM's beliefs conditional on messages satisfy

$$\frac{Pr(true|m)}{Pr(false|m)} = \begin{cases} x_{(+,\cdot)} & \textit{if } m = \{+\} \\ x_{\Phi\setminus(+,+)} & \textit{if } m = \{-\} \\ x_{(+,+)} & \textit{if } m = \{\emptyset\} \\ x_m & \textit{otherwise.} \end{cases}$$

6. *Otherwise, the DM-preferred and the advisor-preferred equilibrium coincide and the equilibrium is characterized by Lemma 9.*

7. *In any DM-preferred equilibrium, the DM achieves her first-best payoff, i.e.* $\overline{\Phi}^*_{HT} = \Phi^{FB}_{DM}$.

Proof: For Parts 1-5, it is straightforward to verify that the DM's strategy is sequentially rational given beliefs. The DM optimally accepts if and only if the likelihood ratio conditional on m exceeds λ_{DM}, as shown in equation (3).

Part 1: σ_M is optimal: since for $\widetilde{h} \neq \{-,-\}$ there is no message available for which the DM would reject, and otherwise, the advisor prefers to reject since $x_{(-,-)} < \lambda_A$ and the message assigned by σ_M leads the DM to reject. σ_A is optimal: In period 2, following $h_1 \in \{(+),(\emptyset)\}$, the advisor stops testing because irrespective of the second outcome the DM will accept. Following $h_1 = (-)$, the advisor tests. If he stopped, the DM would accepts. If he tests the DM will reject if and only if $h_2 = (-,-)$. Testing is optimal since

$$\lambda_A Pr(false|h_1 = (-)) > \lambda_A Pr(false|h_1 = (-))(1-p_F) + Pr(true|h_1 = (-))(1-p_T)$$

$$\lambda_A > \frac{q(1-p_T)^2}{(1-q)p_F^2} = x_{(-,-)}.$$

This also implies that he tests in period 1.

Consider a completely mixed strategy for the advisor as a function of some $\varepsilon \in (0,1)$, denoted by $(\sigma_A^\varepsilon, \sigma_M^\varepsilon)$, which is constructed as follows: for any history of outcomes h:

$$\sigma_A^\varepsilon(h) = \begin{cases} 1-\varepsilon^3 & \text{if } \sigma_A(h) = 1 \\ \varepsilon^3 & \text{if } \sigma_A(h) = 0 \end{cases}$$

and if $\widetilde{h} = \{+,-\}$ then $m = \emptyset$ is sent with probability ε, $m \in \{\{+\},\{+,-\}\}$ is sent with probability ε^3 and $m = \{-\}$ is sent with probability $1-\varepsilon-2\varepsilon^3$. Otherwise, for any h_2 and associated message space M, the message specified by σ_M is sent with probability $1-(|M|-1)\varepsilon^3$ and any other feasible message $m \in M$ is sent with probability ε^3, where $|M|$ denotes the total number of feasible messages given h_2. As $\varepsilon \to 0$, clearly $\left(\sigma_A^\varepsilon, \sigma_M^\varepsilon\right) \to (\sigma_A, \sigma_M)$. Furthermore, for any m, as $\varepsilon \to 0$,

$$\frac{Pr^{\varepsilon}(true|m)}{Pr^{\varepsilon}(false|m)} \to \frac{Pr(true|m)}{Pr(false|m)}, \tag{34}$$

since if $\varepsilon = 0$ $m = \{-\}$ occurs on path following $h_2 = (-,+)$, and $m = \{+\}$ occurs on path following $h_2 = (+,\emptyset)$, and, in the limit as $\varepsilon \to 0$, $m = \{\emptyset\}$ is mostly likely observed following $h_2 = (-,+)$ since this history occurs with a probability of order ε (due to a reporting mistake) whereas all other histories occur with probability of at least order ε^3.

The set of complete lists of Nature's draws leading to acceptance is given by

$$\overline{\Phi}^*_{HT} = \{(+,+),(+,-),(-,+)\}.$$

Part2. The advisor's strategy is optimal by the same reasoning as for Part 1.

Consider a completely mixed strategy for the advisor as a function of some $\varepsilon \in (0,1)$, denoted by $\left(\sigma_A^\varepsilon, \sigma_M^\varepsilon\right)$, which is constructed as follows: for any history of outcomes h:

$$\sigma_A^\varepsilon(h) = \begin{cases} 1-\varepsilon & \text{if } \sigma_A(h) = 1 \\ \varepsilon & \text{if } \sigma_A(h) = 0 \end{cases}$$

and for any h_2 and associated message space M, the message specified by σ_M is sent with probability $1-(|M|-1)\varepsilon^3$ and any other feasible message $m \in M$ is sent with probability ε^3, where $|M|$ denotes the total number of feasible messages given h_2. As $\varepsilon \to 0$, clearly $\left(\sigma_A^\varepsilon, \sigma_M^\varepsilon\right) \to (\sigma_A, \sigma_M)$. Furthermore, for any m, as $\varepsilon \to 0$ (34) holds for the following reasons: if $\varepsilon = 0$ $m = \{-\}$ occurs on path following $h_2 = (-,+)$, and $m = \{\emptyset\}$ occurs on path following $h_2 = (+,\emptyset)$. In the limit as $\varepsilon \to 0$, $m = \{+\}$ is mostly likely observed following $h_2 \in \{(+),(-,+)\}$ since each of these histories occur with a probability of order ε^3 (due to a reporting mistake) whereas all other histories occur with probability of at least order ε^4 because they require both a reporting and a testing mistake.

The set of complete lists of Nature's draws leading to acceptance is given by

$$\overline{\Phi}^*_{HT} = \{(+,+),(+,-),(-,+)\}.$$

Part 3: The advisor's strategy is optimal by the same reasoning as for Part 1 and 2. The equilibrium is a sequential equilibrium by the same reasoning as for Part 2.

Part 4: σ_M is optimal since for $\widetilde{h} = \{+,+\}$ the advisor prefers acceptance since $\lambda_A < x_{(+,+)}$ and the DM accepts if $m = \{+,+\}$. Otherwise, there is no message available for which the DM would accept. σ_A is optimal: In period 2, following $h_1 \in \{(-),(\emptyset)\}$, the advisor stops testing because irrespective of the second outcome the DM will reject. Following $h_1 = (+)$, if he stopped the DM would reject. If he tests the DM will accept if and only if $h_2 = (+,+)$. The advisor tests since (23) holds. This implies that he tests in period 1.

Consider a completely mixed strategy constructed as in Part 2. As $\varepsilon \to 0$, clearly $\left(\sigma_A^\varepsilon, \sigma_M^\varepsilon\right) \to (\sigma_A, \sigma_M)$. Furthermore, for any m, as $\varepsilon \to 0$ (34) holds for the following reasons: If $\varepsilon = 0$ $m = \{-\}$ occurs on path following $h_2 = (-,\emptyset)$, $m = \{+\}$ occurs on path following $h_2 = (+,-)$. In the limit as $\varepsilon \to 0$, $m = \{\emptyset\}$ is mostly likely observed following $h_2 \in \{(\emptyset,\emptyset)\}$ since this history occurs with a probability of order ε^2 (due to two testing mistakes) whereas all other histories occur with probability of at least order ε^3 because they require a reporting mistake.

The set of complete lists of Nature's draws leading to acceptance is given by $\overline{\Phi}^*_{HT} = \{(+,+)\}$.

Part 5. σ_M is optimal: for $\widetilde{h} \in \{\{-\},\{+,-\},\{-,-\}\}$ the advisor prefers rejection since $\lambda_A > x_{(+,+)}$ and the DM rejects if $m = \{-\}$. Otherwise, there is no message available for which the DM would reject. σ_A 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 will reject. Following $h_1 = (+)$, if he stopped the DM would accept, but if he tests then the DM accepts if and only if $h_2 = (+,+)$ and (18) does not hold. Following $h_1 = (\emptyset)$, if he stopped the DM would accept, but if he tests then the DM accepts if and only if $h_2 = (\emptyset,+)$. Testing is optimal since

$$\lambda_A Pr(false) > Pr(true)(1-p_T) + \lambda_A Pr(false)(1-p_F)$$

$$\lambda_A > \frac{q(1-p_T)}{(1-q)\,p_F} = x_{(-,\cdot)}$$

This implies that he tests in period 1.

Consider a completely mixed strategy constructed as in Part 2. As $\varepsilon \to 0$, clearly $\left(\sigma_A^\varepsilon, \sigma_M^\varepsilon\right) \to (\sigma_A, \sigma_M)$. Furthermore, for any m, as $\varepsilon \to 0$ (34) holds for the following reasons: If $\varepsilon = 0$ $m = \{-\}$ occurs on path following $h_2 \in \{(-,\emptyset),(+,-)\}$, $m = \{\emptyset\}$ occurs on path following $h_2 = (+,+)$. In the limit as $\varepsilon \to 0$, $m = \{+\}$ is mostly likely observed following $h_2 \in \{(+,+),(+,-)\}$ since each of these histories occurs with a probability of order ε^3 (due to a reporting mistake) whereas all other histories occur with probability of at least order ε^4 because they require a testing and a reporting mistake.

The set of complete lists of Nature's draws leading to acceptance is given by $\overline{\Phi}^*_{HT} = \{(+,+)\}$.

Part 6: If the DM attains her first-best expected payoff in the Pareto-undominated equilibrium that minimizes her payoff, she must also attain her first-best in the one that maximizes her payoff, i.e. if $\underline{\Phi}^*_{HT} = \Phi^{FB}_{DM}$ then $\overline{\Phi}^*_{HT} = \Phi^{FB}_{DM}$. By Lemma 9, for any $(\lambda_A, \lambda_{DM})$ not covered in Parts 1-5 it holds that $\underline{\Phi}^*_{HT} = \Phi^{FB}_{DM}$.

Part 7: $\overline{\Phi}^*_{HT} = \Phi^{FB}_{DM}$ by the proofs of Parts 1-5 and by Part 6.

A.3 Lemma 1 [Equilibrium]

Lemma 8 shows that there exists a unique sequential equilibrium under observable testing. Lemmas 9 and 10 show that equilibrium exists under hidden testing, and that multiple equilibria can exist, e.g. if $x_{\Phi\setminus(-,-)} < \lambda_{DM} \leq x_{(+,\cdot)}$ and $\lambda_A < x_{(+,-)}$.

A.4 Theorem 1 [DM Payoff Comparison]

The proof of Theorem 1 will use the definition of Φ^* from Section A.1, where Φ^* is the set of complete lists of Nature's draws which are mapped into acceptance on the equilibrium path. Note that Φ^* contains all the information needed to compare expected payoffs since payoffs depend only on the state of the world and the DM's action. The proof will use Φ^*_{OT} for equilibria under observable testing as given by Lemma 8, and $\underline{\Phi}^*_{HT}$ for advisor-preferred equilibria under hidden testing as given by Lemma 9 and $\overline{\Phi}^*_{HT}$ for DM-preferred equilibria under hidden testing as given by Lemma 10. In addition, the proof will use the definition of the DM's first-best acceptance set, denoted by Φ^{FB}_{DM} and defined in Section A.2 to be a subset of the complete lists of Nature's draws which contains all $\phi \in \Phi$ such that the DM prefers to accept if and only if she observed $h_2 = \phi$. Φ^{FB}_{DM} is given by (24).

Region 1: Suppose $x_\Phi < \lambda_{DM} \leq x_{(+,-)}$, then $\Phi^{FB}_{DM} = \{(+,+),(+,-),(-,+)\}$.

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

$$\Phi^*_{OT} = \underline{\Phi}^*_{HT} = \overline{\Phi}^*_{HT} = \Phi^{FB}_{DM}. \tag{35}$$

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

$$\underline{\Phi}^*_{HT} = \{(+,+)\} \subset \Phi^*_{OT} = \{(+,+),(+,-)\} \subset \overline{\Phi}^*_{HT} = \Phi^{FB}_{DM}.$$

(c) If $x_{(+,\cdot)} \leq \lambda_A < x_{(+,+)}$ then $(\lambda_A, \lambda_{DM}) \in B_{II}$ and

$$\Phi^*_{OT} = \{\emptyset\} \subset \underline{\Phi}^*_{HT} = \{(+,+)\} \subset \overline{\Phi}^*_{HT} = \Phi^{FB}_{DM}.$$

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

$$\Phi^*_{OT} = \underline{\Phi}^*_{HT} = \{\emptyset\} \subset \overline{\Phi}^*_{HT} = \Phi^{FB}_{DM}. \tag{36}$$

Region 2: Suppose $max\left\{x_\Phi, x_{(+,-)}\right\} < \lambda_{DM} \leq x_{(+,\cdot)}$, then $\Phi^{FB}_{DM} = \{(+,+)\}$.

