Lying and Lie-detection*

Terry Yin-Chi Tam †

please click for the latest version

Abstract

This paper analyzes strategic interactions between lying and lie-detection, and studies the optimal design for costly lie-detection and its effectiveness. An informed sender wants to persuade an uninformed receiver to take high actions but the receiver wants to match the action with the true state. The sender makes a claim about the true state and the receiver decides whether to incur a cost to inspect the truthfulness of the claim. I show that lie-detection technology is useful in improving the receiver's welfare if and only if the cost of lie-detection is sufficiently low and prior expectation of the state is not too high. The receiver-optimal design of admissible claims leads to an equilibrium with three intervals in the state space, where types in the top interval are induced to make precise and truthful claims about the state, which are mimicked by types in the bottom interval and randomly inspected, while types in the middle interval make a truthful but vague claim that is never inspected. Compared to state verification, lie-detection is shown to be more beneficial to the receiver because it provides incentives for moderate and high types to be truthful. This suggests that fact-checking of politicians' claims is effective in holding them countable and deterring them from lying.

Keywords: Lying, lie-detection, strategic communication

JEL Classification: D82

*I am grateful to Sergei Severinov, Li Hao, Vitor Farinha Luz, Michael Peters, Wei Li and seminar participants at UBC for helpful comments, discussions, and feedback.

†The Oliver Hart Research Center of Contracts and Governance, East China University of Science and Technology, Shanghai, China. Email: terrytamhk2012@gmail.com

1 Introduction

In many situations, an informed party is tempted to misrepresent information when communicating with an uninformed party. Politicians hope to impress voters by exaggerating their past achievements and future policy goals; Companies want to convince consumers to believe that their products are better than what they actually are; Suspects who have committed crimes often refuse to plead guilty. Such tension is widely studied in the literature with various resolutions. Contract theory assumes the uninformed receiver can commit to a menu of outcomes corresponding to the informed sender's report of his private information. The cheap talk literature assumes no such commitment but the sender and the receiver share a degree of common interest in how they would want to respond to the sender's private information. In both cases, lying by the sender about his private information does not occur in equilibrium.

This paper introduces a theory of lying in cheap talk communication where the receiver has access to a costly lie-detecting technology. The informed sender is opportunistic and the uninformed receiver is skeptical. Claims made by the sender are statements regarding the state. Some of them emerge in equilibrium as lies and others as truth [1]. Lie-detection is a broadly used technology in our daily lives. For instance, skeptical voters might go online to fact-check a politician's claim. The innovation in this paper is that lies emerge as equilibrium claims about the state in the face of costly lie-detection. To my knowledge, this is the first paper that develops a framework allowing strategic interaction between lying and lie-detecting, and analyzes tensions between the sender's incentive of lying and the receiver's incentive of inspection. This assumption is motivated by the observation that not all claims are treated equally by an uninformed party. Oftentimes some claims are more suspicious than the others and draw more attention in inspecting them. For instance, if a health product company claims their new drug is proven to reduce any cancer risk by 99 percent, potential consumers might be skeptical and search if such claim is backed by any trust-worthy, independent studies; but if the company makes a mild claim that the drug strengthens immune system, consumers might just take the company's word and do not bother checking it.

The main objective of this paper is to study the "optimal design" for costly lie-detection and its effectiveness in improving the quality of communication. The optimal design codifies the set of

[1] Sobel (2019) points out that lying depends on the existence of accepted meanings for messages. In this paper, a sender's message has the following accepted meaning: "The true state is within Θ", where Θ is some subset of the state space. Naturally, the sender is lying if the true state is not one of the states in Θ.

claims about the state, truthfully or falsely, that are made in the best equilibrium for the uninformed receiver in a cheap talk game with costly inspection. Several major questions regarding lie-detection are addressed:

(1) Under what conditions is lie-detection technology helpful in improving the welfare of the uninformed party?

(2) What are the best lie-detection outcomes for the uninformed party?

(3) How do changes in the technology, in particular, the information generated from an inspection, affect the welfare of the uninformed party?

To study these questions, I analyze a framework with endogenous lying and lie-detection. An informed sender prefers a higher action by an uninformed receiver to a lower action independent of the true state, while the receiver has a quadratic loss function and wants to match the action with the true state. The sender communicates with the receiver by making a claim modeled as a subset of possible states; a claim is truthful if it includes the true state and is a lie otherwise. Before taking the payoff relevant action, the receiver has the option to incur a cost to inspect the truthfulness of the claim. The list of claims made in an equilibrium that maximizes receiver's expected payoff is an optimal design for costly lie-detection.

Lie-detection technology improves the receiver's welfare only if lying occurs in equilibrium. If the sender never lies, the receiver has no incentive to inspect the sender's claims; If there is no inspection, babbling is the only equilibrium as the sender and the receiver share no common interest. I show that there exists equilibrium where lying and lie-detection occur if and only if inspection cost is sufficiently low and prior expectation of the state is not too high. The threshold for prior expectation increases as inspection cost decreases and converges to the upper bound of the state space as inspection cost goes to zero. Intuitively, this result comes from the conflict between the sender's incentive to lie and the receiver's incentive to inspect. Liars aim to convince the receiver that they are better than the average (prior expectation) when they get away with the lie. If the prior expectation is too high, this can happen only when a small number of liars mimic a large number of truth-tellers, but then the claim is not worth inspecting because the sender is too likely to be truthful. This result echoes a common perception that lie-detection is effective when the sender is suspicious, in the sense that there is a substantial difference between the receiver's prior belief and the belief preferred by the sender. For example, the police usually conduct an interrogation only if they believe that the suspect is likely to have committed a crime. When the sender is likely to be "innocent", there is no cost-effective way to separate lies from truths.

Even though full revelation is achievable using lie-detection technology, it is neither sequentially rational nor ex-ante optimal for the receiver because some claims are not worth inspecting. The receiver's benefit from lie-detection can be decomposed into two components. First, lie-detection generates a direct information value by distinguishing liars from truth-tellers which generally provides information about the true state. Second, the sender might stay honest in fear of being caught lying. Therefore, the possibility of lie-detection creates a threat that deters potential liars and facilitates information transmission. This is called the indirect deterrent effect. I show that under the optimal lie-detection policy, the direct information value of inspection is completely offset by the cost of inspection. Improvement of the receiver's ex-ante payoff is driven by the deterrent effect: the receiver is able to elicit information from the sender due to the credible threats of lie-detection. This is perhaps surprising as one might expect an optimal lie-detection design should allow the receiver to acquire as much information as possible. Such intuition turns out to be incorrect because excess amount of information acquired from inspection indicates that the design induces sender to lie too often, and costly inspection takes place more frequently than necessary. This suggests that lie-detection technology better serves as a mean of deterrence than a mean of information acquisition.

The receiver's ex-ante payoff depends on the degree of information transmission facilitated by the indirect deterrence, which is affected by the set of claims available to the sender. Therefore, an optimal lie-detection design involves not only a contingent plan of inspection, but also a set of admissible claims. I show that the optimal design is characterized by three intervals which partition the state space. The sender makes truthful claims when the true state is in the high interval (good types), lies and mimics the high claims when the true state is in the low interval (bad types). These high claims are randomly inspected. In fear of being caught lying and perceived as bad types, the sender in the intermediate interval (moderate types) is deterred from mimicking the high claims. These moderate types pool at a vague yet truthful claim which is not inspected by the receiver. It is optimal for the receiver to give the sender an option of being vague because precise claims require inspections to sustain, while moderate types are not distant enough with each other to justify the cost of inspection. Technically speaking, it is always optimal to pool an interval of types to a single claim and leave it uninspected because the conditional variance of a small enough interval is smaller than the inspection cost.

If the density of the bottom half of the prior distribution is concentrated toward the center, then the optimal design corresponds to a so-called decreasing mimicking mechanism, where the inspected high claims are precise, meaning that the optimal set of admissible claims consists of a vague claim that indicates moderate types and a continuum of precise high claims. The mimicking is decreasing in the sense that liars with lower types mimic truth-tellers with higher types. Inspection probabilities of those claims are chosen so that liars are indifferent between all equilibrium lies. Since liars who make higher claims will be punished by worse posterior beliefs upon lie-detection, the inspection probability need not be increasing in the level of claim. In practice, remaining silent can be interpreted as the vague claim which induces a moderate belief from the receiver. Any claims in attempts to induce better beliefs are required to be precise and will be inspected stochastically.

I study the effect of inspection technology on the receiver's welfare by comparing lie-detection with state-verification. There are substantial differences between lie-detection and state-verification. Under state-verification technology, an inspection reveals the true state of the world. There will be no uncertainty upon inspection. Under lie-detection technology, an inspection returns a binary signal on the truthfulness of the sender's claim. Information learned from an inspection is endogenously determined by the claim made by the sender. Practically, state-verification is a hard skill which requires the receiver to be able to acquire knowledge about the true state, which might not be feasible in some situation. For example, there might not be any objective evidence in the crime scene that provides further information about whether a suspect has committed to the crime. On the other hand, lie-detection can be a soft skill. A competent detective might be able to spot a lie told by the suspect using various interrogation tactics. Studies in psychology and cognitive science have shown possibilities of detecting lies using methods such as asking questions that raise cognitive load (Vrij et al., 2011), measuring brain activities (Christ et al., 2008) and reading micro-expressions (Porter and Ten Brinke, 2006), with nearly 70 percent accuracy (Hartwig and Bond, 2014) and 85 percent accuracy for trained interviewers (Hartwig et al., 2006).

Even if state-verification is feasible, lie-detection technology can yield a higher benefit to the receiver. Assuming the same unit cost of the two technologies, I show that the receiver's welfare is higher under optimal lie-detection design compared with optimal state-verification design. This is because revealing the true state upon inspection removes any strategic uncertainty that can serve as a threat of punishment to potential deviators. Since state-verification leads to an accurate assessment of the true state, there is no credible punishment for the liar thus sender always has the incentive to exaggerate the state to "try his luck". As a result, the deterrence effect is eliminated under state-verification technology and there will not be any informative communication. This result sheds light on the optimal approaches of fact-checking as a tool to combat misinformation in politics. The internet has enabled the public to more easily verify politicians' claims using fact-checking websites such as *FactCheck.org* and *PolitiFact*. A question regarding the socially desired mission of these organizations is whether they should focus on presenting verdicts on politicians' statements (lie-detection) or educating the public about policy-related issues (state-verification). The latter is more informative as verdicts on politicians' statements can be derived from knowledge in policy-related issues. An argument for the former is that simple verdicts cost less time to read and are easier to comprehend, compared with the complex policy-related issues. Another argument for the former is that targeting politicians' statements hold them accountable and deter them from lying. Some studies find evidence that fact-checking reduces lying behaviors of politicians (e.g. Nyhan and Reifler (2015); Lim (2018)). This paper provides a theoretical ground for the deterrence argument and shows that the public can be better off under lie-detection in spite of the ignorance of details in policy-related issues.

The remainder of the paper is organized as follows. Section 2 reviews the related literature on lying and communication. Section 3 presents the model. Section 4 derives necessary and sufficient conditions for the existence of welfare-improving lie-detection mechanisms. Section 5 characterizes the optimal lie-detection mechanism. Section 6 compares lie-detection with state-verification. Section 7 concludes. The proofs are relegated to the Appendices.

2 Related literature

My work is mostly related to the literature of strategic communication with lie-detection. Balbuzanov (2017) analyzes a cheap talk model akin to the setup in Crawford and Sobel (1982), with the addition that a lie of the sender will be detected with an exogenous probability. He shows that given intermediate probability of lie-detection and sufficiently small bias, fully revealing equilibria exist. Dziuda and Salas (2018) study a pure persuasion game with the same lie-detection technology as Balbuzanov (2017) and show that certain refinement criteria lead to a unique equilibrium where moderate types and high types stay honest and low types lie to imitate high types. My findings in optimal mechanisms echo findings from Dziuda and Salas (2018) that moderate types do not exaggerate their types to avoid being mistaken as the low type liars. The key difference between this paper and previous literature is that this paper models lie-detection as a decision of the receiver, where the probability of lie-detection can be chosen conditional on the sender's claim. This allows an analysis of tensions between the sender's incentive of lying and the receiver's incentive of inspection. Under different sets of permissible claims, the resolutions of such tensions result in different degrees of inspection and information revelation, and hence different payoffs. Therefore, this leads to a non-trivial design problem on the optimal set of permissible claims. Jehiel (2019) analyzes an interesting multi-round cheap talk environment where lie can be spotted from the inconsistent messages of a forgetful liar who cannot remember the content of the lie he has told.

Another strand of literature study strategic communication with lying cost for the sender. Kartik, Ottaviani and Squintani (2007) study a strategic communication environment where the sender is upwardly biased and misrepresenting information is costly. They show that when the state space is unbound above, fully separating equilibrium exists in which sender lies and uses inflated language. Kartik (2009) studies a similar environment and shows that when state space is bound above, there is some pooling on the highest message and the degree of information revelation depends on the intensity of lying cost. While lying behaviors arise in equilibrium in both their models and my model, the natures and interpretations of lies are quite different. In my model, lies serve as disguises to confuse the receiver. Liars try to mimic the types they claim to be and the receiver cannot tell them apart without an inspection. In their models, lies serve as inflated languages. The sender tells lie to avoid being mistaken as a worse type and a strategic receiver does not confuse a liar with the type he claims to be. An alternatives interpretation of the models in Kartik, Ottaviani and Squintani (2007), alongside other related works (e.g. Ottaviani and Squintani (2006); Chen (2011)) is that a proportion of receivers naively believes sender's message. The coexistence of strategic and naive receivers imposes an endogenous cost for the sender to overly exaggerate the state since the naive receivers will take it at face value, which is not preferred by a sender whose bias is not too large. In the equilibria of their models, lies are chosen by the sender to balance the induced beliefs of two groups of receivers who interpret messages differently. In my model, lies are chosen to mimic the corresponding truthful senders and confuse the receiver.

For a broader discussion on the role of lying in strategic interactions, Sobel (2019) establishes a general framework of lying with various applications. My model adopts the same definition of lying as in Sobel. His framework does not incorporate the possibilities of lie-detection.

3 The Model

There are a decision-maker (DM) and a sender. DM has to make a decision, but only the sender has the relevant information. Sender privately observes the state of the world, θ, which is distributed according to a continuously differentiable c.d.f. F over the normalized state space $\Theta \equiv [0,1]$, with associated density f. θ is also referred to as the sender's type. For example, θ might represent the quality of the advertised product or the severity of crimes committed by a suspect.

Message: The sender sends a message $m \in \mathcal{M}$ to DM, where $\mathcal{M}$ is the set of all measurable subsets of the state space Θ. A message sent by the sender is interpreted as a statement regarding his type. To provide a few examples, a message $m = [0.3, 0.4]$ is interpreted as the following statement: "my type lies somewhere in between 0.3 and 0.4"; a message $m = \{0.5\} \cup \{0.7\}$ is interpreted as "my type is either 0.5 or 0.7"; a message $m = \Theta$ can be interpreted as to remain silent because it essentially means "Anything is possible".

Costly inspection: DM, after observing m, chooses whether or not to inspect the message with a cost $c > 0$. An inspection reveals the truthfulness of the statement. Formally, if an inspection takes place, DM will receive a binary signal

$$s(m,\theta) = \begin{cases} t & \text{if } \theta \in m \\ l & \text{otherwise} \end{cases} \tag{1}$$

If DM chooses not to inspect, she receives an uninformative signal $s(m,\theta) = u$. The signal t indicates the sender's claim is inspected and confirmed to be truthful; l indicates the sender's claim is inspected and confirmed to be a lie; u indicates the sender's claim is uninspected.

Action: After observing both the message m and the inspection signal s, DM chooses a payoff relevant action $x \in [0,1]$.

Preference: DM has a quadratic loss function $-(x-\theta)^2 - c\mathbf{1}_I$, where $\mathbf{1}_I = 1$ if an inspection took place, $\mathbf{1}_I = 0$ otherwise. The sender has a von Neumann-Morgenstern utility $u(x)$ which is strictly increasing in x. In other words, there is no common interest between DM and the sender. DM wants to take an action that matches the true state, while the sender always prefers a higher action, independent of the state.

The design problem: An mechanism (q, P, X) consists of a message rule $q : \Theta \to \Delta_{\mathcal{M}}$, where $q(.|\theta)$ is type θ's probability distribution over the message space $\mathcal{M}$; an inspection rule $P : \mathcal{M} \to [0,1]$, where $P(m)$ is the probability of inspecting message m; and an action rule $X : \mathcal{M} \times \{t, l, u\} \to [0,1]$, where $X(m,s)$ is the action taken following message m and inspection signal $s \in \{t, l, u\}$.

For expositional clarity, I confine attention to pure message rule in this paper, i.e. each type of sender θ sends a message $m_q(\theta)$ with probability 1. In an Online Appendix, I show that all results can be generalized to allow mixed message rules.

Given a pure message rule m_q, let $\mathcal{M}_q = m_q(\Theta)$ be the set of all on-path messages [2]. For any on-path message $m \in \mathcal{M}_q$, let

$$\Theta_q^t(m) = \{\theta \in \Theta : m_q(\theta) = m \text{ and } \theta \in m\} \tag{2}$$

$$\Theta_q^l(m) = \{\theta \in \Theta : m_q(\theta) = m \text{ and } \theta \notin m\} \tag{3}$$

$$\Theta_q^u(m) = \Theta_q^t(m) \cup \Theta_q^l(m) \tag{4}$$

be the sets of truthful senders, lying senders and senders of m. DM cannot commit to an inspection rule and/or an action rule. They have to be sequentially rational based on a Bayesian updated belief.

Sequentially rational action: Since DM's utility is quadratic, her optimal action equal conditional expectation of the sender's type given the posterior belief, so an action rule X is **sequentially rational** given q if for any $m \in \mathcal{M}_q$ and $s \in \{t, l, u\}$,

$$X(m, s) = E[\Theta_q^s(m)] \tag{5}$$

where $E[\Theta'] \equiv \frac{\int_{\Theta'} \theta dF(\theta)}{Pr(\Theta')}$ denotes the conditional expected type given a set of type $\Theta' \subseteq \Theta$, and $Pr(\Theta') \equiv \int_{\Theta'} dF(\theta)$ denotes the probability of Θ' [3].

After observing the on-path message m and inspection signal s, DM chooses an action to match the conditional expected type of senders who send m and lead to inspection signal s given the message rule q. Instead of blindly taking a message at its face value, a Bayesian, sequentially rational DM updates her belief given the set of equilibrium senders who would pass/fail an inspection, and reacts optimally. When there is no inspection, DM remains aware of the possibility of lying and chooses an action that matches the weighted average type of the equilibrium truth-tellers and liars.

Information value of inspection: Given a message rule q and a sequentially rational action rule X, DM's expected continuation payoff if she inspects an on-path message $m \in \mathcal{M}_q$ is:

$$-w_q(m)Var(\Theta_q^l(m)) - (1 - w_q(m))Var(\Theta_q^t(m))$$

where $w_q(m) = \frac{Pr(\Theta_q^l(m))}{Pr(\Theta_q^u(m))}$ is the conditional probability of sender being a liar given that he sends m, $Var(\Theta') \equiv \frac{\int_{\Theta'}(\theta - E[\Theta'])^2 d\theta}{Pr(\Theta')}$ denotes the conditional variance given Θ'[4]. Upon inspection, DM's

[2]Throughout this paper, I follow the convention and refer to $g(X)$ as $\{y : \exists x \in X \text{ such that } y \in g(x)\}$ for any function or correspondence g and set X within the domain of g.

[3]It is possible that $Pr(\Theta_q^s(m)) = 0$ even if the message m is on-path, if m is sent by a set of types with zero measure but positive density. Therefore, a more precise version of condition (5) is that for any subset of on-path messages $M \subseteq \mathcal{M}_q$, $\int_M \int_{\Theta_q^s(m)} X(m, s)dF(\theta)dm = \int_{\Theta_q^s(M)} \theta dF(\theta)$ for $s \in \{t, l, u\}$. This ensures that DM's action rule is sequentially rational given q almost surely.

[4]Similarly, a more precise condition for $w_q(m)$ is that for any $M \subseteq \mathcal{M}_q$, $\int_M w_q(m) \int_{\Theta_q^l(m)} dF(\theta)dm = Pr(\Theta_q^l(M))$.

expected loss from action imprecision for a message m is the weighted average conditional variance of equilibrium truth-tellers and liars of m. DM's expected continuation payoff if she does not inspect m is:

$$-Var(\Theta_q^u(m)$$

which is the variance of the sender's type conditional on him sending m. Therefore, the information value of inspecting m is the reduction in conditional variance from the binary signal:

$$V_q(m) = w_q(m)(1 - w_q(m))(E[\Theta_q^l(m)] - E[\Theta_q^t(m)])^2 \tag{6}$$

An inspection allows DM to make a better inference on the sender's type and chooses more precise action accordingly. If there is a large difference between the expected type of truth-tellers and liars who send m, the value of differentiating these two groups is high. Besides, an inspection is more informative when the liar to truth-teller ratio is less extreme. If the sender of m is very likely to be on one side, not much information is revealed from an inspection. An inspection rule P is **sequentially rational** given q if for any $m \in \mathcal{M}_q$,

$$P(m) \in \begin{cases} \{0\} & \text{if } c > V_q(m) \\ [0,1] & \text{if } c = V_q(m) \\ \{1\} & \text{if } c < V_q(m) \end{cases} \cdot \tag{7}$$

i.e. inspecting a message is credible only if information value of inspection is no less than cost of inspection.

Sender's optimality: Given inspection rule P and action rule X, type θ sender's expected utility from sending a message m is:

$$EU_{X,P}(m|\theta) = \begin{cases} P(m)u(X(m,t)) + (1 - P(m))u(X(m,u)) & \text{if } \theta \in m \\ P(m)u(X(m,l)) + (1 - P(m))u(X(m,u)) & \text{if } \theta \notin m \end{cases} \tag{8}$$

A pure message rule q is **optimal** given P and X if for any $\theta \in \Theta$ and $m' \in \mathcal{M}_q$ [5],

$$EU_{X,P}(m_q(\theta)|\theta) \geq EU_{X,P}(m'|\theta) \tag{9}$$

[5] Incentive constraints over off-path messages are omitted because sequential rationality put no restriction on the inspections and actions following those messages, so we can without loss of generality let $X(m', s) = 0$ for any off-path message m', and sender will have no incentive to deviate to those messages.

4 Incentive Compatible Mechanisms

This Section defines an incentive compatible mechanism and establishes the necessary and sufficient conditions for the existence of an incentive compatible mechanism where inspections take place with positive probability.

Definition 1 *An mechanism $\Omega \equiv (q, P, X)$ is incentive compatible if P and X are sequentially rational given q and q is optimal given P and X.*

Since the decision-maker has no commitment power, an incentive compatible mechanism requires that DM has no incentive to deviate after any history, so it corresponds to a Perfect Bayesian equilibrium in a game-theoretic approach. The two concepts are interchangeable under this framework. Given an incentive compatible mechanism Ω, DM's ex-ante expected payoff is:

$$EU_{DM}(\Omega) = -\int_{\mathcal{M}_q} \sum_{s=t,l} \int_{\Theta^s_(m)} [(1 - P(m))(X(m, u) - \theta)^2 + P(m)[(X(m, s) - \theta)^2 + c]]dF(\theta)dm \quad (10)$$

Define

$$G_\Omega(x) = \int_{\mathcal{M}_q} \sum_{s=t,l} \int_{\Theta^s_(m)} [(1 - P(m))\mathbf{1}(X(m, u) \leq x) + P(m)\mathbf{1}(X(m, s) \leq x)]dF(\theta)dm \quad (11)$$

be the distribution of induced actions under Ω, and

$$p_\Omega = \int_{\mathcal{M}_q} P(m) \int_{\Theta^u_q(m)} dF(\theta)dm \quad (12)$$

be the ex-ante probability that a sender is inspected under Ω. Sequential rationality of the action rule X implies that

$$EU_{DM}(\Omega) = \int_{[0,1]} x^2 dG_\Omega(x) - cp_\Omega - E[\theta^2] \quad (13)$$

where $E[\theta^2] \equiv \int_\Theta \theta^2 dF(\theta)$. Sender's ex-ante expected payoff is:

$$EU_S(\Omega) = \int_{[0,1]} u(x)dG_\Omega(x) \quad (14)$$

I refers to the pair (G_Ω, p_Ω) as the **induced outcome distribution** of an mechanism. I say two mechanisms Ω and Ω' are **distribution equivalent** if they have the same induced outcome distribution. Since (G_Ω, p_Ω) uniquely determine payoffs in an incentive compatible mechanism, two incentive compatible, distribution equivalent mechanisms induce the same expected payoffs for DM and every type of sender.

Since DM cannot commit to a sub-optimal action rule, the expected value of induced actions must equal the expected value of the state. In fact, the distribution of induced actions G is a mean-preserving contraction of F. A more dispersed G implies a more precise match between the induced actions and the states, and thus a higher expected payoff for DM.

Proposition 1 *For any incentive compatible mechanism* $\Omega = (q, P, X)$ *there exists a distribution equivalent mechanism* $(\hat{q}, \hat{P}, \hat{X})$ *such that for any* $m \in \mathcal{M}_{\hat{q}}$*:*
(i) $\hat{X}(m, t) \geq \hat{X}(m, l)$*, and*
(ii) $m = \Theta^t_{\hat{q}}(m)$*.*

Condition (i) of Proposition 1 provides a natural interpretation of the mechanism: liars pretend to be truth-tellers in the hope of inducing higher actions [6]. Condition (ii) comes from the fact that condensing the statement of a message to include only equilibrium truth-tellers is the most effective design in maintaining incentive compatibility. Under such design, any type who deviates from his equilibrium message to any other on-path message will be identified as a liar, which according to (i), gets a lower expected payoff than if he is identified as a truth-teller. Proposition 1 is useful in analyzing the set of implementable outcome because an outcome distribution is implementable if and only if it can be induced by an incentive compatible mechanism that satisfies the above properties [7]. Unless otherwise stated, any mechanism discussed henceforth satisfies conditions (i) - (ii) of Proposition 1.

Let $\mathcal{M}^0_q = \{m \in \mathcal{M}_q : P(m) = 0\}$ be the set of on-path uninspected messages. Sender's optimality requires that any messages in $\mathcal{M}^0_q$ must induce the same action, for otherwise senders who receive a lower uninspected action will deviate to a higher one. Therefore, we can without loss assume that there

[6] In a model where sender can make a truthful claim and tricks the lie-detector to identify him as a liar (for example, by acting nervous or intentionally failing a test), then condition (i) must holds in any incentive compatible mechanism for any inspected message m, for otherwise equilibrium truth-tellers who act normally and get $X(m, t)$ will deviate to act nervously and get $X(m, l)$.

[7] Note however that oftentimes an implementable outcome distribution can also be induced by other incentive compatible mechanisms. For example, if there exists an on-path message m' which is never inspected, and Θ' is the set of senders of m', then an incentive compatible mechanism that satisfies (ii) requires the statement m' to be a subset of Θ'. However, the mechanism will still be incentive compatible if senders of m' simply " remain silent ", i.e. $m' = \Theta$. By definition, it means every type becomes truth-teller of m', but it has no effect on the sender's incentive because being truthful and lying makes no difference to the outcome when m' is never inspected.

is at most one such message, m_q^0, and all senders of that message are truthful, i.e. $m_q^0 = \Theta_q^u(m) = \Theta_q^t(m) \equiv \Theta_q^0$, where Θ_q^0 is the set of types who are never inspected in equilibrium.Sequential rationality of X requires $X(m_q^0, u) = E[\Theta_q^0]$. Let $\mathcal{M}_q^+ = \{m \in \mathcal{M}_q : P(m) > 0\}$ be the set of messages that are inspected with positive probability. $\mathcal{M}_q^+$ is simply referred to as the **set of inspected messages**. For $\theta \in \Theta$, I say θ is **truthful** if $\theta \in \Theta_q^t(\mathcal{M}_q^+)$; θ is **lying** if $\theta \in \Theta_q^l(\mathcal{M}_q^+)$; θ is **uninspected** if $\theta \in \Theta_q^0$.

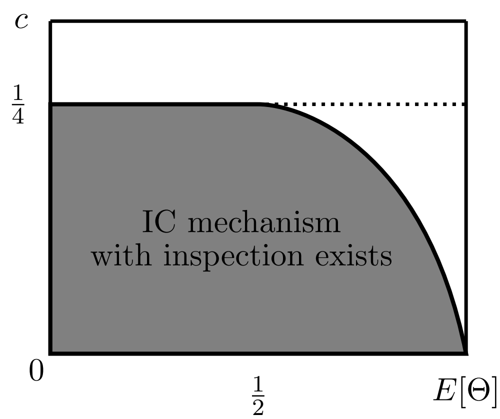
I say Ω is a **mechanism with inspection** if $p_\Omega > 0$, i.e, some on-path messages are inspected with positive probability. The following assumption and proposition establish the necessary and sufficient conditions for the existence of incentive compatible mechanism with inspection.

Assumption 1 *$c < \frac{1}{4}$ and $E[\Theta] \equiv \int_0^1 \theta dF(\theta) < \frac{1}{2} + \sqrt{\frac{1}{4} - c}$.*

Proposition 2 *There exists an incentive compatible mechanism with inspection if and only if Assumption 1 is satisfied.*

The credibility of inspections relies on the existence of both liars and truth-tellers. Upon receiving a message, if DM's interim expectation on the sender's type is extreme (either too high or too low), the information value of an inspection is low because the sender is either very likely to be truth-telling or very likely to be lying, and inspection is non-credible. Now consider an uninspected message m and a randomly inspected message m'. In order to incentivize the liars who send m' to take the risk of being caught, DM's interim expectation on sender's type upon receiving m' must be higher than the interim expectation upon receiving m, so that if liars of m' get away with the lie, they receive a higher payoff than those who send m. Since DM is Bayesian, her interim expectation upon receiving some inspected messages must be higher than the prior expectation, so if the prior expectation is too optimistic, the interim expectation of those messages will be too optimistic for inspection to be credible. It is worth noting that the condition is not symmetric. Even if prior expectation on the sender's type is pessimistic, it is possible to design a mechanism with pessimistic interim belief for the uninspected message and moderate interim beliefs for the inspected messages so that liars of the inspected message are incentivized and inspections are credible. Therefore, the lie-detection technology is useful when the prior expectation is moderate or pessimistic, but not when the prior expectation is optimistic. Figure 1 depicts the region of parameter values in which an incentive compatible mechanism with a positive probability of inspection exists. The threshold of prior expectation such that inspection is incentive compatible is decreasing in cost of inspection, meaning that when the cost is smaller, inspection is incentive compatible for a larger range of optimistic beliefs. When the cost of inspection is small,

Figure 1: Existence of incentive compatible mechanism with inspection



Note: Horizontal axis depicts the expectation on type under prior distribution F; Vertical axis depicts the cost of inspection.

inspection is credible even if conditional expectations given the inspected statements are optimistic and information values of inspection are small. Inspection can therefore facilitate information transmission. As cost goes to 0, the lie-detection technology is useful for almost any prior distribution.

5 Optimal Mechanisms

This section defines optimal mechanism and establishes the properties of optimal mechanisms.

Definition 2 *An mechanism Ω is optimal if it is incentive compatible, and for any incentive compatible mechanism Ω', $EU_{DM}(\Omega) \geq EU_{DM}(\Omega')$.*

An optimal mechanism induces the highest expected payoff to DM among all incentive compatible mechanisms. I focus on analyzing the best mechanism for DM because oftentimes DM's welfare reflects the public interest, for instance consumers and voters who have to make decisions under incomplete information. The optimal mechanism indicates an upper bound to the welfare of the public under lie-detection technology. Besides, the optimal mechanism minimize an weighted average objective of inference error and inspection cost. Therefore, it can be interpreted as the most efficient way of combating misinformation using lie-detection technology. On the other hand, the sender's welfare is sensitive to his risk attitude. However, it is worth-noting that if the sender is risk neutral, he will get the same ex-ante payoff in any incentive compatible mechanism because the mean of the induced action distribution must equal the prior expectation of the state. In such case, the outcome induced by an optimal mechanism is also Pareto-efficient.

Now we derive some properties of an optimal mechanism. For any set of messages $M \in \mathcal{M}$, let $Pr_q(M) = Pr(\Theta_q^u(M))$ be the ex-ante probability of the senders of M under message rule q. We say a property holds **almost everywhere** for a set of messages M if it holds for a subset of messages $M' \subseteq M$ such that $Pr_q(M') = Pr_q(M)$.

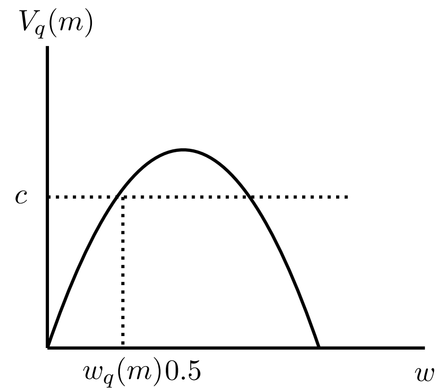
Proposition 3 ***(No direct benefit of inspection)***
In an optimal mechanism Ω, $V_q(m) = c$ almost everywhere for $m \in \mathcal{M}_q^+$.

Proposition 4 ***(Liars are minority)***
In an optimal mechanism Ω, $w_q(m) \leq 0.5$ almost everywhere for $m \in \mathcal{M}_q^+$.