(a) If $max\left\{x_\Phi, x_{(+,-)}\right\} < \lambda_{DM} \leq x_{\Phi\setminus(-,-)}$ and $\lambda_A < x_{(+,-)}$ then $(\lambda_A, \lambda_{DM}) \in W_I$ and

$$\Phi^{FB}_{DM} = \overline{\Phi}^*_{HT} \subset \Phi^*_{OT} = \{(+,+),(+,-)\} \subset \underline{\Phi}^*_{HT} = \{(+,+),(+,-),(-,+)\}.$$

(b) If $x_{\Phi\backslash(-,-)} < \lambda_{DM} \leq x_{(+,\cdot)}$ and $\lambda_A < x_{(+,-)}$ then $(\lambda_A, \lambda_{DM}) \in B_I$ and

$$\Phi^{FB}_{DM} = \overline{\Phi}^*_{HT} = \underline{\Phi}^*_{HT} \subset \Phi^*_{OT} = \{(+,+),(+,-)\}.$$

(c) If $x_{(+,-)} \leq \lambda_A < x_{(+,+)}$ then (35) holds.

(d) If $\lambda_A \geq x_{(+,+)}$ then (36) holds.

Region 3: Suppose $x_{(+,\cdot)} < \lambda_{DM} \leq x_{(+,+)}$, then $\Phi^{FB}_{DM} = (+,+)$. If $\lambda_A < x_{(+,+)}$ then (35) holds. If $\lambda_A \geq x_{(+,+)}$ then (36) holds.
Region 4: Suppose $\lambda_{DM} > x_{(+,+)}$, then $\Phi^{FB}_{DM} = \{\emptyset\}$. For any λ_A, (35) holds.

The DM achieves a strictly lower expected payoff in the advisor-preferred equilibrium under hidden testing than in the unique equilibrium under observable testing if $\Phi^{FB}_{DM} \subseteq \Phi^*_{OT} \subset \underline{\Phi}^*_{HT}$ or $\underline{\Phi}^*_{HT} \subset \Phi^*_{OT} \subseteq \Phi^{FB}_{DM}$, achieves a strictly higher expected payoff if $\Phi^{FB}_{DM} \subseteq \underline{\Phi}^*_{HT} \subset \Phi^*_{OT}$ or $\Phi^*_{OT} \subset \underline{\Phi}^*_{HT} \subseteq \Phi^{FB}_{DM}$ and achieves the same expected payoff if $\Phi^{FB}_{DM} = \underline{\Phi}^*_{HT} = \Phi^*_{OT}$.
Part 1: Hence, the DM achieves a strictly higher expected payoff in the advisor-preferred equilibrium under hidden testing than in the unique equilibrium under observable testing if and only if $(\lambda_A, \lambda_{DM}) \in B_I \cup B_{II}$. By definition, she a weakly higher expected payoff in the DM-preferred than in the advisor-preferred equilibrium.
Part 2: Hence, the DM achieves a strictly lower expected payoff in the advisor-preferred equilibrium under hidden testing than in the unique equilibrium under observable testing if and only if $(\lambda_A, \lambda_{DM}) \in W_I \cup W_{II}$.

A.5 Corollary 1 [Preference Alignment]

By Theorem 1, the DM is weakly better off in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing if and only if both $(\lambda_A, \lambda_{DM}) \notin W_I$ and $(\lambda_A, \lambda_{DM}) \notin W_{II}$. Solve $\sup\,(\lambda_{DM} - \lambda_A)$ s.t. $(\lambda_A, \lambda_{DM}) \in W_I$. The solution is $\lambda_{DM} - \lambda_A = x_{\Phi\backslash(-,-)}$. Next, solve $\sup\,(\lambda_A - \lambda_{DM})$ s.t. $(\lambda_A, \lambda_{DM}) \in W_{II}$. The solution is $\lambda_A - \lambda_{DM} = x_{(+,\cdot)} - x_{\Phi}$. Choose $d = max\left\{x_{(+,\cdot)}, x_{(+,\cdot)} - x_{\Phi}\right\} = max\left\{\frac{q\left(1-(1-p_T)^2\right)}{(1-q)\left(1-p_F^2\right)}, \frac{qp_T}{(1-q)(1-p_F)} - \frac{q}{1-q}\right\}$.

A.6 Proposition 1 [First-Best Benchmark]

Part 1: By the proof of Theorem 1, the complement of Z_{OT}, denoted by Z^c_{OT}, is given by:

$$Z^c_{OT} = W_I \cup W_{II} \cup B_I \cup B_{II} \cup \left\{(\lambda_{DM}, \lambda_A) : x_{\Phi} < \lambda_{DM} < x_{(+,+)}, \lambda_A \geq x_{(+,+)}\right\},$$

and the complement of Z_{HT}, denoted by Z_{HT}^c, is given by:

$$Z_{HT}^c = W_I \cup W_{II} \cup B_I \cup \left\{ (\lambda_{DM}, \lambda_A) : x_\Phi < \lambda_{DM} < x_{(+,+)}, \lambda_A \geq x_{(+,+)} \right\}.$$

Hence, $Z_{OT} = Z_{HT} \cup B_{II}$ and, hence, $Z_{OT} \subset Z_{HT}$.
Part 2: This is shown by Part 7 of Lemma 10.

A.7 Proposition 2 [Advisor Payoff Comparison]

I will use the same procedure as in the proof of Theorem 1. The advisor's first-best acceptance set is denoted by Φ_A^{FB} and given if (25).
Proof: Whenever $\Phi_{OT}^* = \underline{\Phi}_{HT}^* = \overline{\Phi}_{HT}^*$ then the advisor must be as well off in any Pareto-undominated equilibrium under hidden testing as in the unique equilibrium under observable testing. By the proof of Theorem 1 this holds unless either i) $(\lambda_A, \lambda_{DM}) \in W_I \cup W_{II} \cup B_I \cup B_{II}$ or ii) $\lambda_A \geq x_{(+,+)}$ and $\lambda_{DM} \leq x_{(+,+)}$.
Region 1: Suppose $x_\Phi < \lambda_{DM} \leq x_{(+,-)}$.

(a) If $x_{(+,-)} \leq \lambda_A < x_{(+,\cdot)}$ then $(\lambda_A, \lambda_{DM}) \in W_{II}$ and

$$\underline{\Phi}_{HT}^* = \Phi_A^{FB} = \{(+,+)\} \subset \Phi_{OT}^* = \{(+,+), (+,-)\} \subset \overline{\Phi}_{HT}^* = \{(+,+), (+,-), (-,+)\}.$$

(b) If $x_{(+,\cdot)} \leq \lambda_A < x_{(+,+)}$ then $(\lambda_A, \lambda_{DM}) \in B_{II}$ and

$$\Phi_{OT}^* = \{\emptyset\} \subset \underline{\Phi}_{HT}^* = \Phi_A^{FB} = \{(+,+)\} \subset \overline{\Phi}_{HT}^* = \{(+,+), (+,-), (-,+)\}.$$

(c) If $\lambda_A \geq x_{(+,+)}$ then

$$\Phi_{OT}^* = \underline{\Phi}_{HT}^* = \Phi_A^{FB} = \{\emptyset\} \subset \overline{\Phi}_{HT}^* = \{(+,+), (+,-), (-,+)\}.$$

Region 2: Suppose $max\left\{x_\Phi, x_{(+,-)}\right\} < \lambda_{DM} \leq x_{(+,\cdot)}$, then $\Phi_{DM}^{FB} = \{(+,+)\}$.

(a) If $max\left\{x_\Phi, x_{(+,-)}\right\} < \lambda_{DM} \leq x_{\Phi \setminus (-,-)}$ and $\lambda_A < x_{(+,-)}$ then $(\lambda_A, \lambda_{DM}) \in W_I$ and
$$\overline{\Phi}_{HT}^* \subset \Phi_{OT}^* = \{(+,+), (+,-)\} \subset \underline{\Phi}_{HT}^* = \{(+,+), (+,-), (-,+)\} \subseteq \Phi_A^{FB}.$$

(b) If $x_{\Phi \setminus (-,-)} < \lambda_{DM} \leq x_{(+,\cdot)}$ and $\lambda_A < x_{(+,-)}$ then $(\lambda_A, \lambda_{DM}) \in B_I$ and

$$\overline{\Phi}_{HT}^* = \underline{\Phi}_{HT}^* = \{(+,+)\} \subset \Phi_{OT}^* = \{(+,+), (+,-)\} \subset \{(+,+), (+,-), (-,+)\} \subseteq \Phi_A^{FB}$$

(c) If $\lambda_A \geq x_{(+,+)}$ then

$$\Phi^*_{OT} = \underline{\Phi}^*_{HT} = \Phi^{FB}_A = \{\emptyset\} \subset \overline{\Phi}^*_{HT} = \{(+,+)\}.$$

Region 3: Suppose $x_{(+,\cdot)} < \lambda_{DM} \leq x_{(+,+)}$. If $\lambda_A \geq x_{(+,+)}$ then the same applies as in Region 2, Part (c).

The advisor achieves a strictly lower expected payoff in the advisor-preferred equilibrium under hidden testing than in the unique equilibrium under observable testing if $\Phi^{FB}_A \subseteq \Phi^*_{OT} \subset \underline{\Phi}^*_{HT}$ or $\underline{\Phi}^*_{HT} \subset \Phi^*_{OT} \subseteq \Phi^{FB}_A$, achieves a strictly higher expected payoff if $\Phi^{FB}_A \subseteq \underline{\Phi}^*_{HT} \subset \Phi^*_{OT}$ or $\Phi^*_{OT} \subset \underline{\Phi}^*_{HT} \subseteq \Phi^{FB}_A$ and achieves the same expected payoff if $\Phi^{FB}_A = \underline{\Phi}^*_{HT} = \Phi^*_{OT}$. In addition, the advisor achieves a strictly lower expected payoff in the DM-preferred equilibrium under hidden testing than in the unique equilibrium under observable testing if $\Phi^{FB}_A \subseteq \Phi^*_{OT} \subset \overline{\Phi}^*_{HT}$ or $\overline{\Phi}^*_{HT} \subset \Phi^*_{OT} \subseteq \Phi^{FB}_A$, achieves a strictly higher expected payoff if $\Phi^{FB}_A \subseteq \overline{\Phi}^*_{HT} \subset \Phi^*_{OT}$ or $\Phi^*_{OT} \subset \overline{\Phi}^*_{HT} \subseteq \Phi^{FB}_A$ and achieves the same expected payoff if $\Phi^{FB}_A = \overline{\Phi}^*_{HT} = \Phi^*_{OT}$. If $x_{(+,\cdot)} \leq \lambda_A < x_{(+,+)}$ then $(\lambda_A, \lambda_{DM}) \in B_{II}$ and the advisor is strictly worse off in the DM-preferred equilibrium under hidden testing than in the unique equilibrium under observable testing since

$$\begin{aligned} Pr(true) &< Pr(true)(1-p_T)^2 + \lambda_A Pr(false)\left(1-p_F^2\right) \\ q &< q(1-p_T)^2 + \lambda_A (1-q)\left(1-p_F^2\right) \\ x_{\Phi\setminus(-,-)} = \frac{q\left(1-(1-p_T)^2\right)}{(1-q)\left(1-p_F^2\right)} &< \lambda_A. \end{aligned}$$

Part 1: Hence, in any region, the advisor has a weakly lower expected payoff in the DM-preferred equilibrium under hidden testing than in the unique equilibrium under observable testing.

Part 2: Hence, the advisor has a weakly higher expected payoff in the DM-preferred equilibrium under hidden testing than in the unique equilibrium under observable testing if and only if $(\lambda_A, \lambda_{DM}) \in B_I$. By definition, he achieves a weakly higher expected payoff in the advisor-preferred than in the DM-preferred equilibrium. Therefore, there is no other region in which he achieves a strictly higher expected payoff in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing.