The value of lie-detection technology to DM is composed of two parts: direct information value and indirect deterrence effect. Proposition 3 says that direct information value of inspection is offset by the cost of inspection in an optimal mechanism, and the net benefit of inspection comes from its effect on the sender's incentive: some types of sender refrain from making a higher claim because of the possible lie-detection. As a result, some information is transmitted through the messages in the sense that interim expectations of the sender's type upon receiving different messages are different, so DM is able to make a better inference on the sender's type even when inspection does not take place ex-post. Proposition 4 says that for any inspected message in an optimal mechanism, liars are a minority compared with truth-tellers. It is because the role of liars is to sustain moderate liar to truth-teller ratios so that information values are high enough for credible inspections. Such ratios can be achieved by either a minority of liars or a majority of liars. Compared with a mechanism with a majority of liars, a mechanism with a minority of liars means that the expected types of the sender of inspected messages are higher. That creates larger differences between conditional expectations given inspected messages and the uninspected message, which means more information is transmitted through messages under a mechanism with a minority of liars.

Proposition 5 ***(Three-interval structure)***
In an optimal mechanism Ω, there exists $\underline{\theta}_\Omega$ and $\overline{\theta}_\Omega$ such that $0 \leq \underline{\theta}_\Omega < \overline{\theta}_\Omega \leq 1$ and for almost every $\theta \in \Theta$, θ is lying if $\theta < \underline{\theta}_\Omega$; truthful if $\theta > \overline{\theta}_\Omega$; uninspected if $\theta \in [\underline{\theta}_\Omega, \overline{\theta}_\Omega]$.

Figure 2: The optimal proportion of liars



Note: Horizontal axis depicts the proportion of liars in an inspected message m; Vertical axis depicts value of inspecting m. An optimal mechanism minimizes the proportion liars subject to the constraint that $V(m) \geq c$. The expression of $w_q(m)$ is given in (16).

An optimal mechanism has a three-interval configuration such that when the state is above the cutoff $\overline{\theta}_\Omega$, sender is truthful; When the state is below the cutoff $\underline{\theta}_\Omega$, sender lies and claims that the state is somewhere above $\overline{\theta}_\Omega$, such claims are inspected with positive probabilities; When the state is intermediate, sender makes the claim in which DM does not inspect. Such configuration induces disperse inferences upon inspection, which benefit DM the most.

Under the optimal mechanism, low type senders are incentivized to lie in order to justify inspections of the truthful statements made by high type senders. Such inspections prevent moderate type senders from exaggerating their types in fear of getting caught lying and perceived as low types.

5.1 Optimal Mechanism: decreasing mimicking with precise statements

This subsection defines the decreasing mimicking mechanism and establishes the conditions in which such a mechanism is optimal.

For $d \in [2\sqrt{c}, 1]$, define

$$w^-(d) = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{c}{d^2}} \tag{15}$$

which is the smaller root of the equation $w(1-w)d^2 = c$. Given that d is the distance between the conditional expected type of truth-tellers and liars in a message m, $w^-(d)$ is the minimum proportion of liars such that information value of inspecting m is no less than c. This minimum proportion is decreasing in d, meaning that the credibility of inspection can be sustained for a smaller proportion of liars when the distance of conditional expectations is larger. Note that $2\sqrt{c}$ is the minimum required distance such that an inspection can be made credible, and $w^{-}(2\sqrt{c}) = \frac{1}{2}$. Proposition 3 and Proposition 4 imply for any inspected message $m \in \mathcal{M}_q^+$ in an optimal mechanism,

$$w_q(m) = w^{-}(X(m,t) - X(m,l)) \tag{16}$$

so for any inspected message in an optimal mechanism, proportion of liars is uniquely determined by the distance between expected types of truth-tellers and liars. For $x_l \in [0, 1-2\sqrt{c}]$ and $x_t \in [x_l + 2\sqrt{c}, 1]$, define

$$X_u^*(x_t, x_l) = w^{-}(x_t - x_l)x_l + (1 - w^{-}(x_t - x_l))x_t \tag{17}$$

which is the expected type of senders of a message m where x_t is the expected type of truth-tellers, x_l is the expected type of liars, and proportion of liars is minimized subject to DM's incentive constraint of inspection. Since DM is sequentially rational, for any $m \in \mathcal{M}_q^+$ in an optimal mechanism,

$$X(m,u) = X_u^*(X(m,t), X(m,l)) \tag{18}$$

so the induced action when the inspection does not take place is uniquely determined by the expected type of truth-tellers and liars.

Now we define the decreasing mimicking mechanism. Define a pair of cutoffs $\underline{\theta}_d$, $\overline{\theta}_d$ and matching function $\phi_d : [\overline{\theta}_d, 1] \to [0, \underline{\theta}_d]$ as a solution of the following system of differential equation and boundary conditions:

$$\dot{\phi}_d(\theta) = -\frac{w^{-}(\theta - \phi_d(\theta))}{1 - w^{-}(\theta - \phi_d(\theta))}\frac{f(\theta)}{f(\phi_d(\theta))} \tag{19}$$

$$\phi_d(1) = 0 \tag{20}$$

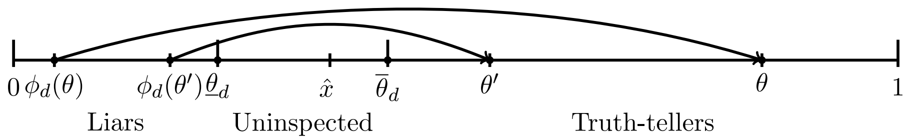
$$\phi_d(\overline{\theta}_d) = \underline{\theta}_d \tag{21}$$

To determine the boundaries $\underline{\theta}_d$ and $\overline{\theta}_d$, first define $\hat{\theta}$ such that

$$\hat{\theta} - \phi_d(\hat{\theta}) = 2\sqrt{c} \tag{22}$$

Lemma 1 *If Assumption 1 is satisfied, then there exists a unique solution $(\hat{\theta}, \phi_d)$ that satisfies conditions (19), (20) and (22). Furthermore, there exists a unique $\overline{\theta}_d \in [\hat{\theta}, 1]$ such that such that for any $\theta \in [\hat{\theta}, 1]$, $\theta < \overline{\theta}_d$ implies $X_u^*(\theta, \phi_d(\theta)) < E[\phi_d(\theta), \theta]$; $\theta > \overline{\theta}_d$ implies $X_u^*(\theta, \phi_d(\theta)) > E[\phi_d(\theta), \theta]$.*

Figure 3: The structure of decreasing mimicking mechanism Ω_d.



Note: In the decreasing mimicking mechanism, types above $\overline{\theta}_d$ make truthful and precise claims, which are mimicked by liars below $\underline{\theta}_d$ according to a decreasing mimicking function. These claims are randomly inspected. Types in between $\underline{\theta}_d$ and $\overline{\theta}_d$ pool at a single claim which is never inspected. $\hat{x}$ denotes the mean of the interval $(\underline{\theta}_d, \overline{\theta}_d)$. If $\hat{x}$ is above the mid-point of this interval, the decreasing mimicking mechanism is optimal.

$\overline{\theta}_d(.)$ represents a decreasing matching function from the truthful interval to the lying interval which specifies the lying pattern in the decreasing mimicking mechanism. Lemma 1 pins down a unique pair of boundaries $(\underline{\theta}_d, \overline{\theta}_d)$ for the two intervals.

Define the **decreasing mimicking mechanism** Ω_d which is characterized by $(\underline{\theta}_d, \overline{\theta}_d, \phi^d)$ defined in conditions (19) - (21) and Lemma 1 such that:

(i) **Intermediate types - Uninspected vague claim**: There is an uninspected message $m_q^0 = [\underline{\theta}_d, \overline{\theta}_d]$ sent by $\theta \in [\underline{\theta}_d, \overline{\theta}_d]$ and $P(m_q^0) = 0$;

(ii) **High types - Randomly inspected, precise claims**: There is a continuum of randomly inspected messages $\mathcal{M}_q^+ = \{m = \{\theta\} : \theta \in (\overline{\theta}_d, 1]\}$, each $m \in \mathcal{M}_q^+$ sent by the truthful type $\theta = m$ and $P(m) \in (0, 1)$;

(iii) **Low types - Liars of the high claims**: Each $m \in \mathcal{M}_q^+$ is sent by a liar $\phi_d(m)$.

The action rule X is determined by sequential rationality. For $m \in \mathcal{M}_q^+$,

$$\begin{aligned} X(m,t) &= m \\ X(m,l) &= \phi_d(m) \\ X(m,u) &= X_u^*(m, \phi_d(m)) \end{aligned} \tag{23}$$

and

$$X(m_q^0, u) = E[\underline{\theta}_d, \overline{\theta}_d] \tag{24}$$

The inspection rule P for $m \in \mathcal{M}_q^+$ is determined by the incentive compatibility conditions of the liars:

$$P(m) = \frac{u(X(m,u)) - u(X(m_q^0, u))}{u(X(m,u)) - u(X(m,l))} \tag{25}$$

Figure 3 depicts the structure of the decreasing mimicking mechanism. Under Ω_d, each truthful type θ makes the precise claim "My type is θ", and each of such claim is mimicked by exactly one type of liar $\phi_d(\theta)$, where $\phi_d(.)$ is decreasing so worse liars tell bigger lies. Upon receiving each of these messages, DM is indifferent between inspecting and not inspecting. The inspection probability is chosen so that liars are indifferent between telling such lies and making the uninspected claim. The optimal mechanism specifies a list of permissible claims the sender is allowed to make: a vague claim that represents moderate states, and a continuum of precise, high claims. Requiring a precise statement for high claims helps make more precise decisions upon inspection. Random inspections of those claims are justified because each of them is made by a low type and a high type. A vague moderate claim pools the moderate types which are not distant enough to be worth inspecting [8].

Lemma 2 *If Assumption 1 is satisfied, the decreasing mimicking mechanism Ω_d is incentive compatible with $0 < \underline{\theta}_d < \overline{\theta}_d < 1$, and $EU_{DM}(\Omega_d) > -\min\{Var(\Theta), c\}$.*

Lemma 2 means that Ω_d is incentive compatible whenever there exists an incentive compatible mechanism with inspection. I say two mechanisms $\Omega = (q, X, P)$ and $\Omega' = (q', X', P')$ are **equal almost everywhere** if for almost every $\theta \in \Theta$ and $s = \{t, l, u\}$, $m_q(\theta) = m_{q'}(\theta)$, $X(m_q(\theta), s) = X'(m_{q'}(\theta), s)$ and $P(m_q(\theta)) = P'(m_{q'}(\theta))$, i.e. sender sends the same messages, induces the same actions and inspected with the same probabilities in the two mechanisms almost surely.

Proposition 6 ***(Optimality of the decreasing mimicking mechanism)*** *Suppose in an optimal mechanism Ω, $E[\underline{\theta}_\Omega, \overline{\theta}_\Omega] > \frac{\underline{\theta}_\Omega + \overline{\theta}_\Omega}{2}$, then Ω and Ω_d are equal almost everywhere.*

When the mean of the uninspected interval is skewed towards its boundary to the truthful interval $\overline{\theta}_\Omega$, the value of inspecting the marginal truthful type around $\overline{\theta}_\Omega$ is small, so extending the uninspected interval to the right would be beneficial. Decreasing matching minimizes the truth-teller to liar ratios and allows the uninspected interval to extend farthest to the right. The condition of Proposition 6

[8] It is worth noting that despite having a list of permissible claims, exogenous enforcement on the sender's obedience is not necessary. It is because there is always a perfect Bayesian equilibrium where any off-path claim is regarded as a signal of the worst state and punished maximally so that the sender will never deviate to any claim out of the list.

is satisfied under a broad class of prior distributions. Two examples are symmetric single peaked distributions and distributions with increasing density.

Remark 1 *Suppose Assumption 1 is satisfied, and either:*
(1) F is symmetric and single peaked, or
(2) $f'(\theta) > 0$ for any $\theta \in [0,1]$,
then in the optimal mechanism Ω, $E[\underline{\theta}_\Omega, \overline{\theta}_\Omega] > \frac{\underline{\theta}_\Omega + \overline{\theta}_\Omega}{2}$.

6 State-verification and lie-detection

In this section, I compare state-verification technology to lie-detection technology, in particular, DM's welfare under the two technologies. Instead of revealing a binary signal as in (1), consider now the true state is revealed upon inspection, so by paying cost c to inspect the message m, DM receives the precise signal

$$s(m,\theta) = \theta \tag{26}$$

If DM chooses not to inspect, she receives an uninformative signal $s(m,\theta) = u$. Under state-verification technology, the sequentially rational action rule for DM is

$$X(m,\theta) = \theta; X(m,u) = E[\Theta_q^u(m)] \tag{27}$$

where $\Theta_q^u(m)$ is the set of senders who send m, and value of verifying m is the conditional variance of the sender's type:

$$V_q(m) = Var(\Theta_q^u(m)) \tag{28}$$

and the sequentially rational inspection rule for DM is

$$P(m) \in \begin{cases} \{0\} & \text{if } c > V_q(m) \\ [0,1] & \text{if } c = V_q(m) \\ \{1\} & \text{if } c < V_q(m) \end{cases} \cdot \tag{29}$$

Type θ sender's expected utility from sending a message m is

$$EU_{X,P}(m|\theta) = P(m)u(\theta) + (1 - P(m))u(X(m,u)) \tag{30}$$

and sender's optimality implies that for any on-path message $m' \in \mathcal{M}_q$,

$$P(m_q(\theta))u(\theta) + (1 - P(m_q(\theta)))u(X(m_q(\theta), u) \geq P(m')u(\theta) + (1 - P(m'))u(X(m', u) \qquad (31)$$

where $m_q(\theta)$ is the message sent by θ under the mechanism. I will show that there are only two kinds of incentive compatible mechanism under costly state-verification,

Uninformative mechanism: $P(m) = 0$ and $X(m, u) = E[\Theta]$ for any $m \in \mathcal{M}_q$, and

State-verifying mechanism: $P(m) = 1$ and $X(m, \theta) = \theta$ for any $m \in \mathcal{M}_q$.

Proposition 7 ***(No informative communication under state-verification.)***
Under costly state-verification technology, if $c > Var(\Theta)$, only the uninformative mechanism is incentive compatible; if $c < Var(\Theta)$, only the state-verifying mechanism is incentive compatible.

The ability to reveal the state precisely upon an inspection completely eliminates any incentive for the sender to transmit information. Gain from state-verification technology comes solely from the direct information value. It is contrary to the lie-detection technology, which benefits DM by manipulating the sender's incentive to transmit information. Such manipulation is possible because the nature of lie-detection creates a strategic uncertainty to DM: even if she spots a lie, she does not reveal the true type of the liar and has to decide the action base on equilibrium inference. This could benefit DM in an ex-ante sense because the sender might be deterred from deviation in fear of being mistaken as a worse type than what he actually is, and such a deterrence effect facilitates informative. However, if DM reveals the true state from an inspection, this deterrence will not be credible, and there will be no reason for the sender to stay honest. As a result, revealing more information from inspection eliminates voluntary information transmission from the sender. The following Proposition shows that learning more from inspection reduces DM's payoff. With Proposition 7, DM's ex-ante payoff under costly state-verification technology is

$$EU^s_{DM} = -\min\{Var(\Theta), c\} \qquad (32)$$

Proposition 8 ***(DM is better off under lie-detection technology than state-verification.)***
Let Ω^* be the optimal mechanism under lie-detection technology. Then under any inspection cost and distribution, $EU_{DM}(\Omega^*) \geq EU^s_{DM}$. Furthermore, if $c < Var(\Theta)$, then $EU_{DM}(\Omega^*) > EU^s_{DM}$.

This result provides a theoretical foundation for the emphasis on expert's integrity, instead of the objective information. By neglecting further information about the truth (other than the information that determines whether the sender is lying), the decision-maker is able to impose a credible threat that whoever being caught lying will be perceived poorly, regardless of the sender's true type. Therefore, even though there is no common interest between experts and decision-makers, some types of experts refrain from making higher claims in fear of being perceived as a worse type than they actually are.

7 Conclusion

Lying and lie-detection emerge from the opportunism of informed parties and the skepticism of uninformed parties. I establish a framework that allows analyses on the strategic interaction between lying and lie-detection, and characterize the optimal lie-detection policy. The results suggest that optimal lie-detection works as a credible deterrence tool. Low types are induced to lie so that inspections are justified, which deter higher types from lying. Under certain conditions, such optimal equilibrium can be achieved by allowing the sender to choose among a vague moderate claim and a continuum of precise high claims. This provides a direction for efficient allocation of resources in combating misinformation in various aspects such as politics and product advertising.

Several potential extensions are worth mentioning. As the first attempt in the literature to study endogenous lying and costly lie-detection, I restrict attention to the setting of single round communication and lie-detection. In some applications, the sender and the receiver can conduct multiple rounds of communication and inspect lie-detection, before a final decision is made by the receiver. For instance, the police can ask the suspect multiple questions and conduct lie-detection for each claim made by the suspect. Dziuda and Salas (2018) show that the receiver prefers to commit to a single round communication when the probability of lie-detection is exogenously high, because anticipating the second chance of communication makes the sender more likely to lie. It might appear that this effect is strengthened when lie-detection is costly as the receiver has to pay the cost of inspection in each round. A formal analysis is required for such an argument. Another potential extension is to allow a certain degree of common interest between the sender and the receiver, such as biased sender as in Crawford and Sobel (1981). It is not clear whether having a sender with a smaller bias would benefit the receiver when lie-detection is possible. On one hand, sender with smaller bias is willing to reveal more precise information, as suggested by the standard cheap talk model. On the other hand, when bias is small, there is no way to induce the sender to tell big lies. This hinders the formation of credible inspection. Without inspection, the sender might be tempted to tell small lies, which impede informative communication. The analysis of these opposing effects may present interesting avenues for future research.

References

Balbuzanov, I. (2017), "Lies and Consequences: The Effect of Lie Detection on Communication Outcomes" *mimeo.*

Chen, Y. (2011), "Perturbed communication games with honest senders and naive receivers," *Journal of Economic Theory*, 20(1), 146(2), 401-424.

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E. and K. B. McDermott (2008), "The contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses," *Cerebral cortex*, 19(7), 1557-1566.

Crawford, V. P. and J. Sobel (2003), "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 1431-1451.

Dziuda, W. and C. Salas (2018), "Communication with detectable deceit," *Available at SSRN*, 3234695.

Hartwig, M. and C. F. Bond Jr (2014), "Lie detection from multiple cues: A meta-analysis," *Applied Cognitive Psychology*, 28(5), 661-676.

Hartwig, M., Granhag, P. A., Strömwall, L. A. and O. Kronkvist (2006), "Strategic use of evidence during police interviews: When training to detect deception works," *Law and human behavior*, 30(5), 603-619.

Jehiel, P. (2019), "Communication with Forgetful Liars."

Kartik, N. (2009), "Strategic communication with lying costs," *The Review of Economic Studies*, 76(4), 1359-1395.

Kartik, N., Ottaviani, M. and F. Squintani (2007), "Credulity, lies, and costly talk," *Journal of Economic theory*, 134(1), 93-116.

Lim, C. (2018), "Can Fact-checking Prevent Politicians from Lying?"

Nyhan, B. and J. Reifler (2015), "The effect of fact-checking on elites: A field experiment on US state legislators," *American Journal of Political Science*, 59(3), 628-640.

Ottaviani, M. and F. Squintani (2006), “Naive audience and communication bias,” *International Journal of Game Theory*, 35(1), 129-150.

Porter, M. and F. Ten Brinke (2008), “Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions,” *Psychological science*, 19(5), 508-514.

Sobel, J. (2019), “Lying and Deception in Games.”

Vrij, A., Granhag, P. A., Mann, S. and S. Leal (2011), “Outsmarting the liars: Toward a cognitive lie detection approach,” *Current Directions in Psychological Science*, 20(1), 28-32.

Appendix A

This Appendix provides proofs of Propositions 1 - 8 and Lemmas 1 - 2.

Proof of Proposition 1:

Fix an incentive compatible mechanism Ω. Let $M_1 = \{m \in \mathcal{M}_q : X(m,l) > X(m,t)\}$ be the set of on-path messages such that the induced action of liars is higher than the induced action of truth-tellers. Define a mortified mechanism $\hat{\Omega}$ such that for each $m \in \mathcal{M}_q/M_1$, the set of senders remain unchanged but they now send the transformed message $T(m) = \Theta_q^t(m)$. For each $m \in M_1$, the set of senders remain unchanged but they now send the transformed message $T(m) = \Theta_q^l(m)$. For inspection probabilities, let $\hat{P}(T(m)) = P(m)$ for each $m \in \mathcal{M}_q$.

The set of on-path messages of the modified mechanism $\hat{\Omega}$ is $\mathcal{M}_{\hat{q}} = T(\mathcal{M}_q)$. The sequentially rational actions for $\hat{\Omega}$ are $\hat{X}(T(m),s) = X(m,s)$ for $m \in \mathcal{M}_q/M_1$ and $s = t,l,u$; $\hat{X}(T(m),t) = X(m,l)$, $\hat{X}(T(m),l) = X(m,t)$ and $\hat{X}(T(m),u) = X(m,u)$ for $m \in M_1$. It is straight-forward that for all $m \in \mathcal{M}_q$ $\hat{X}(T(m),t) \geq \hat{X}(T(m),l)$ and $T(m) = \Theta_q^t(T(m))$. Therefore, condition (i) and (ii) are satisfied in the mortified mechanism $\hat{\Omega}$. Furthermore, since the induced actions remain unchanged for every type of sender, so $\hat{\Omega}$ and Ω are distribution equivalent.

To see that $\hat{\Omega}$ is incentive compatible, note that for $m \in \mathcal{M}_q/M_1$, $w_q(m) = w_(\hat{q})(T(m))$, and for $m \in M_1$, $w_q(m) = 1 - w_(\hat{q})(T(m))$. Therefore, for any $m \in \mathcal{M}_q$, $V_q(m) = w_q(m)(1-w_q(m))(X(m,t) - X(m,l))^2 = w_(\hat{q})(T(m))(1-w_(\hat{q})(T(m)))(\hat{X}(T(m),t) - \hat{X}(T(m),l))^2 = V_{\hat{q}}(T(m))$, thus (7) remains satisfied in $\hat{\Omega}$. To check incentive constraints (9), note that the equilibrium payoff of each type of sender remain unchanged, i.e. $EU_{\hat{X},\hat{P}}(m_{\hat{q}}(\theta)|\theta) = EU_{X,P}(m_q(\theta)|\theta)$. By the definition of the mortified set of message, any type θ would be identified as a liar of any on-path message other than its equilibrium message, i.e. $\theta \notin m'$ for any $m' \in T(\mathcal{M}_q)$ and $m' \neq m_{\hat{q}}(\theta)$. This combined with the fact that $\hat{X}(T(m),t) \geq \hat{X}(T(m),l)$ imply $EU_{\hat{X},\hat{P}}(T(m')|\theta) \leq EU_{X,P}(m'|\theta)$ for any $m' \in \mathcal{M}_q$. Therefore, $EU_{\hat{X},\hat{P}}(m_{\hat{q}}(\theta)|\theta) = EU_{X,P}(m_q(\theta)|\theta) \geq EU_{X,P}(m'|\theta) \geq EU_{\hat{X},\hat{P}}(T(m')|\theta)$ for any $\theta \in \Theta$ and $m' \in \mathcal{M}_q$, where the first inequality holds by incentive compatibility of the original mechanism, thus (9) is satisfied in the modified mechanism. Therefore, we conclude that $\hat{\Omega}$ is incentive compatible.

Q.E.D.

Proof of Proposition 2:

"Only if":

Let $\Omega \equiv (q, P, X)$ be an incentive compatible mechanism where $p_\Omega > 0$, then the set of inspected messages $\mathcal{M}_q^+$ has positive measure, so (5) implies that for almost every $m \in \mathcal{M}_q^+$, $X(m,t) < 1$ and $X(m,l) > 0$, so $V_q(m) = w_q(m)(1-w_q(m))(X(m,t)-X(m,l))^2 < \frac{1}{4}$. Since (6) and (7) imply $V_q(m) \geq c$, it must be the case that $c \leq V_q(m) < \frac{1}{4}$.

By (6) and the definition of $w^-(.)$, we have that $w^-(X(m,t)-X(m,l)) = \min\{w \in [0,1] : V_q(m) \geq c\}$. Since for any $m \in \mathcal{M}_q^+$, $V_q(m) \geq c$, so $w_q(m) \geq w^-(X(m,t)-X(m,l))$, and thus $X(m,u) = w_q(m)X(m,l) + (1-w_q(m))X(m,t) \leq X_u^*(X(m,t), X(m,l))$. Since X is sequentially rational, $X(m,t) = E[\Theta_q^t(m)] \leq 1$ and $X(m,t) = E[\Theta_q^l(m)] \geq 0$, with the inequalities hold strictly if $Pr(\Theta_q^u(m)) > 0$. Therefore, $X(m,u) \leq X_u^*(X(m,t), X(m,l)) \leq X_u^*(1,0) = \frac{1}{2} + \sqrt{\frac{1}{4} - c}$, where the second inequality holds since by Lemma 6 $X_u^*(X(m,t), X(m,l))$ is strictly increasing in $X(m,t)$ and strictly decreasing in $X(m,l)$, and it hold strictly if $Pr(\Theta_q^u(m)) > 0$.

Since $\Theta_q^0 \cup \Theta_q^u(\mathcal{M}_q^+) = \Theta$, so $Pr(\Theta_q^0) = 1 - P(\Theta_q^u(\mathcal{M}_q^+))$ and

$$(1 - P(\Theta_q^u(\mathcal{M}_q^+)))E[\Theta_q^0] + P(\Theta_q^u(\mathcal{M}_q^+))E[\Theta_q^u(\mathcal{M}_q^+)] = E[\Theta] \tag{33}$$

. Since X is sequentially rational,

$$\begin{aligned} P(\Theta_q^u(\mathcal{M}_q^+))E[\Theta_q^u(\mathcal{M}_q^+)] =& \int_{\mathcal{M}_q^+} X(m,u) \int_{\Theta_q^u(m)} dF(\theta)dm \\ \leq& X_u^*(1,0) \int_{\mathcal{M}_q^+} \int_{\Theta_q^u(m)} dF(\theta)dm \\ =& Pr(\Theta_q^u(\mathcal{M}_q^+))(\frac{1}{2} + \sqrt{\frac{1}{4} - c}) \end{aligned} \tag{34}$$

where the inequality holds strictly if $Pr(\Theta_q^u(\mathcal{M}_q^+)) > 0$. If $Pr(\Theta_q^0) = 0$, then $P(\Theta_q^u(\mathcal{M}_q^+)) = 1$ and (33) and (34) imply $E[\Theta] = E[\Theta_q^u(\mathcal{M}_q^+)] < \frac{1}{2} + \sqrt{\frac{1}{4} - c}$. If $Pr(\Theta_q^0) > 0$, then Lemma 3 implies that for any $m \in \mathcal{M}_q^+$, $E[\Theta_q^0] = X(m_q^0, u) < X(m,u)$, so by (34) $E[\Theta_q^0] < E[\Theta_q^u(\mathcal{M}_q^+)] \leq \frac{1}{2} + \sqrt{\frac{1}{4} - c}$, then (33) implies $E[\Theta] < \frac{1}{2} + \sqrt{\frac{1}{4} - c}$.

"If":

The decreasing mechanism Ω_d defined at (19) - (24) is an example, where Lemma 2 implies that if Assumption 1 is satisfied, then Ω_d is incentive compatible with $0 < \underline{\theta}_d < \overline{\theta}_d < 1$, and thus $p_{\Omega_d} = \int_{\overline{\theta}_d}^1 P(\{\theta\})[f(\theta) + \dot{\phi}_d(\theta)f(\phi_d(\theta))]d\theta > 0$.

Q.E.D.

Proof of Proposition 3: Suppose contrary to the claim, there exists a positive measure set inspected messages $M_1 \subseteq \mathcal{M}_q^+$ such that $V_q(m) \neq c$ for all $m \in M_1$. Since $P(m) > 0$, sequential rationality

(7) then requires that $V_q(m) > c$ and $P(m) = 1$, and (b) of Lemma 3 implies that for all $m \in M_1$, $X(m,l) = \hat{x}$, where $\hat{x} = X(m_q^0, u)$ if m_q^0 exists, $\hat{x} = \max_{m'} X(m', l)$ otherwise. Therefore, we have $E[\Theta_q^l(m)] = E[\Theta_q^l(M_1)] = \hat{x}$ for all $m \in M_1$.

Denote $\hat{\Theta} = \Theta_q^l(M_1)$ be the set of liars who send $m \in M_1$ in the original mechanism. For $m \in M_1$, denote $\hat{w}(m) = w^-(X(m,t) - \hat{x})$ be the smallest weight on liars such that value of inspection is no less than c. Since for all $m \in M_1$, $V_q(m) > c$, so $w_q(m) > \hat{w}(m)$. Now define $\hat{p} = \int_{M_1} \int_{\Theta_q^t(m)} \frac{\hat{w}(m)}{1-\hat{w}(m)} dF(\theta)dm$, which is the total minimum measure of liars required to match with truth-tellers of $m \in M_1$ such that value of inspection is no less than c. We have $\hat{p} < Pr(\hat{\Theta}) = \int_{M_1} \int_{\Theta_q^t(m)} \frac{w_q(m)}{1-w_q(m)} dF(\theta)dm$.

Assign an arbitrary strict ranking $r : M_1 \to \mathbb{R}$ to the message set M_1. Then for any $m \in M^l$, let

$$z^-(m) = \frac{1}{Pr(\Theta_q^l(M_1))} \int_{m' \in M_1 : r(m') < r(m)} \frac{\hat{w}(m')}{1 - \hat{w}(m')} \int_{\Theta_q^t(m')} dF(\theta)dm' \tag{35}$$

$$z^+(m) = \frac{1}{Pr(\Theta_q^l(M_1))} \int_{m' \in M_1 : r(m') = r(m)} \frac{\hat{w}(m')}{1 - \hat{w}(m')} \int_{\Theta_q^t(m')} dF(\theta)dm' \tag{36}$$

be the cumulative required fraction of liars.

For any positive measure set of types $\hat{\Theta}$, define the mean-preserving division $\hat{\Theta}(z) = \hat{\Theta} \cap [\underline{\theta}(z), \overline{\theta}(z)]$ such that $\underline{\theta}(z)$ and $\overline{\theta}(z)$ solve

$$Pr(\hat{\Theta}(z)) = zPr(\hat{\Theta}) \tag{37}$$

$$E[\hat{\Theta}(z)] = E[\hat{\Theta}] \tag{38}$$

Define an modified messaging and action rules $\hat{q}, \hat{X}$ where other things remain unchanged, except the set of messages M_1. The uninspected message is modified to $m_{\hat{q}}^0 = m_q^0 \cup (\hat{\Theta}/\hat{\Theta}(\frac{\hat{p}}{Pr(\hat{\Theta})}))$, where $\hat{\Theta}/\hat{\Theta}(\frac{\hat{p}}{Pr(\hat{\Theta})})$ is a mean-preserving division of $\hat{\Theta}$ with mean $\hat{x}$ and measure $Pr(\hat{\Theta}) - \hat{p}$. For $m \in M_1$, the set of truth-tellers remain unchanged, while the set of liars is modified to $\Theta_{\hat{q}}^l(m) = \hat{\Theta}(z^+(m))/int(\hat{\Theta}(z^-(m)))$, a mean preserving division of $\hat{\Theta}$ where $int(X)$ is the interior of set X, so that $E[\Theta_{\hat{q}}^l(m)] = \hat{x}$ and the set has measure $\frac{\hat{w}(m)}{1-\hat{w}(m)} \int_{\Theta_q^t(m)} dF(\theta)$.

The sequentially rational actions for the modified uninspected messages $m_{\hat{q}}^0$ is

$$\hat{X}(m_{\hat{q}}^0, u) = E[\hat{\Theta}(\underline{z})] = \hat{x} \tag{39}$$

and for $m \in M_1$,

$$\hat{X}(m,t) = X(m,t)$$

$$\hat{X}(m,l) = X(m,l) = \hat{x} \tag{40}$$
$$\hat{X}(m,u) = \hat{w}(m)\hat{x} + (1-\hat{w}(m))X(m,t)$$

where $\hat{X}(m,u) > \hat{x}$, so $(\hat{q},\hat{X})$ satisfies (a) of Lemma 3. Furthermore, by the definition of $\Theta^l_{\hat{q}}(m)$ for $m \in M_l$, we have

$$w_{\hat{q}}(m) = \hat{w}(m) = w^-(\hat{X}(m,t) - \hat{X}(m,l)) \tag{41}$$

and thus

$$V_{\hat{q}}(m) = c \tag{42}$$

so $(\hat{q},\hat{X})$ satisfies (b) of Lemma 3. Therefore, there exists $\hat{P}$ such that $\hat{\Omega} = (\hat{q},\hat{P},\hat{X})$ is incentive compatible.

Under the modified mechanism $\hat{\Omega}$, the sequentially rational actions remain unchanged for every type, but the ex-ante probability of inspection is reduced by $Pr(\hat{\Theta}) - \hat{p} > 0$. Therefore, $EU_{DM}(\hat{\Omega}) > EU_{DM}(\Omega)$, contradicts that Ω is an optimal mechanism.

Q.E.D.

Proof of Proposition 4: By Proposition 3 $V_q(m) = w_q(m)(1-w_q(m))(X(m,t)-X(m,l))^2 = c$ for $m \in \mathcal{M}^+_q$, ans since $X(m,u) = w_q(m)X(m,l) + (1-w_q(m)X(m,t))$, we have $(X(m,t) - X(m,u))((X(m,u) - X(m,l))) = c$. Suppose contrary to the claim, $w_q(m) > 0.5$ some positive measure set of messages in $\mathcal{M}^+_q$, which implies $X(m,t)-X(m,u) > (X(m,u)-X(m,l))$, take a positive measure set of message $M^+ \in \mathcal{M}^+_q$ such that $X(m,t)-X(m,u) > (X(m,u)-X(m,l))+\delta$ for some $\delta > 0$. Then for any $\epsilon > 0$ there exists a positive measure set of message $M^+_\epsilon \subseteq M^+$ such that for any $m, m' \in M^+_\epsilon$ and $s = t,l,u$, $|X(m,s) - X(m',s)| < \epsilon$ and $X(m,t)-X(m,u)+\delta < X(m,u)-X(m,l)$.

Let $\Theta^l_\epsilon = \Theta^l_q(M^+_\epsilon)$ and $\Theta^t_\epsilon = \Theta^t_q(M^+_\epsilon)$ be the aggregate set of truth-tellers and liars of M^+_ϵ, and $Pr^l_\epsilon = Pr(\Theta^l_q(M^+_\epsilon))$ and $Pr^t_\epsilon = Pr(\Theta^t_q(M^+_\epsilon))$ be the measure of the two sets. Let $E^l_\epsilon = E[\Theta^l_\epsilon]$, $E^t_\epsilon = E[\Theta^t_\epsilon]$ and $E^u_\epsilon = E[\Theta^l_\epsilon \cup \Theta^t_\epsilon]$ be the corresponding expected values of the sets. Note that $|E^s_\epsilon - X(m,s)| < \epsilon$ for any $m \in M^+_\epsilon$ and $s = t,l,u$, so we have

$$E^t_\epsilon - E^u_\epsilon > E^u_\epsilon - E^l_\epsilon + \delta - 2\epsilon \tag{43}$$
$$|(E^t_\epsilon - E^u_\epsilon)(E^u_\epsilon - E^l_\epsilon) - c| < 4\epsilon^2 \tag{44}$$

Let $\hat{E}$ be the larger root of $(E^t_\epsilon - \hat{E})(\hat{E} - E^l_\epsilon) - c = 0$. (43) and (44) imply that for small enough ϵ, $E^t_\epsilon - \hat{E} < \hat{E} - E^l_\epsilon$ and $\hat{E} > E^u_\epsilon + \delta$. Fix any $m \in \mathcal{M}^+_q$ and Let $\underline{u} = P(m)u(X(m,l)) + (1-P(m))u(X(m,u))$

be the expected payoff of the liars, Let $\hat{x} = u^{-1}(\underline{u})$ be its certainty equivalence. Note that Proposition 1 implies any liar can mimic the payoff of any other liar, so incentive compatibility means all liars receive the same payoff, and $\hat{x} = X(m_q^0, u)$ if an uninspected message m_q^0 exists. Lemma 3 implies $X(m,l) \leq \hat{x} < X(m,u)$ for any $m \in M_\epsilon^+$, so we have

$$E_\epsilon^l \leq \hat{x} < E_\epsilon^u < \hat{E} - \delta < E_\epsilon^t - \delta \tag{45}$$

Let z_l, z_t solve

$$z_l Pr_\epsilon^l E_\epsilon^l + z_t Pr_\epsilon^t E_\epsilon^t = (z_l Pr_\epsilon^l + z_t Pr_\epsilon^t)\hat{x} \tag{46}$$

$$(1 - z_l)Pr_\epsilon^l E_\epsilon^l + (1 - z_t)Pr_\epsilon^t E_\epsilon^t = [(1 - z_l)Pr_\epsilon^l + (1 - z_t)Pr_\epsilon^t]\hat{E} \tag{47}$$

Since $Pr_\epsilon^l E_\epsilon^l + Pr_\epsilon^t E_\epsilon^t = (Pr_\epsilon^l + Pr_\epsilon^t)E_\epsilon^u$, so (45) means $z_l \in (0,1)$ and $z_t \in [0,1)$

For any positive measure set of types $\hat{\Theta}$, define the mean-preserving division $\hat{\Theta}(z) = \hat{\Theta} \cap [\underline{\theta}(z), \overline{\theta}(z)]$ such that $\underline{\theta}(z)$ and $\overline{\theta}(z)$ solve

$$Pr(\hat{\Theta}(z)) = zPr(\hat{\Theta}) \tag{48}$$

$$E[\hat{\Theta}(z)] = E[\hat{\Theta}] \tag{49}$$

We divide the liar set Θ_ϵ^l into $\Theta_\epsilon^l(z_l)$ and $\Theta_\epsilon^l/\Theta_\epsilon^l(z_l)$, and truthful set Θ_ϵ^t into $\Theta_\epsilon^t(z_t)$ and $\Theta_\epsilon^t/\Theta_\epsilon^t(z_t)$. The mean-preserving divisions implies $E[\Theta_\epsilon^l(z_l)] = E[\Theta_\epsilon^l/\Theta_\epsilon^l(z_l)] = E_\epsilon^l$ and $E[\Theta_\epsilon^t(z_t)] = E[\Theta_\epsilon^t/\Theta_\epsilon^t(z_t)] = E_\epsilon^t$. From (46) and (47) we have $E[\Theta_\epsilon^l(z_l) \cup \Theta_\epsilon^t(z_t)] = \hat{x}$ and $E[\Theta_\epsilon^l/\Theta_\epsilon^l(z_l) \cup \Theta_\epsilon^t/\Theta_\epsilon^t(z_t)] = \hat{E}$.

Now define an modified mechanism $\hat{\Omega} = (\hat{q}, \hat{P}, \hat{X})$ where other things remain unchanged, except the set of messages M_ϵ^+ is off-path and an message $\hat{m} = \Theta_\epsilon^t/\Theta_\epsilon^t(z_t)$ is added with $\hat{q}(\hat{m}|\theta) = 1$ for $\theta \in \Theta_\epsilon^l/\Theta_\epsilon^l(z_l) \cup \Theta_\epsilon^t/\Theta_\epsilon^t(z_t)$. The uninspected message m_q^0 (if exists) is modified to $m_{\hat{q}}^0 = m_q^0 \cup \Theta_\epsilon^l(z_l) \cup \Theta_\epsilon^t(z_t)$ with $\hat{q}(m_{\hat{q}}^0|\theta) = 1$ for $\theta \in \Theta_\epsilon^l(z_l) \cup \Theta_\epsilon^t(z_t)$.

The sequentially rational actions for the modified messages $\hat{m}$ and $m_{\hat{q}}^0$ are $\hat{X}(\hat{m}, t) = E_\epsilon^t$, $\hat{X}(\hat{m}, l) = E_\epsilon^l$, $\hat{X}(\hat{m}, u) = \hat{E}$, $\hat{X}(m_{\hat{q}}^0, u) = \hat{x}$. By (45) we still have $\hat{X}(m,l) \leq \hat{X}(m_{\hat{q}}^0, u) < \hat{X}(m,u)$ for all $m \in \mathcal{M}_{\hat{q}}^+$, so (a) in Lemma 3 is satisfied. For the newly added inspected message $\hat{m}$, $(\hat{X}(m_{\hat{q}}^0, t) - \hat{X}(m_{\hat{q}}^0, u))(\hat{X}(m_{\hat{q}}^0, u) - \hat{X}(m_{\hat{q}}^0, l)) = (E_\epsilon^t - \hat{E})(\hat{E} - E_\epsilon^l) = c$, so (b) in Lemma 3 is satisfied. Therefore there exists $\hat{P}$ such that $\hat{\Omega}$ is incentive compatible.

To compare DM's exante payoffs, let G_Ω^u and $G_{\hat{\Omega}}^u$ be the distribution of uninspected induced actions of the two mechanism defined in (56). By sequential rationality the two distributions have the same mean $\int_0^1 xdG^u_{\hat{\Omega}}(x) = \int_0^1 xdG^u_{\Omega}(x) = \int_\Theta \theta dF(\theta)$ and they differ only by actions induced by the set $\Theta^l_\epsilon \cup \Theta^t_\epsilon$. In the original mechanism Ω, a type in $\Theta^l_\epsilon \cup \Theta^t_\epsilon$ sends some $m \in M^+_\epsilon$ with induced action $X(m,u)$ where $|X(m,u) - E^u_\epsilon| < \epsilon$; In the modified mechanism $\hat{\Omega}$, a type in $\Theta^l_\epsilon \cup \Theta^t_\epsilon$ send either $\hat{m}$ or $m^0_{\hat{q}}$ with induced action either $\hat{X}(m^0_{\hat{q}}, u) = \hat{x}$ or $\hat{X}(\hat{m}, u) = \hat{E}$. (45) implies that for small enough ϵ, $X(m^0_{\hat{q}}, u) < X(m,u) < \hat{X}(\hat{m}, u)$ for any $m \in M^+_\epsilon$. Therefore, $G^u_{\hat{\Omega}}$ is a mean-preserving spread of G^u_Ω, which means $\int_{[0,1]} x^2 dG^u_{\hat{\Omega}}(x) > \int_{[0,1]} x^2 dG^u_\Omega(x)$, then (55) implies $EU_{DM}(\hat{\Omega}) > EU_{DM}(\Omega)$, contradicts that Ω is an optimal mechanism.

Q.E.D.

Proof of Proposition 5: By Lemma 16 the statement is true if $Pr(\Theta^0_q) > 0$. Now suppose there is an optimal mechanism Ω in which $Pr(\Theta^0_q) = 0$, then Proposition 3 implies $V_q(m) = c$ for almost every $m \in \mathcal{M}^+_q$, then Lemma 4 implies $EU_{DM}(\Omega) = EU^I_{DM}(\Omega) \leq -c$. Since Ω is incentive compatible, Assumption 1 holds, but then Lemma 2 implies that the decreasing mechanism Ω_d is also incentive compatible, with $EU_{DM}(\Omega_d) > c \geq EU_{DM}(\Omega)$, so Ω with $Pr(\Theta^0_q) = 0$ cannot be optimal. *Q.E.D.*

Proof of Lemma 1: Since $w^-(d) \in (0, \frac{1}{2}]$ is well-defined and positive for any $d \geq 2\sqrt{c}$, and $c \leq \frac{1}{4}$ means $2\sqrt{c} \leq 1$, so $\dot{\phi}_d(1) = -\frac{w^-(1)}{1-w^-(1)}\frac{f(1)}{f(0)}$ is well-defined, and for any $\theta \leq 1$ in which $\theta - \phi_d(\theta) \geq 2\sqrt{c}$, $\dot{\phi}_d(\theta)$ is well-defined and negative, which means $\frac{d(\theta - \phi_d(\theta))}{d\theta} > 1$. Therefore, there exists unique solutions ϕ_d and $\hat{\theta} \in (0, 1]$ that satisfy (19), (20) and (22).

To show the second part of the statement, for any $\theta \in [\hat{\theta}, 1]$,

$$\frac{dX^*_u(\theta, \phi_d(\theta))}{d\theta} = \frac{\partial X^*_u(\theta, \phi_d(\theta))}{\partial \theta} + \frac{\partial X^*_u(\theta, \phi_d(\theta))}{\partial \phi_d(\theta)}\dot{\phi}_d(\theta) > 0 \tag{50}$$

where the inequality holds by Lemma 6 and $\dot{\phi}_d(\theta) < 0$. Also,

$$\begin{aligned}
&\frac{dE[\phi_d(\theta), \theta]}{d\theta} \\
=&\frac{\partial E[\phi_d(\theta), \theta]}{\partial \theta} + \frac{\partial E[\phi_d(\theta), \theta]}{\partial \phi_d(\theta)}\dot{\phi}_d(\theta) \\
=&\frac{f(\theta)(\theta - E[\phi_d(\theta), \theta])}{Pr([\phi_d(\theta), \theta])} - \frac{f(\phi_d(\theta))(\phi_d(\theta) - E[\phi_d(\theta), \theta])}{Pr([\phi_d(\theta), \theta])}\dot{\phi}_d(\theta) \\
=&\frac{f(\theta)}{Pr([\phi_d(\theta), \theta])(1 - w^-(\theta - \phi_d(\theta)))}[(1 - w^-(\theta - \phi_d(\theta))\theta + w^-(\theta - \phi_d(\theta))\phi_d(\theta) - E[\phi_d(\theta), \theta]] \\
=&\frac{f(\theta)}{Pr([\phi_d(\theta), \theta])(1 - w^-(\theta - \phi_d(\theta)))}[X^*_u(\theta, \phi_d(\theta)) - E[\phi_d(\theta), \theta]]
\end{aligned} \tag{51}$$

Where the third equality holds by (19), the fourth equality holds by (54), and the last equality holds by (18). Since (50) and (51) means when $X^*_u(\theta, \phi_d(\theta)) \leq E[\phi_d(\theta), \theta]$, $X^*_u(\theta, \phi_d(\theta)) - E[\phi_d(\theta), \theta]$ is strictly increasing in θ, therefore for any $\theta' < \theta$,

$$X_u^*(\theta, \phi_d(\theta)) \leq E[\phi_d(\theta), \theta] \Rightarrow X_u^*(\theta', \phi_d(\theta')) < E[\phi_d(\theta'), \theta'] \tag{52}$$

By assumption of the Lemma, $X_u^*(1, \phi_d(1)) = X_u^*(1, 0) = \frac{1}{2} + \sqrt{\frac{1}{4} - c} > E[\Theta] \equiv E[0, 1] = E[\phi_d(1), 1]$. Now we consider two cases.

Case 1: $X_u^*(\hat{\theta}, \phi_d(\hat{\theta})) > E[\phi_d(\hat{\theta}), \hat{\theta}]$**:** then (52) and $X_u^*(1, \phi_d(1)) > E[\phi_d(1), 1]$ imply that $X_u^*(\theta, \phi_d(\theta)) > E[\phi_d(\theta), \theta]$ for any $\theta \in [\hat{\theta}, 1]$, then $\overline{\theta}_d = \hat{\theta}$.

Case 2: $X_u^*(\hat{\theta}, \phi_d(\hat{\theta})) \leq E[\phi_d(\hat{\theta}), \hat{\theta}]$**:** then since $X_u^*(1, \phi_d(1)) > E[\phi_d(1), 1]$, there exists $\overline{\theta}_d \in [\hat{\theta}, 1)$ such that $X_u^*(\overline{\theta}_d, \phi_d(\overline{\theta}_d)) = E[\phi_d(\overline{\theta}_d), \overline{\theta}_d]$. (52) implies that $X_u^*(\theta, \phi_d(\theta)) < E[\phi_d(\theta), \theta]$ for $\theta < \overline{\theta}_d$ and $X_u^*(\theta, \phi_d(\theta)) > E[\phi_d(\theta), \theta]$ for $\theta > \overline{\theta}_d$. ■

Proof of Lemma 2: To show that Ω_d is incentive compatible, for $m \in \mathcal{M}_q^+$,

$$\begin{aligned} \frac{w_q(m)}{1 - w_q(m)} &= \lim_{\epsilon} \frac{Pr([\phi_d(m+\epsilon), \phi_d(m-\epsilon)])}{Pr([m-\epsilon, m+\epsilon])} \\ &= \frac{-\dot{\phi}_d(m) f(\phi_d(m))}{f(m)} = \frac{w^-(m - \phi_d(m))}{1 - w^-(m - \phi_d(m))} \end{aligned} \tag{53}$$

where the first equality holds because of the continuously decreasing message rule, and the third equality holds by (19). Therefore, $w_q(m) = w^-(m - \phi_d(m))$ and $V_q(m) = c$, thus condition (b) of Lemma 3 is satisfied.

$X(m_q^0, u)$, $X(m, t)$ and $X(m, l)$ for $m \in \mathcal{M}_q^+$ are clearly sequentially rational given the message rule. For $m \in \mathcal{M}_q^+$,

$$\begin{aligned} X(m, u) = X_u^*(m, \phi_d(m)) &= w^-(m - \phi_d(m))\phi_d(m) + (1 - w^-(m - \phi_d(m)))m \\ &= w_q(m) X(m, l) + (1 - w_q(m)) X(m, t) \end{aligned}$$

is also sequentially rational. For any $m \in \mathcal{M}_q^+ = (\overline{\theta}_d, 1]$,

$$\begin{aligned} X(m, u) = X_u^*(m, \phi_d(m)) > E[\phi_d(m), m] &> E[\phi(\overline{\theta}_d), \overline{\theta}_d] \\ &= X(m_q^0, u) > \underline{\theta}_d > \phi_d(m) = X(m, l) \end{aligned}$$

where the first and second inequalities holds because by the definition of $\overline{\theta}_d$, $X_u^*(\theta, \phi_d(\theta)) > E[\phi_d(\theta), \theta]$ for any $\theta \in (\overline{\theta}_d, 1]$, then by (51) $E[\phi_d(\theta), \theta]$ is strictly increasing for $\theta \in (\overline{\theta}_d, 1]$. Therefore, condition (a) of Lemma 3 is satisfied, and thus Ω_d is incentive compatible with the inspection rule specified by (25). Since $c < \frac{1}{4}$ implies $2\sqrt{c} < 1 = 1 - \phi_d(1)$, so $\hat{\theta} < 1$, and $E[\Theta] < \frac{1}{2} + \sqrt{\frac{1}{4} - c}$ implies $X_u^*(1, \phi_d(1)) > E[0, 1]$, so $\overline{\theta}_d < 1$. Then $\underline{\theta}_d = \phi_d(\overline{\theta}_d) > 0$ because ϕ_d is strictly decreasing.

Since $V_q(m) = c$ for any $m \in \mathcal{M}_q^+$, Lemma 4 implies $EU_{DM}(\Omega_d) = EU^U_{DM}(\Omega_d) = EU^I_{DM}(\Omega_d)$. To show that $EU_{DM}(\Omega_d) > -c$,

$$\begin{aligned}
&EU_{DM}(\Omega_d) = EU^I_{DM}(\Omega_d)\\
=&Pr([\underline{\theta}_d,\overline{\theta}_d])E[\underline{\theta}_d,\overline{\theta}_d]^2 + \int_{[0,\underline{\theta}_d)\cup(\overline{\theta}_d,1]}(\theta^2 - c)dF(\theta) - E[\theta^2]\\
=&-Pr([\underline{\theta}_d,\overline{\theta}_d])Var([\underline{\theta}_d,\overline{\theta}_d]) - (1 - Pr([\underline{\theta}_d,\overline{\theta}_d]))c\\
=&-c + Pr([\underline{\theta}_d,\overline{\theta}_d])[c - Var([\underline{\theta}_d,\overline{\theta}_d])]\\
>&-c + Pr([\underline{\theta}_d,\overline{\theta}_d])[c - (\overline{\theta}_d - E[\underline{\theta}_d,\overline{\theta}_d])(E[\underline{\theta}_d,\overline{\theta}_d] - \underline{\theta}_d)]\\
\geq&-c
\end{aligned}$$

where the first inequality holds because of Bhatia–Davis inequality, and the second inequality holds because definition of $\overline{\theta}_d$ implies either $X^*_u(\overline{\theta}_d,\underline{\theta}_d) = E[\underline{\theta}_d,\overline{\theta}_d]$, which means $(\overline{\theta}_d - E[\underline{\theta}_d,\overline{\theta}_d])(E[\underline{\theta}_d,\overline{\theta}_d] - \underline{\theta}_d) = c$, or $\overline{\theta}_d - \underline{\theta}_d = 2\sqrt{c}$, which means $(\overline{\theta}_d - E[\underline{\theta}_d,\overline{\theta}_d])(E[\underline{\theta}_d,\overline{\theta}_d] - \underline{\theta}_d) \leq c$.
To show that $EU_{DM}(\Omega_d) > -Var(\Theta)$,

$$\begin{aligned}
&EU_{DM}(\Omega_d) = EU^U_{DM}(\Omega_d)\\
=&Pr([\underline{\theta}_d,\overline{\theta}_d])E[\underline{\theta}_d,\overline{\theta}_d]^2 + \int_{(\overline{\theta}_d}^1 X(\{\theta\},u)^2(f(\theta) + \dot{\phi}_d(\theta)f(\phi_d(\theta)))d\theta - E[\theta^2]\\
=&Pr([\underline{\theta}_d,\overline{\theta}_d])E[\underline{\theta}_d,\overline{\theta}_d]^2 + \int_{\overline{\theta}_d}^1 X(\{\theta\},u)^2(f(\theta) + \dot{\phi}_d(\theta)f(\phi_d(\theta)))d\theta - (Var(\Theta) + E[\Theta]^2)\\
=&Pr([\underline{\theta}_d,\overline{\theta}_d])(E[\underline{\theta}_d,\overline{\theta}_d] - E[\Theta])^2 + \int_{\overline{\theta}_d}^1 (X(\{\theta\},u) - E[\Theta])^2(f(\theta) + \dot{\phi}_d(\theta)f(\phi_d(\theta)))d\theta - Var(\Theta)\\
>&-Var(\Theta)
\end{aligned}$$

where the last equality holds because of sequential rationality of X, and the inequality holds because $c < \frac{1}{4}$ implies $2\sqrt{c} < 1 = 1 - \phi_d(1)$ and $E[\Theta] < \frac{1}{2} + \sqrt{\frac{1}{4} - c}$ implies $X^*_u(1,\phi_d(1)) > E[0,1]$, so $\overline{\theta}_d < 1$.
■

Proof of Proposition 6: For an optimal mechanism Ω where $E[\underline{\theta}_\Omega,\overline{\theta}_\Omega] > \frac{\underline{\theta}_\Omega+\overline{\theta}_\Omega}{2}$, by Lemma 13 for almost every $m \in \mathcal{M}_q^+$, $\Theta^t_q(m) = \{X(m,s)\}$ and $\Theta^l_q(m) = \{X(m,l)\}$ are two singleton sets, so there exist $\overline{\Theta} \subseteq [\overline{\theta}_\Omega, 1]$ where $Pr(\overline{\Theta}) = Pr([\overline{\theta}_\Omega, 1])$, $\underline{\Theta} \subseteq [0,\underline{\theta}_\Omega]$ where $Pr(\underline{\Theta}) = Pr([0,\underline{\theta}_\Omega])$ and a bijective matching function $\phi : \overline{\Theta} \to \underline{\Theta}$ such that for $m \in \overline{\Theta}$, $\Theta^t_q(m) = \{m\}$ and $\Theta^l_q(m) = \{\phi(m)\}$. By Lemma 14,

$$m_1 > m_2 \iff X(m_1,t) > X(m_2,t)$$

$$\iff X(m_1, l) < X(m_2, l) \iff \phi(m_1) < \phi(m_2)$$

so ϕ is a strictly decreasing function. Since Ω is optimal, $w_q(m) = w^-(m - \phi(m))$, so (53) implies that ϕ must the a solution of (19), so $\phi = \phi_d$ with almost identical domain.

Lemma 16 implies $(\overline{\theta}_\Omega - E[\underline{\theta}_\Omega, \overline{\theta}_\Omega])(E[\underline{\theta}_\Omega, \overline{\theta}_\Omega] - \underline{\theta}_\Omega) = c$, which combines with $E[\underline{\theta}_\Omega, \overline{\theta}_\Omega] > \frac{\underline{\theta}_\Omega + \overline{\theta}_\Omega}{2}$ implies that $X_u^*(\overline{\theta}_\Omega, \underline{\theta}_\Omega) = E[\underline{\theta}_\Omega, \overline{\theta}_\Omega]$. By Lemma 1 there exists a unique $\overline{\theta}_d$ such that $X_u^*(\overline{\theta}_d, \phi(\overline{\theta}_d)) = E[\phi_d(\overline{\theta}_d), \overline{\theta}_d]$, so $\overline{\theta}_\Omega = \overline{\theta}_d$ and $\underline{\theta}_\Omega = \phi(\overline{\theta}_d) = \underline{\theta}_d$.

Therefore, Ω and Ω_d are both characterized by the two cutoffs $\overline{\theta}_d, \underline{\theta}_d$ and matching function ϕ_d, which is defined for almost the same set of types, thus they are equal almost everywhere.

Q.E.D.

Proof of Proposition 7: We claim that for any $m, m' \in \mathcal{M}_q$ such that $P(m) < 1$ and $P(m') < 1$, $X(m,u) = X(m',u)$ and $P(m) = P(m')$. Suppose contrary to the claim, $X(m',u) > X(m,u)$. Then since $X(m,u) = E[\Theta_q^u(m)]$, there exists $\overline{\theta} \geq X(m,u)$ such that $m_q(\overline{\theta}) = m$. Since (30) implies $[1 - P(m)][X(m,u) - \overline{\theta}] \geq [1 - P(m')][X(m',u) - \overline{\theta}]$, thus $P(m) > P(m')$. Similarly, there exists $\underline{\theta} \leq X(m,u)$ such that $m_q(\underline{\theta}) = m$, and (30) implies $[1 - P(m)][X(m,u) - \underline{\theta}] \geq [1 - P(m')][X(m',u) - \underline{\theta}]$, thus $P(m) < P(m')$, contradiction. Therefore, it must be the case that $X(m,u) = X(m',u)$, which by (30) implies that $P(m) = P(m')$.

We claim that if there exists $\tilde{m} \in \mathcal{M}_q$ such that $P(\tilde{m}) = 1$, then $P(m) = 1$ for almost every $m \in \mathcal{M}_q$. Suppose Contrary to the claim, there exists positive measure subset $\hat{\Theta} \subseteq \Theta$ such that $P(m_q(\theta)) < 1$ for any $\theta \in \hat{\Theta}$, then the first claim implies that $X(m_q(\theta), u) = E[\hat{\Theta}]$, and since $Pr(\hat{\Theta}) > 0$, there exists $\theta \in \hat{\Theta}$ such that $\theta > E[\hat{\Theta}] = X(m_q(\theta), u)$, but then $[1 - P(m_q(\theta)][X(m_q(\theta), u) - \theta] < 0 = [1 - P(\tilde{m})][X(\tilde{m}, u) - \theta]$, $m_q(\theta)$ is not optimal for θ, contradiction.

The first two claims implies that $P(m) = \hat{P}$ for almost every $m \in \mathcal{M}_q$, where $hatP$ is a constant, and if $hatP < 1$, $X(m,u) = E[\Theta]$ for almost every $m \in \mathcal{M}_q$. Since $P(m) = P(m')$, (28) and (29) imply that $v_q(m) = V_q(m') = Var(\Theta_q^u(m)) = Var(\Theta_q^u(m')) = Var(\Theta)$. Therefore, If $c > Var(\Theta)$, then $c > v_q(m)$ and $P(m) = 0$ for almost every $m \in \mathcal{M}_q$, and the mechanism is uninformative; if $c < Var(\Theta)$, then $c < v_q(m)$ and $P(m) = 1$ for almost every $m \in \mathcal{M}_q$, the mechanism is state-verifying.

Q.E.D.

Proof of Proposition 8: Suppose $c \geq Var(\Theta)$, then $EU_{DM}^s = Var(\Theta)$. Consider an uninformative mechanism Ω under lie-detection technology, where $m_q(\theta) = m_q^0 = \Theta$ for any $\theta \in \Theta$ and $P(m_q^0) = 0$, $X(m_q^0, u) = E[\Theta]$. It is clear that such mechanism is incentive compatible, and $EU_{DM}(\Omega) = Var(\Theta) = EU_{DM}^s$. Therefore, $EU_{DM}(\Omega^*) \geq EU_{DM}(\Omega) = EU_{DM}^s$.//

Suppose $c < Var(\Theta)$, then $c < Var(\Theta) < (1 - E\Theta)(E(\Theta) - 0) \leq \frac{1}{4}$, where the second inequality holds by Bhatia–Davis inequality. Since $c < (1 - E\Theta)(E(\Theta) - 0)$, we have $\frac{1}{2} - \sqrt{\frac{1}{4} - c} < E(\Theta) < \frac{1}{2} + \sqrt{\frac{1}{4} - c}$, thus Assumption 1 is satisfied, and by Lemma 2, Ω_d is incentive compatible with $EU_{DM}(\Omega_d) < c$. Therefore, $EU_{DM}(\Omega^*) \geq EU_{DM}(\Omega_d) > c = EU_{DM}^s$.

Q.E.D.

Appendix B

This Appendix provides results that facilitate proofs in Appendix A.

Below are some definitions and terminologies that are useful for proving the results. For $d \in [2\sqrt{c}, 1]$, let

$$h(d) = \frac{w^-(d)}{1 - w^-(d)} \tag{54}$$

be the required liar to truth-teller ratio to maintain incentive for DM to inspect a message. It can be verified that $h(.)$ is a strictly decreasing and strictly convex function, with $\lim_{d \to 2\sqrt{c}} h'(d) = -\infty$ and $\lim_{d \to 2\sqrt{c}} h''(d) = +\infty$.

We say θ is **essentially revealed upon inspection** in Ω if $x^d_\Omega(\theta) = \theta$.

The following Proposition establishes the necessary and sufficient conditions of an incentive compatible mechanism.

Lemma 3 *Let q and X be a pair of message and action rule that satisfies (i) - (ii) of Proposition 1, X satisfies DM's sequential rationality (5) given q, and let m^0_q be the (potentially non-exist) uninspected message. Then there exists an inspection rule $P(.)$ on the set of inspected messages M^+_q such that (q, P, X) is incentive compatible if and only if for any $m, m' \in \mathcal{M}^+_q$:*

(a) $X(m, l) \leq X(m^0_q, u) < X(m', u)$ if m^0_q exists; $X(m, l) < X(m', u)$ otherwise;

(b) $w_q(m)(1-w_q(m))(X(m,t)-X(m,l))^2 \begin{cases} = c & \text{if } X(m,l) < X(m^0_q,u) \\ \geq c & \text{if } X(m,l) = X(m^0_q,u) \end{cases}$; If m^0_q does not exists, replace $X(m^0_q, u)$ with $\sup_{m' \in \mathcal{M}^+_q} X(m', l)$.

In particular, $P(m) = \frac{u(X(m,u)) - u(x^0_q)}{u(X(m,u)) - u(X(m,l))}$, where $x^0_q = X(m^0_q, u)$ if m^0_q exists; $x^0_q \in [\sup_{m' \in \mathcal{M}^+_q} X(m', l), \inf_{m'' \in \mathcal{M}^+_q} X(m'', u)]$ if m^0_q does not exist and $v_q(m) = c$ for all $m \in \mathcal{M}^+_q$; $x^0_q = \max_{m' \in \mathcal{M}^+_q} X(m', l)$ otherwise.

Proof of Lemma 3: Given (i) of Proposition 1 we have $X(m, t) \geq X(m, l)$, and for any $m \in \mathcal{M}^+_q$, incentive compatibility requires $X(m, t) > X(m, l)$, for otherwise value of inspection $V_q(m) = w_q(m)(1 - w_q(m))(X(m, t) - X(m, l))^2 = 0$, violating DM's sequential rationality. (5) then implies that $X(m, t) > X(m, u) > X(m, l)$ for any $m \in \mathcal{M}^+_q$. (i) of Proposition 1 then imply that if m^0_q exists, $P(m)u(X(m, l)) + (1 - P(m))u(X(m, u)) = P(m^0_q)u(X(m^0_q, l)) + (1 - P(m^0_q))u(X(m^0_q, u)) = u(X(m^0_q, u))$, which means $P(m) = \frac{u(X(m,u)) - u(X(m^0_q,u))}{u(X(m,u)) - u(X(m,l))}$. Since sender's utility $u(.)$ is strictly increasing, there exists such $P(m) \in (0, 1]$ if and only if $X(m, l) \leq X(m^0_q, u) < X(m, u)$, which holds for all $m, m' \in \mathcal{M}^+_q$, so $X(m, l) \leq X(m^0_q, u) < X(m', u)$. If $X(m, l) < X(m^0_q, u)$, it must be $P(m) \in (0, 1)$, so sequentially rational inspection requires $V_q(m) = c$; If $X(m,l) = X(m_q^0, u)$, $P(m) = 1$, sequentially rational inspection requires $V_q(m) \geq c$.