Part 3: Hence, the advisor has a strictly higher expected payoff in the advisor-preferred equilibrium under hidden testing than in the unique equilibrium under observable testing if and only if $(\lambda_A, \lambda_{DM}) \in B_{II} \cup W_I \cup W_{II}$. By definition, he achieves a weakly higher expected payoff in the advisor-preferred than in the DM-preferred equilibrium. Therefore, there is no other region in which he achieves a strictly higher expected payoff in some Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing.

A.8 Lemma 5 [Observable Testing Equilibrium]

Suppose testing is observable. Denote by x_j the posterior likelihood ratio that the hypothesis is true conditional on observing j excess positive outcomes, i.e. $h_n = (y,z)$ where $y-z=j$ and $j \in \{-N,...,N\}$. Denote by $x(n)$ the likelihood ratio at the end of period n. Define

$$l \equiv sup\left\{j|j \in \{0,...,N-1\},x_j < \lambda_{DM}\right\} \tag{37}$$

$$r \equiv sup\left\{j|j \in \mathbb{Z},x_j < \lambda_A\right\} \tag{38}$$

i.e. l (r) is the highest number of excess positive outcomes such that the DM (advisor) prefers to reject. 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.

Part 1. Choose $\underline{\lambda}_A = x_{2(l+1)-N}$. Suppose $\lambda_A < \underline{\lambda}_A$, i.e. $r+1 \leq 2(l+1)-N$. Note that this implies that the advisor is more enthusiastic than the DM, since $2(l+1)-N \leq l$. The DM's optimal strategy is to accept if and only if $h_N = (y,z)$ where $y-z \geq l+1$ since $\lambda_{DM} \leq x_{l+1}$, as shown by (3). In Claims 1a-d) below, I show that the advisor's optimal strategy is as follows: test in period $n+1$ if and only if, at the end of period n, $x_{l+1-(N-n)} \leq x(n) \leq x_l$ for $n \in \{0,...,N-1\}$.

Claim 1a *If the posterior likelihood ratio at the end of period n is equal to x_{l+1}, i.e. $x(n) = x_{l+1}$, then the advisor does not test in period $n+1$ for $n \in \{l+1,...,N-1\}$.*

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

Claim 1b *If $x(n) \geq x_{l+2}$ then the advisor does not test in period $n+1$ for $n \in \{l+2,...,N-1\}$.*

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

Claim 1c *If $x(n) \leq x_{l-(N-n)}$ then the advisor does not test in period $n+1$ for $n \in \{0,...,N-1\}$.*

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

Claim 1d *If $x_{l+1-(N-n)}\leq x(n)\leq x_l$ then the advisor tests for $n\in\{0,...,N-1\}$.*

Proof: For any $x_{l+1-(N-n)}\leq x(n)\leq x_l$, there are realizations of the remaining $N-n$ outcomes such that at some future period $n^{'}>n$ it holds that $x\left(n^{'}\right)=x_{l+1}$. Hence, if the advisor tests in any period $1+n$ if $x_{l+1-(N-n)}\leq x(n)\leq x_l$, then eventually the advisor will stop either because Claim 1a applies or because Claim 1c applies or because he runs out of tests. If Claim 1a applies, the DM accepts. If Claim 1c 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 $x_{l+1-(N-n)}\leq x(n)\leq x_l$, then the DM rejects. Therefore, testing is pivotal to the DM's choice if and only if x_{l+1} is reached. The advisor prefers acceptance at x_{l+1} since $\lambda_A\leq x_{r+1}<x_{l+1}$. Hence, there is a strict upside but no downside to testing.

Part 2. Suppose $\lambda_A>\lambda_{DM}$. Choose $\overline{\lambda}_A=x_k$, where

$$k=inf\left\{g|g\in\mathbb{Z}_{>0},g\geq l+1+\frac{N-l+1}{2}\right\}. \tag{39}$$

Suppose $\lambda_A>\overline{\lambda}_A$. Note that this implies that the advisor is more reluctant than the DM since $\lambda_{DM}\leq x_{l+1}\leq\overline{\lambda}_A$ since $k\geq l+1$. As in Part 1, the DM's optimal strategy is to accept if and only if $h_N=(y,z)$ where $y-z\geq l+1$. I will show that the advisor's optimal strategy is to test in any period $n+1$ for $n\in\{0,...,N-1\}$ if and only if at the end of period n it holds that $x_{l+1}\leq x(n)\leq x_{l+(N-n)}$.

Claim 2a *If $x(n)=x_k$ then the advisor does not test in period $n+1$ for $n\in\{k,...,N-1\}$.*

Proof: First, note that given likelihood ratio x_k the DM prefers to accept since $x_k>x_{l+1}$. Second, I will show that if $x(n)=x_k$ then no realization of the remaining test outcomes can lead the DM to prefer rejection. If all remaining $N-n$ outcomes are negative then $x(N)=x_{k-(N-n)}$. This is the lowest possible value of $x(N)$ given $x(n)=x_k$. Independent of n, the DM prefers to accept at the lowest possible value of $x(N)$. This is because $\lambda_{DM}\leq x_{l+1}\leq x_{k-(N-n)}$ since $k\geq\frac{N+l+1}{2}$ and $n\geq k$ for $x(n)=x_k$ to be feasible. Therefore, there is no reason to test.

Claim 2b *If $x(n)\geq x_{k+1}$ then the advisor does not test in period $n+1$ for $n\in\{k+1,...,N-1\}$.*

Proof: Whatever the advisor's action in any period $n+1$ for $n\in\{k+1,...,N-1\}$, if $x(n)>x_k$, then by Claim 2a it must be that in any period $n^{'}>n$, $x\left(n^{'}\right)\leq x_k$. This implies that the DM always accepts since $\lambda_{DM}\leq x_l<x_k$. Therefore, there is no reason to test.

Claim 2c *If $x(n)\leq x_{l-(N-n)}$ then the advisor does not test in period $n+1$ for $n\in\{0,...,N-1\}$.*

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

Claim 2d *If $x_{l-(N-n)}\leq x(n)\leq x_l$ then the advisor does not test for $n\in\{0,...,N-1\}$.*

Proof: First, I will show that if the advisor tests in any period $n+1$ if $x_{l-(N-n)}\leq x(n)\leq x_l$ then with a strictly positive probability the DM will accept when the advisor prefers to reject. If the advisor tests in any period $n+1$ if $x_{l-(N-n)} \leq x(n) \leq x_l$ then eventually either Claim 2c applies and the DM rejects or at some future period $n^{'} > n$ it holds that $x\left(n^{'}\right) \geq x_{l+1}$. I have not yet specified what the advisor will do in period $n+1$ if $x_{l+1} \leq x(n) \leq x_{k-1}$. 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 $x_{l+1} \leq x(n) \leq x_{k-1}$ then $x_{l+1} \leq x(N) \leq x_{k-1}$ and the DM accepts but the advisor prefers to reject since $\lambda_{DM} \leq x(N) < \overline{\lambda}_A$. If he were to test at any $n+1$ if $x_{l+1} \leq x(n) \leq x_{k-1}$ then eventually either Claim 2a applies and the DM accepts or Claim 2c will apply and the DM rejects. If Claim 2a applies then the advisor prefers to reject since $x_k \leq \overline{\lambda}_A$. Second, if the advisor does not test in in any period $n+1$ if $x_{l-(N-n)} \leq x(n) \leq x_l$ then $x(N) \leq x_l$ and the DM rejects since $x_l < \lambda_{DM}$. Lastly, note that if the DM rejects then the advisor also prefers to reject given that the advisor is more reluctant than the DM. This implies that testing has no upside but a strict downside.

Claim 2e *The advisor tests if* $x_{l+1} \leq x(n) \leq x_{k-1}$ *for* $n \in \{l+1,...,N-1\}$.

Proof: For any likelihood ratio $x_{l+1} \leq x(n) \leq x_{k-1}$, the advisor prefers to reject since $x_{k-1} < x_k = \overline{\lambda}_A$. If he does not test then $x_{l+1} \leq x(N) \leq x_{k-1}$ and the DM accepts. If he tests in any period $n+1$ then eventually he stops either because Claim 2a applies, or Claim 2d applies or the runs out of tests. If Claim 2a applies or if he runs out of tests then the DM accepts, but the advisor prefers to reject since $x(N) \leq x_k < \lambda_A$. If Claim 2d applies, the DM rejects and the advisor prefers to reject. Therefore, testing has a strict upside but no downside.

To show the equilibria in Part 1 and Part 2 are sequential equilibria, consider the fully mixed strategy for the advisor in which he tests with probability $1-\varepsilon$ in period $n+1$ whenever the equilibrium strategy specified that he tests, and otherwise he tests with probability ε, where $\varepsilon \in (0,1)$. As $\varepsilon \to 0$ this strategy converges to the advisor's equilibrium strategy. Beliefs at any history h_N are constructed using Bayes' rule and, therefore, the beliefs formed based on the completely mixed strategy must be equal to the equilibrium system of beliefs for any ε.

A.9 Lemma 6 [Hidden Testing: Advisor-preferred Equilibrium]

Suppose testing is hidden. Denote by $\widetilde{h}_n = (y,z)$ the unordered history of outcomes with exactly y positive and z negative outcomes at the end of period n, where $y,z \in \mathbb{Z}_{\geq 0}$. Denote by $m = (\widetilde{y},\widetilde{z})$ a report of exactly $\widetilde{y}$ positive and $\widetilde{z}$ negative outcomes, where $\widetilde{y} \in \{0,...,y\}$ and $z \in \{0,...,z\}$. Define l and r by (37) and (38).

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

$$\frac{Pr\left(true|m=\left(\widetilde{y},\widetilde{z}\right)\right)}{Pr\left(false|m=\left(\widetilde{y},\widetilde{z}\right)\right)}=\begin{cases}\frac{q\sum_{s=\widetilde{y}}^{s=N-\widetilde{z}}\binom{N}{s}p^{s}(1-p)^{N-s}}{(1-q)\sum_{s=\widetilde{y}}^{s=N-\widetilde{z}}\binom{N}{s}(1-p)^{s}p^{N-s}} & \text{if } \widetilde{y}\geq\overline{y} \text{ and } \widetilde{z}\geq 0,\\ \frac{q\sum_{s=max\{\widetilde{z},\overline{z}\}}^{s=N-\widetilde{y}}\binom{N}{s}p^{N-s}(1-p)^{s}}{(1-q)\sum_{s=max\{\widetilde{z},\overline{z}\}}^{s=N-\widetilde{y}}\binom{N}{s}(1-p)^{N-s}p^{s}} & \text{otherwise.}\end{cases}$$

which means that if the advisor reports $\widetilde{y}\geq\overline{y}$ and $\widetilde{z}\geq 0$, the DM believes that at least $\widetilde{y}$ and at most $N-\widetilde{z}$ in N tests were positive, otherwise, the DM believes that at least $max\{\widetilde{z},\overline{z}\}$ and at most $N-\widetilde{y}$ in N tests were negative. Given $m=(\widetilde{y},\widetilde{z})$, the DM accepts if and only if $\widetilde{y}\geq min\left\{\overline{y},\widehat{Y}\left(\widetilde{z}\right)\right\}$, where

$$\widehat{Y}\left(\widetilde{z}\right)\equiv inf\left\{j|j\in\mathbb{N},\frac{q\sum_{s=max\{\widetilde{z},\overline{z}\}}^{s=N-j}\binom{N}{s}p^{N-s}\left(1-p\right)^{s}}{\left(1-q\right)\sum_{s=max\{\widetilde{z},\overline{z}\}}^{s=N-j}\binom{N}{s}\left(1-p\right)^{N-s}p^{s}}\geq\lambda_{DM}\right\}. \tag{40}$$

The advisor's strategy is

$$\sigma_{n}\left(\widetilde{h}_{n}=(y,z)\right)=\begin{cases}0 & \text{if } y\geq\overline{y}\\ & \text{or if } z\geq\overline{z},\\ 1 & \text{otherwise.}\end{cases} \tag{41}$$

and

$$m=\begin{cases}(0,z) & \text{if } y-z\leq r,\\ (y,0) & \text{otherwise.}\end{cases} \tag{42}$$

Claim 1: *The advisor's reporting strategy is optimal.*
Proof: For any $\widetilde{h}_N=(y,z)$, reporting $m=(\widetilde{y},\widetilde{z})=(0,z)$ always leads to rejection. Therefore, if the advisor prefers rejection at $\widetilde{h}_N=(y,z)$, i.e. if $x(N)=x_{y-z}\leq x_r<\lambda_A$ or $y-z\leq r$, then it is optimal to report $m=(\widetilde{y},\widetilde{z})=(0,z)$. For any $\widetilde{h}_N=(y,z)$, if reporting $m=(\widetilde{y},\widetilde{z})=(y,0)$ does not lead to acceptance then no message will. Therefore, if the advisor prefers acceptance it is optimal to report $m=(\widetilde{y},\widetilde{z})=(y,0)$.
Claim 2: *The advisor's testing strategy is optimal.*
Proof: First, if $z\geq\overline{z}$ I will show that the advisor prefers to reject irrespective of the realization of the remaining outcomes. At $\widetilde{h}_n=(y,z)$, $x(n)=x_{(n-z)-z}$ and if all remaining $N-n$ outcomes are positive then $x(N)=x_{n-2z+(N-n)}=x_{N-2z}$. This is the highest possible likelihood ratio at the end of period N. If $z\geq\overline{z}$ then the advisor prefers to reject even at the highest possible likelihood ratio since $\lambda_A>x_r\geq x_{N-2z}$. Given his reporting strategy, if he stops testing he reports $m=(z,0)$ and the DM rejects. Hence, there is no reason to test. Second, if $y\geq\overline{y}$ I will show that the advisor prefers to accept irrespective of the realization of the remaining outcomes. At $\widetilde{h}_n=(y,z)$, $x(n)=x_{y-(n-y)}$ and if all remaining $N-n$ outcomes are negative then $x(N)=x_{2y-n-(N-n)}=x_{2y-N}$. This is the

lowest possible likelihood ratio at the end of period N. If $y \geq \bar{y}$ then the advisor prefers to accept even at the lowest possible likelihood ratio since $\lambda_A \leq x_{l+1} \leq x_{2y-N}$. Given his reporting strategy, if he stops testing he reports $m = (y, 0)$ 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 $y \geq \bar{y}$ or $z \geq \bar{z}$ and in each case the DM acts in his interest.

Claim 3: *The DM's strategy is optimal given her beliefs.*

Proof: Given $\widetilde{y} \geq \bar{y}$, the DM always accepts because, as shown in Claim 2, 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 $y \geq \widehat{Y}(\widetilde{z})$ by the definition of $\widehat{Y}(\widetilde{z})$ in (40).

Claim 4: *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* μ^k *denotes the beliefs derived from strategy profile* σ^k *using Bayes' rule.*

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

$$\sigma_n\left(\widetilde{h}_n = (y,z)\right) = \begin{cases} \varepsilon^N & \text{if } y \geq \bar{y} \\ & \text{or if } z \geq \bar{z}, \\ 1-\varepsilon^N & \text{otherwise,} \end{cases}$$

The advisor deviates from the equilibrium reporting strategy (42) as follows. If $y - z \leq r$ he fails to report each negative outcome independently with probability ε and reports any strictly positive number of positive outcomes with probability ε. If $y - z \geq r$ the advisor fails to report each positive outcome independently with probability ε^N and reports any strictly positive number of negative outcome with probability ε.

Clearly, the advisor's mixed strategy converges to his equilibrium strategy as $\varepsilon \to 0$. I will show that the DM's beliefs also converge. At $\varepsilon = 0$ the message $m = (\bar{y}, 0)$ follows a history in which the advisor stops testing as soon as $\widetilde{y} = \bar{y}$ and reports all positive outcomes. Therefore, at $\varepsilon = 0$ this message rises if and only if at least $\bar{y}$ outcomes in N tests are positive. By the same reasoning, at $\varepsilon = 0$ the message $m = (0, \bar{z})$ follows a history in which when the advisor stops testing as soon as $z = \bar{z}$ and reports all negative outcomes. Therefore, at $\varepsilon = 0$ this message rises if and only if at least $\bar{z}$ outcomes in N tests are negative.

Any message where $\widetilde{y} \geq \bar{y}$ and $\widetilde{z} \geq 0$, in the limit as $\varepsilon \to 0$, follows a history at which the advisor stops as soon as $y = \widetilde{y}$, reports all positive outcomes and a subset of negative outcomes. Since it must be that the advisor prefers to accept given $y \geq \bar{y}$ by Claim 2, reporting all positive outcomes is in line with the equilibrium reporting strategy. Note that the advisor must deviate from the equilibrium testing strategy to achieve $\widetilde{y} > \bar{y}$. The lowest number of deviations to allow for $\widetilde{y} \geq \bar{y}$

arises if the advisor stops as soon as $y = \widetilde{y}$. Any deviation to report some number of negative outcomes is equally likely. Hence, as $\varepsilon \to 0$ this message arises if and only if at least $\widetilde{y}$ and at most $N - \widetilde{z}$ outcomes in N tests are positive. By the same reasoning, any message where $\widetilde{y} \geq 0$ and $\widetilde{z} \geq \overline{z}$, in the limit as $\varepsilon \to 0$, follows a history in which the advisor stops testing as soon as $z = \widetilde{z}$, reports all negative outcomes and some subset of positive outcomes. Hence, as $\varepsilon \to 0$ this message arises if and only if at least $\widetilde{z}$ and at most $N - \widetilde{y}$ outcomes in N tests are negative.

Any message where $\widetilde{y} < \overline{y}$ and $\widetilde{z} < \overline{z}$, in the limit as $\varepsilon \to 0$, follows a history in which the advisor stops testing as soon as $z = \overline{z}$, fails to report $\overline{z} - \widetilde{z}$ negative outcomes and reports a subset of positive outcomes. To see why, suppose the advisor follows the equilibrium testing strategy. Then by period N, either $y = \overline{y}$ or $z = \overline{z}$. If $y = \overline{y}$, then to send this message the advisor must omit $\overline{y} - \widetilde{y}$ positive outcomes, which occurs with a probability of $\varepsilon^{N(\overline{y} - \widetilde{y})}$, and if $\widetilde{z} > 0$ then he must report a subset of negative outcomes, which occurs with probability ε. If $z = \overline{z}$, then to send this message the advisor must omit $\overline{z} - \widetilde{z}$ negative outcomes, which occurs with a probability of $\varepsilon^{\overline{z} - \widetilde{z}}$, and if $\widetilde{y} > 0$ then he must report a subset of positive outcomes, which occurs with probability ε. Since $N > \overline{z} - \widetilde{z}$, it must be that $N(\overline{y} - \widetilde{y}) > \overline{z} - \widetilde{z}$. Therefore, even if $\widetilde{z} = 0$ and $0 < \widetilde{y} < \overline{y}$, it is more likely that $z = \overline{z}$ than $y = \overline{y}$ since $\varepsilon^{\overline{z}}\varepsilon > \varepsilon^{N(\overline{y} - \widetilde{y})}$. In addition, a single deviation from the advisor's testing strategy occurs with probability ε^N, and therefore, is less likely than the event that the advisor followed the equilibrium testing strategy and misreported. Hence, as $\varepsilon \to 0$ this message arises if and only if at least $\overline{z}$ and at most $N - \widetilde{y}$ outcomes in N tests are negative.

Claim 5: *This is an advisor-preferred equilibrium.*

By period N, either $y = \overline{y}$ and the DM accepts or $z = \overline{z}$ and the DM rejects. By Claim 2, 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. Using the definition of the first-best acceptance set from Section A.2, for any realization ϕ acceptance is chosen if and only if $\phi \in \Phi_A^{FB}$.

Part 2: The advisor is more enthusiastic, i.e. $r + 1 \leq l$. It is helpful to define

$$\widehat{Z}(\widetilde{y}) \equiv sup \left\{ j | j \in \mathbb{N}_{\geq 0}, \lambda_{DM} \leq \frac{q \sum_{s=\widetilde{y}}^{s=N-j} \binom{N}{s} p^s (1-p)^{N-s}}{(1-q) \sum_{s=\widetilde{y}}^{s=N-j} \binom{N}{s} (1-p)^s p^{N-s}} \right\}, \tag{43}$$

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

First, if $max\{Y, \overline{y}\} = \overline{y}$, which implies $min\{\overline{z}, N - Y + 1\} = \overline{z}$ then the following is an advisor-preferred sequential equilibrium: the DM's beliefs are given by

$$\frac{Pr(true|m=(\widetilde{y},\widetilde{z}))}{Pr(false|m=(\widetilde{y},\widetilde{z}))}=\begin{cases}\frac{q\sum_{s=\widetilde{y}}^{s=N-\widetilde{z}}\binom{N}{s}p^s(1-p)^{N-s}}{(1-q)\sum_{s=\widetilde{y}}^{s=N-\widetilde{z}}\binom{N}{s}(1-p)^s p^{N-s}} & \text{if } \widetilde{y}\geq Y \text{ and } \widetilde{z}\geq 0,\\ \frac{q\sum_{s=max\{\widetilde{z},\bar{z}\}}^{s=N-\widetilde{y}}\binom{N}{s}p^{N-s}(1-p)^s}{(1-q)\sum_{s=max\{\widetilde{z},\bar{z}\}}^{s=N-\widetilde{y}}\binom{N}{s}(1-p)^{N-s}p^s} & \text{otherwise.}\end{cases}$$

Given $m=(\widetilde{y},\widetilde{z})$, the DM accepts if and only if the advisor reports $\widetilde{y}\geq Y$ and $\widetilde{z}\leq\widehat{Z}(\widetilde{y})$, where The advisor's strategy is given by (41) and (42).

Second, if $max\{Y,\bar{y}\}=Y$, which implies $min\{\bar{z},N-Y+1\}=N-Y+1$ then the following is a sequential equilibrium: the DM's beliefs are given by

$$\frac{Pr(true|m=(\widetilde{y},\widetilde{z}))}{Pr(false|m=(\widetilde{y},\widetilde{z}))}=\begin{cases}\frac{q\sum_{s=\widetilde{y}}^{s=N-\widetilde{z}}\binom{N}{s}p^s(1-p)^{N-s}}{(1-q)\sum_{s=\widetilde{y}}^{s=N-\widetilde{z}}\binom{N}{s}(1-p)^s p^{N-s}} & \text{if } \widetilde{y}\geq\bar{y} \text{ and } \widetilde{z}\geq 0\\ & \text{or if } \bar{z}>\widetilde{z}\geq N-Y+1 \text{ and } r+\widetilde{z}+1\leq\widetilde{y}<\bar{y},\\ \frac{q\sum_{s=\widetilde{z}}^{s=N-\widetilde{y}}\binom{N}{s}p^{N-s}(1-p)^s}{(1-q)\sum_{s=\widetilde{z}}^{s=N-\widetilde{y}}\binom{N}{s}(1-p)^{N-s}p^s} & \text{if } \bar{z}>\widetilde{z}\geq N-Y+1 \text{ and } \widetilde{y}<min\{\bar{y},r+\widetilde{z}+1\}\\ & \text{or if } \widetilde{z}\geq\bar{z} \text{ and } \widetilde{y}<\bar{y},\\ \frac{q\sum_{s=N-Y+1}^{s=N-\widetilde{y}}\binom{N}{s}p^{N-s}(1-p)^s}{(1-q)\sum_{s=N-Y+1}^{s=N-\widetilde{y}}\binom{N}{s}(1-p)^{N-s}p^s} & \text{if } \widetilde{z}<N-Y+1 \text{ and } \widetilde{y}<\bar{y}.\end{cases}$$

Given $m=(\widetilde{y},\widetilde{z})$, the DM accepts if and only if the advisor reports $\widetilde{y}\geq Y$ and $\widetilde{z}\leq\widehat{Z}(\widetilde{y})$. The advisor's testing strategy is

$$\sigma_n\left(\widetilde{h}_n=(y,z)\right)=\begin{cases}0 & \text{if } y\geq Y\\ & \text{or if } z\geq N-Y+1,\\ 1 & \text{otherwise.}\end{cases}$$

and his reporting strategy is given by (42).

Claim 1: *The advisor reporting strategy is optimal.*

Proof: See Claim 1 in Part 1 above.