If m_q^0 does not exist, then for any $m, m' \in \mathcal{M}_q^+$, $P(m)u(X(m,l)) + (1 - P(m))u(X(m,u)) = P(m')u(X(m',l)) + (1 - P(m'))u(X(m',u)) = u(X(m',u))$, which can be achieved with some strictly positive $P(.)$ if and only if $\sup_{m' \in \mathcal{M}_q^+} X(m',l) < \inf_{m'' \in \mathcal{M}_q^+} X(m'',u)$. Sequentially rational inspection requires $V_q(m) = c$ for any m such that $P < 1$, which must be the case when $X(m,l) < \sup_{m' \in \mathcal{M}_q^+} X(m',l)$. $V_q(m) \geq c$ and $P(m) = 1$ is allowed if and only if $X(m,l) = \max_{m' \in \mathcal{M}_q^+} X(m',l)$. *Q.E.D.*[2mm]

Given an mechanism Ω, DM's expected payoff of Ω when every messages are ex-post uninspected and actions $X(m,u)$ are induced is

$$\begin{aligned} EU_{DM}^U(\Omega) &= \int_{\mathcal{M}_q^+} Pr(\Theta_q^u(m))X(m,u)^2 dm + Pr(\Theta_q^u(m_q^0))X(m_q^0,u)^2 - E[\theta^2] \\ &= \int_{[0,1]} x^2 dG_\Omega^u(x) - E[\theta^2] \end{aligned} \tag{55}$$

Where

$$G_\Omega^u(x) = \int_{\mathcal{M}_q} \int_{\Theta_q^u(m)} \mathbf{1}(X(m,u) \leq x) dF(\theta) dm \tag{56}$$

is the distribution of induced actions when messages are ex-post uninspected; DM's expected payoff when every messages in $\mathcal{M}_q^+$ are ex-post inspected and actions $X(m,t)$ ($X(m,l)$) are induced when sender is truthful (lying) is

$$\begin{aligned} EU_{DM}^I(\Omega) &= \int_{\mathcal{M}_q^+} Pr(\Theta_q^u(m))[w_q(m)X(m,l)^2 + (1 - w_q(m))X(m,t)^2 - c]dm \\ &\quad + Pr(\Theta_q^u(m_q^0))X(m_q^0,u)^2 - E[\theta^2] \\ &= (1 - Pr(\Theta_q^u(m_q^0))) \int_{[0,1]} (x^2 - c) dG_\Omega^i(x) + Pr(\Theta_q^u(m_q^0))X(m_q^0,u)^2 - E[\theta^2] \end{aligned} \tag{57}$$

where

$$G_\Omega^i(x) = \frac{1}{1 - Pr(\Theta_q^0)} \int_{\mathcal{M}_q^+} \sum_{s=t,l} \int_{\Theta_q^s(m)} \mathbf{1}(X(m,s) \leq x)] dF(\theta) dm \tag{58}$$

is the distribution of actions induced by messages with positive probability of inspection, when those messages are ex-post inspected.

Lemma 4 *Let Ω be an mechanism such that X satisfies (5) given q, and $V_q(m) = c$ almost everywhere for $m \in \mathcal{M}_q^+$, then $EU_{DM}(\Omega) = EU_{DM}^U(\Omega) = EU_{DM}^I(\Omega)$.*

Proof of Lemma 4: From equation (10),

$$
\begin{aligned}
& EU_{DM}(\Omega) \\
=& -\int_{\Theta}\int_{\mathcal{M}_q} q(m|\theta)[(1-P(m))(X(m,u)-\theta)^2 + P(m)\sum_{s=t,l}\mathbf{1}(\theta\in\Theta_s(m))[(X(m,s)-\theta)^2+c]]dmdF(\theta) \\
=& -\int_{\mathcal{M}_q^+}[(1-P(m))\int_{\Theta_q^u(m)}(X(m,u)-\theta)^2dF(\theta) + P(m)\sum_{s=t,l}\int_{\Theta_q^s(m)}[(X(m,s)-\theta)^2+c]dF(\theta)]dm \\
& -\int_{\Theta_q^u(m_q^0)}(X(m_q^0,u)-\theta)^2dF(\theta) \\
=& \int_{\mathcal{M}_q^+}[(1-P(m))\int_{\Theta_q^u(m)}dF(\theta)X(m,u)^2 + P(m)\sum_{s=t,l}\int_{\Theta_q^s(m)}dF(\theta)[X(m,s)^2-c]]dm \\
& +\int_{\Theta_q^u(m_q^0)}dF(\theta)X(m_q^0,u)^2 - E[\theta^2] \\
=& \int_{\mathcal{M}_q^+}Pr(\Theta_q^u(m))[(1-P(m))X(m,u)^2 + P(m)[w_q(m)X(m,l)^2+(1-w_q(m))X(m,t)^2-c]]dm \\
& + Pr(\Theta_q^u(m_q^0))X(m_q^0,u)^2 - E[\theta^2]
\end{aligned}
\tag{59}
$$

where the second equality holds because of (2) - (4), the third equality holds because (5) implies $-\int_{\Theta_q^s(m)}(X(m,s)-\theta)^2dF(\theta) = \int_{\Theta_q^s(m)}X(m,s)^2dF(\theta) - \int_{\Theta_q^s(m)}\theta^2dF(\theta)$ for each $(m,s)\in\mathcal{M}_q\times\{t,l,u\}$; the forth equality holds because $w_q(m)Pr(\Theta_q^u(m)) = Pr(\Theta_q^l(m))$ and $(1-w_q(m))Pr(\Theta_q^u(m)) = Pr(\Theta_q^t(m))$.

Since $X(m,u) = w_q(m)X(m,l)+(1-w_q(m))X(m,t)$, so for any $m\in\mathcal{M}_q^+$, $w_q(m)X(m,l)^2+(1-w_q(m))X(m,t)^2 - X(m,u)^2 = w_q(m)(1-w_q(m))(X(m,t)-X(m,l))^2 \equiv V_q(m)$. Therefore, $V_q(m)=c$ implies that

$$
X(m,u)^2 = w_q(m)X(m,l)^2 + (1-w_q(m))X(m,t)^2 - c \tag{60}
$$

holds almost everywhere for $m\in\mathcal{M}_q^+$, thus

$$
\begin{aligned}
EU_{DM}(\Omega) =& \int_{\mathcal{M}_q^+}Pr(\Theta_q^u(m))X(m,u)^2dm + Pr(\Theta_q^u(m_q^0))X(m_q^0,u)^2 - E[\theta^2] \\
=& EU_{DM}^U(\Omega)
\end{aligned}
$$

and

$$
\begin{aligned}
EU_{DM}(\Omega) =& \int_{\mathcal{M}_q^+}Pr(\Theta_q^u(m))[w_q(m)X(m,l)^2+(1-w_q(m))X(m,t)^2-c]dm \\
& + Pr(\Theta_q^u(m_q^0))X(m_q^0,u)^2 - E[\theta^2]
\end{aligned}
$$

$$=EU^I_{DM}(\Omega)$$

■

Lemma 5 *For any $\epsilon > 0$, suppose there is a mechanism Ω such that X satisfies (5) given q, $w_q(m) = w^-(X(m,t) - X(m,l))$ almost everywhere for $m \in \mathcal{M}_q^+$, $\sup_{m\in\mathcal{M}_q^+} X(m,l) < X(m_q^0, u) + \epsilon$ and $\inf_{m\in\mathcal{M}_q^+} X(m,u) > X(m_q^0,u) - \epsilon$, then there exists an incentive compatible mechanism $\hat{\Omega}$ such that $EU_{DM}(\hat{\Omega}) > EU_{DM}(\Omega) - 4\epsilon^2$.*

Proof of Lemma 5: Let $M^u = \{m \in \mathcal{M}_q^+ : X(m_q^0, u) - \epsilon < X(m,u) \leq X(m_q^0,u)\}$ be the set of inspected messages that violate $X(m_q^0,u) < X(m,u)$. Define $\hat{x}$ that solves

$$Pr(\Theta_q^0 \cup \Theta_q^u(M^u))(\hat{x} - E[\Theta_q^0 \cup \Theta_q^u(M^u)]) = \int_{m\in\mathcal{M}_q^+ : X(m,l)\geq\hat{x}} Pr(\Theta_q^l(m))(X(m,l) - \hat{x})dm \tag{61}$$

If $\sup_{m\in\mathcal{M}_q^+} X(m,l) > E[\Theta_q^0 \cup \Theta_q^u(M^u)]$, and $\hat{x} = E[\Theta_q^0 \cup \Theta_q^u(M^u)]$ otherwise. Let $M^l = \{m \in \mathcal{M}_q^+ : X(m,l) \geq \hat{x}\}$. Since X satisfies (5), we have

$$E[\Theta_q^0 \cup \Theta_q^u(M^u) \cup \Theta_q^l(M^l)] = \hat{x} \in (X(m_q^0,u) - \epsilon, X(m_q^0,u) + \epsilon) \tag{62}$$

and

$$X(m,u) > \hat{x} > X(m,l) \text{ for any } m \in M_q^+/(M^u \cup M^l) \tag{63}$$

Now for each $m \in M^l$ define $\hat{w} = w^-(X(m,t) - \hat{x})$ be the new required weight of liars given that the liar-induced action of m is $\hat{x}$. Since $w^-(.)$ is a decreasing function, $w_q(m) = w^-(X(m,t) - X(m,l))$ and $X(m,l) \geq \tilde{x}$ for $m \in M^l$, so $\hat{w}(m) \leq w_q(m)$. Let $\hat{p} = \int_{M^l} \frac{\hat{w}(m)}{1-\hat{w}(m)} \int_{\Theta_q^t(m)} dF(\theta)dm$ be the total mass of liars required to be pooled with truth-tellers in M^l, given that liar-induced actions are $\hat{x}$. $\hat{w}(m) \leq w_q(m)$ implies $\hat{p} \leq Pr(\Theta_q^l(M^l)) = \int_{M^l} \frac{w_q(m)}{1-w_q(m)} \int_{\Theta_q^t(m)} dF(\theta)dm.$

Let $\hat{\Theta} = \Theta_q^0 \cup \Theta_q^u(M^u) \cup \Theta_q^l(M^l)$ be the pool of modifying types. Let $\underline{z} = \frac{Pr(\hat{\Theta})-\hat{p}}{Pr(\hat{\Theta})}$. Assign an arbitrary strict ranking $r : M^l \to \mathbb{R}$ to the message set M^l. Then for any $m \in M^l$, let

$$z^-(m) = \underline{z} + \frac{1}{Pr(\hat{\Theta})} \int_{m'\in M^l : r(m')<r(m)} \frac{\hat{w}(m')}{1-\hat{w}(m')} \int_{\Theta_q^t(m')} dF(\theta)dm' \tag{64}$$

$$z^+(m) = z^-(m) + \frac{1}{Pr(\hat{\Theta})} \frac{\hat{w}(m)}{1-\hat{w}(m)} \int_{\Theta_q^t(m)} dF(\theta) \tag{65}$$

Define an modified messaging and action rules $\hat{q}, \hat{X}$ where other things remain unchanged, except the set of messages M^u is off-path; the uninspected message is modified to $m_{\hat{q}}^0 = \hat{\Theta}(\underline{z})$ with the set of sender identical to the statement, where $\hat{\Theta}(\underline{z})$ is a mean preserving division of $\hat{\Theta}$ so that $E[\hat{\Theta}(\underline{z}] = E[\hat{\Theta}] = \hat{x}$ and $Pr(\hat{\Theta}(\underline{z})) = \underline{z}Pr(\hat{\Theta}) = Pr(\hat{\Theta}) - \hat{p}$; For any $m \in M^l$, the set of truth-tellers remain unchanged, while the set of liars is modified to $\Theta^l_{\hat{q}}(m) = \hat{\Theta}(z^+(m))/int(\hat{\Theta}(z^-(m)))$, a mean preserving division of $\hat{\Theta}$ where $int(X)$ is the interior of set X, so that $E[\Theta^l_{\hat{q}}(m)] = \hat{x}$ and the set has measure $\frac{\hat{w}(m)}{1-\hat{w}(m)} \int_{\Theta^t_q(m)} dF(\theta)$.

The sequentially rational actions for the modified uninspected messages $m^0_{\hat{q}}$ is

$$\hat{X}(m^0_{\hat{q}}, u) = E[\hat{\Theta}(\underline{z})] = \hat{x} \tag{66}$$

and for $m \in M^l$,

$$\begin{aligned} \hat{X}(m,t) &= X(m,t) \\ \hat{X}(m,l) &= E[\Theta^l_{\hat{q}}(m)] = \hat{x} \\ \hat{X}(m,u) &= \hat{w}(m)\hat{x} + (1-\hat{w}(m))X(m,t) \end{aligned} \tag{67}$$

where $X(m,t) > X(m,l) \geq \hat{x}$. (63), (66) and (67) imply that

$$\hat{X}(m,u) > \hat{X}(m^0_{\hat{q}}, u) \leq \hat{X}(m,l) \text{ for any } m \in M^+_{\hat{q}} \tag{68}$$

so $(\hat{q}, \hat{X})$ satisfies (a) of Lemma 3. Furthermore, by the definition of $\Theta^l_{\hat{q}}(m)$ for $m \in M^l$, we have

$$w_{\hat{q}}(m) = \hat{w}(m) = w^-(\hat{X}(m,t) - \hat{X}(m,l)) \tag{69}$$

and thus

$$V_{\hat{q}}(m) = c \tag{70}$$

so $(\hat{q}, \hat{X})$ satisfies (b) of Lemma 3. Therefore, there exists $\hat{P}$ such that $\hat{\Omega} = (\hat{q}, \hat{P}, \hat{X})$ is incentive compatible.

To compare DM's ex-ante payoffs, since $V_q(m) = V_{\hat{q}}(m) = c$, (60) holds for both mechanisms, thus (59) implies

$$\begin{aligned} EU_{DM}(\Omega) = &\int_{\mathcal{M}^+_q/M^l} Pr(\Theta^u_q(m))X(m,u)^2 dm + Pr(\Theta^0_q)X(m^0_q,u)^2 - E[\theta^2] \\ &+ \int_{M^l} [Pr(\Theta^l_q(m))X(m,l)^2 + Pr(\Theta^t_q(m))X(m,t)^2 - Pr(\Theta^u_q(m))c]dm \end{aligned} \tag{71}$$

and

$$EU_{DM}(\hat{\Omega}) = \int_{\mathcal{M}^+_q/(M^l \cup M^u)} Pr(\Theta^u_q(m))X(m,u)^2 dm + Pr(\hat{\Theta}(\underline{z}))\hat{x}^2 - E[\theta^2]$$

$$+\int_{M^l}[Pr(\hat{\Theta}^l_q(m))\hat{x}^2 + Pr(\Theta^t_q(m))X(m,t)^2 - Pr(\hat{\Theta}^u_q(m))c]dm \quad (72)$$

so

$$\begin{aligned}
&EU_{DM}(\Omega) - EU_{DM}(\hat{\Omega}) \\
=&\int_{M^u} Pr(\Theta^u_q(m))X(m,u)^2dm + Pr(\Theta^0_q)X(m^0_q,u)^2 + \int_{M^l} Pr(\Theta^l_q(m))(X(m,l)^2 - c)dm \\
&- Pr(\hat{\Theta}(\underset{\sim}{z}))\hat{x}^2 - \int_{M^l} Pr(\hat{\Theta}^l_q(m))(\hat{x}^2 - c)dm \\
=&\int_{M^u} Pr(\Theta^u_q(m))X(m,u)^2dm + Pr(\Theta^0_q)X(m^0_q,u)^2 + \int_{M^l} Pr(\Theta^l_q(m))X(m,l)^2dm \\
&- Pr(\hat{\Theta})\hat{x}^2 + c\int_{M^l}[Pr(\hat{\Theta}^l_q(m)) - Pr(\Theta^l_q(m))]dm \\
\leq&\int_{M^u} Pr(\Theta^u_q(m))X(m,u)^2dm + Pr(\Theta^0_q)X(m^0_q,u)^2 + \int_{M^l} Pr(\Theta^l_q(m))X(m,l)^2dm - Pr(\hat{\Theta})\hat{x}^2 \\
=&\int_{M^u} Pr(\Theta^u_q(m))[X(m,u)^2 - \hat{x}^2]dm + \int_{M^l} Pr(\Theta^l_q(m))[X(m,l)^2 - \hat{x}^2]dm \\
&+ Pr(\Theta^0_q)[X(m^0_q,u)^2 - \hat{x}^2] \\
=&\int_{M^u} Pr(\Theta^u_q(m))[X(m,u) - \hat{x}]^2dm + \int_{M^l} Pr(\Theta^l_q(m))[X(m,l) - \hat{x}]^2dm \\
&+ Pr(\Theta^0_q)[X(m^0_q,u) - \hat{x}]^2 \\
<&4Pr(\hat{\Theta})\epsilon^2 \leq 4\epsilon^2
\end{aligned} \quad (73)$$

where the first inequality holds because $\int_{M^l} Pr(\hat{\Theta}^l_q(m)) - Pr(\Theta^l_q(m))dm = \hat{p} - Pr(\Theta^l_q(M^l)) \leq 0$; the third equality holds because $\hat{\Theta} = \Theta^0_q \cup \Theta^u_q(M^u) \cup \Theta^l_q(M^l)$; the fourth equality holds because (5) implies that $\int_{M^u} Pr(\Theta^u_q(m))X(m,u)dm + \int_{M^l} Pr(\Theta^l_q(m))X(m,l)dm + Pr(\Theta^0_q)X(m^0_q,u) = E[\hat{\Theta}] = \hat{x}$; the second inequality holds because $|X(m,s) - X(m^0_q)| < \epsilon$ for $m \in M^s$, $s \in \{u,l\}$ by definitions of M^l, M^u and $|\hat{x} - X(m^0_q)| < \epsilon$ by (62); the last inequality holds because $Pr(\hat{\Theta}) \leq 1$. ■

Lemma 6 *$\frac{dx^*_u(x_t,x_l)}{dx_t} > 0$ and $\frac{dx^*_u(x_t,x_l)}{dx_l} < 0$.*

Proof of Lemma 6:

$$\begin{aligned}
\frac{dx^*_u(x_t,x_l)}{dx_t} =&1 - w^-(x_t - x_l) - \frac{dw^-(x_t - x_l)}{dx_t}[x_t - x_l] \\
=&\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{c}{(x_t - x_l)^2}} + (\frac{1}{4} - \frac{c}{(x_t - x_l)^2})^{-0.5}\frac{c}{(x_t - x_l)^2} \\
>&0
\end{aligned}$$

$$\begin{aligned}\frac{dx_u^*(x_t,x_l)}{dx_l} =& w^-(x_t-x_l) - \frac{dw^-(x_t-x_l)}{dx_l}[x_t-x_l] \\ =& \frac{1}{2} - \sqrt{\frac{1}{4}-\frac{c}{(x_t-x_l)^2}} - (\frac{1}{4}-\frac{c}{(x_t-x_l)^2})^{-0.5}\frac{c}{(x_t-x_l)^2} \\ =& (\frac{1}{4}-\frac{c}{(x_t-x_l)^2})^{-0.5}[\frac{1}{2}\sqrt{\frac{1}{4}-\frac{c}{(x_t-x_l)^2}} - (\frac{1}{4}-\frac{c}{(x_t-x_l)^2}) - \frac{c}{(x_t-x_l)^2}] \\ =& (\frac{1}{4}-\frac{c}{(x_t-x_l)^2})^{-0.5}[\frac{1}{2}\sqrt{\frac{1}{4}-\frac{c}{(x_t-x_l)^2}} - \frac{1}{4}] \\ <& 0\end{aligned}$$

where the last inequality holds because $\sqrt{\frac{1}{4}-\frac{c}{(x_t-x_l)^2}} < \frac{1}{2}$. ■

Lemma 7 *Let* $(x_0, x_1, x_2, \hat{x}, \delta) \in [0,1]^5$*. If* $x_1 > \hat{x}$*,* $\hat{x} > \frac{x_1+x_2}{2}$ *and* $(x_1-x_0)(x_0-x_2)-(x_1-\hat{x})(\hat{x}-x_2) > \delta > 0$*, then* $\hat{x} > x_0 + \delta$*.*

Proof of Lemma 7: First we have

$$(\hat{x}-x_0)(\hat{x}+x_0-\mu_1-\mu_2) = (x_1-x_0)(x_0-x_2)-(x_1-\hat{x})(\hat{x}-x_2) > \delta \tag{74}$$

It must be the case that $\hat{x} > x_0$, for otherwise $x_0 \geq \hat{x} > \frac{x_1+x_2}{2}$ would implies that LHS of (74) is non-positive. $x_1 > \hat{x}$ implies that $\hat{x}+x_0-\mu_1-\mu_2)1$, therefore $\hat{x}-x_0 > \delta$. ■

For an Ω, let $\overline{\theta}^0 = inf\{\theta : Pr([0,\theta]\cap\Theta_q^0) = Pr(\Theta_q^0)\}$ and $\underline{\theta}^0 = sup\{\theta : Pr([\theta,1]\cap\Theta_q^0) = Pr(\Theta_q^0)\}$ be the probabilistic upper bound and lower bound of the set of uninspected types, $\mu = E[\Theta_q^0]$ be the mean of Θ_q^0.

Lemma 8 *In an optimal mechanism* Ω*,* $(\overline{\theta}^0-\mu)(\mu-\underline{\theta}^0) \leq c$ *.*

Proof of Lemma 8: Suppose contrary to the claim, $(\overline{\theta}^0-\mu)(\mu-\underline{\theta}^0) > c$, then there exists $\delta > 0$ such that $(\overline{\theta}^0-\mu)(\mu-\underline{\theta}^0) - c > \delta$, and for any small enough $\epsilon > 0$ there exist $\Theta_1, \Theta_2 \subseteq \Theta_q^0$ such that $Pr(\Theta_1) = Pr(\Theta_2) = \epsilon$ and

$$(\mu_1-\mu)(\mu-\mu_2) - c > \delta \tag{75}$$

where $\mu_1 \equiv E[\Theta_1]$ and $\mu_2 \equiv E[\Theta_2]$. Let $\hat{w} = w^-(\mu_1-\mu_2)$, where $w^-(.)$ is the minimum liar weight function defined in (15), which is well defined because (75) implies that $\mu_1-\mu_2 > 2\sqrt{c}$. Create a mean-preserving division of Θ_2, $\Theta_2(\frac{\hat{w}}{1-\hat{w}})$, so $E[\Theta_2(\frac{\hat{w}}{1-\hat{w}})] = E[\Theta_2]$ and $Pr(\Theta_2(\frac{\hat{w}}{1-\hat{w}})) = \frac{\hat{w}}{1-\hat{w}}Pr(\Theta_2)$.

Now define an modified mechanism $\hat{\Omega} = (\hat{q}, \hat{P}, \hat{X})$ where other things remain unchanged, except the uninspected message is modified to $m_{\hat{q}}^0 = \Theta_{\hat{q}}^0 \equiv \Theta_q^0/(\Theta_1 \cup \Theta_2(\frac{\hat{w}}{1-\hat{w}}))$, with the set of sender identical to the statement $\Theta_{\hat{q}}^0$. An inspected message $\hat{m} = \Theta_1$ is added, with truthful senders $\Theta_{\hat{q}}^t(\hat{m}) = \Theta_1$ and lying senders $\Theta_{\hat{q}}^l(\hat{m}) = \Theta_2(\frac{\hat{w}}{1-\hat{w}})$.

The sequentially rational actions for the modified message $\hat{m}$ are

$$\hat{X}(\hat{m}, t) = E[\Theta_1] = \mu_1$$
$$\hat{X}(\hat{m}, l) = E[\Theta_2(\frac{\hat{w}}{1-\hat{w}})] = \mu_2$$
$$\hat{X}(\hat{m}, u) = E[\Theta_1 \cup \Theta_2(\frac{\hat{w}}{1-\hat{w}})] = \hat{w}\mu_2 + (1-\hat{w})\mu_1$$

the last equality holds because $\frac{Pr(\Theta_2(\frac{\hat{w}}{1-\hat{w}}))}{Pr(\Theta_2(\frac{\hat{w}}{1-\hat{w}}))+Pr(\Theta_1)} = \frac{Pr(\Theta_2(\frac{\hat{w}}{1-\hat{w}}))}{Pr(\Theta_2(\frac{\hat{w}}{1-\hat{w}}))+Pr(\Theta_2)} = \hat{w}$. The sequentially rational action for the modified message $m_{\hat{q}}^0$ is

$$\hat{X}(m_{\hat{q}}^0, u) = E[\Theta_q^0/(\Theta_1 \cup \Theta_2(\frac{\hat{w}}{1-\hat{w}}))]$$

The definition of $\hat{w}$ implies that $V_q(\hat{m}) = \hat{w}(1-\hat{w})(\mu_1 - \mu_2)^2 = c$, so (b) of Lemma 3 is satisfied for $\hat{m}$. The above equality also implies

$$(\mu_1 - \hat{X}(\hat{m}, u))(\hat{X}(\hat{m}, u) - \mu_2) = c \tag{76}$$

and since $\hat{w} < \frac{1}{2}$, we have $\hat{X}(\hat{m}, u) > \frac{\mu_1+\mu_2}{2}$, so Lemma 7, (75) and (76) imply

$$\hat{X}(\hat{m}, u) > \mu + \delta \tag{77}$$

By sequential rationality,

$$Pr(\Theta_1 \cup \Theta_2(\frac{\hat{w}}{1-\hat{w}}))\hat{X}(\hat{m}, u) + Pr(\Theta_q^0/(\Theta_1 \cup \Theta_2(\frac{\hat{w}}{1-\hat{w}})))\hat{X}(m_{\hat{q}}^0, u) = Pr(\Theta_q^0)\mu \tag{78}$$

where $Pr(\Theta_1 \cup \Theta_2(\frac{\hat{w}}{1-\hat{w}})) = \epsilon + \frac{\hat{w}}{1-\hat{w}}\epsilon = \frac{\epsilon}{1-\hat{w}}$. Rearranging (78) yields

$$\begin{aligned}\hat{X}(m_{\hat{q}}^0, u) =& \mu_0 - \frac{\epsilon}{(1-\hat{w})(Pr(\Theta_q^0) - \frac{\epsilon}{1-\hat{w}})}(\hat{X}(\hat{m}, u) - \mu) \\ >& \mu_0 - \frac{2\epsilon}{Pr(\Theta_q^0) - 2\epsilon}\end{aligned} \tag{79}$$

because $\hat{w} < 0.5$ and $\hat{X}(\hat{m}, u) - \mu < 1$. Now we have

$$\hat{X}(m_{\hat{q}}^0, u) < \mu < \hat{X}(\hat{m}, u) \tag{80}$$

and for any unmodified on-path message $m \in \mathcal{M}_q^+$, incentive compatibility of the original mechanism means $\hat{X}(m, u) = X(m, u) > \mu > \hat{X}(m_{\hat{q}}^0, u)$, so

$$\inf_{m \in M_{\hat{q}}^+} \hat{X}(m, u) > \hat{X}(m_{\hat{q}}^0, u) \tag{81}$$

Since $\hat{X}(\hat{m},l)=\mu_2<\hat{w}\mu_2+(1-\hat{w})\mu_1=\hat{X}(m^0_{\hat{q}},u)$ and $\sup_{m\in\mathcal{M}^+_q}X(m,l)\le\mu$ by incentive compatibility of the original mechanism, so (79) implies

$$\sup_{m\in M^+_{\hat{q}}} X(m,l)<\hat{X}(m^0_{\hat{q}},u)+\frac{2\epsilon}{Pr(\Theta^0_q)-2\epsilon} \tag{82}$$

Since the original mechanism is optimal, $w_q(m)=w^-(X(m,t)-X(m,l))$ hold almost everywhere for unmodified message $m\in\mathcal{M}^+_q$. For the modified message, $w_{\hat{q}}(\hat{m})=w^-(\mu_1-\mu_2)=w^-(\hat{X}(\hat{m},t)-\hat{X}(\hat{m},l))$, so we have

$$w_{\hat{q}}(m)=w^-(\hat{X}(m,t)-\hat{X}(m,l)) \tag{83}$$

hold almost everywhere at $\mathcal{M}^+_{\hat{q}}$.

To compare DM's payoffs, since $V_q(m)=V_{\hat{q}}(m)=c$, by Lemma 4, $EU_{DM}(\Omega)=EU^U_{DM}(\Omega)$ and $EU_{DM}(\hat{\Omega})=EU^U_{DM}(\hat{\Omega})$, so

$$\begin{aligned}
&EU_{DM}(\hat{\Omega})-EU_{DM}(\Omega)\\
=&\frac{\epsilon}{1-\hat{w}}\hat{X}(\hat{m},u)^2+[Pr(\Theta^0_q)-\frac{\epsilon}{1-\hat{w}}]\hat{X}(m^0_{\hat{q}},u)^2-Pr(\Theta^0_q)(\mu)^2\\
=&\frac{\epsilon}{1-\hat{w}}[\hat{X}(\hat{m},u)-\mu]^2+[Pr(\Theta^0_q)-\frac{\epsilon}{1-\hat{w}}][\hat{X}(m^0_{\hat{q}},u)-\mu]^2\\
>&\frac{\epsilon}{1-\hat{w}}\delta^2>2\delta^2\epsilon
\end{aligned} \tag{84}$$

where the second equality holds because $\frac{\epsilon}{1-\hat{w}}\hat{X}(\hat{m},u)+[Pr(\Theta^0_q)-\frac{\epsilon}{1-\hat{w}}]\hat{X}(m^0_{\hat{q}},u)-Pr(\Theta^0_q)(\mu)=0$ by (78); the first inequality holds because of (77), and the last inequality hold because $\hat{w}<0.5$. Finally, given (81) - (83), Lemma 5 implies that there exist an incentive compatible mechanism $\tilde{\Omega}$ such that $EU_{DM}(\tilde{\Omega})>EU_{DM}(\hat{\Omega})-4(\frac{2\epsilon}{Pr(\Theta^0_q)-2\epsilon})^2$, then by (84) $EU_{DM}(\tilde{\Omega})-EU_{DM}(\Omega)>2\delta^2\epsilon-4(\frac{2\epsilon}{Pr(\Theta^0_q)-2\epsilon})^2>0$ for small enough ϵ, but it contradicts that Ω is an optimal mechanism. Therefore, we conclude that $(\overline{\theta}^0-\mu)(\mu-\underline{\theta}^0)\le c$ in an optimal mechanism. ■

For an mechanism Ω such that $Pr(\Theta^0_q)>0$, let $\underline{\theta}^t=\sup\{\theta:Pr([\theta,1]\cap\Theta^t_q(\mathcal{M}^+_q))=Pr(\Theta^t_q(\mathcal{M}^+_q)\}$ and $\overline{\theta}^l_q=sup\{\theta:Pr([0,\theta]\cap\Theta^l(\mathcal{M}^+_q))=Pr(\Theta^l_q(\mathcal{M}^+_q))\}$ be the probabilistic lower bound of the set of inspected truthful types and probabilistic upper bound of the set of inspected lying types.

Lemma 9 *In an optimal mechanism Ω, $\underline{\theta}^t\ge\overline{\theta}^0$.*

Proof of Lemma 9: Suppose contrary to the claim, $\underline{\theta}^t<\overline{\theta}^0$, then there exists $\delta>0$ such that $\underline{\theta}^t<\overline{\theta}^0-\delta$. Let $M_t=\{m\in\mathcal{M}^+_q:\exists\theta\in\Theta^t_q(m)$ such that $\theta<\overline{\theta}^0-\delta\}$ be the set of messages containing truthful types below $\overline{\theta}^0 - \delta$, and $\Theta_t = \{\theta : \exists m \in M_t \text{ such that } \theta \in \Theta_q^t(m), \theta < \overline{\theta}^0 - \delta \text{ and } \theta \leq X(m,t)\}$ be the set of such truthful types with a weaker higher induced action. We have $Pr(\Theta_t) > 0$ and there exist positive measure subset of uninspected types $\Theta_0 \subseteq \Theta_q^0$ such that for any $\theta_t \in \Theta_t$ and $\overline{\mu} \equiv E[\Theta_0]$,

$$\overline{\mu} - \theta_t > \delta \tag{85}$$

Denote $a = \frac{Pr(\Theta_t)}{Pr(\Theta_q^t(M_t))} > 0$, and let $M_t^a = \cap M \subseteq M_t : \frac{Pr(\Theta_t \cap \Theta_q^t(M))}{Pr(\Theta_q^t(M))} \geq a$ be the subset containing every message in M_t where proportion of truthful types within Θ_t is no less than a. Since $Pr(M_t^a) > 0$, for any $\epsilon > 0$ there is positive measure subset $M' \subseteq M_t^a$ such that for any $m, m' \in M'$ and $s \in \{t, l, u\}$,

$$|X(m,s) - X(m',s)| < \epsilon \tag{86}$$

and $Pr(M') \leq \epsilon$ (If every $m \in M_t^a$ is such that $Pr(m) > \epsilon$, then we can take any of such message m and divide it into two messages m_1, m_2 with exact same induced actions and $Pr(m_1) = \epsilon$, $Pr(m_2) = Pr(m) - \epsilon$ and take $M' = \{m_1\}$). Therefore, M' satisfies $Pr(M') \in (0, \epsilon)$, $\frac{Pr(\Theta' \cap \Theta_q^t(M'))}{Pr(\Theta_q^t(M'))} \geq a$.
Let $\Theta' = \Theta_q^t(M') \cap \Theta_t$ be the set of truthful types in M' that satisfies (85) and $\theta < X(m,t)$ for $\theta \in \Theta_q^t(m) \cap \Theta'$. Since $M' \subseteq M_t^a$, we have

$$Pr(\Theta') \geq a Pr(\Theta_q^t(M')) \tag{87}$$

Let $\Theta^l = \Theta_q^l(M')$, $\Theta^t = \Theta_q^t(M')$ and $\Theta^u = \Theta^l \cup \Theta^t$ be the aggregate set of truth-tellers, liars and senders of M' ;$E^u = E[\Theta^u]$, $E^t = E[\Theta^t]$, $E^l = E[\Theta^l]$, $E' = E[\Theta']$ be their corresponding expected values, and $z^t = \frac{Pr(\Theta^t)}{Pr(\Theta^u)}$, $z^l = \frac{Pr(\Theta^l)}{Pr(\Theta^u)} = 1 - z^t$, $z' = \frac{Pr(\Theta')}{Pr(\Theta^u)}$ be their corresponding ratios of measure to set of senders Θ^u.Since $\Theta' \subseteq \Theta_t$,

$$E' < \overline{\mu} - \delta \tag{88}$$

and

$$E' \leq E^t \tag{89}$$

Since the original mechanism is optimal, we have $(X(m,t) - X(m,u))(X(m,u) - X(m,l)) = c$ and $w_q(m) = w^-(X(m,t) - X(m,l)) \leq 0.5$ almost everywhere at M', so $|E^s - X(m,s)| < \epsilon$ for any $m \in M'$ and $s = t, l, u$ imply

$$|(1 - z^t)z^t(E^t - E^l)^2 - c| \equiv |(E^t - E^u)(E^u - E^l) - c| < 4\epsilon^2 \tag{90}$$

$$z^t \geq 0.5 \tag{91}$$

For any $z \in [0, \infty)$, let $E^t(z) = \frac{z^t E^t - z' E' + z\overline{\mu}}{z^t - z' + z}$ be the expected value of the set $(\Theta^t_q(M')/\Theta') \cup \Theta_z$, where Θ_z is a set with expected value $\overline{\mu}$ and measure $zPr(\Theta^u)$, and $E^u(z) = \frac{z^t E^t - z' E' + z\overline{\mu} + (1-z^t)E^l}{1 - z' + z}$ be the expected values of $(\Theta^u_q(M')/\Theta') \cup \Theta_z$. Define $\hat{z}$ that solves

$$\frac{(1-z^t)(z^t - z' + z)}{(1-z'+z)^2}(E^t(z) - E^l)^2 \equiv (E^t(z) - E^u(z))(E^u(z) - E^l) = c \tag{92}$$

$$z^t + z - z' \geq 1 - z^t \tag{93}$$

By definition of $E^t(z)$,

$$\begin{aligned} E^t(z) - E^t =& \frac{z}{z^t}(\overline{\mu} - E') - \frac{z - z'}{z^t}(E^t(z) - E') \\ \geq& \frac{z}{z^t}\delta - \frac{z - z'}{z^t}(E^t(z) - E') \\ \geq& a\delta - \frac{z - z'}{z^t}(E^t(z) - E') \end{aligned} \tag{94}$$

where the first inequality holds by (88)., the second inequality holds by (87). When $z = z'$, $E^t(z) - E^t \geq a\delta$, and $(1-z^t)z^t = \frac{1-z^t)(z^t-z'+z)}{(1-z'+z)^2}$, so for small enough ϵ, $\frac{(1-z^t)(z^t-z'+z)}{(1-z'+z)^2}(E^t(z) - E^l)^2 > (1-z^t)z^t(E^t - E^l + a\delta)^2 > (1-z^t)z^t(E^t - E^l)^2 - 4\epsilon^2 > c$; On the other hand, $\lim_{z\to\infty} \frac{(1-z^t)(z^t-z'+z)}{(1-z'+z)^2}(E^t(z) - E^l)^2 \to 0 < c$. Therefore, for small enough ϵ there exists $\hat{z} \in (z', \infty)$ that is solution to (92) and (113).