Claim 2: *The advisor's testing strategy is optimal.*

Proof: If $z\geq N-Y+1$ then by period N it cannot be that $y\geq Y$. Therefore, there cannot be a feasible message for which the DM accepts and, hence, there is no reason to test. In addition, by the argument in the proof of Claim 1 in Part 1 above, it follows that the advisor's preferred action is not affected by future outcomes if $z\geq\bar{z}$ or $y\geq\bar{y}$. If $z\geq\bar{z}$, the advisor can always induce the DM to reject. Hence, there is no reason to test if $z\geq\bar{z}$. If $y\geq\bar{y}$ the advisor can induce the DM to accept if and only if $y\geq Y$. Hence, there is no reason to test if $y\geq max\{Y,\bar{y}\}$. 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 3: *The DM's strategy is optimal given her beliefs.*

Proof: By the definition of Y in (12), the DM accepts if $m = (\widetilde{y}, 0)$ where $\widetilde{y} \geq Y$. Given $\widetilde{y} \geq Y$, the DM accepts if and only if $\widetilde{z} \leq \widehat{Z}(\widetilde{y})$ by definition of $\widehat{Z}(\widetilde{y})$ in (43). Suppose $max\{Y, \bar{y}\} = \bar{y}$. Then given $\widetilde{y} < Y$, the DM believes that at least $\bar{z}$ outcomes in N tests are negative. If exactly $\bar{z}$ outcomes in N tests are negative then the advisor prefers to reject by Claim 2, and since the advisor is more enthusiastic, it must be that the DM would also optimally reject. Suppose $max\{Y, \bar{y}\} = Y$. Then given $\widetilde{y} < Y$, the DM believes that at least z outcomes in N tests are negative, where $z > N - Y + 1$. Knowing that at least Y outcomes in N tests are positive raises her posterior belief above her prior by the definition of Y. Since her expected posterior must equal her prior, the complementary event that at least $N - Y + 1$ 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 z outcomes in N tests are negative where $z > N - Y + 1$.

Claim 4: *There exists a sequence of completely mixed strategies* $\left\{\sigma^k\right\}_{k=1}^{\infty}$, *with* $\lim_{k\to\infty} \sigma^k = \sigma$, *such that the system of beliefs* $\mu = \lim_{k\to\infty} \mu^k$, *where* μ^k *denotes the beliefs derived from strategy profile* σ^k *using Bayes' rule.*

Proof: First, suppose that $max\{Y, \bar{y}\} = \bar{y}$, which implies $min\{\bar{z}, N - Y + 1\} = \bar{z}$. Consider the same completely mixed strategy for the advisor with $\varepsilon \in (0,1)$ as in Part 1 above. Clearly, the advisor's mixed strategy converges to his equilibrium strategy as $\varepsilon \to 0$. The DM's beliefs converge by the same reasoning as in Part 1 above.

Second, suppose that $max\{Y, \bar{y}\} = Y$, which implies $min\{\bar{z}, N - Y + 1\} = N - Y + 1$. Consider the following completely mixed strategy for the advisor with $\varepsilon \in (0,1)$:

$$\sigma_n\left(\widetilde{h}_n = (y,z)\right) = \begin{cases} \varepsilon^N & \text{if } y \geq Y \\ & \text{or if } z \geq N - Y + 1, \\ 1 - \varepsilon^N & \text{otherwise.} \end{cases}$$

If $y - z \geq r + 1$ the advisor fails to report each positive outcome independently with probability ε^N and reports any strictly positive number of negative outcomes with probability ε. If $y - z \leq r$ the advisor fails to report each negative outcome independently with probability ε and reports any strictly positive number of positive outcomes with probability ε.

Clearly, the advisor's mixed strategy converges to his equilibrium strategy as $\varepsilon \to 0$. I will show that the DM's beliefs also converge. Any message where $\widetilde{y} \geq Y$ and $\widetilde{z} \geq 0$, in the limit as $\varepsilon \to 0$, follows a history at which the advisor stopped as soon as $y = \widetilde{y}$, reported all positive outcomes and a subset of negative outcomes. Since it must be that the advisor prefers to accept given $y \geq Y \geq \bar{y}$ by Claim 2, reporting all positive outcomes is in line with the equilibrium reporting strategy. Note that the advisor must deviated from the equilibrium testing strategy to achieve $\widetilde{y} > Y$. The lowest number of deviations to allow for $\tilde{y} \geq Y$ arises if the advisor stopped as soon as $y = \tilde{y}$. Any deviation to report some number of negative outcomes is equally likely. Hence, as $\varepsilon \to 0$ this message rises if and only if at least $\tilde{y}$ and at most $N - \tilde{z}$ outcomes in N tests are positive. Any message where $Y \geq \tilde{y} \geq \bar{y}$ and $\tilde{z} \geq 0$, in the limit as $\varepsilon \to 0$, follows a history at which the advisor stopped as soon as $z = N - Y + 1$, reported all positive outcomes and a subset of negative outcomes. No deviation from the equilibrium testing strategy is necessary. In addition, since it must be that the advisor prefers to accept given $y \geq \bar{y}$ by Claim 2, reporting all positive outcomes is in line with the equilibrium reporting strategy. Therefore, the only deviation is to report some number of negative outcomes if $\tilde{z} > 0$ and this kind of deviation is always necessary to generate a message with $\tilde{y} > 0$ and $\tilde{z} > 0$. Hence, as $\varepsilon \to 0$ this message rises if and only if at least $\tilde{y}$ and at most $N - \tilde{z}$ outcomes in N tests are positive.

Any message where $\bar{z} > \tilde{z} \geq N - Y + 1$ and $r + \tilde{z} + 1 \leq \tilde{y} < \bar{y}$, in the limit as $\varepsilon \to 0$, follows a history in which the advisor stops testing as soon as $z = \tilde{z}$, reports all positive outcomes and some subset of negative outcomes. Note that the advisor must deviated from the equilibrium testing strategy to achieve $\tilde{z} > N - Y + 1$. The lowest number of deviations to allow for $\tilde{z} > N - Y + 1$ arises if the advisor stopped as soon as $z = \tilde{z}$. Suppose the advisor preferred to accept, i.e. $y - z \geq r + 1$. Then the only necessary deviation from the equilibrium reporting strategy is that a subset of negative outcomes is reported. If the advisor had preferred to reject then he would have needed to deviate from the equilibrium reporting strategy 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 rises if and only if at least $\tilde{z}$ and at most $N - \tilde{y}$ outcomes in N tests are positive. Any message where $\tilde{z} \geq \bar{z}$ (which implies $\tilde{y} < \bar{y}$) or where $\bar{z} > \tilde{z} \geq N - Y + 1$ and $\tilde{y} < min\{\bar{y}, r + \tilde{z} + 1\}$, in the limit as $\varepsilon \to 0$, follows a history in which the advisor stops testing as soon as $z = \tilde{z}$, reports all negative outcomes and some subset of positive outcomes. Again, the lowest number of deviations to allow for $\tilde{z} > N - Y + 1$ arises if the advisor stopped as soon as $z = \tilde{z}$. If $\tilde{z} \geq \bar{z}$ or if $\bar{z} > \tilde{z} \geq N - Y + 1$ and $\tilde{y} - \tilde{z} \leq r$, the only necessary deviation from the equilibrium reporting strategy is that a subset of positive outcomes is reported and this kind of deviation is always necessary to generate a message with $\tilde{y} > 0$ and $\tilde{z} > 0$. Hence, as $\varepsilon \to 0$ this message rises if and only if at least $\tilde{z}$ and at most $N - \tilde{y}$ outcomes in N tests are negative.

Any message where $\tilde{y} < \bar{y}$ and $N - Y + 1 > \tilde{z}$, in the limit as $\varepsilon \to 0$, follows a history in which the advisor stops testing as soon as $z = N - Y + 1$, fails to report $N - Y + 1 - \tilde{z}$ negative outcomes and reports some subset of positive outcomes. Suppose the advisor follows the equilibrium testing strategy. Then by period N either $z = N - Y + 1$ or $y = Y$. If $y = Y$ then the advisor prefers to accept and would have to fail to report $Y - \tilde{y}$ positive outcomes, which happens with probability $\varepsilon^{N(Y - \tilde{y})}$. If $z = N - Y + 1$ then if he preferred to accept he would have to fail to report $y - \tilde{y}$ where $y < Y$ positive outcomes, which happens with probability $\varepsilon^{N(y - \tilde{y})}$, or if he preferred to reject he would have to fail to report $N - Y + 1 - \tilde{z}$ negative outcomes, which happens with probability $\varepsilon^{N - Y + 1 - \tilde{z}}$.

Since $N-Y+1-\widetilde{z}>N(y-\widetilde{y})>N(Y-\widetilde{y})$, it is most likely that $z=N-Y+1$ and the advisor fails to report $N-Y+1-\widetilde{z}$ negative outcomes. In addition, a single deviation from the advisor's testing strategy occurs with probability ε^N, and therefore, is less likely than the event that the advisor followed the equilibrium testing strategy and misreported. Hence, as $\varepsilon \to 0$ this message rises if and only if at least $N-Y+1$ and at most $N-\widetilde{y}$ outcomes in N tests are negative.

Claim 5: *This is an advisor-preferred equilibrium.*

By period N, either $y = max\{Y,\bar{y}\}$ and the DM accepts or $z = min\{N-Y+1,\bar{z}\}$ and the DM rejects. If $max\{Y,\bar{y}\} = \bar{y}$ then the advisor achieves his first-best payoff by the same reasoning as in Claim 5 of Part 1. If $max\{Y,\bar{y}\} = Y$, then given $y = Y \geq \bar{y}$ the advisor prefers to accept by Claim 2. However, if $z = N-Y+1$ the advisor may prefer to accept. To increase the advisor's payoff the DM would have to accept if $\widetilde{y} \geq y'$ for some $y' < Y$, but if the DM were to accept for $\widetilde{y} \geq y'$ then the advisor would stop testing as soon as $y = y'$ and send $m = \left(y',0\right)$. Given $m = \left(y',0\right)$ the DM would only infer that at least y' outcomes in N tests were positive. But then it is optimal for the DM to reject given the definition of Y.

A.10 Proposition 3 [DM Payoff Comparison]

Part 1a. Define

$$B_1 \equiv \{(\lambda_A,\lambda_{DM}) : \lambda_A < x_{N-2}, \bar{x} < \lambda_{DM} \leq x_{N-1}\}, \tag{44}$$

where

$$\bar{x} \equiv \frac{q\sum_{s=N-1}^{s=N}\binom{N}{s}p^s(1-p)^{N-s}}{(1-q)\sum_{s=N-1}^{s=N}\binom{N}{s}(1-p)^s p^{N-s}}, \tag{45}$$

i.e. $\bar{x}$ is the likelihood ratio conditional on at least $N-1$ outcomes in N tests being positive. Note that if $N=2$ then $B_1 = B_{II}$.

Suppose $(\lambda_A,\lambda_{DM}) \in B_1$. Applying Part 1 of the proof of Lemma 5, given $l = N-2$ it must be that $\underline{\lambda}_A = x_{N-2}$. Since $\lambda_A < x_{N-2} = \underline{\lambda}_A$, under observable testing, the advisor stops as soon as the likelihood ratio satisfies $x(n) = x_{N-1}$ or $x(n) = x_{n-2}$. If $x(N) = x_{N-1}$ then the DM accepts since $\lambda_{DM} \leq x_{N-1}$, otherwise the DM rejects.

Applying Part 2 of the proof of Lemma 6,

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

since

$$x_N = \frac{qp^N}{(1-q)(1-p)^N} > \lambda_{DM} > \frac{q\sum_{s=N-1}^{s=N}\binom{N}{s}p^s(1-p)^{N-s}}{(1-q)\sum_{s=N-1}^{s=N}\binom{N}{s}(1-p)^s p^{N-s}} \equiv \bar{x}. \tag{46}$$

Hence, $Z \equiv N-Y+1 = 1$. Given $\lambda_A < x_{N-2}$, then $r < N-3$ and hence $\bar{y} < N-1$. Therefore,

$max\{\bar{y}, Y\} = Y$. Hence, the advisor test 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 $x_{-1} < \lambda_A$ or $m = (0, 0)$ if $\lambda_A \leq x_{-1}$ and the DM rejects. Suppose the complete list of Nature's N draws is fixed. Under hidden testing, the DM accepts if and only if all N test outcomes are positive. Under observable testing, the DM accepts if and only if the first $N-1$ in N test outcomes are positive. Therefore, she accepts under observable testing even if the first $N-1$ outcomes are positive and the final outcome is negative. Ideally, the DM optimally accepts if all N outcomes are positive since $\lambda_{DM} < x_N$, but not if $N-1$ outcomes in N tests are positive since $x_{(N-1)-1} = x_{N-2} < \lambda_{DM}$. Therefore, the DM is strictly better off under hidden than observable testing. Recall the definition of the DM's first-best acceptance set as the set of complete lists of Nature's draws for which the DM optimally accepts from Section A.2. The above argument shows that the DM achieves her first-best acceptance set under hidden testing.

Part 1b. Define

$$W_1 \equiv \left\{(\lambda_A, \lambda_{DM}) : \lambda_A < x_{N-4}, \bar{\bar{x}} < \lambda_{DM} \leq \bar{x}\right\}, \tag{47}$$

where $\bar{x}$ is defined by (45) and

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

i.e. $\bar{\bar{x}}$ is the likelihood ratio conditional on at least $N-2$ outcomes in N tests being positive.

Suppose $(\lambda_A, \lambda_{DM}) \in W_1$. Under observable testing, the proof of Part 1 applies since $x_{N-2} < \bar{\bar{x}}$ and $\bar{x} < x_{N-1}$, hence, $l = N-2$ and $\lambda_A < \underline{\lambda}_A = x_{N-2}$. Under hidden testing,

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

since

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

Hence, $Z \equiv N - Y + 1 = 2$. Since $\lambda_A \leq x_{N-4}$, then $r < N-5$ and hence $\bar{y} < N-2$. Therefore, $max\{\bar{y}, Y\} = Y$. Hence, the advisor test 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 is negative, the advisor reports $m = (0, 2)$ if $x_{-2} < \lambda_A$ or $m = (0, 0)$ if $\lambda_A \leq x_{-2}$ and the DM rejects.

Suppose that the complete list of Nature's N draws is fixed. If exactly $N-1$ in N tests are positive and the single negative outcome is drawn in some period before the final period, then the DM accepts under hidden but rejects under observable testing, otherwise the DM chooses the same action under observable and hidden testing. If only $N-1$ outcomes in N tests are positive the DM prefers to reject since $x_{(N-1)-1} = x_{N-2} < \lambda_{DM}$. Hence, the DM must be strictly better off under hidden testing.

Part 2a. Define

$$B_2 \equiv \{(\lambda_A, \lambda_{DM}) : \lambda_A > x_k, x_0 < \lambda_{DM} \leq x_N\}, \tag{48}$$

where k is given by (39) in the proof of Lemma 5. Suppose $(\lambda_A, \lambda_{DM}) \in B_2$. Under observable testing, Part 2 of Lemma 5 shows that the advisor does not test and the DM always rejects. Under hidden testing, apply Part 1 of Lemma 6. Given $\lambda_A > x_k$, then $r \geq k$ and $\bar{y} \geq \frac{k+1+N}{2}$ and $\bar{z} \leq \frac{N-k}{2}$. In period N, either $y = \bar{y}$ and the advisor reports $m = (\bar{y}, 0)$ and the DM accepts, or $z = \bar{z}$ and the advisor reports $m = (0, \bar{z})$ and the DM rejects.

As shown in the proof of Part 1 of Lemma 6, given any complete list of Nature's draws, under hidden testing the DM accepts if and only if the advisor prefers to accept. Given that the advisor is more reluctant, if the advisor prefers to accept it is optimal for the DM to accept. Hence, the DM must be strictly better off under hidden testing than under observable testing.

Part 2b. Define

$$W_2 \equiv \{(\lambda_A, \lambda_{DM}) : x_{N-2} < \lambda_A \leq x_{N-1}, x_{N-3} < \lambda_{DM} \leq x_{N-2}\}. \tag{49}$$

Suppose $(\lambda_A, \lambda_{DM}) \in W_2$. Under observable testing, I will show that the advisor will test in any period $n+1$ if $x_{n-j} < x(n) < x_{N-1}$. Given $x(n) = x_{n-2}$, the advisor prefers to reject irrespective of the remaining outcomes, because even if the remaining $N-n$ outcomes are positive, $x(N) = x_{n-2+(N-n)} = x_{N-2} < \lambda_A$. The DM will reject if $x(N) \leq x_{N-3}$. Therefore, if $x(n) = x_{n-2} \leq x_{N-3}$, there is no reason to test. Given $x(n) = x_{N-1}$, 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 $\lambda_{DM} \leq x_{N-2}$. Therefore, there is no reason to test if $x(n) = x_{N-1}$. Given $x(n) \leq x_{N-2}$ and $x(n) > x_{n-2}$ or if $x(n) \leq x_{N-2}$ and $x(n) = x_{n-2} \geq x_{n-2}$, there is a reason to test. If the advisor continues testing, the remaining outcomes either lead to $x(N) = x_{N-1}$ and the DM accepts, or they lead to $x(N) = x_{N-3}$ and the DM rejects. In both cases the DM's action is in line with the advisor's interest. If the advisor stopped at $x(n) = x_{N-2}$ the DM would accept whereas the advisor would prefer to reject. If the advisor stopped at $x_{n-2} < x(n) < x_{N-2}$ the remaining outcomes would still be pivotal to the advisor's preferred action and, therefore, he would forgo the chance to learn more.

Under hidden testing, apply Part 1 of Lemma 6. Given $r = N-1$, $\bar{y} = N$ and $\bar{z} = 1$. In period N, either $y = N$ and the advisor reports $m = (N, 0)$ and the DM accepts, or $z = 1$ and the advisor reports $m = (0, 1)$ and the DM rejects.

Suppose that the complete list of Nature's N draws is fixed. Under observable testing, the DM accepts if and only if the first $N-1$ tests were positive. If only $N-2$ tests in the first $N-1$ periods were positive then $x(N-2) = x_{n-2} \leq x_{N-3}$ since $x(N-2) = x_{(N-3)-1}$ and $x_{n-2} = x_{N-2-2}$, hence the DM rejects. Under hidden testing, the DM accepts if and only if all N tests were positive. The DM is strictly better off under observable testing since she optimally accepts if $N-1$ in N tests are positive given $\lambda_{DM} \leq x_{N-2}$.

A.11 Proposition 5 [First-Best Benchmark]

This follows immediately from the proof of Part 1a of Proposition 3.

A.12 Proposition 4 [Preference Alignment]

Claim 1: *If $\lambda_{DM} - \lambda_A \geq \bar{x}$ where $\bar{x}$ is defined by (45) the DM is weakly better off in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing.*

Proof: I will show that the DM is weakly better off in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing for $\lambda_{DM} \geq \bar{x}$ and $\lambda_{DM} > \lambda_A$. Note that $\lambda_{DM} > x_0$ by the assumption that the DM optimally rejects at the prior and, hence, Claim 1 follows. If $\lambda_A < x_{N-2}$ and $\bar{x} < \lambda_{DM} \leq x_{N-1}$, then the DM is strictly better off under hidden testing by the proof of Part 1 of Proposition 3. If $x_{N-1} < \lambda_{DM} \leq x_N$ and $\lambda_{DM} > \lambda_A$ then it is an advisor-preferred equilibrium for the advisor to run all N tests and for the DM to accept if and only if all N tests are positive, irrespective of whether testing is hidden or observable. In addition, $x_{N-2} < \lambda_A < \lambda_{DM} \leq x_{N-1}$ then there is no conflict of interest between the DM and the advisor and, therefore, the DM achieves the same expected payoff whether testing is hidden or observable. If $\lambda_A > x_N$, the advisor prefers to reject independent of the outcome of N tests and, therefore, it is an advisor-preferred equilibrium for the advisor not to test and for the DM to reject irrespective of whether testing is hidden or observable. If $\lambda_A \leq x_N$ and $\lambda_{DM} > x_N$ it is always part of the equilibrium that the DM rejects independent of the advisor's report and, hence, the DM has the same expected payoff whether testing is hidden or observable.

Claim 2: *If $\lambda_A - \lambda_{DM} > inf\left\{j | j \in \mathbb{Z}_{>0}, j \geq \frac{N+3}{2}\right\}$ the DM is weakly better off in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing.*

Proof: First, assume $\lambda_{DM} \leq x_N$. If $\lambda_A > x_k$ where k is defined by (39) then by the proof of of Part 3 of proposition 3, the DM is strictly better off under in any Pareto-undominated equilibrium under hidden testing than in the unique equilibrium under observable testing. k is a function of λ_{DM}. $x_k - \lambda_{DM}$ is maximized if $x_0 < \lambda_{DM} \leq x_1$ i.e. if $l = 0$, since $k = inf\left\{j | j \in \mathbb{Z}_{>0}, j \geq \frac{N+3}{2}\right\} > 0$ given

$l=0$ and k increases less than one for one with l. Hence, if $\lambda_A - \lambda_{DM} > inf\left\{j|j \in \mathbb{Z}_{>0}, j \geq \frac{N+3}{2}\right\}$ then it must be that $\lambda_A > x_k$. Second, if $\lambda_A > \lambda_{DM} > x_N$ it is always part of the equilibrium that the DM rejects independent of the advisor's report and, hence, the DM has the same expected payoff whether testing is hidden or observable.
To prove Proposition 4 choose $d = \max\left\{\bar{x}, inf\left\{j|j \in \mathbb{Z}_{>0}, j \geq \frac{N+3}{2}\right\}\right\}$.

A.13 Proposition 6 [Advisor Payoff Comparison]

Suppose $(\lambda_A, \lambda_{DM}) \in B_1$ where B_1 is defined by (44). Suppose that the complete list of Nature's N draws is fixed. As shown by the proof of Part 1 of Proposition 3, under hidden testing, the DM accepts if and only if all N test outcomes are positive, and under observable testing, the DM accepts if and only if the first $N-1$ in N test outcomes are positive. Since the advisor prefers to accept if exactly $N-1$ in N outcomes are positive given $\lambda_A < x_{(N-1)-1}$, the advisor is strictly better off under observable testing.

A.14 Proposition 7 [Delegation]

Part 1. Suppose the advisor is more enthusiastic than the DM and fix the complete list of Nature's N draws to be ϕ. Recall the definition of player i's first-best acceptance set Φ_i^{FB} as the set of complete lists of Nature's draws for which player i optimally accepts from Section A.2, where $i \in \{A, DM\}$. When the advisor has decision-making authority, then rejection is chosen in equilibrium if and only if $\phi \notin \Phi_A^{FB}$ by payoff maximization.
Suppose the advisor is more enthusiastic. When the advisor does not have decision making authority and testing is hidden, I will show that rejection is always chosen if $\phi \notin \Phi_A^{FB}$ and acceptance is always chosen if $\phi \in \Phi_{DM}^{FB}$. Since the advisor is more enthusiastic, $\phi \notin \Phi_A^{FB} \Rightarrow \phi \notin \Phi_{DM}^{FB}$. Hence, the DM rejects in any Pareto-undominated equilibrium under hidden testing, because if this were not the case then the advisor has a profitable deviation to test more and to report all outcomes. In addition, $\phi \in \Phi_{DM}^{FB} \Rightarrow \phi \in \Phi_A^{FB}$. Hence, DM accepts in any Pareto-undominated equilibrium under hidden testing, because if this were not the case then the advisor has a profitable deviation to test more and to report all outcomes. From this it follows that the DM can never be strictly better off when he delegates decision-making authority to the advisor than in the advisor-preferred equilibrium under hidden testing.
Suppose the advisor is more reluctant. By the proof of Lemma 9 for $N=2$ and by the proof of Lemma 6 for $N>2$, in the advisor-preferred equilibrium under hidden testing, for any ϕ a more reluctant advisor accepts if and only if $\phi \in \Phi_A^{FB}$. It follows that the DM can never be strictly better off when he delegates decision-making authority to the advisor than in the advisor-preferred equilibrium under hidden testing.

Part 2. Suppose $(\lambda_A, \lambda_{DM}) \in B_2$ where B_2 is defined by (48). As shown in the proof of Part 1 above, in the advisor-preferred equilibrium under hidden testing the DM achieves the same expected payoff as under delegation. Then the statement follows by the proof of Part 2a of Proposition 3.
Part 3. Suppose $(\lambda_A, \lambda_{DM}) \in W_2$ where W_2 is defined by (49). As shown in the proof of Part 1 above, in the advisor-preferred equilibrium under hidden testing the DM achieves the same expected payoff as under delegation. Then the statement follows by the proof of Part 2b of Proposition 3.