We claim that for small enough ϵ, $E^t(\hat{z}) - E^t \geq \epsilon^{\frac{1}{3}}$. Suppose $E^t(\hat{z}) - E^t < \epsilon^{\frac{1}{3}}$, then (94) implies $a\delta - \frac{\hat{z}-z'}{z^t}(E^t(\hat{z}) - E') < \epsilon^{\frac{1}{3}}$, and since $E^t(\hat{z}) - E' \leq 1$, we have $\hat{z} - z' > z^t(a\delta - \epsilon^{\frac{1}{3}})$. Now $(1-z^t)z^t - \frac{(1-z^t)(z^t-z'+\hat{z})}{(1-z'+\hat{z})^2} = \frac{(1-z^t)(\hat{z}-z')[(\hat{z}-z')z^t+2z^t-1]}{(1-z'+\hat{z})^2} > \frac{(1-z^t)z^t}{(1-z'+z)^2}(\hat{z} - z')^2 > \frac{(1-z^t)z^t}{(1+z^t(a\delta-\epsilon^{\frac{1}{3}}))^2}(z^t(a\delta - \epsilon^{\frac{1}{3}}))^2$; and $(E^t(\hat{z}) - E^l)^2 - (E^t - E^l)^2 < (E^t - E^l + \epsilon^{\frac{1}{3}})^2 - (E^t - E^l)^2 = 2\epsilon^{\frac{1}{3}}(E^t - E^l) - \epsilon^{\frac{2}{3}}$. Therefore, for small enough ϵ, $\frac{(1-z^t)(z^t-z'+z)}{(1-z'+z)^2}(E^t(z) - E^l)^2 < [(1-z^t)z^t - \frac{(1-z^t)z^t}{(1+z^t(a\delta-\epsilon^{\frac{1}{3}}))^2}(z^t(a\delta - \epsilon^{\frac{1}{3}}))^2][(E^t - E^l)^2 + 2\epsilon^{\frac{1}{3}}(E^t - E^l) - \epsilon^{\frac{2}{3}}] < (1-z^t)z^t(E^t - E^l)^2 - 4\epsilon^2 < c$, but it contradicts to (92), so for small enough ϵ, we have

$$E^t(\hat{z}) - E^t \geq \epsilon^{\frac{1}{3}} \tag{95}$$

and by definition of $E^u(z)$ and E^u, we have $E^u(\hat{z}) - E^u = \hat{z}(\overline{\mu} - E') - (\hat{z} - z')(E^u(z) - E') = z^t(E^t(\hat{z}) - E^t) + (\hat{z} - z')(E^t(\hat{z}) - E^u(\hat{z})) > z^t(E^t(\hat{z}) - E^t)$, then (91) and (95 imply

$$E^u(\hat{z}) - E^u \geq \frac{1}{2}\epsilon^{\frac{1}{3}} \tag{96}$$

Define a modified truth-tellers set $\hat{\Theta}^t = (\Theta^t_q(M')/\Theta') \cup \Theta_0(\frac{Pr(\Theta^u)}{Pr(\Theta_0)}\hat{z})$, where $\Theta_0(\frac{Pr(\Theta^u)}{Pr(\Theta_0)}\hat{z})$ is a mean-preserving division of Θ_0 so that $E[\Theta_0(\frac{Pr(\Theta^u)}{Pr(\Theta_0)}\hat{z})] = \overline{\mu}$ and $Pr(\Theta_0(\frac{Pr(\Theta^u)}{Pr(\Theta_0)}\hat{z})) = \hat{z}Pr(\Theta^u)$. Let $\hat{\Theta}^u =$

$\hat{\Theta}^t \cup \Theta^l$ be the modified set of senders. As a results, we have

$$E[\hat{\Theta}^t] = \frac{z^t Pr(\Theta^u)E^t - z' Pr(\Theta^u)E' + \hat{z} Pr(\Theta^u)\overline{\mu}}{(z^t - z' + \hat{z})Pr(\Theta^u)} = E^t(\hat{z}) \tag{97}$$

$$Pr(\hat{\Theta}^t) = (z^t + \hat{z} - z')Pr(\Theta^u) = Pr(\Theta^t) + (\hat{z} - z'))Pr(\Theta^u) \tag{98}$$

$$E[\hat{\Theta}^u] = E^u(\hat{z}) \tag{99}$$

$$Pr(\hat{\Theta}^u) = Pr(\Theta^u) + (\hat{z} - z'))Pr(\Theta^u) \tag{100}$$

Define the modified uninspected set $\Theta^0_{\hat{q}} = (\Theta^0_q / \cup \Theta')/\Theta_0(\frac{Pr(\Theta^u)}{Pr(\Theta_0)}\hat{z})$, so that

$$Pr(\Theta^0_{\hat{q}}) = Pr(\Theta^0_q) - (\hat{z} - z'))Pr(\Theta^u) \tag{101}$$

and since $\Theta^0_{\hat{q}} \cup \hat{\Theta}^u = \Theta^0_q \cup \Theta^u$, we have

$$Pr(\Theta^0_{\hat{q}})E[\Theta^0_{\hat{q}}] + Pr(\hat{\Theta}^u)E^u(\hat{z}) = Pr(\Theta^0_q)E[\Theta^0_q] + Pr(\Theta^u)E^u \tag{102}$$

and thus

$$\begin{aligned} E[\Theta^0_{\hat{q}}] - E[\Theta^0_q] &= \frac{Pr(\Theta^u)}{Pr(\Theta^0_q)}(E^u(\hat{z}) - E^u) + \frac{(\hat{z} - z')Pr(\Theta^u)}{Pr(\Theta^0_q)}(E^u(\hat{z}) - E[\Theta^0_{\hat{q}}]) \\ &\leq \frac{1 + z^{max} - z'}{Pr(\Theta^0_q)} Pr(\Theta^u) \end{aligned} \tag{103}$$

where the inequality holds for z^{max} equals the larger root of $\frac{(1-z^t)(z^t - z' + z)}{(1-z'+z)^2} = c$, so (92) implies $\hat{z} \leq z^{max}$.

Now define an modified message and action rules $\hat{q}, \hat{X}$ where other things remain unchanged, except the set of messages M' is off-path; The uninspected message is modified to $m^0_{\hat{q}} = \Theta^0_{\hat{q}}$ with the set of senders identical to the statement, and an message $\hat{m} = \hat{\Theta}^t$ is added with the set of truthful senders $\Theta^t_{\hat{q}}(\hat{m}) = \hat{\Theta}^t$, and the set of lying senders $\Theta^l_{\hat{q}}(\hat{m}) = \Theta^l$.

The sequentially rational actions for the modified message $\hat{m}$ are

$$\hat{X}(\hat{m}, t) = E[\hat{\Theta}^t] = E^t(\hat{z})$$

$$\hat{X}(\hat{m}, l) = E[\Theta^l] = E^l$$

$$\hat{X}(\hat{m}, u) = E[\hat{\Theta}^t \cup \Theta^l) = E^u(\hat{z})$$

The sequentially rational action for the modified uninspected message $m^0_{\hat{q}}$ is

$$\hat{X}(m^0_{\hat{q}}, u) = E[\Theta^0_{\hat{q}}]$$

For small enough ϵ, $\hat{X}(\hat{m},u) \geq E^t + \epsilon^{\frac{1}{3}} > \inf_{m\in M'} X(m,u) - \epsilon + \epsilon^{\frac{1}{3}} > \inf_{m\in M'} X(m,u)$, and for any unmodified on-path message $m \in \mathcal{M}_q^+$, incentive compatibility of the original mechanism means $\hat{X}(m,u) = X(m,u) > \mu > \hat{X}(m^0_{\hat{q}},u)$, so

$$\inf_{m\in M^+_{\hat{q}}} \hat{X}(m,u) > \hat{X}(m^0_{\hat{q}},u) \tag{104}$$

Since $\hat{X}(\hat{m},l) = E^l \leq \sup_{m\in M'} X(m,l)$ and $\sup_{m\in \mathcal{M}^+_q} X(m,l) \leq \mu$ by incentive compatibility of the original mechanism, so (103) implies

$$\sup_{m\in M^+_{\hat{q}}} X(m,u) < \hat{X}(m^0_{\hat{q}},u) + \frac{1+z^{max}-z'}{Pr(\Theta^0_q)} Pr(\Theta^u) \tag{105}$$

Since the original mechanism is optimal, $w_q(m) = w^-(X(m,t) - X(m,l))$ hold almost everywhere for unmodified message $m \in \mathcal{M}^+_q/M'$. For the modified message, (92) and (113) imply $w_{\hat{q}}(\hat{m}) = w^-(E^t(\hat{z}) - E^l) = w^-(\hat{X}(\hat{m},t) - \hat{X}(\hat{m},l))$, so we have

$$w_{\hat{q}}(m) = w^-(\hat{X}(m,t) - \hat{X}(m,l)) \tag{106}$$

hold almost everywhere at $\mathcal{M}^+_{\hat{q}}$.

To compare DM's payoffs, since $V_q(m) = V_{\hat{q}}(m) = c$, by Lemma 4, $EU_{DM}(\Omega) = EU^U_{DM}(\Omega)$ and $EU_{DM}(\hat{\Omega}) = EU^U_{DM}(\hat{\Omega})$, so

$$\begin{aligned}
&EU_{DM}(\hat{\Omega}) - EU_{DM}(\Omega)\\
=&Pr(\Theta^0_{\hat{q}})E[\Theta^0_{\hat{q}}]^2 + Pr(\hat{\Theta}^u)E^u(\hat{z})^2 - Pr(\Theta^0_q)E[\Theta^0_q]^2 - Pr(\Theta^u)(E^u)^2\\
=&-Pr(\Theta^0_q)[(E[\Theta^0_q] - E[\Theta^0_{\hat{q}}])^2 + 2E[\Theta^0_{\hat{q}}](E[\Theta^0_q] - E[\Theta^0_{\hat{q}}])]\\
&+Pr(\Theta^u)[(E^u(\hat{z}) - E^u)^2 + 2E^u(E^u(\hat{z}) - E^u)]\\
&+(\hat{z}-z')Pr(\Theta^u)[(E^u(\hat{z}) - E[\Theta^0_{\hat{q}}])^2 + 2E[\Theta^0_{\hat{q}}](E^u(\hat{z}) - E[\Theta^0_{\hat{q}}])]\\
=&-Pr(\Theta^0_q)(E[\Theta^0_q] - E[\Theta^0_{\hat{q}}])^2 + Pr(\Theta^u)(E^u(\hat{z}) - E^u)^2 + (\hat{z}-z')Pr(\Theta^u)(E^t(\hat{z}) - E[\Theta^0_{\hat{q}}])^2\\
>&-Pr(\Theta^0_q)(\frac{1+z^{max}-z'}{Pr(\Theta^0_q)}Pr(\Theta^u))^2 + Pr(\Theta^u)\frac{1}{4}\epsilon^{\frac{2}{3}}
\end{aligned} \tag{107}$$

Where the second equality holds by (100) and (101); the third equality holds by (102) and $E^u(\hat{z}) > E^u > E[\Theta^0_{\hat{q}}]$; the inequality holds by (96) and (103) for small enough ϵ.

Finally, given (104) - (106), Lemma 5 implies that there exists an incentive compatible mechanism $\tilde{\Omega}$ such that $EU_{DM}(\tilde{\Omega}) > EU_{DM}(\hat{\Omega}) - 4(\frac{1+z^{max}-z'}{Pr(\Theta^0_q)}Pr(\Theta^u))^2$, then by (107), for small enough ϵ,

$$EU_{DM}(\tilde{\Omega}) - EU_{DM}(\Omega)$$

$$>Pr(\Theta^u)[\frac{1}{4}\epsilon^{\frac{2}{3}} - Pr(\Theta_q^0)(\frac{1+z^{max}-z'}{Pr(\Theta_q^0)})^2 Pr(\Theta^u) - 4(\frac{1+z^{max}-z'}{Pr(\Theta_q^0)})^2 Pr(\Theta^u)]$$
$$\geq Pr(\Theta^u)[\frac{1}{4}\epsilon^{\frac{2}{3}} - Pr(\Theta_q^0)(\frac{1+z^{max}-z'}{Pr(\Theta_q^0)})^2 \epsilon - 4(\frac{1+z^{max}-z'}{Pr(\Theta_q^0)})^2 \epsilon] > 0$$

where the second inequality holds because $Pr(\Theta_q^0) = Pr(M') \leq \epsilon$, but it contradicts that Ω is an optimal mechanism. Therefore, we conclude that $\underline{\theta}^t) \geq \overline{\theta}^0$ in an optimal mechanism. ■

Lemma 10 *In an optimal mechanism Ω, $w_q(m) < 0.5$ almost everywhere for $m \in \mathcal{M}_q^+$.*

Proof of Lemma 10: Suppose Contrary to the claim, there exist a positive measure set of messages $M \subseteq \mathcal{M}_q^+$ such that for any $m \in M$, $w_q(m) = 0.5$, i.e. $X(m,u) = \frac{X(m,t)+X(m,l)}{2}$, then $V_q(m) = (X(m,t)-X(m,u))(X(m,u)-X(m,l)) = c$ implies that $X(m,t)-X(m,u) = X(m,u)-X(m,l) = \sqrt{c}$. By Lemma 3, $X(m,u) > X(m_q^0,u) \equiv \mu$, so $X(m,l) > \mu - \sqrt{c}$. Therefore, there exists $\delta > 0$ and a positive measure set of message $M' \subseteq \mathcal{M}_q^+$ such that for any $m \in M'$,

$$(\mu - X(m,l))^2 < c - \delta \tag{108}$$

Now we consider two cases.

Case 1: there is a message $m' \in M'$ such that $Pr(m') \equiv a > 0$.

Let $\Theta^l = \Theta_q^l(m')$, $\Theta^t = \Theta_q^t(m')$ and $\Theta^u = \Theta^l \cup \Theta^t$ be the set of truth-tellers, liars and senders of m' ;$E^u = E[\Theta^u]$, $E^t = E[\Theta^t]$, $E^l = E[\Theta^l]$, be their corresponding expected values. Then we have

$$Pr(\Theta^l) = Pr(\Theta^t) = \frac{a}{2} \tag{109}$$
$$E^t - E^u = E^u - E^l = \sqrt{c} \tag{110}$$
$$(\mu - E^l)^2 < c - \delta \tag{111}$$

For $\epsilon \in (0,1]$ and $s \in \{t,l\}$, let $\Theta_\epsilon^s = \Theta^s/\Theta^s(1-\epsilon)$ the outer ring mean-preserving division of Θ^s so that $E[\Theta_\epsilon^s] = E^s$, $Pr(\Theta_\epsilon^s) = \frac{a}{2}\epsilon$. Since Θ_ϵ^s and $\Theta^s/\Theta_\epsilon^s$ induce the same actions $X(m',s)$ with the same weight of truthful and lying types, we can without loss separate them into two messages m_ϵ and $m_{1-\epsilon}$. Now for $z \in [0,1]$, let $\theta_\epsilon^l(z) = \inf \theta : Pr(\Theta_\epsilon^l \cap [\theta,1]) = zPr(\Theta^l)$ be the $1-z$ percentile type in Θ_ϵ^l. Let $E_\epsilon^l(z) = E[\Theta_\epsilon^l \cap [0, \theta_\epsilon^l(z))]$ be is the expected value for the bottom $1-z$ percentile types in Θ_ϵ^l. Note that $E_\epsilon^l(z) = \frac{\int_0^z \theta_\epsilon^l(z')dz'}{1-z}$, $E_\epsilon^l(0) = E^l$ and $\frac{dE_\epsilon^l(z)}{dz} = E_\epsilon^l(z) - \theta_\epsilon^l(z)$. Now define $\hat{z}_\epsilon$ that solves

$$\frac{1-z}{(2-z)^2}(E^t - E_\epsilon^l(z))^2 = c \tag{112}$$
$$z \in (0,1) \tag{113}$$

To show that there exists such solution, When $z = 1$, $\frac{1-z}{(2-z)^2}(E^t - E^l_\epsilon(z))^2 = 0 < c$; when $z = 0$, $\frac{1-z}{(2-z)^2}(E^t - E^l_\epsilon(z))^2 = \frac{1}{4}(E^t - E^l_\epsilon(0))^2 = \frac{1}{4}(E^t - E^l)^2 = c$, where the last equality holds because of (110). Since

$$\begin{aligned}\lim_{z\to 0^+} d[\frac{1-z}{(2-z)^2}(E^t - E^l_\epsilon(z))^2]/dz &= -\frac{1}{4}2(E^t - E^l)\lim_{z\to 0^+}\frac{dE^l_\epsilon(z)}{dz} \\ &= \sqrt{c}(\theta^l_\epsilon(0) - E^l) \\ &= \sqrt{c}(\sup\Theta^l - E^l) > 0, \end{aligned} \tag{114}$$

such $\hat{z}_\epsilon \in (0,1)$ exists for any $\epsilon \in (0,1]$. Now Define a modified inspected liar set $\hat{\Theta}^l = \Theta^l_\epsilon \cap [0, \theta^l_\epsilon(\hat{z}_\epsilon))$ so that

$$Pr(\hat{\Theta}^l) = \frac{a}{2}(1-\hat{z}_\epsilon)\epsilon \tag{115}$$

$$E[\hat{\Theta}^l] = E^l_\epsilon(\hat{z}_\epsilon) \tag{116}$$

$$\frac{Pr(\hat{\Theta}^l)Pr(\Theta^t_\epsilon)}{(Pr(\hat{\Theta}^l) + Pr(\Theta^t_\epsilon))^2}(E^t - E[\hat{\Theta}^l])^2 = \frac{1-\hat{z}_\epsilon}{(2-\hat{z}_\epsilon)^2}(E^t - E^l_\epsilon(\hat{z}_\epsilon))^2 = c \tag{117}$$

Define the modified uninspected set $\Theta^0_{\hat{q}} = \Theta^0_q \cup (\Theta^l_\epsilon \cap [\theta^l_\epsilon(\hat{z}_\epsilon), 1])$, so that

$$Pr(\Theta^0_{\hat{q}}) = Pr(\Theta^0_q) + \frac{a}{2}\hat{z}_\epsilon\epsilon \tag{118}$$

$$E[\Theta^0_{\hat{q}}] = \frac{Pr(\Theta^0_q)\mu + \frac{a}{2}\hat{z}_\epsilon\epsilon\overline{E}^l}{Pr(\Theta^0_q) + \frac{a}{2}\hat{z}_\epsilon\epsilon} \tag{119}$$

where $\overline{E}^l \equiv E[\Theta^l_\epsilon \cap [\theta^l_\epsilon(\hat{z}_\epsilon), 1]] > E^l$ is the expected value for the top z percentile types in Θ^l_ϵ.

Now define the modified message and action rules $\hat{q}, \hat{X}$ where other things remain unchanged, except the messages m_ϵ is modified to message $\hat{m} = \Theta^t$ with the set of truthful senders $\Theta^t_{\hat{q}}(\hat{m}) = \Theta^t$, and the set of lying senders $\Theta^l_{\hat{q}}(\hat{m}) = \hat{\Theta}^l$; The uninspected message is modified to $m^0_{\hat{q}} = \Theta^0_{\hat{q}}$ with the set of senders identical to the statement.

The sequentially rational actions for the modified message $\hat{m}$ are

$$\begin{aligned}&\hat{X}(\hat{m}, t) = E[\Theta^t] = E^t \\ &\hat{X}(\hat{m}, l) = E[\hat{\Theta}^l] = E^l_\epsilon(\hat{z}_\epsilon) \\ &\hat{X}(\hat{m}, u) = E[\hat{\Theta}^t \cup \Theta^l)] = x^*_u(E^t, E^l_\epsilon(\hat{z}_\epsilon)) > x^*_u(E^t, E^l) = X(m', u)\end{aligned} \tag{120}$$

where the second equality of (120) holds by (117); the inequality holds by Lemma 6 and $E^l_\epsilon(\hat{z}_\epsilon) < E^l$; the last equality holds by optimality of Ω.

The sequentially rational action for the modified uninspected message $m^0_{\hat{q}}$ is

$$\hat{X}(m^0_{\hat{q}}, u) = E[\Theta^0_{\hat{q}}] = X(m^0_q, u) - \frac{a\hat{z}_\epsilon}{2Pr(\Theta^0_q)}(\hat{X}(m^0_{\hat{q}}, u) - \overline{E}^l)\epsilon \tag{121}$$

Since $\Theta^0_{\hat{q}} \cup \hat{\Theta}^l = \Theta^0_q \cup \Theta^l_\epsilon$, we have $Pr(\Theta^0_{\hat{q}}) - Pr(\Theta^0_q) = Pr(\Theta^l_\epsilon) - Pr(\hat{\Theta}^l) = \frac{a}{2}\hat{z}_\epsilon\epsilon$ and

$$Pr(\Theta^0_q)(\mu - E[\Theta^0_{\hat{q}}]) + Pr(\hat{\Theta}^l)(E^l - E^l_\epsilon(\hat{z}_\epsilon)) - \frac{a}{2}\hat{z}_\epsilon\epsilon(E[\Theta^0_{\hat{q}}] - E^l) = 0 \tag{122}$$

Since the original mechanism is optimal, $w_q(m) = w^-(X(m,t) - X(m,l))$ hold almost everywhere for unmodified message $m \in \mathcal{M}^+_q/m_\epsilon$. For the modified message, (117) implies $w_{\hat{q}}(\hat{m}) = w^-(E^t - E^l(\hat{z}_\epsilon)) = w^-(\hat{X}(\hat{m},t) - \hat{X}(\hat{m},l))$, so we have

$$w_{\hat{q}}(m) = w^-(\hat{X}(m,t) - \hat{X}(m,l)) \tag{123}$$

hold almost everywhere at $\mathcal{M}^+_{\hat{q}}$.

To compare DM's payoffs, since $V_q(m) = V_{\hat{q}}(m) = c$, by Lemma 4, $EU_{DM}(\Omega) = EU^I_{DM}(\Omega)$ and $EU_{DM}(\hat{\Omega}) = EU^I_{DM}(\hat{\Omega})$, so

$$\begin{aligned}
&EU_{DM}(\hat{\Omega}) - EU_{DM}(\Omega) \\
=&Pr(\Theta^0_{\hat{q}})E[\Theta^0_{\hat{q}}]^2 + Pr(\hat{\Theta}^l)(E^l_\epsilon(\hat{z}_\epsilon)^2 - c) - Pr(\Theta^0_q)\mu^2 - Pr(\Theta^u)((E^l)^2 - c) \\
=&-Pr(\Theta^0_q)(\mu^2 - E[\Theta^0_{\hat{q}}]^2) - Pr(\hat{\Theta}^l)((E^l)^2 - E^l_\epsilon(\hat{z}_\epsilon)^2) + \frac{a}{2}\hat{z}_\epsilon\epsilon(E[\Theta^0_{\hat{q}}]^2 - (E^l)^2) + \frac{a}{2}\hat{z}_\epsilon\epsilon c \\
=&-Pr(\Theta^0_q)(\mu^2 - E[\Theta^0_{\hat{q}}]^2 - 2E[\Theta^0_{\hat{q}}](\mu - E[\Theta^0_{\hat{q}}])) - Pr(\hat{\Theta}^l)((E^l)^2 - E^l_\epsilon(\hat{z}_\epsilon)^2 - 2E[\Theta^0_{\hat{q}}]((E^l - E^l_\epsilon(\hat{z}_\epsilon))) \\
&+\frac{a}{2}\hat{z}_\epsilon\epsilon(E[\Theta^0_{\hat{q}}]^2 - (E^l)^2 - 2E[\Theta^0_{\hat{q}}](E[\Theta^0_{\hat{q}}] - E^l)) + \frac{a}{2}\hat{z}_\epsilon\epsilon c \\
=&-Pr(\Theta^0_q)(\mu - E[\Theta^0_{\hat{q}}])^2 - Pr(\hat{\Theta}^l)((E^l - E^l_\epsilon(\hat{z}_\epsilon))(E^l + E^l_\epsilon(\hat{z}_\epsilon) - 2E[\Theta^0_{\hat{q}}])) \\
&+\frac{a}{2}\hat{z}_\epsilon\epsilon(c - (E[\Theta^0_{\hat{q}}] - E^l)^2) \\
>&-Pr(\Theta^0_q)(\frac{a\hat{z}_\epsilon}{2Pr(\Theta^0_q)}(\hat{X}(m^0_{\hat{q}}, u) - \overline{E}^l)\epsilon)^2\epsilon^2 + \frac{a}{2}\hat{z}_\epsilon\epsilon\delta
\end{aligned} \tag{124}$$

where the third equality holds by (122); the inequality holds because of (111), (121) and $E[\Theta^0_{\hat{q}}] \approx \mu > E^l > E^l_\epsilon(\hat{z}_\epsilon)$ for small enough ϵ.

Case 2: $Pr(m) = 0$ **for all** $m \in M'$**:**

If almost every $m, m' \in M'$ induces the same actions $X(m,s) = X(m',s) = \overline{X}^s$, then it is without loss to pool them into a same message with positive measure, and case 1 applies. Now if there exist $\delta' > 0$ and two positive measure subsets of messages $M'_1 \subseteq M'$ and $M'_2 \subseteq M'$ such that for any $(m_1, m_2) \in M'_1 \times M'_2$ and $s \in t, l, u$,

$$X(m_1, s) - X(m_2, s) > \delta' \tag{125}$$

then for any $\epsilon > 0$ there exist two positive measure subsets $M_1'' \subseteq M_1'$ and $M_2'' \subseteq M_2'$ such that for any $i = 1, 2$, $m, m' \in M_i''$ and $s \in t, l, u$,

$$|X(m, s) - X(m', s)| < \epsilon \tag{126}$$

and $Pr(M_1'') = zPr(M_2'') \leq \epsilon$ for any $z \in (0, \infty)$. For $i = 1, 2$ and $s \in t, l, u$, let $\Theta_i^s = \Theta_q^s(M_i'')$ be the aggregate sets of truthful senders, lying senders and senders of M_i'', and $E_i^s = E[\Theta_i^s]$ be their corresponding expected value. Since $w_q(m) = 0.5$ for any $m \in M_i$, we have

$$Pr(\Theta_i^l) = Pr(\Theta_i^t) \tag{127}$$

$$E_i^t - E_i^u = E_i^u - E_i^l = \sqrt{c} \tag{128}$$

$$(\mu - E_i^l)^2 < c - \delta \tag{129}$$

and thus

$$E_1^t - E_2^t = E_1^l - E_2^l > \delta' \tag{130}$$

For $z \in (0, \infty)$, let $\hat{\Theta}^t(z) = \Theta_1^t \cup \Theta_2^t$ be the aggregate set of truthful type, and $E^t(z) = E[\hat{\Theta}^t(z)] = \Theta_2^t + \frac{z}{1+z}(\Theta_1^t - \Theta_2^t)$. Define $\hat{z}$ that solves

$$\frac{1+z}{(2+z)^2}(E^t(z) - E_2^l)^2 = c \tag{131}$$

$$z \in (0, \infty) \tag{132}$$

To show that there exists such solution, When $z \to \infty$, $\frac{1+z}{(2+z)^2}(E^t(z) - E_2^l)^2 \to 0 < c$; when $z = 0$, $\frac{1+z}{(2+z)^2}(E^t(z) - E_2^l)^2 = \frac{1}{4}(E_2^t - E_2^l)^2 = \frac{1}{4}(E^t - E^l)^2 = c$.Since

$$\begin{aligned}\lim_{z \to 0^+} d[\frac{1+z}{(2+z)^2}(E^t(z) - E_2^l)^2]/dz &= -\frac{1}{4}2(E_2^t - E_2^l) \lim_{z \to 0^+} \frac{dE^t(z)}{dz} \\ &= \sqrt{c}(E_1^t - E_2^t) \\ &> \sqrt{c}\delta' > 0,\end{aligned} \tag{133}$$

such $\hat{z}_\epsilon \in (0, 1)$ exists. Now take a set of messages M_1'' such that $Pr(M_1'') = \hat{\Theta}^t Pr(M_2'') \leq \epsilon$. Denote $b = Pr(M_2'')$ and let $\hat{\Theta}^t(\hat{z})$ be the modified inspected truth-teller set so that

$$Pr(\hat{\Theta}^t) = (1 + \hat{z})\frac{b}{2} \tag{134}$$

$$E[\hat{\Theta}^t] = E^t(\hat{z}) = \Theta_2^t + \frac{\hat{z}}{1+\hat{z}}(\Theta_1^t - \Theta_2^t) \tag{135}$$

$$\frac{Pr(\Theta_2^l)Pr(\hat{\Theta}^t)}{(Pr(\Theta_2^l)+Pr(\hat{\Theta}^t)^2}(E[\hat{\Theta}^t]-E_2^l)^2=\frac{1+\hat{z}}{(2+\hat{z})^2}(E^t(z)-E_2^l)^2=c \tag{136}$$

Define the modified uninspected set $\Theta_{\hat{q}}^0=\Theta_q^0\cup\Theta_1^l$, so that

$$Pr(\Theta_{\hat{q}}^0)=Pr(\Theta_q^0)+\hat{z}\frac{Pr(M_2'')}{2} \tag{137}$$

$$E[\Theta_{\hat{q}}^0]=\frac{Pr(\Theta_q^0)\mu+\hat{z}\frac{Pr(M_2'')}{2}E_1^l}{Pr(\Theta_q^0)+\hat{z}\frac{b}{2}} \tag{138}$$

Now define the modified message and action rules $\hat{q},\hat{X}$ where other things remain unchanged, except the set of messages $M_1''\cup M_2''$ is off-path; an inspected message $\hat{m}=\hat{\Theta}^t(\hat{z})$ is added with the set of truthful senders $\Theta_{\hat{q}}^t(\hat{m})=\hat{\Theta}^t(\hat{z})$, and the set of lying senders $\Theta_{\hat{q}}^l(\hat{m})=\Theta_2^l$; The uninspected message is modified to $m_{\hat{q}}^0=\Theta_{\hat{q}}^0$ with the set of senders identical to the statement.

The sequentially rational actions for the modified message $\hat{m}$ are

$$\begin{aligned}
&\hat{X}(\hat{m},t)=E[\hat{\Theta}^t(\hat{z})]=E^t(\hat{z})\\
&\hat{X}(\hat{m},l)=E[\hat{\Theta}^l]=E_2^l\\
&\hat{X}(\hat{m},u)=E[\hat{\Theta}^t\cup\Theta^l)]=x_u^*(E^t(z),E_2^l>x_u^*(E_2^t,E_2^l)=E_2^u
\end{aligned} \tag{139}$$

where the second equality of (120) holds by (136); the inequality holds by Lemma 6 and $E^t(\hat{z})>E_2^t$; the last equality holds by optimality of Ω.

The sequentially rational action for the modified uninspected message $m_{\hat{q}}^0$ is

$$\hat{X}(m_{\hat{q}}^0,u)=E[\Theta_{\hat{q}}^0]=X(m_q^0,u)-\frac{\hat{z}b}{2Pr(\Theta_q^0)}(\hat{X}(m_{\hat{q}}^0,u)-E_1^l) \tag{140}$$

Since $\Theta_{\hat{q}}^0=\Theta_q^0\cup\Theta_1^l$, $\hat{\Theta}^t(\hat{z})=\Theta_1^t\cup\Theta_2^t$ and $Pr(\Theta_{\hat{q}}^0)-Pr(\Theta_q^0)=Pr(\hat{\Theta}^t)-Pr(\Theta_2^t)=Pr(\Theta_1^t)=Pr(\Theta_1^l)=\frac{\hat{z}}{2}b$,

$$Pr(\Theta_q^0)(\mu-E[\Theta_{\hat{q}}^0])-\frac{\hat{z}}{2}b(E[\Theta_{\hat{q}}^0]-E_1^l)=0 \tag{141}$$

$$\int_{m\in M_i''}(E_i^s-X(m,s))\int_{\Theta_q^s(m)}dF(\theta)dm=0 \text{ for } i=1,2; s=t,l \tag{142}$$

$$(1+\hat{z})E^t(\hat{z})-E_2^t-\hat{z}E_1^t=0 \tag{143}$$

Since the original mechanism is optimal, $w_q(m)=w^-(X(m,t)-X(m,l))$ hold almost everywhere for unmodified message $m\in\mathcal{M}_q^+/m_\epsilon$. For the modified message, (136) implies $w_{\hat{q}}(\hat{m})=w^-(E^t-E^l(\hat{z}_\epsilon))=w^-(\hat{X}(\hat{m},t)-\hat{X}(\hat{m},l))$, so

$$w_{\hat{q}}(m)=w^-(\hat{X}(m,t)-\hat{X}(m,l)) \tag{144}$$

hold almost everywhere at $\mathcal{M}^+_{\hat{q}}$.

To compare DM's payoffs, since $V_q(m) = V_{\hat{q}}(m) = c$, by Lemma 4, $EU_{DM}(\Omega) = EU^I_{DM}(\Omega)$ and $EU_{DM}(\hat{\Omega}) = EU^I_{DM}(\hat{\Omega})$, so

$$\begin{aligned}
&EU_{DM}(\hat{\Omega}) - EU_{DM}(\Omega)\\
=&Pr(\Theta^0_{\hat{q}})E[\Theta^0_{\hat{q}}]^2 + Pr(\hat{\Theta}^t)(E^t(\hat{z})^2 - c) + Pr(\Theta^l_2)((E^l_2)^2 - c)\\
&- Pr(\Theta^0_q)\mu^2 - \sum_{i=1,2}\sum_{s=t,l}\int_{M''_i}(X(m,s)^2 - c)\int_{\Theta^s_q(m)} dF(\theta)dm\\
=&-Pr(\Theta^0_q)(\mu^2 - E[\Theta^0_{\hat{q}}]^2) - \frac{b}{2}\hat{z}((E^l_1)^2 - E[\Theta^0_{\hat{q}}]^2) - \frac{b}{2}((1+\hat{z})E^t(\hat{z})^2 - (E^t_2)^2 - \hat{z}(E^t_1)^2)\\
&- \sum_{i=1,2}\sum_{s=t,l}\int_{M''_i}(X(m,s)^2 - (E^s_i)^2)\int_{\Theta^s_q(m)} dF(\theta)dm + \frac{b}{2}\hat{z}c\\
=&-Pr(\Theta^0_q)(\mu - E[\Theta^0_{\hat{q}}])^2 - \frac{b}{2}\hat{z}(E^l_1 - E[\Theta^0_{\hat{q}}])^2 - \frac{b}{2}(\frac{\hat{z}}{1+\hat{z}}(E^t_1 - E^t_2)^2)\\
&- \sum_{i=1,2}\sum_{s=t,l}\int_{M''_i}(X(m,s) - E^s_i)^2\int_{\Theta^s_q(m)} dF(\theta)dm + \frac{b}{2}\hat{z}c\\
\geq&-Pr(\Theta^0_q)(\frac{\hat{z}b}{2Pr(\Theta^0_q)}(\hat{X}(m^0_{\hat{q}},u) - E^l_1))^2 - b(1+\hat{z})\frac{1}{4}\epsilon^2 + \frac{b}{2}\hat{z}[c - \frac{1}{1+\hat{z}}(E^t_1 - E^t_2)^2 - (E[\Theta^0_{\hat{q}}] - E^l_1)^2]\\
>&-Pr(\Theta^0_q)(\frac{\hat{z}}{2Pr(\Theta^0_q)}(\hat{X}(m^0_{\hat{q}},u) - E^l_1))^2 b^2 - b(1+\hat{z})\frac{1}{4}\epsilon^2 + \frac{b}{2}\hat{z}[c - (\mu - E^l_2)^2]\\
>&-Pr(\Theta^0_q)(\frac{\hat{z}}{2Pr(\Theta^0_q)}(\hat{X}(m^0_{\hat{q}},u) - E^l_1))^2 b^2 - b(1+\hat{z})\frac{1}{4}\epsilon^2 + \frac{b}{2}\hat{z}\delta
\end{aligned} \tag{145}$$

where the third equality holds by (141)-(143); the first inequality holds because of (140), (126) and Popoviciu's inequality; the second inequality holds for small ϵ because $E^t_1 - E^t_2 = E^l_1 - E^l_2$ and $\mu \approx E[\Theta^0_{\hat{q}}] > E^l_1 > E^l_2$; the last inequality holds by (129).

Finally for Case $j = 1, 2$, (121) and (140) imply that $X(m^0_q, u) - \hat{X}(m^0_{\hat{q}}, u) = K_j(\epsilon)$ where

$$K_1(\epsilon) = \frac{a\hat{z}_\epsilon}{2Pr(\Theta^0_q)}(\hat{X}(m^0_{\hat{q}}, u) - \overline{E}^l)\epsilon$$

$$K_2(\epsilon) = \frac{\hat{z}}{2Pr(\Theta^0_q)}(\hat{X}(m^0_{\hat{q}}, u) - E^l_1)b \leq \frac{\hat{z}}{2Pr(\Theta^0_q)}(\hat{X}(m^0_{\hat{q}}, u) - E^l_1)\epsilon$$

so Lemma 5 implies that there exists an incentive compatible mechanism $\tilde{\Omega}$ such that $EU_{DM}(\tilde{\Omega}) > EU_{DM}(\hat{\Omega}) - 4K_j(\epsilon)^2$, then by (124) and (145) $EU_{DM}(\tilde{\Omega}) - EU_{DM}(\Omega) > \frac{a}{2}\hat{z}_\epsilon\epsilon\delta - [4 + Pr(\Theta^0_q)](\frac{a\hat{z}_\epsilon}{2Pr(\Theta^0_q)}(\hat{X}(m^0_{\hat{q}}, u) - \overline{E}^l)\epsilon)^2$ for case 1 and $EU_{DM}(\tilde{\Omega}) - EU_{DM}(\Omega) > \frac{b}{2}\hat{z}\delta - b(1+\hat{z})\frac{1}{4}\frac{1}{4}\epsilon^2 - [4 + Pr(\Theta^0_q)][(\frac{\hat{z}}{2Pr(\Theta^0_q)}(\hat{X}(m^0_{\hat{q}}, u) - E^l_1))]^2 b^2$ for case 2, where $b \leq \epsilon$ goes to 0 as $\epsilon \to 0$, so in both cases $EU_{DM}(\tilde{\Omega}) - EU_{DM}(\Omega) > 0$ for small enough ϵ, but it contradicts that Ω is an optimal mechanism. Therefore, we conclude that $w_q(m) < 0.5$ almost everywhere for $m \in \mathcal{M}^+_q$ in an optimal mechanism. ■

Lemma 11 *Suppose for an optimal mechanism* Ω, $Pr(\Theta_q^0) > 0$ *and* $\mu - \overline{\theta}^l > \underline{\theta}^t - \mu$, *then* $\underline{\theta}^t \in (\mu, \mu + \sqrt{c})$ *and for almost every* $\theta \in (\underline{\theta}^t, \mu + \sqrt{c}) \cap \Theta_q^t(\mathcal{M}_q^+)$, θ *is essentially revealed upon inspection in* Ω, *and* $\Theta_q^t(m_q(\theta)) = \{\theta\}$.

Proof of Lemma 11: First We claim that $\underline{\theta}^t = \overline{\theta}^0 > \mu$, where the inequality holds by definition and $Pr(\Theta_q^0) > 0$. Suppose $\underline{\theta}^t \neq \overline{\theta}^0$, then by Lemma 9 $\underline{\theta}^t > \overline{\theta}^0$, and since $\Theta_q^t(\mathcal{M}_q^+) \cup \Theta_q^t(\mathcal{M}_q^+) \cup \Theta_q^0 = [0,1]$, we have $\underline{\theta}^t = \overline{\theta}^l$, but that contradicts $\mu - \overline{\theta}^l > \underline{\theta}^t - \mu > 0$. Now $\underline{\theta}^t = \overline{\theta}^0$ and $\Theta_q^t(\mathcal{M}_q^+) \cup \Theta_q^t(\mathcal{M}_q^+) \cup \Theta_q^0 = [0,1]$ imply $\overline{\theta}^l \geq \underline{\theta}^0$. Then $\underline{\theta}^t = \overline{\theta}^0$, $\overline{\theta}^l \geq \underline{\theta}^0$, $\mu - \overline{\theta}^l > \underline{\theta}^t - \mu > 0$ and Lemma 8 imply that $\underline{\theta}^t - \mu < \sqrt{c}$.

Now we show that for almost every $\theta \in (\underline{\theta}^t, \mu + \sqrt{c}) \cap \Theta_q^t(\mathcal{M}_q^+)$, θ is essentially revealed upon inspection in Ω. Suppose on the contrary, there exist a positive measure set $\Theta_t \subseteq (\underline{\theta}^t, \mu + \sqrt{c}) \cap \Theta_q^t(\mathcal{M}_q^+)$ such that for any $\theta \in \Theta_t$, $x_\Omega^d(\theta) = X(m_q(\theta), t) \neq \theta$. Let $\Theta_t' = \{\theta \in \Theta_t : X(m_q(\theta), t) < \theta\}$. Since $X(.)$ satisfies (5), Θ_t' must have positive measure. For any $m \in m_q(\Theta_t')$, Lemma 3 implies that $X(m, u) > \mu$, and Lemma 10 implies $w_q(m) < 0.5$, so there exist $\delta > 0$ and a positive measure subset $\Theta_t'' \subseteq \Theta_t'$ such that for any $\theta \in \Theta_t''$ and $m \in m_q(\Theta_t'')$,

$$(\theta - \mu)^2 < c - \delta \tag{146}$$

$$\theta < x_\Omega^d(\theta) - \delta \tag{147}$$

$$X(m, u) > \mu + \delta \tag{148}$$

$$w_q(m) < 0.5 - \delta \tag{149}$$

Denote $a = \frac{Pr(\Theta_t'')}{Pr(\Theta_q^t(m_q(\Theta_t'')))} > 0$, and let $M_t^a = \cap M \subseteq m_q(\Theta_t'') : \frac{Pr(\Theta_t'' \cap \Theta_q^t(M))}{Pr(\Theta_q^t(M))} \geq a$ be the subset containing every message in $m_q(\Theta_t'')$ where proportion of truthful types within Θ_t'' is no less than a. Since $Pr(M_t^a) > 0$, for any $\epsilon > 0$ there is positive measure subset $M' \subseteq M_t^a$ such that for any $m, m' \in M'$ and $s \in \{t, l, u\}$,

$$|X(m, s) - X(m', s)| < \epsilon \tag{150}$$

and $Pr(M') \leq \epsilon$ Denote $b = Pr(M') \equiv Pr(\Theta_q^u(M'))$. Let $\Theta' = \Theta_q^t(M') \cap \Theta_t''$ be the set of truthful types in M' that satisfies (146) and(147). Since $M' \subseteq M_t^a$, we have

$$Pr(\Theta') \geq aPr(\Theta_q^t(M')) \tag{151}$$

Let $\Theta^l = \Theta_q^l(M')$, $\Theta^t = \Theta_q^t(M')$ and $\Theta^u = \Theta^l \cup \Theta^t$ be the aggregate set of truth-tellers, liars and senders of M'; $\Theta_{ex} = \Theta_q^t(M')/\Theta'$ be the set of truth-tellers excluding those in Θ'; $E^u = E[\Theta^u]$, $E^t = E[\Theta^t]$,

$E^l = E[\Theta^l]$, $E' = E[\Theta']$ and $E_{ex} = E[\Theta_{ex}]$ be their corresponding expected values, so that

$$E^t = \frac{z'E' + z_{ex}E_{ex}}{z' + z_{ex}} \tag{152}$$

and $z^t = \frac{Pr(\Theta^t)}{b}$, $z^l = \frac{Pr(\Theta^l)}{b} = 1 - z^t$, $z' = \frac{Pr(\Theta')}{b}$ and $z_{ex} = \frac{Pr(\Theta_{ex})}{b}$ be their corresponding ratios of measure to set of senders Θ^u. Since $\Theta' \subseteq \Theta''_t$ and $M' \in M^a_t$,

$$(E' - \mu)^2 < c - \delta \tag{153}$$

$$E' + \delta < E^t < E_{ex} \tag{154}$$

$$z' \geq az^t \tag{155}$$

Since the original mechanism is optimal, we have $(X(m,t) - X(m,u))(X(m,u) - X(m,l)) = c$ and $w_q(m) = w^-(X(m,t) - X(m,l)) \leq 0.5$ almost everywhere at M', so $|E^s - X(m,s)| < \epsilon$ for any $m \in M'$ and $s = t, l, u$ and (149) imply

$$|(1 - z^t)z^t(E^t - E^l)^2 - c| \equiv |(E^t - E^u)(E^u - E^l) - c| < 4\epsilon^2 \tag{156}$$

$$z^t > 0.5 + \delta \tag{157}$$

Therefore $z^l = 1 - z^t \in (\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{c+4\epsilon^2}{(E^t-E^l)^2}}, \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{c-4\epsilon^2}{(E^t-E^l)^2}})$, so

$$\frac{z^l}{z^t} = \frac{w^-(E^t - E^l)}{1 - w^-(E^t - E^l)} + k_1(\epsilon)\epsilon^2 = h(E^t - E^l) + k_1(\epsilon)\epsilon^2 \tag{158}$$

where $k_1(\epsilon)$ is a bounded function.

For any $r \in [0,1]$, let $E^t(r) = \frac{z^tE^t - rz'E' - \epsilon z_{ex}E_{ex}}{z^t - rz' - \epsilon z_{ex}}$ be the expected value of the set $(\Theta^t_q(M')/(\Theta'_r) \cup \Theta_{ex,\epsilon})$, where Θ'_r is a set with expected value E' and measure $rPr(Theta')$, and $\Theta_{ex,\epsilon}$ is a set with expected value E_{ex} and measure $\epsilon Pr(\Theta_{ex})$. Define $\hat{r} \in (0,1)$ that solves

$$h(E_{ex} - E^l)\epsilon z_{ex} + h(E^t(r) - E^l)(z^t - rz' - \epsilon z_{ex}) = z^l \tag{159}$$

To show that such solution exists for small enough ϵ, $E^t(1) = E_{ex} > E^t$, and $h(E_{ex} - E^l)\epsilon z_{ex} + h(E^t(r) - E^l)(z^t - rz' - \epsilon z_{ex}) = h(E_{ex} - E^l)(z^t - z') < h(E^t - E^l)z^t \to z^l$, where the first inequality holds as $h(.)$ is a decreasing function, and (158) implies $\lim_{\epsilon \to 0} h(E^t - E^l)z^t - z^l = 0$, so when $r = 1$, LHS of (159) is smaller than z^l for small ϵ; On the other end, $E^t(0) = E^t - \frac{\epsilon z_{ex}}{z^t - \epsilon z_{ex}}(E_{ex} - E^t) < E^t < E_{ex}$ and $E^t = \frac{z_{ex}}{z^t}\epsilon E^{ex} + (1 - \frac{z_{ex}}{z^t}\epsilon)E^t(0)$, so by second order taylor expansion, $h(E^t - E^l) = h(E^t(0) - E^l) + h'(E^t(0) - E^l)\frac{z_{ex}}{z^t}(E_{ex} - E^t(0))\epsilon + h''(\tilde{E} - E^l)(\frac{z_{ex}}{z^t}(E_{ex} - E^t(0))\epsilon)^2$, where $\tilde{E} \in [E^t(0), E_{ex}]$. Then

$$(1 - \frac{z_{ex}}{z^t}\epsilon)h(E^t(0) - E^l) + \frac{z_{ex}}{z^t}\epsilon h(E_{ex} - E^l) - h(E^t - E^l)$$

$$
\begin{aligned}
=&h(E^t(0)-E^l)+\frac{z_{ex}}{z^t}\epsilon[h(E_{ex}-E^l)-h(E^t(0)-E^l)]-h(E^t-E^l)\\
=&\frac{z_{ex}}{z^t}\epsilon[h(E_{ex}-E^l)-h(E^t(0)-E^l)-(E_{ex}-E^t(0))h'(E^t(0)-E^l)]\\
&-h''(\tilde{E}-E^l)(\frac{z_{ex}}{z^t}(E_{ex}-E^t(0))\epsilon)^2\\
\equiv&k_2(\epsilon)\epsilon
\end{aligned}
\tag{160}
$$

where $\lim_{\epsilon\to 0}k_2(\epsilon)>0$ because $h(.)$ is strictly convex and $E_{ex}>E^t\approx E^t(0)$, so $h(E_{ex}-E^l)-h(E^t(0)-E^l)-(E_{ex}-E^t(0))h'(E^t(0)-E^l)$ is bound away from 0. Therefore, $h(E_{ex}-E^l)\epsilon z_{ex}+h(E^t(r)-E^l)(z^t-rz'-\epsilon z_{ex})=z^t[(1-\frac{z_{ex}}{z^t}\epsilon)h(E^t(0)-E^l)+\frac{z_{ex}}{z^t}\epsilon h(E_{ex}-E^l)]=z^t[h(E^t-E^l)+k_2(\epsilon)\epsilon]=z^l+k_1(\epsilon)\epsilon^2+z^tk_2(\epsilon)\epsilon>z^l$ for small enough ϵ, so when $r=0$, LHS of (159) is larger than z^l for small ϵ, thus there exists $\hat{r}\in(0,1)$ such that (159) is satisfied. Furthermore, we have $\lim_{\epsilon\to 0}\hat{r}=0$, for otherwise $\lim_{\epsilon\to 0}E^t(\hat{r})>E^t$, and $\lim_{\epsilon\to 0}h(E_{ex}-E^l)\epsilon z_{ex}+h(E^t(\hat{r})-E^l)(z^t-\hat{r}z'-\epsilon z_{ex})=\lim_{\epsilon\to 0}h(E^t(\hat{r})-E^l)(z^t-\hat{r}z')<\lim_{\epsilon\to 0}h(E^t-E^l)z^t=z^l$, contradicting (159).

Now let $\Theta_{ex,\epsilon}=\Theta_{ex}(\epsilon)$ the modified upper set of inspected truth-teller, where $\Theta_{ex}(\epsilon)$ is a mean-preserving division of Θ_{ex} so that

$$Pr(\Theta_{ex,\epsilon})=\epsilon Pr(\Theta_{ex})=\epsilon z_{ex}b \tag{161}$$

$$E[\Theta_{ex,\epsilon}]=E_{ex} \tag{162}$$

Let $\Theta^0_{\hat{q}}=\Theta^0_q\cup\Theta'(\hat{r})$ be the modified uninspected set, where $\Theta'(\hat{r})$ is a mean-preserving division of Θ' so that

$$Pr(\Theta'(\hat{r}))=\hat{r}Pr(\Theta')=\hat{r}z'b \tag{163}$$

$$E[\Theta'(\hat{r})]=E' \tag{164}$$

and thus

$$Pr(\Theta^0_{\hat{q}})=Pr(\Theta^0_q)+\hat{r}z'b \tag{165}$$

$$E[\Theta^0_{\hat{q}}]=\frac{Pr(\Theta^0_q)\mu+\hat{r}z'bE'}{Pr(\Theta^0_q)+\hat{r}z'b} \tag{166}$$

Let $\hat{\Theta}=(\Theta'/\Theta'(\hat{r}))\cup(\Theta_{ex}/\Theta_{ex,\epsilon})$ be the modified lower set of inspected truth-teller, so that

$$Pr(\hat{\Theta})=(1-\epsilon)Pr(\Theta_{ex})+(1-\hat{r})Pr(\Theta')=[z^t-\epsilon z_{ex}-\hat{r}z')]b \tag{167}$$

$$E[\hat{\Theta}]=\frac{z^tE^t-rz'E'-\epsilon z_{ex}E_{ex}}{z^t-rz'-\epsilon z_{ex}}=E^t(r) \tag{168}$$

Let $\overline{z}^l = h(E_{ex} - E^l)\epsilon z_{ex}$ be the required share of liars for the modified upper set of inspected truth-teller, and $\underline{z}^l = h(E^t(\hat{r}) - E^l)(z^t - \hat{r}z' - \epsilon z_{ex})$ be the the required share of liars for the modified lower set of inspected truth-teller. (159) implies that $\overline{z}^l + \underline{z}^l = z^l$. Let $\overline{\Theta}^l = \Theta^l(\frac{\overline{z}^l}{z^l})$ and $\underline{\Theta}^l = \Theta^l / \overline{\Theta}^l$ be the mean-preserving divisions of Θ^l so that $E[\overline{\Theta}^l] = E[\underline{\Theta}^l] = E^l$, $Pr(\overline{\Theta}^l) = h(E_{ex} - E^l)\epsilon z_{ex} b$ and $Pr(\underline{\Theta}^l) = h(E^t(\hat{r}) - E^l)(z^t - \hat{r}z' - \epsilon z_{ex})b$. By the definition of $h(.)$,

$$\frac{Pr(\overline{\Theta}^l)Pr(\Theta_{ex,\epsilon})}{(Pr(\overline{\Theta}^l) + Pr(\Theta_{ex,\epsilon}))^2}(E_{ex} - E^l)^2 = c \tag{169}$$

$$\frac{Pr(\underline{\Theta}^l)Pr(\hat{\Theta})}{(Pr(\underline{\Theta}^l) + Pr(\hat{\Theta}))^2}(E^t(\hat{r}) - E^l)^2 = c \tag{170}$$

Now define the modified message and action rules $\hat{q}, \hat{X}$ where other things remain unchanged, except the set of messages M' is off-path; an upper inspected message $m_{ex} = \Theta_{ex,\epsilon}$ is added with the set of truthful senders $\Theta^t_{\hat{q}}(m_{ex}) = \Theta_{ex,\epsilon}$, and the set of lying senders $\Theta^l_{\hat{q}}(m_{ex}) = \overline{\Theta}^l$; an lower inspected message $\hat{m} = \hat{\Theta}$ is added with the set of truthful senders $\Theta^t_{\hat{q}}(\hat{m}) = \hat{\Theta}$, and the set of lying senders $\Theta^l_{\hat{q}}(\hat{m}) = \underline{\Theta}^l$; The uninspected message is modified to $m^0_{\hat{q}} = \Theta^0_{\hat{q}}$ with the set of senders identical to the statement.

The sequentially rational actions for the upper modified message m_{ex} are

$$\begin{aligned} \hat{X}(m_{ex}, t) &= E_{ex} \\ \hat{X}(m_{ex}, l) &= E^l \\ \hat{X}(m_{ex}, u) &= x^*_u(E_{ex}, E^l) > x^*_u(E^t, E^l) > \mu \end{aligned} \tag{171}$$

The sequentially rational actions for the lower modified message $\hat{m}$ are

$$\begin{aligned} \hat{X}(\hat{m}, t) &= E^t(\hat{r}) \\ \hat{X}(\hat{m}, l) &= E^l \\ \hat{X}(\hat{m}, u) &= x^*_u(E^t(\hat{r}), E^l_2) \approx x^*_u(E^t, E^l) > \mu \end{aligned} \tag{172}$$

where the approximation of (172) holds as $\epsilon \to 0$, and the inequality holds by (148). The sequentially rational action for the modified uninspected message $m^0_{\hat{q}}$ is

$$\hat{X}(m^0_{\hat{q}}, u) = X(m^0_q, u) + \frac{\hat{r}z'b}{Pr(\Theta^0_q) + \hat{r}z'b}(E' - X(m^0_q, u)) \tag{173}$$

Since the original mechanism is optimal, $w_q(m) = w^-(X(m,t) - X(m,l))$ hold almost everywhere for unmodified message $m \in \mathcal{M}^+_q / m_\epsilon$. For the modified message, (169) and (170) imply $w_{\hat{q}}(m_{ex}) ==$

$w^-(E_{ex}-E^l)$ and $w_{\hat{q}}(\hat{m}) = w^-(E^t(\hat{r})-E^l)$, so

$$w_{\hat{q}}(m) = w^-(\hat{X}(m,t)-\hat{X}(m,l)) \tag{174}$$

hold almost everywhere at $\mathcal{M}^+_{\hat{q}}$.

To compare DM's payoffs, since $V_q(m) = V_{\hat{q}}(m) = c$, by Lemma 4, $EU_{DM}(\Omega) = EU^I_{DM}(\Omega)$ and $EU_{DM}(\hat{\Omega}) = EU^I_{DM}(\hat{\Omega})$, so

$$\begin{aligned}
&EU_{DM}(\hat{\Omega}) - EU_{DM}(\Omega)\\
=&(Pr(\Theta^0_q)+\hat{r}z'b)E[\Theta^0_{\hat{q}}]^2 + \epsilon z_{ex}b(E^2_{ex}-c) + (z^t-\epsilon z_{ex}-\hat{r}z')b(E^t(\hat{r})^2-c) + z^l b((E^l)^2-c)\\
&-Pr(\Theta^0_q)\mu^2 - \sum_{s=t,l}\int_{M'}(X(m,s)^2-c)\int_{\Theta^s_q(m)}dF(\theta)dm\\
=&-[Pr(\Theta^0_q)\mu^2+\hat{r}z'b(E')^2-(Pr(\Theta^0_q)+\hat{r}z'b)E[\Theta^0_{\hat{q}}]^2]+\hat{r}z'bc\\
&+b[\epsilon z_{ex}E^2_{ex}+\hat{r}z'(E')^2+(z^t-\epsilon z_{ex}-\hat{r}z')E^t(\hat{r})^2-z^t(E^t)^2]\\
&-\sum_{s=t,l}\int_{M'}(X(m,s)^2-(E^s)^2)\int_{\Theta^s_q(m)}dF(\theta)dm\\
=&-Pr(\Theta^0_q)(E[\Theta^0_{\hat{q}}]-\mu)^2+\hat{r}z'b[c-(E'-E[\Theta^0_{\hat{q}}])^2]\\
&+b[\epsilon z_{ex}(E_{ex}-E^t)^2+\hat{r}z'(E^t-E')^2+(z^t-\epsilon z_{ex}-\hat{r}z')(E^t-E^t(\hat{r}))^2]\\
&-\sum_{s=t,l}\int_{M'}(X(m,s)-(E^s))^2\int_{\Theta^s_q(m)}dF(\theta)dm\\
>&-Pr(\Theta^0_q)(\frac{\hat{r}z'b}{Pr(\Theta^0_q)+\hat{r}z'b}(E'-X(m^0_q,u)))^2+\hat{r}z'b\delta+\frac{1}{4z_{ex}}a^2\delta^2b\epsilon-\frac{1}{4}b\epsilon^2
\end{aligned} \tag{175}$$

where the third equality holds by sequential rational actions ; the inequality holds because of (173), $(E'-E[\Theta^0_{\hat{q}}])^2 < (E'-\mu)^2 < c-\delta$, $b\epsilon z_{ex}(E_{ex}-E^t)^2 = b\epsilon z_{ex}[\frac{z'}{z_{ex}}(E^t-E')]^2 > b\epsilon z_{ex}[\frac{az^t}{z_{ex}}\delta]^2 > b\epsilon z_{ex}[\frac{0.5a}{z_{ex}}\delta]^2$, (126) and Popoviciu's inequality.

(173) implies that $\hat{X}(m^0_{\hat{q}},u)-X(m^0_q,u) = \frac{z'}{Pr(\Theta^0_q)+\hat{r}z'b}(E'-X(m^0_q,u))\hat{r}b$, so Lemma 5 implies that there exists an incentive compatible mechanism $\tilde{\Omega}$ such that $EU_{DM}(\tilde{\Omega}) > EU_{DM}(\hat{\Omega}) - 4(\frac{z'}{Pr(\Theta^0_q)+\hat{r}z'b}(E'-X(m^0_q,u))\hat{r}b)^2$, then by (175) $EU_{DM}(\tilde{\Omega})-EU_{DM}(\Omega) > \hat{r}z'b\delta+\frac{1}{4z_{ex}}a^2\delta^2b\epsilon-\frac{1}{4}b\epsilon^2-[4+Pr(\Theta^0_q)](\frac{z'}{Pr(\Theta^0_q)+\hat{r}z'b}(E'-X(m^0_q,u))\hat{r}b)^2$, where $b \le \epsilon$ goes to 0 as $\epsilon \to 0$, so $EU_{DM}(\tilde{\Omega}) - EU_{DM}(\Omega) > 0$ for small enough ϵ, but it contradicts that Ω is an optimal mechanism. Therefore, we have shown that for almost every $\theta \in (\underline{\theta}^t, \mu+\sqrt{c}) \cap \Theta^t_q(\mathcal{M}^+_q)$, θ is essentially revealed upon inspection in an optimal mechanism Ω.

Finally, suppose there is a positive measure set $\Theta_t \in (\underline{\theta}^t, \mu+\sqrt{c}) \cap \Theta^t_q(\mathcal{M}^+_q)$, such that for $\theta \in \Theta_t$, $x^d_\Omega(\theta) = \theta$ but $\Theta^t_q(m_q(\theta)) \neq \{\theta\}$, then there is a positive measure set of truthful types $\Theta' =$

$\Theta_q^t(m_q(\Theta_t))/\Theta_t$ who pool with Θ_t but not in Θ'. Then since $x_\Omega^d(\theta) = \theta$ for $\theta \in \Theta_t$, sequential rationality of X means there exists positive measure set $\Theta'_- = \{\theta \in \Theta' : \theta < X_\Omega^d(\theta)\}$, but it means that for $\theta \in \Theta'_-$, $\theta < X_\Omega^d(\theta) < \mu + \sqrt{c}$ is not essentially revealed upon inspection, contradicts to the first statement.

■

Lemma 12 *Suppose there is a mechanism Ω such that $Pr(\Theta_q^0) > 0$, $\underline{\theta_t} \in (\mu, \mu+\sqrt{c})$, X is sequentially rational given q, and there is a message $\hat{m} \in \mathcal{M}_q^+$ such that conditions (a) and (b) of Lemma 3 and $w_q(m) = w^-(X(m,t) - X(m,l))$ are satisfied almost everywhere for $m \in \mathcal{M}_q^+/\hat{m}$ and Properties stated in Lemma 11 are satisfied for $\Theta_q^t(\mathcal{M}_q^+\hat{m})$, except that $Pr(\hat{m}) > 0$, $X(\hat{m},t) > \inf_{m\in\mathcal{M}+_q} X(m,t)$, $X_u^*(X(\hat{m},t), X(\hat{m},l)) > \mu$ and $h_q(\hat{m}) \equiv \frac{Pr(\Theta_q^l(\hat{m}))}{Pr(\Theta_q^t(\hat{m}))} < h(X(\hat{m},t) - X(\hat{m},l))$.*
Let $\hat{b} = (1 - \frac{h_q(\hat{m})}{h(X(\hat{m},t)-X(\hat{m},l))})Pr(\Theta_q^t(\hat{m}))$. Then for any small enough $\hat{b}$ there exists an incentive compatible mechanism $\hat{\Omega}$ such that $EU_{DM}^I(\hat{\Omega}) > EU_{DM}^I(\Omega) + k(\hat{b})\hat{b}$, where $\lim_{\hat{b}\to 0} k(\hat{b}) > 0$.