A.15 Proof of Lemma 7 [DM Commitment]

Suppose the advisor is more enthusiastic (reluctant). Under observable testing, the DM optimally commits to accept (reject) if and only if the advisor has conducted N tests and $\lambda_{DM} \leq x(N)$ ($\lambda_{DM} > x(N)$). Under hidden testing, the DM optimally commits to accept (reject) if and only if the advisor discloses N tests and reports $m = (y, N-y)$ where $\lambda_{DM} \leq x_{y-(N-y)}$ ($\lambda_{DM} > x_{y-(N-y)}$). Under either regime, there is no reason for the advisor to stop testing in any period $n < N$ and under hidden testing there is no reason to hide any outcomes, because if the advisor stops or hides outcomes then the DM rejects (accepts). If the advisor continues and reports all outcomes, then the DM accepts (rejects) if and only if $\lambda_{DM} \leq x(N)$ ($\lambda_{DM} > x(N)$). Given $\lambda_{DM} \leq x(N)$ ($\lambda_{DM} > x(N)$) then also $\lambda_A \leq x(N)$ ($\lambda_A > x(N)$) since $\lambda_A < \lambda_{DM}$ ($\lambda_A > \lambda_{DM}$). Hence, if additional tests are pivotal to the DM's choice then this choice is in line with the advisor's interest.

A.16 Corollary 2 [Commitment vs. Hidden Testing]

Recall the definition of the DM's first-best acceptance set as the set of complete lists of Nature's draws for which the DM optimally accepts from Section A.2. By Proposition 7 the DM achieves her first-best acceptance set when she has commitment power.
Part 1. If $N \geq 2$ then by the proof of Part 1a of Proposition 3, the DM achieves her first-best acceptance set under hidden testing, but not under observable testing if $(\lambda_A, \lambda_{DM}) \in B_1$.
Part 2. Apply the proof of Part 2a of Proposition 3. If $(\lambda_A, \lambda_{DM}) \in B_2$, the DM is strictly better off in the advisor-preferred equilibrium under hidden testing than in the unique equilibrium under observable testing. In addition, in the advisor-preferred equilibrium under hidden testing the DM accepts if and only if the advisor prefers to accept given any realization of the N draws by Nature. $\pi_{DM}(NC, HT) < \pi_{DM}(C)$ arises if the DM's first-best acceptance set is strictly larger than the advisor's first-best acceptance set.

A.17 Proposition 8 [Advisor Commitment]

Under observable testing, the advisor does not strictly benefit from being able to commit to a testing strategy. The DM chooses an action after the advisor has completed testing and the advisor has no private information.

Part 1. Suppose the advisor is more enthusiastic than the DM, i.e. $\lambda_{DM} > \lambda_A$, and the advisor has no commitment power. Fix the complete list of Nature's N draws to be ϕ. As shown by the proof of Proposition 7, the DM accepts if $\phi \in \Phi_A^{FB}$, where Φ_A^{FB} is the first-best acceptance set as defined in Section A.2. In addition, the proof of Proposition 7 shows that the DM chooses the advisor's preferred action if $\phi \in \Phi_{DM}^{FB}$ or $\phi \notin \Phi_A^{FB}$. Therefore, the advisor would gain from commitment power if it led the DM to accept given ϕ where $\phi \in \Phi_A^{FB}$ and $\phi \notin \Phi_{DM}^{FB}$. Hence, the DM can never be strictly better off if the advisor has commitment power.

Part 2. If the advisor is more reluctant than the DM, i.e. $\lambda_{DM} < \lambda_A$, then he achieves his first-best expected payoff in the advisor-preferred equilibrium under hidden testing, as shown for $N > 2$ in the proof of Lemma 6. Hence, if he could commit, then he would commit to the testing and reporting strategy chosen in the advisor-preferred equilibrium under hidden testing without commitment. Therefore, if the advisor could commit then the DM would achieve the same expected payoff as in the advisor-preferred equilibrium under hidden testing without commitment. For $N > 2$, Proposition 3 shows that if $(\lambda_A, \lambda_{DM}) \in B_2$ the DM is strictly better off in the advisor-preferred equilibrium under hidden testing than in the unique equilibrium under observable testing.

B Online Appendix

B.1 Proof of Proposition 9 [Simultaneous Testing]

Let x_j denote the posterior likelihood ratio conditional on observing j more positive than negative outcomes. Let $x(n)$ denote the posterior likelihood ratio after n tests. Let y denote the total number of positive outcomes.

Part 1. Suppose testing is observable, and $x_l < \lambda_{DM} \leq x_{l+1}$ and $x_r < \lambda_A \leq x_{r+1}$ where l and r are defined by (37) and (38), and $r+1 \leq l$, i.e. if all tests were observable the advisor would accept for a strictly larger set of outcomes. I will show that if l is even (odd) the advisor chooses the largest odd (even) number of tests n that satisfies $n \leq N$. Hence, given any even (odd) l, if N is odd (even) the advisor chooses $n = N$ and, therefore, the DM can never be strictly better off under hidden testing for any value of r.

As a first step, I will show that if l is even (odd) then the advisor never chooses an even (odd) number of tests. To see why, suppose the advisor has already committed to run n tests, where $n \leq N-1$. An additional test only affects the advisor's expected loss 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 in n tests there were l more positive than negative outcomes, i.e. $x(n) = x_l$. 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 in n tests there were $l+1$ more positive than negative outcomes, i.e. $x(n) = x_{l+1}$. Since the advisor prefers accept if there are l more positive than negative outcomes, the advisor is strictly better off with an additional test in the first situation, but strictly worse off in the second situation. If l 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 l is even (odd) and n is odd (even). As a second step, I will show that if the advisor benefits from running an additional two tests at some n then he will benefit 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 in n tests there were $l-1$ more positive than negative outcomes or a total of $\frac{n+(l-1)}{2}$ positive outcomes, i.e. $x(n) = x_{l-1}$. 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 in n tests there were $l+1$ more positive than negative outcomes or a total of $\frac{n+l+1}{2}$ positive outcomes, i.e. $x(n) = x_{l+1}$. The advisor will be strictly better off by conducting these test if and only if

$$\lambda_A Pr(false)\left\{Pr\left(y = \frac{n+l+1}{2}|false\right)p^2 - Pr\left(y = \frac{n+(l-1)}{2}|false\right)(1-p)^2\right\}$$
$$+Pr(true)\left[Pr\left(y = \frac{n+(l-1)}{2}|true\right)p^2 - Pr\left(\frac{n+l+1}{2}|true\right)(1-p)^2\right] > 0 \quad (50)$$

where y denotes the total number of positive outcomes in n tests and

$$Pr(y = j|true) = \binom{n}{j}(1-p)^{n-j}p^j,$$

$$Pr(y = j|false) = \binom{n}{j}(1-p)^j p^{n-j}.$$

Substituting into (50):

$$\begin{aligned}
&\lambda_A (1-q)(1-p)^{\frac{n+l+1}{2}} (p)^{\frac{n-(l-1)}{2}} 2\frac{n!}{\left(\frac{n+l+1}{2}-1\right)!\left(\frac{n-l-1}{2}\right)!}\left[\frac{1}{(n+l+1)}p-\frac{1}{(n-l+1)}(1-p)\right] \\
&\qquad +q(p)^{\frac{n+l+1}{2}}(1-p)^{\frac{n-l+1}{2}} 2\frac{n!}{\left(\frac{n+l+1}{2}-1\right)!\left(\frac{n-l+1}{2}\right)!}\left[\frac{1}{n-l+1}p-\frac{1}{n+l+1}(1-p)\right]>0 \Leftrightarrow \\
&\qquad \lambda_A (1-q)(1-p)^l (1-p)^{\frac{n-l+1}{2}} (p)^{\frac{n-l+1}{2}}\left[\frac{1}{(n+l+1)}p-\frac{1}{(n-l+1)}(1-p)\right] \\
&\qquad +qp^l p^{\frac{n-l+1}{2}}(1-p)^{\frac{n-l+1}{2}}\left[\frac{1}{n-l+1}p-\frac{1}{n+l+1}(1-p)\right]>0 \Leftrightarrow \\
&\qquad \lambda_A (1-q)(1-p)^l\left[(n-l+1)p-(n+l+1)(1-p)\right] \\
&\qquad +qp^l\left[(n+l+1)p-(n-l+1)(1-p)\right]>0.
\end{aligned} \tag{51}$$

The LHS increases with n since $p>1-p$. Hence, if (51) 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 $l-1$ or fewer tests, because even if all $l-1$ tests were positive, the DM would still reject. However, he must strictly benefit from running $l+1$ 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 $n=l-1$ 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 testing is observable, and $x_0<\lambda_{DM}\le x_1$ and $x_{N-1}<\lambda_A\le x_N$, i.e. if all tests were observable the DM would accept if and only if the number of positive outcomes strictly exceeds the number of negative outcomes and the advisor would accept if and only if all N outcomes were positive.

I will show that under observable testing, it is optimal for the advisor not to conduct any tests, whereas under hidden testing the advisor will conduct all N tests and disclose outcomes if and only if all N outcomes are positive. Hence, the DM must be strictly better off under hidden testing.

First, suppose testing is observable. I will show that if l is even (odd) then 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.. Since the advisor prefers reject if there is an equal number of positive and negative outcomes, the advisor is strictly better off with an additional test in the situation in which there is one more positive than negative outcome after n tests, i.e. $x(n)=x_1$, but strictly worse off in the situation in which there is an equal number of positive and negative outcomes, i.e. $x(n)=x_0$. Given l even (odd), the first situation can only arise if n is odd (even) and the second 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.

Consider only even n. By an analogous argument as in Part 1.), if the advisor benefits from running an additional two tests at some n then he will benefit from running the largest even number of tests such that $n \leq N$. In particular, 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 in n tests there as many positive as negative outcomes, i.e. $x(n) = x_0$. 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 in n tests there were two more positive than negative outcomes, i.e. $x(n) = x_2$. The advisor will be strictly better off by conducting these test if and only if

$$\begin{aligned} &\lambda_A Pr(false)\left\{Pr\left(y=\frac{n+2}{2}|false\right)p^2 - Pr\left(y=\frac{n}{2}|false\right)(1-p)^2\right\} \\ &+Pr(true)\left[Pr\left(y=\frac{n}{2}|true\right)p^2 - Pr\left(\frac{n+2}{2}|true\right)(1-p)^2\right] > 0 \end{aligned} \tag{52}$$

$$\begin{aligned} &\lambda_A(1-q)(1-p)^{\frac{n+2}{2}}(p)^{\frac{n}{2}}2\frac{n!}{\left(\frac{n+2}{2}-1\right)!\left(\frac{n-2}{2}\right)!}\left[\frac{1}{(n+2)}p - \frac{1}{(n)}(1-p)\right] \\ &+q(p)^{\frac{n+2}{2}}(1-p)^{\frac{n}{2}}2\frac{n!}{\left(\frac{n+2}{2}-1\right)!\left(\frac{n}{2}\right)!}\left[\frac{1}{n}p - \frac{1}{n+2}(1-p)\right] > 0 \Leftrightarrow \\ &\lambda_A(1-q)(1-p)\left[np-(n+2)(1-p)\right] \\ &+qp\left[(n+2)p-(n)(1-p)\right] > 0. \end{aligned} \tag{53}$$

The LHS increases with n since $p > 1-p$. Hence, if (53) holds at some n then it must hold at any larger n. Therefore, if the advisor runs any tests at all then he will run all N tests.

Given the DM accepts if and only if there are more positive than negative outcomes, the advisor prefers no test to N tests if and only if

$$Pr(true) < Pr(false)\lambda_A Pr\left(y > \frac{N}{2}|false\right) + Pr(true)Pr\left(y \leq \frac{N}{2}|true\right)$$

$$\frac{qPr\left(y > \frac{N}{2}|true\right)}{(1-q)Pr\left(y > \frac{N}{2}|false\right)} < \lambda_A$$

This must hold since for $N > 2$, having exactly $N-1$ positives in N tests is a stronger indication of the hypothesis being true than having strictly more positive than negative outcomes:

$$\frac{qPr\left(y>\frac{N}{2}|true\right)}{(1-q)Pr\left(y>\frac{N}{2}|false\right)}<x_{N-1}\equiv\frac{qp^{N-1}}{(1-q)(1-p)^{N-1}},$$

and $x_{N-1}<\lambda_A$.