Proof of Lemma 12: Let $z^t = \frac{Pr(\Theta_q^t(\hat{m}))}{Pr(\hat{m})}$ and $z^l = \frac{Pr(\Theta_q^l(\hat{m}))}{Pr(\hat{m})}$, be the share of truthful types and lying types in $\hat{m}$, $z_r^t = \frac{z^l}{h(X(\hat{m},t)-X(\hat{m},l))}$ and $z_e^t = z^t - z_r^t$ be the required share of truthful types and excess share of truthful types. We have

$$z_e^t = z^t(1 - \frac{h_q(\hat{m})}{h(X(\hat{m},t) - X(\hat{m},l))}) > 0 \tag{176}$$

Let $\Theta^l = \Theta_q^l(\hat{m})$ be the set of liars in $\hat{m}$, $\Theta_r^t = \Theta_q^t(\hat{m})(\frac{z_r^t}{z^t})$ and $\Theta_e^t = \Theta_q^t(\hat{m})/\Theta_q^t(\hat{m})(\frac{z_r^t}{z^t})$ be the mean-preserving divisions of $\Theta_q^t(\hat{m})$ so that

$$E[\Theta_r^t] = E[\Theta_e^t] = E[\Theta_q^t(\hat{m})] = X(\hat{m},t) \tag{177}$$

$$Pr(\Theta_r^t) = \frac{z_r^t}{z^t}Pr(\Theta_q^t(\hat{m})) = \frac{Pr(\Theta_q^l(\hat{m}))}{h(X(\hat{m},t) - X(\hat{m},l))} \tag{178}$$

$$Pr(\Theta_e^t) = \frac{z_e^t}{z^t}Pr(\Theta_q^t(\hat{m})) = (1 - \frac{h_q(\hat{m})}{h(X(\hat{m},t) - X(\hat{m},l))})Pr(\Theta_q^t(\hat{m})) = \hat{b} \tag{179}$$

Let $\underline{\Theta}_t = \{\theta \in \Theta_q^t(\mathcal{M}_q^+/\hat{m}) : \theta < X(\hat{m},t)\}$. $\underline{\Theta}_t$ has positive measure because $X(\hat{m},t) > \inf_{m\in\mathcal{M}+_q} X(m,t)$. Then by Lemma 11, there exists $\delta > 0$ and a positive measure subset $\underline{\Theta}'_t \subseteq \underline{\Theta}_t$ such that for any $\theta \in \underline{\Theta}'_t$

,

$$(\theta - \mu)^2 < c - \delta \tag{180}$$

$$x_\Omega^d(\theta) = \theta \tag{181}$$

$$\Theta_q^t(m_q(\theta)) = \{\theta\} \tag{182}$$

Let $M_t = m_q(\underline{\Theta}'_t)$ be the set of messages sent by truthful types $\underline{\Theta}'_t$. (182) and sequential rationality of X imply that for any $m \in M_t$, $\Theta^t_q(m) = \{X(m,t)\}$. By definition of M_t, for any $m \in M_t$, $X(\hat{m},t) > X(m,t)$, and since $h(X(m,t) - X(m,l))$ is well-defined in Ω, $h(X(\hat{m},t) - X(m,l))$ is also well-defined with $h(X(\hat{m},t) - X(m,l)) < h(X(m,t) - X(m,l))$. Since $Pr(M_t) > 0$, for small enough $\hat{b}$, we have $\int_{m\in M_t} \frac{1}{h(X(\hat{m},t)-X(m,l))} \int_{\Theta^l_q(m)} dF(\theta)dm \geq \hat{b}$, and because M_t is a collection of zero measure messages, so there exists $M'_t \subseteq M_t$ such that

$$\int_{m\in M'_t} \frac{1}{h(X(\hat{m},t) - X(m,l))} \int_{\Theta^l_q(m)} dF(\theta)dm = \hat{b} \tag{183}$$

Assign an arbitrary strict ranking $r : M'_t \to \mathbb{R}$ to the message set M'_t. Then for any $m \in M'_t$, let

$$z(m) = \frac{1}{\hat{b}} \int_{m'\in M'_t:r(m')\leq r(m)} \frac{1}{h(X(\hat{m},t) - X(m',l))} \int_{\Theta^t_q(m')} dF(\theta)dm' \tag{184}$$

be the cumulative required share of truthful types in Θ^t_e to pair with the liars in M'_t .

Define an modified messaging and action rules $\hat{q}, \hat{X}$ where other things remain unchanged, except the message $\hat{(}m)$ is modified to $T(\hat{m}$, the set of truthful senders $\Theta^t_{\hat{q}}(T(\hat{m})) = \Theta^t_r$ and the set of lying senders $\Theta^l_{\hat{q}}(T(\hat{m})) = \Theta^l$; for each $m \in M'_t$, m is modified to $T(m)$ with the set of truthful senders $\Theta^t_{\hat{q}}(T(m)) = \Theta^t_e(z(m))\ int(\Theta^t_e(z(m)))$ and the set of lying senders $\Theta^l_{\hat{q}}(T(m)) = \Theta^l_q(m)$, where $\Theta^t_e(z(m))\ int(\Theta^t_e(z(m)))$ is the boundary set of a mean preserving division of $\Theta^t_e(z(m))$ so that $E[\Theta^t_{\hat{q}}(T(m))] = E[\Theta^t_e] = X(\hat{m},t)$ and the set has measure $\frac{1}{h(X(\hat{m},t)-X(m,l))} \int_{\Theta^t_q(m)} dF(\theta)$; The uninspected message is modified to $m^0_{\hat{q}}$ with the set of senders $\Theta^0_q \cup \Theta^t_q(M'_t)$.

The sequentially rational actions for $T(\hat{m})$ are

$$\begin{aligned}
\hat{X}(T(\hat{m}),t) &= X(\hat{m},t) \\
\hat{X}(T(\hat{m}),l) &= X(\hat{m},l) \\
\hat{X}(T(\hat{m}),u) &= x^*_u(X(\hat{m},t),X(\hat{m},l)) > \mu
\end{aligned} \tag{185}$$

For each $m \in M'_t$, The sequentially rational actions for $T(m)$ are

$$\begin{aligned}
\hat{X}(T(m),t) &= X(\hat{m},t) \\
\hat{X}(T(m),l) &= X(m,l) \\
\hat{X}(T(m),u) &= x^*_u(X(\hat{m},t),X(m,l)) > x^*_u(X(m,t),X(m,l)) > \mu
\end{aligned} \tag{186}$$

where the first inequality of (215) holds because $X(\hat{m},t) > X(m,t)$. The sequentially rational action for the modified uninspected message $m^0_{\hat{q}}$ is

$$\hat{X}(m^0_{\hat{q}},u) = X(m^0_q,u) + \frac{1}{Pr(\Theta^0_q) + Pr(\Theta^t_q(M'_t))} \int_{M'_t} (X(m,t) - X(m^0_q,u)) \int_{\Theta^t_q(m)} dF(\theta)dm$$

$$< X(m_q^0, u) + \frac{Pr(\Theta_q^t(M_t'))}{Pr(\Theta_q^0) + Pr(\Theta_q^t(M_t'))} \tag{187}$$

where the inequality holds because $X(m,t) - X(m_q^0, u) \in (0,1)$. In the original mechanism, $w_q(m) = w^-(X(m,t) - X(m,l))$ hold almost everywhere for unmodified message $m \in \mathcal{M}_q^+/(M_t' \cup \hat{m})$. For the modified message, (178) and (184) imply $w_{\hat{q}}(T(\hat{m})) == w^-(X(\hat{m},t) - X(\hat{m},l))$ and $w_{\hat{q}}(T(m)) = w^-(X(\hat{m},t) - X(m,t))$, so

$$w_{\hat{q}}(m) = w^-(\hat{X}(m,t) - \hat{X}(m,l)) \tag{188}$$

hold almost everywhere at $\mathcal{M}_{\hat{q}}^+$. $w_q(m) = w^-(X(m,t) - X(m,l))$ almost everywhere for $m \in M_t'$ implies that $Pr(\Theta_q^t(M_t')) = \int_{m \in M_t'} \frac{1}{h(X(m,t)-X(m,l))} \int_{\Theta_q^l(m)} dF(\theta) dm$, combined with (183) means

$$Pr(\Theta_q^t(M_t')) = \frac{\int_{m \in M_t'} \frac{1}{h(X(m,t)-X(m,l))} \int_{\Theta_q^l(m)} dF(\theta) dm}{\int_{m \in M_t'} \frac{1}{h(X(\hat{m},t)-X(m,l))} \int_{\Theta_q^l(m)} dF(\theta) dm} \hat{b} \in (h(1)\hat{b}, \hat{b}) \tag{189}$$

where the upper bound holds because $X(\hat{m},t) > X(m,t)$ and $h(.)$ is a decreasing function, the lower bound holds because $w_q(m) > 0.5$ means $h(X(m,t) - X(m,l)) < 1$ and $h(X(\hat{m},t) - X(m,l)) > h(1)$.

To compare DM's payoffs,

$$\begin{aligned}
& EU_{DM}^I(\hat{\Omega}) - EU_{DM}^I(\Omega) \\
=& (Pr(\Theta_q^0) + Pr(\Theta_q^t(M_t')))E[\Theta_{\hat{q}}^0]^2 \\
& - Pr(\Theta_q^0)\mu^2 - \int_{M_t'} (X(m,t)^2 - c) \int_{\Theta_q^t(m)} dF(\theta) dm \\
=& - Pr(\Theta_q^0)(E[\Theta_{\hat{q}}^0] - \mu)^2 + \int_{M_t'} c - (X(m,t) - E[\Theta_{\hat{q}}^0])^2 \int_{\Theta_q^t(m)} dF(\theta) dm \\
>& - Pr(\Theta_q^0)(\frac{Pr(\Theta_q^t(M_t'))}{Pr(\Theta_q^0) + Pr(\Theta_q^t(M_t'))})^2 + Pr(\Theta_q^t(M_t'))\delta
\end{aligned} \tag{190}$$

where the second equality holds by sequential rational actions ; the inequality holds because of (187), (180) and (181). Since $V_q(m) = V_{\hat{q}}(m) = c$, by Lemma 4, $EU_{DM}(\hat{\Omega}) = EU_{DM}^U(\hat{\Omega}) = EU_{DM}^I(\hat{\Omega})$, and (187) implies that $\hat{X}(m_{\hat{q}}^0, u) - X(m_q^0, u) < \frac{Pr(\Theta_q^t(M_t'))}{Pr(\Theta_q^0)+Pr(\Theta_q^t(M_t'))}$, so Lemma 5 implies that there exists an incentive compatible mechanism $\tilde{\Omega}$ such that $EU_{DM}^I(\tilde{\Omega}) > EU_{DM}^I(\hat{\Omega}) - 4(\frac{Pr(\Theta_q^t(M_t'))}{Pr(\Theta_q^0)+Pr(\Theta_q^t(M_t'))})^2$, then by (216) $EU_{DM}^I(\tilde{\Omega}) - EU_{DM}^I(\Omega) > Pr(\Theta_q^t(M_t'))\delta - [4 + Pr(\Theta_q^0)](\frac{Pr(\Theta_q^t(M_t'))}{Pr(\Theta_q^0)+Pr(\Theta_q^t(M_t'))})^2$, and by (189), $EU_{DM}^I(\tilde{\Omega}) - EU_{DM}^I(\Omega) = k(\hat{b})\hat{b}$, where $\lim_{\hat{b} \to 0} k(\hat{b}) > h(1)\delta > 0$.

■

Lemma 13 *Suppose for an optimal mechanism Ω, $Pr(\Theta_q^0) > 0$ and $\mu - \overline{\theta}^l > \underline{\theta}^t - \mu$, then for almost every $\theta \in \Theta_q^u(\mathcal{M}_q^+)$, θ is essentially revealed upon inspection in Ω, and for almost every $m \in \mathcal{M}_q^+$ and $s = t, l$, $\Theta_q^s(m) = \{X(m,s)\}$.*

Proof of Lemma 13: We will show that for almost every lying types $\theta \in \Theta^l_q(\mathcal{M}^+_q)$, θ is essentially revealed upon inspection in Ω, the proof for truthful types is symmetrical and omitted.

Suppose on the contrary, there exist a positive measure set $\Theta_l \subseteq \Theta^l_q(\mathcal{M}^+_q)$ such that for any $\theta \in \Theta_l$, $x^d_\Omega(\theta) = X(m_q(\theta), l) \neq \theta$. Let $\Theta'_l = \{\theta \in \Theta_t : X(m_q(\theta), l) > \theta\}$. Since $X(.)$ satisfies (5), Θ'_l must have positive measure. For any $m \in m_q(\Theta'_l)$,Lemma 10 implies $w_q(m) < 0.5$, so there exist $\delta > 0$ and a positive measure subset $\Theta''_l \subseteq \Theta'_t$ such that for any $\theta \in \Theta''_l$ and $m \in m_q(\Theta''_l)$,

$$X(m,t) > \inf_{m' \in \mathcal{M}^+_q} X(m',t) \tag{191}$$

$$\theta < x^d_\Omega(\theta) - \delta \tag{192}$$

$$X(m,u) > \mu + \delta \tag{193}$$

$$w_q(m) < 0.5 - \delta \tag{194}$$

Denote $a = \frac{Pr(\Theta''_l)}{Pr(\Theta^l_q(m_q(\Theta''_l)))} > 0$, and let $M^a_l = \cap M \subseteq m_q(\Theta''_l) : \frac{Pr(\Theta''_l \cap \Theta^l_q(M))}{Pr(\Theta^l_q(M))} \geq a$ be the subset containing every message in $m_q(\Theta''_t)$ where proportion of truthful types within Θ''_l is no less than a. Since $Pr(M^a_l) > 0$, for any $\epsilon > 0$ there is positive measure subset $M' \subseteq M^a_l$ such that for any $m, m' \in M'$ and $s \in \{t, l, u\}$,

$$|X(m,s) - X(m',s)| < \epsilon \tag{195}$$

and $Pr(M') \leq \epsilon$ Denote $b = Pr(M') \equiv Pr(\Theta^u_q(M'))$. Let $\Theta' = \Theta^l_q(M') \cap \Theta''_l$ be the set of lying types in M' that satisfies (192). Since $M' \subseteq M^a_l$, we have

$$Pr(\Theta') \geq aPr(\Theta^l_q(M')) \tag{196}$$

Let $\Theta^l = \Theta^l_q(M')$, $\Theta^t = \Theta^t_q(M')$ and $\Theta^u = \Theta^l \cup \Theta^t$ be the aggregate set of truth-tellers, liars and senders of M'; $\Theta_{ex} = \Theta^l_q(M')/\Theta'$ be the set of liars excluding those in Θ'; $E^u = E[\Theta^u]$, $E^t = E[\Theta^t]$, $E^l = E[\Theta^l]$, $E' = E[\Theta']$ and $E_{ex} = E[\Theta_{ex}]$ be their corresponding expected values, so that

$$E^l = \frac{z'E' + z_{ex}E_{ex}}{z' + z_{ex}} \tag{197}$$

and $z^t = \frac{Pr(\Theta^t)}{b}$, $z^l = \frac{Pr(\Theta^l)}{b} = 1 - z^t$, $z' = \frac{Pr(\Theta')}{b}$ and $z_{ex} = \frac{Pr(\Theta_{ex})}{b}$ be their corresponding ratios of measure to set of senders Θ^u. Since $\Theta' \subseteq \Theta''_l$ and $M' \in M^a_l$,

$$E^t > \inf_{m' \in \mathcal{M}^+_q} X(m',t) \tag{198}$$

$$E' + \delta > E^l > E_{ex} \tag{199}$$

$$z' \geq az^l \tag{200}$$

Since the original mechanism is optimal, we have $(X(m,t)-X(m,u))(X(m,u)-X(m,l)) = c$ and $w_q(m) = w^-(X(m,t)-X(m,l)) \leq 0.5$ almost everywhere at M', so $|E^s - X(m,s)| < \epsilon$ for any $m \in M'$ and $s = t, l, u$ and (194) imply

$$|(1-z^t)z^t(E^t-E^l)^2 - c| \equiv |(E^t-E^u)(E^u-E^l) - c| < 4\epsilon^2 \tag{201}$$

$$z^t > 0.5 + \delta \tag{202}$$

Therefore $z^l = 1 - z^t \in (\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{c+4\epsilon^2}{(E^t-E^l)^2}}, \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{c-4\epsilon^2}{(E^t-E^l)^2}})$, so

$$\frac{z^t}{z^l} = \frac{1-w^-(E^t-E^l)}{w^-(E^t-E^l)} + k_1(\epsilon)\epsilon^2 = J(E^t-E^l) + k_1(\epsilon)\epsilon^2 \tag{203}$$

where $J(.) = \frac{1}{h(.)}$ is the required truth-teller to liar ratio, and $k_1(\epsilon)$ is a bounded function.

For any $\epsilon \in [0,1]$, let $E^l(\epsilon) = \frac{z^l t E^l - \epsilon z_{ex} E_{ex}}{z^l - \epsilon z_{ex}}$ be the expected value of the set $(\Theta_q^l(M')/\Theta_{ex,\epsilon})$, where $\Theta_{ex,\epsilon}$ is a set with expected value E_{ex} and measure $\epsilon Pr(\Theta_{ex})$.

Since $E^l = \frac{z_{ex}}{z^l}\epsilon E^{ex} + (1 - \frac{z_{ex}}{z^l}\epsilon)E^l(\epsilon)$, by second order taylor expansion, $J(E^t - E^l) = J(E^t - E^l(\epsilon)) + J'(E^t - E^l(\epsilon))\frac{z_{ex}}{z^l}(E^l(\epsilon) - E_{ex})\epsilon + J''(E^t - \tilde{E})(\frac{z_{ex}}{z^l}(E^l(\epsilon) - E_{ex})\epsilon)^2$, where $\tilde{E} \in [E_{ex}, E^l(\epsilon)]$. Then

$$\begin{aligned}
&(1 - \frac{z_{ex}}{z^l}\epsilon)J(E^t - E^l(\epsilon)) + \frac{z_{ex}}{z^t}\epsilon J(E^t - E_{ex}) - J(E^t - E^l) \\
=&J(E^t - E^l(\epsilon)) + \frac{z_{ex}}{z^l}\epsilon[J(E^t - E_{ex}) - J(E^t - E^l(\epsilon))] - J(E^t - E^l) \\
=&\frac{z_{ex}}{z^t}\epsilon[J(E^t - E_{ex}) - J(E^t - E^l(\epsilon)) - (E^l(\epsilon) - E_{ex})J'(E^t - E^l(\epsilon))] \\
&- J''(E^t - \tilde{E})(\frac{z_{ex}}{z^t}(E_{ex}) - E^l(\epsilon))\epsilon)^2 \\
\equiv& - k_2(\epsilon)\epsilon
\end{aligned} \tag{204}$$

where $\lim_{\epsilon\to 0} k_2(\epsilon) > 0$ because $J(.)$ is strictly concave and $E_{ex} < E^l \approx E^l(\epsilon))$, so $J(E^t - E_{ex}) - J(E^t - E^l(\epsilon)) - (E_{ex}) - E^l(\epsilon))J'(E^t - E^l(\epsilon))$ is negative and bound away from 0. Therefore, (203) and (204) imply

$$\begin{aligned}
(z^l - z_{ex}\epsilon)J(E^t - E^l(\epsilon)) &= z^l J(E^t - E^l) - z_{ex}\epsilon J(E^t - E_{ex}) - z^l k_2(\epsilon)\epsilon \\
&= z^t - z_{ex}\epsilon J(E^t - E_{ex}) - z^l\epsilon[k_2(\epsilon) - k_1(\epsilon)\epsilon]
\end{aligned} \tag{205}$$

Now let $\Theta_{ex,\epsilon}^l = \Theta_{ex}(\epsilon)$ be the set of liars for the modified lower message, where $\Theta_{ex}(\epsilon)$ is a mean-preserving division of Θ_{ex} so that

$$Pr(\Theta_{ex,\epsilon}^l) = \epsilon Pr(\Theta_{ex}) = \epsilon z_{ex} b \tag{206}$$

$$E[\Theta^l_{ex,\epsilon}] = E_{ex} \tag{207}$$

Let $z^t_{ex} = J(E^t - E_{ex})\epsilon z_{ex}$ be the required share of truth-tellers for the modified lower message, and $\Theta^t_{ex,\epsilon} = \Theta^t(\frac{z^t_{ex}}{z^t})$ be the set of truth-tellers for the modified lower message, where $\Theta^t(\frac{z^t_{ex}}{z^t})$ is a mean-preserving division of Θ^t so that

$$Pr(\Theta^t_{ex,\epsilon}) = \frac{z^t_{ex}}{z^t} Pr(\Theta^t) = J(E^t - E_{ex})\epsilon z_{ex} b \tag{208}$$

$$E[\Theta^t_{ex,\epsilon}] = E^t \tag{209}$$

Let $\hat{\Theta}^l = (\Theta^l/(\Theta_{ex}/\Theta_{ex,\epsilon})$ be the set of liars for the modified upper message, so that

$$Pr(\hat{\Theta}^l) = (1-\epsilon)Pr(\Theta_{ex}) + Pr(\Theta') = [z^l - \epsilon z_{ex})]b \tag{210}$$

$$E[\hat{\Theta}] = \frac{z^l E^l - \epsilon z_{ex} E_{ex}}{z^l - \epsilon z_{ex}} = E^l(\epsilon) \tag{211}$$

Let $\hat{\Theta}^t = \Theta^t/\Theta^t_{ex,\epsilon}$ be the set of truth-tellers for the modified upper message, which is the residual from $\Theta^t_{ex,\epsilon}$ so that

$$Pr(\hat{\Theta}^t) = (z^t - J(E^t - E_{ex})\epsilon z_{ex})b \tag{212}$$

$$E[\Theta^t_{ex,\epsilon}] = E^t \tag{213}$$

Now define the modified message and action rules $\hat{q}, \hat{X}$ where other things remain unchanged, except the set of messages M' is off-path; an lower inspected message m_{ex} is added with the set of truthful senders $\Theta^t_{\hat{q}}(m_{ex}) = \Theta^t_{ex,\epsilon}$, and the set of lying senders $\Theta^l_{\hat{q}}(m_{ex}) = \Theta^l_{ex,\epsilon}$; an upper inspected message $\hat{m}$ is added with the set of truthful senders $\Theta^t_{\hat{q}}(\hat{m}) = \hat{\Theta}^t$, and the set of lying senders $\Theta^l_{\hat{q}}(\hat{m}) = \hat{\Theta}^l$.

The sequentially rational actions for the message m_{ex} are

$$\begin{aligned}
\hat{X}(m_{ex}, t) &= E^t \\
\hat{X}(m_{ex}, l) &= E_{ex} \\
\hat{X}(m_{ex}, u) &= x^*_u(E^t, E_{ex}) > x^*_u(E^t, E^l) > \mu
\end{aligned} \tag{214}$$

The sequentially rational actions for the message $\hat{m}$ are

$$\begin{aligned}
\hat{X}(\hat{m}, t) &= E^t \\
\hat{X}(\hat{m}, l) &= E^l(\epsilon) \\
x^*_u(E^t, E^l(\epsilon)) &\approx x^*_u(E^t, E^l) > \mu
\end{aligned} \tag{215}$$

where the approximation of (215) holds as $\epsilon \to 0$, and the inequality holds by (193).

To compare DM's payoffs,

$$
\begin{aligned}
& EU^I_{DM}(\hat{\Omega}) - EU^I_{DM}(\Omega) \\
=& \epsilon z_{ex} b(E^2_{ex} - c) + (z^l - \epsilon z_{ex}) b(E^l(\epsilon)^2 - c) + z^t b((E^t)^2 - c) \\
& - \sum_{s=t,l} \int_{M'} X(m,s)^2 - c \int_{\Theta^s_q(m)} dF(\theta) dm \\
=& \epsilon z_{ex} b(E_{ex} - E^l)^2 + (z^l - \epsilon z_{ex}) b(E^l - E^l(\epsilon))^2 \\
& - \sum_{s=t,l} \int_{M'} (X(m,s) - E^s)^2 \int_{\Theta^s_q(m)} dF(\theta) dm \\
>& \epsilon z_{ex} b(E_{ex} - E^l)^2 + (z^l - \epsilon z_{ex}) b(E^l - E^l(\epsilon))^2 - \frac{1}{4}\epsilon^2 b
\end{aligned}
\tag{216}
$$

where the second equality holds by sequential rational actions ; the inequality holds because of (195) and Popoviciu's inequality. By definition of z^t_{ex} and $J(.)$, $w_{\hat{q}}(m_{ex}) = w^-(E^t - E_{ex})$. Now For $\hat{m}$,

$$
\begin{aligned}
h_{\hat{q}}(\hat{m}) = \frac{Pr(\hat{\Theta}^l)}{Pr(\hat{\Theta}^t)} &= \frac{z^l - z_{ex}\epsilon}{z^t - J(E^t - E_{ex}) z_{ex}\epsilon} \\
&= \frac{z^l - z_{ex}\epsilon}{(z^l - z_{ex}\epsilon) J(E^t - E^l(\epsilon)) + z^l \epsilon [k_2(\epsilon) - k_1(\epsilon)\epsilon]}
\end{aligned}
\tag{217}
$$

and since $h(E^t - E^l(\epsilon)) - \frac{1}{J(E^t - E^l(\epsilon))}$,

$$
\begin{aligned}
\hat{b} \equiv (1 - \frac{h_{\hat{q}}(\hat{m}))}{h(E^t - E^l(\epsilon))}) Pr(\hat{\Theta}^t) &= h_q(\hat{m}) \frac{z^l}{z^l - z_{ex}\epsilon} (z^t - J(E^t - E^l(\epsilon)) z_{ex}\epsilon) b(k_2(\epsilon) - k_1(\epsilon)\epsilon)\epsilon \\
&\approx h(E^t - E^l) z^t k_2(\epsilon)\epsilon b
\end{aligned}
\tag{218}
$$

for small ϵ. Since $\lim_{\epsilon \to 0} \hat{b} = 0$, so Lemma 12, (198) and (215) imply that for small enough ϵ there exists an incentive compatible mechanism $\tilde{\Omega}$ such that $EU^I_{DM}(\tilde{\Omega}) > EU^I_{DM}(\hat{\Omega}) > k_{\hat{b}}\hat{b} > 0$, and thus for $EU_{DM}(\tilde{\Omega}) - EU_{DM}(\Omega) > 0$ for small enough ϵ, but it contradicts that Ω is an optimal mechanism. Therefore, we have shown that for almost every $\theta \in \Theta^u_q(\mathcal{M}^+_q)$, θ is essentially revealed upon inspection in an optimal mechanism Ω.

Finally, suppose there is a positive measure set $M' \subseteq \mathcal{M}^+_q)$ and $s = t, l$ such that for $m \in M'$, $\Theta^s_q(m) \neq \{X(m,s)\}$, then there is a positive measure set Θ' such that $m_q(\theta) \neq \theta$, which contradicts that θ is essentially revealed upon inspection.

■

Lemma 13 implies in an optimal mechanism where $Pr(\Theta_q^0) > 0$ and $\mu - \overline{\theta}^l > \underline{\theta}^t - \mu$, every inspected type is separated upon inspection, so there exists an matching function from the set of truth-tellers to the set of liars $\phi_q : \Theta_q^t(\mathcal{M}_q^+) \to \Theta_q^l(\mathcal{M}_q^+)$ such that for $m_q(\theta) = m_q(\theta')$ if and only if $\phi_q(\theta) = \theta'$.

Lemma 14 *Suppose for an optimal mechanism Ω, $Pr(\Theta_q^0) > 0$ and $\mu - \overline{\theta}^l > \underline{\theta}^t - \mu$, then $(\underline{\theta}^t - \mu)(\mu - \overline{\theta}^l) \geq c$ and for any $m_1, m_2 \in \mathcal{M}_q^+$, $X(m_1, t) > X(m_2, t)$ if and only if $X(m_1, l) < X(m_2, l)$.*

Proof of Lemma 14: We consider two cases that contrary to the claim.

Case 1: $(\underline{\theta}^t - \mu)(\mu - \overline{\theta}^l) \geq c$ **and there exists positive measure sets** $M_1, M_2 \subseteq \mathcal{M}_q^+$ **such that** $M_1 \cap M_2 = \emptyset$ **and for all** $(m_1, m_2) \in (M_1, M_2)$**,** $X(m_1, t) \geq X(m_2, t)$ **and** $X(m_1, l) \geq X(m_2, l)$:

By Lemma 13 every inspected types is separating, then for $m_1 \neq m_2$ and $s = t, l$, $X(m_1, s) \neq X(m_2, s)$, so it must be the case that for any $(m_1, m_2) \in (M_1, M_2)$ $X(m_1, t) > X(m_2, t)$ and $X(m_1, l) > X(m_2, l)$. Therefore, there exist $\delta > 0$ such that for any $\epsilon > 0$ there are positive measure subsets $M_1' \subseteq M_1$ and $M_2' \subseteq M_2$ such that for any $m_i \in M_i'$, $i = 1, 2$ and $s = t, l$,

$$X(m_1, s) > X(m_2, s) > \delta \text{ for } (m_1, m_2) \in (M_1, M_2) \tag{219}$$

$$|X(m, s) - X(m', s)| < \epsilon \text{ for } m, m' \in M_i, i = 1, 2 \tag{220}$$

$$X(m_2, t) > \inf_{m' \in \mathcal{M}_q^+} X(m', t) \tag{221}$$

and $\frac{1}{r} Pr(M_2') = Pr(M_1') \leq \epsilon$ for any $r \in (0, 1)$. Denote $b = Pr(M_1')$. For $i \in 1, 2$ and $s \in t, l, u$, let $\Theta_i^s = \Theta_q^s(M_i')$ be the aggregate sets of truthful senders, lying senders and senders of M_i'', and $E_i^s = E[\Theta_i^s]$ be their corresponding expected value and $z_i^s = \frac{Pr(\Theta_i^s)}{Pr(\Theta_i^u)}$ be their corresponding ratios of measure to set of senders Θ_i^u.

Since the original mechanism is optimal, we have $(X(m, t) - X(m, u))(X(m, u) - X(m, l)) = c$ and $w_q(m) = w^-(X(m, t) - X(m, l)) \leq 0.5$ almost everywhere at M_i', so $|E^s - X(m, s)| < \epsilon$ for any $m \in M'$ and $s = t, l$ and implies

$$\frac{z_i^l}{z_i^t} = \frac{w^-(E_i^t - E_i^l)}{1 - w^-(E_i^t - E_i^l)} + k_1(\epsilon)\epsilon^2 = h(E_i^t - E_i^l) + k_i(\epsilon)\epsilon^2 \tag{222}$$

where $k_i(\epsilon)$ is a bounded function. Define $\hat{r} = \frac{z_1^t}{z_2^l} h(E_1^t - E_2^l)$, and take the sets M_1, M_2 such that $Pr(M_2') = \hat{r} Pr(M_1')$, so that

$$\frac{Pr(\Theta_2^l)}{Pr(\Theta_1^t)} = \frac{\hat{r} z_2^l}{z_1^t} = h(E_1^t - E_2^l) \tag{223}$$

Now define the modified message and action rules $\hat{q}, \hat{X}$ where other things remain unchanged, except the set of messages $M_1' \cup M_2'$ is off-path; an inspected message m_{ex} is added with the set of truthful senders $\Theta^t_{\hat{q}}(m_{ex}) = \Theta^t_1$, and the set of lying senders $\Theta^l_{\hat{q}}(m_{ex}) = \Theta_2'$; an inspected message $\hat{m}$ is added with the set of truthful senders $\Theta^t_{\hat{q}}(\hat{m}) = \Theta^t_2$, and the set of lying senders $\Theta^l_{\hat{q}}(\hat{m}) = \Theta^t_1$.

The sequentially rational actions for the lower modified message m_{ex} are

$$\begin{aligned}
&\hat{X}(m_{ex}, t) = E^t_1 \\
&\hat{X}(m_{ex}, l) = E^l_2 \\
&\hat{X}(m_{ex}, u) = x^*_u(E^t_1, E^l_2) > x^*_u(E^t_1, E^l_1) > \mu
\end{aligned} \tag{224}$$

The sequentially rational actions for the upper modified message $\hat{m}$ are

$$\begin{aligned}
\hat{X}(\hat{m}, t) &= E^t_2 \\
\hat{X}(\hat{m}, l) &= E^l_1 \\
x^*_u(E^t, E^l(\epsilon)) &= x^*_u(E^t_2, E^l_1) > \mu
\end{aligned} \tag{225}$$

where the inequality holds because $(E^t_2 - \mu)(\mu - E^l_1) > (\underline{\theta}^t - \mu)(\mu - \overline{\theta}^l) \geq c$.

To compare DM's payoffs,

$$\begin{aligned}
&EU^I_{DM}(\hat{\Omega}) - EU^I_{DM}(\Omega) \\
=&b \sum_{s=t,l} [(z^s_1)^2 - c] + \hat{r} b \sum_{s=t,l} [(z^s_2)^2 - c] \\
&- \sum_{i=1,2} \sum_{s=t,l} \int_{M_i'} [X(m,s)^2 - c] \int_{\Theta^s_q(m)} dF(\theta) dm \\
=&- \sum_{i=1,2} \sum_{s=t,l} \int_{M'} (X(m,s) - E^s_i)^2 \int_{\Theta^s_q(m)} dF(\theta) dm \\
>&- (1 + \hat{r}) \frac{1}{4} \epsilon^2 b
\end{aligned} \tag{226}$$

where the second equality holds by sequential rational actions ; the inequality holds because of (220) and Popoviciu's inequality. By (223), $w_{\hat{q}}(m_{ex}) = w^-(E^t_1 - E^l_2)$. Now For $\hat{m}$,

$$\begin{aligned}
h_{\hat{q}}(\hat{m}) = \frac{Pr(\Theta^l_1)}{Pr(\Theta^t_2)} &= \frac{z^l_1}{\hat{r} z^t_2} \\
&= \frac{z^l_1 z^l_2}{z^t_1 z^t_2} \frac{1}{h(E^t_1 - E^l_2)}
\end{aligned}$$

$$= \frac{h(E_1^t - E_1^l)h(E_2^t - E_2^l)}{h(E_1^t - E_2^l)} + g(\epsilon)\epsilon^2 \tag{227}$$

where the second equality holds by definition of $\hat{r}$, and the last equality holds by (222), where $g(\epsilon)$ is a bounded function. Since $h(.)$ is a strictly convex function and $E_1^t - E_2^l - \delta > \max_i E_i^t - E_i^l \geq \min_i E_i^t - E_i^l > E_2^t - E_1^l + \delta$,

$$\begin{aligned} \hat{b} &\equiv (1 - \frac{h_{\hat{q}}(\hat{m})}{h(E_2^t - E_1^l)})Pr(\Theta_2^t) \\ &\approx (1 - \frac{h(E_1^t - E_1^l)h(E_2^t - E_2^l)}{h(E_1^t - E_2^l)h(E_2^t - E_1^l)})\hat{r}z_2^t b \end{aligned} \tag{228}$$

for small ϵ. Since $\lim_{\epsilon \to 0} b = 0$, so Lemma 12, (221) and (225) imply that for small enough ϵ there exists an incentive compatible mechanism $\tilde{\Omega}$ such that $EU_{DM}^I(\tilde{\Omega}) > EU_{DM}^I(\hat{\Omega}) > k_{\hat{b}}\hat{b} > 0$, and thus for $EU_{DM}(\tilde{\Omega}) - EU_{DM}(\Omega) > 0$ for small enough ϵ, but it contradicts that Ω is an optimal mechanism.