Finally, suppose testing is hidden. Then the advisor would commit to run all N tests, because even if testing is sequential and his strategy space is richer, there is no downside to choosing to run N tests. By the same argument as in Lemma 6, the advisor fully discloses outcomes if and only if he himself would accept, i.e. he fully discloses if and only if $x(N)=x_N$ since $x_{N-1}<\lambda_A\le x_N$.

B.2 Infinite Horizon with Testing Cost

The following modifications are made to the model in Section 3: There are infinitely many discrete periods. The DM's payoff function is unchanged, but the advisor's payoff function is modified to include a cost $c>0$ for each test run. Under hidden testing, in each period n, the advisor first privately chooses to test ($a=1$) or not ($a=0$). In addition, at the end of period n, he chooses to send a message ($r=m$) or not ($r=0$), where $m\in M_n$ and the message space is the unordered history at the end of period n, i.e. $M_n=\mathscr{P}\left(\widetilde{h}_n\right)$. After receiving a message, the DM chooses $\tau\in\{accept,reject\}$. The DM does not know in which period the message was sent.[40] A testing strategy for the advisor is $\sigma_A:H_n\longrightarrow\{0,1\}$ for $n=0,1,\ldots$. A reporting strategy for the advisor is $\sigma_M:H_n\longrightarrow 0\times M_n$ for $n=1,2,\ldots$. A strategy for the DM is $\sigma_A:M\longrightarrow\{accept,reject\}$, where $M=\cup_{n=1}^{\infty}M_n$. The solution concept is an advisor-preferred PBE. Under observable testing, in each period n, the advisor chooses publicly whether to test or not and then the DM chooses whether to wait ($\tau=0$), or to accept ($\tau=accept$) or to reject ($\tau=reject$). A strategy for the advisor is $\sigma_A:H_n\longrightarrow\{0,1\}$ for $n=0,1,\ldots$. A strategy for the DM is $\sigma_A:H_n\longrightarrow\{0,accept,reject\}$. The solution concept is a PBE.[41] Any unordered history $\widetilde{h}_n$ can be characterized by $\widetilde{h}_n=(x,y)$, where x denotes the number of positive outcomes and y denotes the number of negative outcomes at the end of period n. Denote number of excess positive outcomes by $k=x-y$ where $k\in\mathbb{Z}$. Denote any unordered history with k excess positive outcomes by h^k and denote the posterior likelihood ratio by $\frac{Pr\left(true|h^k\right)}{Pr\left(false|h^k\right)}\equiv x_k$.

Insurance Effect First, I show that the insurance effect does not exist. Under observable testing, the DM's optimal strategy is to wait if and only if the advisor tested in the current period.

[40] This assumption is made to insure that the DM does not know how many tests have been run.

[41] In this model, multiple equilibria exist under observable testing. To select a unique equilibrium I assume that the advisor tests in the current period if he is indifferent between testing in the current or in a later period and that the DM takes her final decision in the current period if she is indifferent between taking the final decision in the current or in a later period.

Otherwise, the DM optimally rejects if and only if there are l or fewer excess positive outcomes, where $l \equiv sup\left\{j \in \mathbb{Z} | x_j < \lambda_{DM}\right\}$, independent of the current period. Since the number of excess positive outcomes is sufficient to determine the DM's final decision and to form posterior beliefs about the state, the advisor's optimal testing strategy is a function only of the number of excess positive outcomes.

Under hidden testing, the advisor's first-best decision rule will be implemented in equilibrium. By the same reasoning as used in the proof of Lemma 6, if he reveals only positive outcomes at histories at which he wants to accept and only negative outcomes at histories at which he wants to reject, the DM always acts in his interest. This implies that if the advisor stops testing when there are r or fewer excess positive outcomes then the DM rejects and otherwise she accepts, where $r \equiv sup\left\{j \in \mathbb{Z} | x_j < \lambda_A\right\}$. Hence, the advisor's optimal testing strategy is again a function only of the number of excess positive outcomes.

In order for the DM to be better off under hidden than observable testing, there must be some realizations of Nature's draws such that acceptance is chosen in equilibrium. A necessary condition for acceptance to be chosen in equilibrium is that the advisor tests at any history where the number of excess positive outcomes k satisfies $0 \le k \le r$, i.e. for any k satisfying $0 \le k \le r$ the advisor's continuation value of testing exceeds his continuation value of stopping. Suppose this were the case. Under observable testing, his continuation value of stopping at k excess positive outcomes is the same as under hidden testing for $k \le l$ or $k > r$, and it is lower for $l < k \le r$. Therefore, the advisor must also prefer testing to not testing at k excess positive outcomes satisfying $0 \le k \le r$ when testing is observable. Hence, the insurance effect cannot exist.

Skepticism Effect Define two thresholds of c:

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

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

I will show that for any q, there exists a p such that $\overline{c} > \underline{c}$. In addition, if cost c satisfies $\overline{c} > c > \underline{c}$ then I will show that there is a combination of preference parameters λ_A and λ_{DM} such that the skepticism effect exists.

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 $\overline{c} = \frac{q(1-p)p}{q(1-p)+(1-q)p}$, where $\overline{c} > \underline{c}$ since both $\frac{\partial \overline{c}}{\partial p} < 0$ 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.

Suppose p is sufficiently large, the cost satisfies $\overline{c} > c > \underline{c}$, and $\lambda_A = 0$ and

$$\overline{x} \equiv \frac{q\,(p + p\,(1-p))}{(1-q)\,((1-p) + (1-p)\,p)} < \lambda_{DM} < \frac{q\left(p^2 + p^2\,(1-p) + p^2\,(1-p)^2\right)}{(1-q)\left((1-p)^2 + (1-p)^2\,p + p^2\,(1-p)^2\right)} \equiv \widehat{x}.$$

Under observable testing, the DM's optimal strategy is to wait if and only if the advisor tested in the current period. Otherwise, the DM optimally accepts if and only if at least one excess positive outcome was found, independent of the current period. Since the number of excess positive outcomes is sufficient to determine the DM's final decision and to form posterior beliefs about the state, the advisor's optimal testing strategy is a function only of the number of excess positive outcomes. The advisor optimally stops testing when he has found one or more excess positive outcomes, since the DM accepts and this is the advisor's preferred action independent of the state. In addition, the advisor optimally stops testing when he has found one or more excess negative outcomes. Given his posterior beliefs after finding one excess negative outcome, i.e. $x = x_{-1}$, the expected cost of testing outweighs the expected benefit. In particular, even if all future test outcomes were favorable for the advisor then two tests are needed to convince the DM to accept, but the advisor is not willing to pay for two tests since

$$\begin{aligned} -Pr\,(true|x_{-1})\left(1-p^2\right) - 2c &< -Pr\,(true|x_{-1}) \\ Pr\,(true|x_{-1})\,p^2 &< 2c \\ \frac{q\,(1-p)\,p^2}{q\,(1-p) + (1-q)\,p} &< 2c, \end{aligned} \tag{54}$$

which is implied by $c > \underline{c}$. Finally, given that the advisor optimally stops after either one excess positive or one excess negative outcome, he optimally tests when he has as many positive as negative outcomes since

$$\begin{aligned} -Pr\,(true|x_0)\,(1-p) - c &> -Pr\,(true|x_0) \\ Pr\,(true|x_0)\,p &> c \\ qp &> c, \end{aligned}$$

which is also implied by $c > \underline{c}$. Next, suppose testing is hidden. I will show that it cannot be part of an equilibrium for the DM to accept if and only if at least one positive outcome was reported. To see why, suppose the DM were to accept if and only if at least one positive outcome was reported. The advisor's optimal reporting strategy is to hide any negative outcomes and to report any positive outcomes whenever he has stopped testing. The advisor optimally stops testing when he has found at least one positive outcome (independent of how many negative outcomes he has found), because then the DM acts in the advisor's interest. In addition, the advisor optimally stops testing when he has found two negative outcomes, i.e. $x = x_{-2}$, because his expected cost of testing outweighs his expected benefit. In particular, even if all future test outcomes were favorable for the advisor then one test is needed to convince the DM, but the advisor is not willing to pay for even one test since

$$\begin{aligned} -Pr(true|x_{-2})(1-p) - c &< -Pr(true|x_{-2}) \\ Pr(true|x_{-2})\, p &< c \\ \frac{q(1-p)^2 p}{q(1-p)^2 + (1-q)p^2} &< c, \end{aligned} \tag{55}$$

which is also implied by $c > \underline{c}$. By contrast, the advisor optimally continues testing when he has found one negative outcome since

$$\begin{aligned} -Pr(true|x_{-1})(1-p) - c &> -Pr(true|x_{-1}) \\ Pr(true|x_{-1})\, p &> c \\ \frac{q(1-p)p}{q(1-p) + (1-q)p} &> c, \end{aligned} \tag{56}$$

which is implied by $\overline{c} > c$. If he continues testing after one negative, then he also optimally starts testing since he only requires one additional positive outcome to convince the DM and his posterior belief that the state is true is higher than in (56). The advisor's optimal testing and reporting strategies imply that the DM observes a report of one positive outcome either if the first outcome was positive, or if the first outcome was negative and the second positive. The DM optimally rejects such a report given $\overline{x} < \lambda_{DM}$. Next, I will show that there exists an equilibrium in which the DM accepts if and only if at least two positive outcomes were reported. Suppose the DM followed this strategy. The advisor then optimally hides any negative outcomes and reports any positive outcomes whenever he stops testing. In addition, the advisor optimally stops testing when he has found at least two positive outcome, because then the DM acts in the advisor's interest. Further, the advisor optimally stops testing if he has found one negative and no positive outcomes. This is because even if all future test outcomes were favorable for the advisor then two tests are needed to convince the DM, but the advisor is not willing to pay for two tests since (54) holds. In addition, he optimally stops if he has found three negatives and one positive outcome. This is because even if all future test outcomes were favorable for the advisor then one test is needed to convince the DM, but the advisor is not willing to pay for one test since (55) holds. However, it is optimally for him to start testing since

$$
\begin{aligned}
&-q\left(1-p+p(1-p)^3\right)\\
&-(q(1-p)+(1-q)(p))c\\
&-\left(qp^2+(1-q)(1-p)^2\right)2c\\
&-\left(q(1-p)p^2+(1-q)(1-p)^2p\right)3c\\
&-\left(q(1-p)^2p+(1-q)(1-p)p^2\right)4c > -q\\
&\Leftrightarrow\\
&\frac{3p^2q-3p^3q+p^4q}{2-p^3-q+3pq-3p^2q+2p^3q} > c \qquad (57)
\end{aligned}
$$

which is implied by $\overline{c} > c$. Given (57), it must also be optimal to continue testing if he has one positive and one negative or if he has one positive and no negative. In either case, he only needs one additional positive outcome and his posterior belief of the state being true is at least as high as when he started testing. Finally, he will continue testing after two negative outcomes and one positive outcome since (56) holds. This implies that the DM accepts conditional on $h_2 = (+,+)$ or $h_3 = (+,-,+)$ or $h_4 = (+,-,-,+)$. In equilibrium, the DM's posterior likelihood ratio conditional on observing two positive outcomes is given by $\widehat{x}$ and since $\lambda_{DM} < \widehat{x}$, it is indeed optimal for the DM to accept. Hence, in the equilibrium under hidden testing the DM accepts at a higher likelihood ratio than in the equilibrium under observable testing since

$$x_1 \equiv \frac{qp}{(1-q)(1-p)} < \widehat{x}.$$

References

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

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.

A. J. Davis. Prosecutors Who Intentionally Break the Law . *Criminal Law Brief*, 1(1):16, 2006.

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. Sorensen. Strategic Sample Selection. CEPR Discussion Papers 12202, C.E.P.R. Discussion Papers, Aug 2017a.

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

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.

S. Epstein. *Impure Science: AIDS, Activism, and the Politics of Knowledge*. University of California Press, 1998.

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.

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

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.

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.

A. Pollack. 'Viagra for Women' is Backed by an F.D.A. Panel. *The New York Times*, page A1, June 2015.

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