Case 2: $(\underline{\theta}^t - \mu)(\mu - \overline{\theta}^l) < c$:

Then there exists positive measure $\Theta^l \subseteq \Theta_q^l(\mathcal{M}_q^+)$ such that for any $\theta \in \Theta_l$, $(\underline{\theta}^t - \mu)(\mu - \theta) < c$. Incentive compatibility implies $(\phi^{-1}(\theta) - \mu)(\mu - \theta) > c$, where $\phi^{-1}(\theta)$ is the truth-teller who match with θ, then there exists $\delta_1 > 0$ and positive measure subset $\Theta_i^l \subseteq \Theta^l$ such that for all $\theta \in \Theta_i^l$, $(\phi^{-1}(\theta) - \delta_1 - \mu)(\mu - \theta) > c$. Now for each $\theta \in \Theta_i^l$, define $R(\theta)$ be such that $(R(\theta) - \mu)(\mu - \theta) = c$, So we have

$$\underline{\theta}^t < R(\theta) < \phi^{-1}(\theta) - \delta_1 \tag{229}$$

By Lemma 9, $R(\theta) > \underline{\theta}^t \geq \overline{\theta}^0$, so $R(\theta) \in \Theta_q^t(\mathcal{M}_q^+)$. Let $\Theta_{ii}^l = \{\theta \in \Theta_i^l : R(\theta) \notin \phi^{-1}(\Theta_l')\}$ to remove any liars type with $R(\theta)$ duplicate with $\phi^{-1}(\theta')$ for other types in Θ_i^l. Second inequality of (229) implies that Θ_{ii}^l has positive surplus. Incentive compatibility implies $(R(\theta) - \mu)(\mu - \phi(R(\theta))) > c$, so there exists $\delta_2 > 0$ and positive measure subset $\Theta_{iii}^l \subseteq \Theta_{ii}^l$ such that for all $\theta \in \Theta_{iii}^l$, $(R(\theta) - \mu)(\mu - \phi(R(\theta)) + \delta_2) > c$, So we have

$$\phi(R(\theta)) < \theta - \delta_2 \tag{230}$$

let $\delta = min\{\delta_1, \delta_2\}$, and pick a positive measure subset $\Theta_{iv}^l \subseteq \Theta_{iii}^l$ such that for any $\theta, \theta' \in \Theta_{iv}^l$, $\max\{|\theta - \theta'|, |R(\theta) - R(\theta')|, |\phi^{-1}(\Theta_l) - \phi^{-1}(\Theta_l')|, |\phi(R(\theta)) - \phi(R(\theta'))|\} < \epsilon\frac{\delta}{4}$, so that by (229) and (230), for any $\theta, \theta' \in \Theta_{iv}^l$,

$$R(\theta') < \phi^{-1}(\theta) - \frac{\delta}{2} \tag{231}$$

$$\phi(R(\theta')) < \theta - \frac{\delta}{2} \tag{232}$$

Now let $M_1 = m_q(\Theta^l_{iv})$ and $M_2 = m_q(R(\Theta^l_{iv}))$, so for every $(m_1, m_2) \in (M_1, M_2)$,

$$X(m_2, t) < X(m_1, t) - \frac{\delta}{2} \tag{233}$$

$$X(m_2, l) < X(m_1, l) - \frac{\delta}{2} \tag{234}$$

and for $\theta \in \Theta^l_q(M_1)$ there exists $\theta' \in \Theta^t_q(M_2)$ such that

$$(\theta' - \mu)(\mu - \theta) = c \tag{235}$$

Now for $i = 1, 2$, $s = t, l$, let $\Theta^s_i = \Theta^s_q(M_i)$ be the truthful and lying sets, and for $\epsilon \in (0, 1)$, $\theta^s_i(\epsilon) = \{\theta \in \Theta^s_i : Pr(\Theta^s_i \cap [0, \theta]) = \epsilon Pr(\Theta^s_i)\}$ be the $\epsilon - th$ percentile of Θ^s_i. Define $M_1(\epsilon) = m_q(\Theta^l_1 \cap [0, \theta^l_1(\epsilon^2)])$ be the subset of messages of M_1 where the liars are on the bottom $\epsilon^2 - th$ percentile, and $M_2(\epsilon) = m_q(\Theta^t_2 \cap [0, \theta^l_1(1 - \epsilon^2)])$ be the subset of messages of M_2 where the truth-tellers are on the top $\epsilon^2 - th$ percentile, so that for small enough ϵ,

$$|X(m, s) - X(m', s)| < \epsilon \text{ for } m, m' \in M_i(\epsilon), i = 1, 2 \tag{236}$$

$$X(m_2, t) > \inf_{m' \in \mathcal{M}^+_q} X(m', t) \tag{237}$$

$$(E^t_2(\epsilon) - \mu)(\mu - E^l_1(\epsilon)) > c \tag{238}$$

where $E^s_i(\epsilon) = E[\Theta^s_q(M_i(\epsilon))]$, which implies $X*_u(E^t_2(\epsilon), E^l_1(\epsilon)) > \mu$. Then we can derive a contradiction using the same argument as Case 1.

■

Lemma 15 *Suppose for an optimal mechanism Ω, $Pr(\Theta^0_q) > 0$ and $\underline{\theta}^t - \mu \geq \mu - \overline{\theta}^l$, then $\overline{\theta}^l \leq \underline{\theta}^0$.*

Proof of Lemma 15: Suppose contrary to the claim, $\overline{\theta}^l > \underline{\theta}^0$, then it must be the case that $\mu - \overline{\theta}^l \leq \sqrt{c}$, for otherwise Lemma 9 and $\Theta^t_q(\mathcal{M}^+_q) \cup \Theta^t_q(\mathcal{M}^+_q) \cup \Theta^0_q = [0, 1]$ imply that $\underline{\theta}^t = \overline{\theta}^0$, thus $(\overline{\theta}^0 - \mu)(\mu - \underline{\theta}^0) = (\underline{\theta}^t - \mu)(\mu - \underline{\theta}^0) > (\underline{\theta}^t - \mu)(\mu - \overline{\theta}^l) > c$, which contradicts Lemma 8.

Now since $\overline{\theta}^l > \underline{\theta}^0$, there exist $\delta > 0$, a positive measure subset of uninspected types $\Theta_0 \subseteq \Theta^0_q$ and a positive set of lying types $\Theta_l \subseteq \Theta^l_q(\mathcal{M}^+_q)$ such that for any $\theta \in \Theta_l$, $\underline{\mu} \equiv E[\Theta_0]$ and $m \in m_q(\Theta_l)$

$$\theta - \underline{\mu} > \delta \tag{239}$$

$$\theta \geq x^d_\Omega(\theta_l) \tag{240}$$

$$X(m,u) > \mu + \delta \tag{241}$$

$$X(m,l) < \mu - \delta \tag{242}$$

$$w_q(m) < 0.5 - \delta \tag{243}$$

Let $M_l = m_q(\Theta_l)$ be the set of messages sent by those lying types, $a = \frac{Pr(\Theta_l)}{Pr(\Theta_q^l(M_l))} > 0$, and $M_l^a = \cap M \subseteq M_l : \frac{Pr(\Theta_l \cap \Theta_q^l(M))}{Pr(\Theta_q^l(M))} \geq a$ be the subset containing every message in M_l where proportion of lying types within Θ_l is no less than a. Since $Pr(M_t^a) > 0$, for any $\epsilon > 0$ there is positive measure subset $M' \subseteq M_t^a$ such that for any $m, m' \in M'$ and $s \in \{t, l, u\}$ and $\theta \in \Theta_l \cap \Theta_q^l(M_l^a)$,

$$|X(m,s) - X(m',s)| < \epsilon \tag{244}$$

$$\theta > \mu - \sqrt{c} - \epsilon \tag{245}$$

and $Pr(M') \equiv b \leq \epsilon$.

For $s \in \{t, l, u\}$, let $\Theta^s = \Theta_q^s(M')$, $E^s[\Theta^s]$ and $z^s = \frac{Pr(\Theta^s)}{b}$ be the aggregate set of truth-tellers, liars and senders of M', their corresponding expected values and ratios of measure to set of senders Θ^u. Let $\Theta' = \Theta_q^l(M') \cap \Theta_l$ be the set of lying types in M' that satisfies (239) and (240), and $\Theta_{ex} = \Theta^l / \Theta'$ be the set of lying types who send M' but not in Θ'; $E' = E[\Theta']$, $E_{ex} = E[\Theta_{ex}]$, $z' = \frac{Pr(\Theta')}{b}$ and $z_{ex} = \frac{Pr(\Theta_{ex})}{b}$ be their corresponding expected values and ratios of measure to set of senders Θ^u. By (239), (240), (245) and $M' \subseteq M_l^a$,

$$E' > \underline{\mu} + \delta \tag{246}$$

$$E' \geq E^l \geq E_{ex} \tag{247}$$

$$E' \geq \mu - \sqrt{c} - \epsilon \tag{248}$$

$$z' \geq a z^l \tag{249}$$

Since the original mechanism is optimal, we have $(X(m,t) - X(m,u))(X(m,u) - X(m,l)) = c$ and $w_q(m) = w^-(X(m,t) - X(m,l)) \leq 0.5$ almost everywhere at M', so $|E^s - X(m,s)| < \epsilon$ for any $m \in M'$ and $s = t, l, u$ imply

$$\frac{z^l}{z^t} = \frac{w^-(E^t - E^l)}{1 - w^-(E^t - E^l)} + k_1(\epsilon)\epsilon^2 = h(E^t - E^l) + k_1(\epsilon)\epsilon^2 \tag{250}$$

$$\frac{z^t}{z^l} = \frac{1 - w^-(E^t - E^l)}{w^-(E^t - E^l)} + k_1(\epsilon)\epsilon^2 = J(E^t - E^l) + k_2(\epsilon)\epsilon^2 \tag{251}$$

where $k_1(\epsilon)$ and $k_2(\epsilon)$ are bounded functions.

Now we consider two cases.

Cases 1: $\lim_{\epsilon\to 0} E' - E^l = 0$**:**

Let $\hat{r} = h(E^t - \underline{\mu})\frac{z^t}{z^l}$. Since $lim_{\epsilon\to 0}E^l - \underline{\mu} = lim_{\epsilon\to 0}E' - \underline{\mu} > \delta > 0$ and $h'(.) < 0$, (250) imply that

$$lim_{\epsilon\to 0}\hat{r} = lim_{\epsilon\to 0}\frac{h(E^t - \underline{\mu})}{h(E^t - E^l)} \in (0,1) \tag{252}$$

Let $\hat{\Theta}^l = \Theta_0(\hat{r}z^l b\frac{1}{Pr(\Theta_0)})$ be the modified set of liars, where $\Theta_0(\hat{r}z^l b\frac{1}{Pr(\Theta_0)})$ is a mean-preserving division of Θ_0 so that

$$Pr(\hat{\Theta}^l) = \hat{r}z^l b = \hat{r}Pr(\Theta^l) \tag{253}$$

$$E[\hat{\Theta}^l] = \underline{\mu} \tag{254}$$

Let $\Theta^0_{\hat{q}} = (\Theta^0_q/\hat{\Theta}^l) \cup \Theta^l$ be the modified uninspected set.

Now define the modified message ans action rules $\hat{q}, \hat{X}$ where other things remain unchanged, except the set of messages M' is off-path; an inspected message $\hat{m}$ is added with the set of truthful senders $\Theta^t_{\hat{q}}(\hat{m}) = \Theta^t$ and the set of lying senders $\Theta^l_{\hat{q}}(\hat{m}) = \hat{\Theta}^l$; The uninspected message is modified to $m^0_{\hat{q}}$ with the set of senders $\Theta^0_{\hat{q}}$.

The sequentially rational actions for $\hat{m}$ are

$$\begin{aligned}
&\hat{X}(\hat{m}, t) = E^t \\
&\hat{X}(\hat{m}, l) = \underline{\mu} \\
&\hat{X}(\hat{m}, t) = x^*_u(E^t, \underline{\mu}) > x^*_u(E^t, E^l) > \mu
\end{aligned} \tag{255}$$

where the first inequality holds for small enough ϵ because $\lim_{\epsilon\to 0} E^l - \underline{\mu} > 0$, and the second inequality holds by optimality of Ω.

The sequentially rational action for the modified uninspected message $m^0_{\hat{q}}$ is

$$\begin{aligned}
\hat{X}(m^0_{\hat{q}}, u) &= \mu + \frac{\hat{r}z^l b}{Pr(\Theta^0_q) + (1-\hat{r})z^l b}(\mu - \underline{\mu}) - \frac{z^l b}{Pr(\Theta^0_q) + (1-\hat{r})z^l b}(\mu - E^l) \\
&= \mu + k_3(\epsilon)b
\end{aligned} \tag{256}$$

where $k_3(\epsilon)$ is a bounded function.

By (253) and the definition of $\hat{r}$, $Pr(\hat{\Theta}^l) = h(E^t - \underline{\mu})Pr(\Theta^t)$, so $w_{\hat{q}}(\hat{m}) = w^-(E^t - \underline{\mu}) = w^-(\hat{X}(\hat{m}, t) - \hat{X}(\hat{m}, l))$, so

$$w_{\hat{q}}(m) = w^-(\hat{X}(m, t) - \hat{X}(m, l)) \tag{257}$$

hold almost everywhere at $\mathcal{M}_{\hat{q}}^{+}$.

To compare DM's payoffs, since $V_q(m) = V_{\hat{q}}(m) = c$, by Lemma 4, $EU_{DM}(\Omega) = EU_{DM}^{I}(\Omega)$ and $EU_{DM}(\hat{\Omega}) = EU_{DM}^{I}(\hat{\Omega})$, so

$$\begin{aligned}
&EU_{DM}(\hat{\Omega}) - EU_{DM}(\Omega)\\
=&(Pr(\Theta_q^0) + (1-\hat{r})z^l b)E[\Theta_{\hat{q}}^0]^2 + \hat{r}z^l b[\underline{\mu}^2 - c] + z^t b[(E^t)^2 - c]\\
&- Pr(\Theta_q^0)\mu^2 - \sum_{s=t,l}\int_{M'}(X(m,s)^2 - c)\int_{\Theta_q^s(m)} dF(\theta)dm\\
=&- Pr(\Theta_q^0)(\mu - E[\Theta_{\hat{q}}^0])^2 + \hat{r}z^l b[(E[\Theta_{\hat{q}}^0] - \underline{\mu})^2 - c] - z^l b[(E[\Theta_{\hat{q}}^0] - E^l)^2 - c]\\
&- \sum_{s=t,l}\int_{M'}(X(m,s) - E^s)^2\int_{\Theta_q^s(m)} dF(\theta)dm\\
>&- Pr(\Theta_q^0)k_3(\epsilon)^2 b^2 - (1-\hat{r})z^l b[(E[\Theta_{\hat{q}}^0] - E^l)^2 - c]\\
&+ \hat{r}z^l b[(E[\Theta_{\hat{q}}^0] - \underline{\mu})^2 - (E[\Theta_{\hat{q}}^0] - E^l)^2] - \frac{1}{4}\epsilon^2 b\\
>&- Pr(\Theta_q^0)k_3(\epsilon)^2 b^2 - (1-\hat{r})z^l b[(E[\Theta_{\hat{q}}^0] - E^l)^2 - c]\\
&+ \hat{r}z^l b(E^l - \underline{\mu})^2 - \frac{1}{4}\epsilon^2 b\\
>&\hat{r}z^l b\delta^2 - k_4(\epsilon)\epsilon b
\end{aligned} \tag{258}$$

where the second equality holds by sequential rational actions ; the first inequality holds because of (256), (244) and Popoviciu's inequality; the last inequality holds for small enough ϵ and a bounded function $k_4(\epsilon)$ because $b \le \epsilon$, so $Pr(\Theta_q^0)k_3(\epsilon)^2 b^2 \le Pr(\Theta_q^0)k_3(\epsilon)^2\epsilon b$; (248) implies $\lim_{\epsilon\to 0}(E[\Theta_{\hat{q}}^0] - E^l)^2 - c = \lim_{\epsilon\to 0}(\mu - E')^2 - c \le 0$, and (246) implies $\lim_{\epsilon\to 0}(E^l - \underline{\mu})^2 = \lim_{\epsilon\to 0}(E' - \underline{\mu})^2 > \delta^2$.

Since $V_q(m) = V_{\hat{q}}(m) = c$, by Lemma 4, $EU_{DM}(\hat{\Omega}) = EU_{DM}^{U}(\hat{\Omega}) = EU_{DM}^{I}(\hat{\Omega})$, and (256) implies that $\hat{X}(m_{\hat{q}}^0, u) - X(m_q^0, u) < k_3(\epsilon)b$, so Lemma 5 implies that there exists an incentive compatible mechanism $\tilde{\Omega}$ such that $EU_{DM}(\tilde{\Omega}) > EU_{DM}(\hat{\Omega}) - 4(k_3(\epsilon)b)^2$, then by (258) $EU_{DM}(\tilde{\Omega}) - EU_{DM}(\Omega) > \hat{r}z^l b\delta^2 - k_4(\epsilon)\epsilon b - 4(k_3(\epsilon)b)^2 > 0$ for small enough ϵ, but it contradicts that Ω is an optimal mechanism.

Case 2:$\lim_{\epsilon\to 0} E' - E^l > 0$:

For $r \in [0,1]$, define $E_1^l(r) = E^l + r\epsilon^{\frac{1}{3}}(E' - E_{ex})$ and $E_2^l(r) = E^l - r\epsilon^{\frac{1}{3}}(E' - E_{ex}) - \epsilon(E' - \underline{\mu})$. Define $\hat{r} \in (0,1)$ that solves

$$J(E^t - E_1^l(r)) + J(E^t - E_2^l(r)) = 2\frac{z^t}{z^l} \tag{259}$$

To show that such solution exists for small enough ϵ, $E_1^l(0) = E^l$ and $E_2^l(0) = E^l - \epsilon(E' - \underline{\mu}) < E^l$, by

(246) and (251),

$$\begin{aligned}
&\lim_{\epsilon\to 0}\frac{1}{\epsilon}[J(E^t-E_1^l(0))+J(E^t-E_2^l(0))-2\frac{z^t}{z^l}]\\
=&\lim_{\epsilon\to 0}\frac{1}{\epsilon}[J(E^t-E^l)+J(E^t-E^l+\epsilon(E'-\underline{\mu}))-2(J(E^t-E^l)+k_2(\epsilon)\epsilon^2)]\\
=&J'(E^t-E^l)(E'-\underline{\mu})>J'(E^t-E^l)\delta>0 \qquad (260)
\end{aligned}$$

On the other side, $J(E^t-E_1^l(1))=J(E^t-E^l)-J'(E^t-E^l)(\epsilon^{\frac{1}{3}}(E'-E_{ex}))+J''(E^t-\tilde{E}_1)(\epsilon^{\frac{1}{3}}(E'-E_{ex}))^2$ and $J(E^t-E_1^2(1))=J(E^t-E^l)+J'(E^t-E^l)(\epsilon^{\frac{1}{3}}(E'-E_{ex})+\epsilon(E'-\epsilon))+J''(E^t-\tilde{E}_2)(\epsilon^{\frac{1}{3}}(E'-E_{ex})+\epsilon(E'-\epsilon))^2$, where $E_1^l(1)>\tilde{E}_1>E^l>\tilde{E}_2>E_1^2(1)$. Therefore, $J(E^t-E_1^l(1))+J(E^t-E_2^l(1))=2J(E^t-E^l)+J'(E^t-E^l)\epsilon(E'-\epsilon)+J''(E^t-\tilde{E}_1)(\epsilon^{\frac{1}{3}}(E'-E_{ex}))^2+J''(E^t-\tilde{E}_2)(\epsilon^{\frac{1}{3}}(E'-E_{ex})+\epsilon(E'-\epsilon))^2$, thus

$$\begin{aligned}
&\lim_{\epsilon\to 0}\frac{1}{\epsilon^{\frac{2}{3}}}[J(E^t-E_1^l(1))+J(E^t-E_2^l(1))-2\frac{z^t}{z^l}]\\
=&\lim_{\epsilon\to 0}\frac{1}{\epsilon^{\frac{2}{3}}}[J(E^t-E_1^l(1))+J(E^t-E_2^l(1))-2(J(E^t-E^l)+k_2(\epsilon)\epsilon^2)]\\
=&h''(E^t-E^l)(E'-E_{ex})^2>0 \qquad (261)
\end{aligned}$$

where the inequality holds because $J(.)$ is strictly concave and $E'-E_{ex}>E'-E^l>0$, thus for small enough $\epsilon>0$, there exists $\hat{r}\in(0,1)$ such that (259) is satisfied.

Divide Θ_{ex} into two mean-preserving divisions $\Theta_{ex,1}$ and $\Theta_{ex,2}$ so that $E[\Theta_{ex,1}]=E[\Theta_{ex,2}]=E_{ex}$, $Pr(\Theta_{ex,1})=\frac{1}{2}[Pr(\Theta_{ex})-z^lb\hat{r}\epsilon^{\frac{1}{3}}]$ and $Pr(\Theta_{ex,2})=\frac{1}{2}[Pr(\Theta_{ex})+z^lb\hat{r}\epsilon^{\frac{1}{3}}]$; Divide Θ' into three mean-preserving divisions Θ'_0, Θ'_1 and Θ'_2 so that $E[\Theta'_0]=E[\Theta'_1]=E[\Theta'_2]=E'$, $Pr(\Theta'_0)=\frac{1}{2}z^lb\epsilon$, $Pr(\Theta'_1)=\frac{1}{2}[Pr(\Theta')+z^lb\hat{r}\epsilon^{\frac{1}{3}}]$ and $Pr(\Theta'_2)=\frac{1}{2}[Pr(\Theta')-z^lb\hat{r}\epsilon^{\frac{1}{3}}-z^lb\epsilon]$; Divide Θ^t into two mean-preserving divisions $\hat{\Theta}_1^t$ and $\hat{\Theta}_2^t$ so that $E[\hat{\Theta}_1^t]=E[\hat{\Theta}_2^t]=E^t$, $Pr(\hat{\Theta}_1^t)=\frac{1}{2}J(E^t-E_1^l(\hat{r}))Pr(\Theta^l)$ and $Pr(\hat{\Theta}_2^t)=\frac{1}{2}J(E^t-E_2^l(\hat{r}))Pr(\Theta^l)$; Divide a mean-preserving set $\Theta_{0,2}$ from Θ_0 so that $E[\Theta_{0,2}]=\underline{\mu}$ and $Pr(\Theta_{0,2})=\frac{1}{2}z^lb\epsilon$.

Let $\hat{\Theta}_1^l=\Theta_{ex,1}\cup\Theta'_1$ be the set of liars for the modified upper message, so that

$$\begin{aligned}
Pr(\hat{\Theta}_1^l)&=\frac{1}{2}[Pr(\Theta_{ex})-z^lb\hat{r}\epsilon^{\frac{1}{3}}]+\frac{1}{2}[Pr(\Theta')+z^lb\hat{r}\epsilon^{\frac{1}{3}}]\\
&=\frac{1}{2}z^lb \qquad (262)\\
E[\hat{\Theta}_1^l]&=\frac{E_{ex}[Pr(\Theta_{ex})-z^lb\hat{r}\epsilon^{\frac{1}{3}}]+E'[Pr(\Theta')+z^lb\hat{r}\epsilon^{\frac{1}{3}}]}{z^lb}\\
&=E^l+\hat{r}\epsilon^{\frac{1}{3}}(E'-E_{ex})=E_1^l(\hat{r}) \qquad (263)
\end{aligned}$$

Let $\hat{\Theta}^l_2 = \Theta_{ex,2} \cup \Theta'_2 \cup \Theta_{0,2}$ be the set of liars for the modified lower message, so that

$$\begin{aligned} Pr(\hat{\Theta}^l_2) &= \frac{1}{2}[Pr(\Theta_{ex}) + z^l b\hat{r}\epsilon^{\frac{1}{3}}] + \frac{1}{2}[Pr(\Theta') - z^l b\hat{r}\epsilon^{\frac{1}{3}} - z^l b\epsilon] + \frac{1}{2}z^l b\epsilon \\ &= \frac{1}{2}z^l b \end{aligned} \tag{264}$$

$$\begin{aligned} E[\hat{\Theta}^l_1] &= \frac{E_{ex}[Pr(\Theta_{ex}) + z^l b\hat{r}\epsilon^{\frac{1}{3}}] + E'[Pr(\Theta') - z^l b\hat{r}\epsilon^{\frac{1}{3}} - z^l b\epsilon] + \underline{\mu} z^l b\epsilon}{z^l b} \\ &= E^l - \hat{r}\epsilon^{\frac{1}{3}}(E' - E_{ex}) - \epsilon(E' - \underline{\mu}) = E^l_2(\hat{r}) \end{aligned} \tag{265}$$

Let $\Theta^0_{\hat{q}} = (\Theta^0_q / \Theta_{0,2}) \cup \Theta'_0$ be the modified uninspected set.

Now define the modified message ans action rules $\hat{q}, \hat{X}$ where other things remain unchanged, except the set of messages M' is off-path; an upper inspected message m_1 is added with the set of truthful senders $\Theta^t_{\hat{q}}(m_1) = \hat{\Theta}^t_1$ and the set of lying senders $\Theta^l_{\hat{q}}(m_1) = \hat{\Theta}^l_1$; an lower inspected message m_2 is added with the set of truthful senders $\Theta^t_{\hat{q}}(m_2) = \hat{\Theta}^t_2$ and the set of lying senders $\Theta^l_{\hat{q}}(m_2) = \hat{\Theta}^l_2$; The uninspected message is modified to $m^0_{\hat{q}}$ with the set of senders $\Theta^0_{\hat{q}}$

The sequentially rational actions for m_1 are

$$\begin{aligned} &\hat{X}(m_1, t) = E^t \\ &\hat{X}(m_1, l) = E^l_1(\hat{r}) \\ &\hat{X}(m_1, t) = x^*_u(E^t, E^l_1(\hat{r})) \approx x^*_u(E^t, E^l) > \mu \end{aligned} \tag{266}$$

where the approximation holds for small enough ϵ because $\lim_{\epsilon \to 0} E^l_1(\hat{r}) - E^l = 0$, and the second inequality holds by optimality of Ω.

The sequentially rational actions for m_2 are

$$\begin{aligned} &\hat{X}(m_2, t) = E^t \\ &\hat{X}(m_2, l) = E^l_2(\hat{r}) \\ &\hat{X}(m_2, t) = x^*_u(E^t, E^l_2(\hat{r})) > x^*_u(E^t, E^l) > \mu \end{aligned} \tag{267}$$

The sequentially rational action for the modified uninspected message $m^0_{\hat{q}}$ is

$$\hat{X}(m^0_{\hat{q}}, u) = \mu + \frac{\epsilon z^l b}{Pr(\Theta^0_q)}(E' - \underline{\mu}) = k_5(\epsilon) b\epsilon \tag{268}$$

where $k_5(\epsilon)$ is a bounded function. Since for $i = 1, 2$, $Pr(\hat{\Theta}^t_i) = J(E^t - E^l_i(\hat{r})) Pr(\hat{\Theta}^l_i)$, so $w_{\hat{q}}(m_i) = w^-(E^t - E^l_i(\hat{r})) = w^-(\hat{X}(m_i, t) - \hat{X}(m_i, l))$, so

$$w_{\hat{q}}(m) = w^-(\hat{X}(m, t) - \hat{X}(m, l)) \tag{269}$$

hold almost everywhere at $\mathcal{M}^+_{\hat{q}}$.

To compare DM's payoffs, since $V_q(m) = V_{\hat{q}}(m) = c$, by Lemma 4, $EU_{DM}(\Omega) = EU^I_{DM}(\Omega)$ and $EU_{DM}(\hat{\Omega}) = EU^I_{DM}(\hat{\Omega})$, so

$$
\begin{aligned}
&EU_{DM}(\hat{\Omega}) - EU_{DM}(\Omega)\\
=&Pr(\Theta^0_q)E[\Theta^0_{\hat{q}}]^2 + \frac{1}{2}z^l b\sum_{i=1,2}[E^l_i(\hat{r})^2 - c] + z^t b[(E^t)^2 - c]\\
&- Pr(\Theta^0_q)\mu^2 - \sum_{s=t,l}\int_{M'}(X(m,s)^2 - c)\int_{\Theta^s_q(m)} dF(\theta)dm\\
=&- Pr(\Theta^0_q)(\mu - E[\Theta^0_{\hat{q}}])^2 - \frac{1}{2}z^l b\epsilon(E' - E[\Theta^0_{\hat{q}}])^2 + \frac{1}{2}z^l b\epsilon(\underline{\mu} - E[\Theta^0_{\hat{q}}])^2\\
&+ \frac{1}{2}z^l b\hat{r}\epsilon^{\frac{1}{3}}\sum_{i=1,2}(E^l_i(\hat{r}) - E^l)^2 + \frac{1}{2}z^l b\epsilon(E' - E^l)^2 - \frac{1}{2}z^l b\epsilon(\underline{\mu} - E^l)^2\\
&- \sum_{s=t,l}\int_{M'}(X(m,s) - E^s)^2\int_{\Theta^s_q(m)} dF(\theta)dm\\
>&- Pr(\Theta^0_q)[k_5(\epsilon)b\epsilon]^2 - \frac{1}{2}z^l b\epsilon(E' - E[\Theta^0_{\hat{q}}])^2 + \frac{1}{2}z^l b\epsilon(\underline{\mu} - E[\Theta^0_{\hat{q}}])^2\\
&+ \frac{1}{2}z^l b\hat{r}\epsilon^{\frac{1}{3}}\sum_{i=1,2}(E^l_i(\hat{r}) - E^l)^2 + \frac{1}{2}z^l b\epsilon(E' - E^l)^2 - \frac{1}{2}z^l b\epsilon(\underline{\mu} - E^l)^2 - \frac{1}{4}\epsilon^2 b\\
>&- Pr(\Theta^0_q)[k_5(\epsilon)b\epsilon]^2 + z^l b\epsilon(E[\Theta^0_{\hat{q}}] - E^l)(E' - \underline{\mu}) - \frac{1}{4}\epsilon^2 b\\
>&- Pr(\Theta^0_q)[k_5(\epsilon)b\epsilon]^2 + z^l b\epsilon\delta^2 - \frac{1}{4}\epsilon^2 b\\
>&z^l b\epsilon\delta^2 - k_6(\epsilon)b\epsilon^2
\end{aligned}
\tag{270}
$$

where the second equality holds by sequential rational actions ; the first inequality holds because of (268), (244) and Popoviciu's inequality; the third inequality holds because (242) and (246) imply that $E[\Theta^0_{\hat{q}}] - E^l > \mu - E^l > \delta$ and $E' - \underline{\mu} > \delta$; the last inequality holds for a bounded function $k_6(\epsilon)$.

Since $V_q(m) = V_{\hat{q}}(m) = c$, by Lemma 4, $EU_{DM}(\hat{\Omega}) = EU^U_{DM}(\hat{\Omega}) = EU^I_{DM}(\hat{\Omega})$, and (268) implies that $\hat{X}(m^0_{\hat{q}}, u) - X(m^0_q, u) < k_5(\epsilon)b\epsilon$, so Lemma 5 implies that there exists an incentive compatible mechanism $\tilde{\Omega}$ such that $EU_{DM}(\tilde{\Omega}) > EU_{DM}(\hat{\Omega}) - 4(k_5(\epsilon)b\epsilon)^2$, then by (258) $EU_{DM}(\tilde{\Omega}) - EU_{DM}(\Omega) > z^l b\epsilon\delta^2 - k_6(\epsilon)b\epsilon^2 - 4(k_5(\epsilon)b\epsilon)^2 > 0$ for small enough ϵ, but it contradicts that Ω is an optimal mechanism.

■

Lemma 16 *In an optimal mechanism Ω with $Pr(\Theta^0_q) > 0$, $\overline{\theta}^l = \underline{\theta}^0$ and $\underline{\theta}^t = \overline{\theta}^0$. Furthermore, if $\underline{\theta}^t - \mu < \mu - \overline{\theta}^l$, then $(\underline{\theta}^t - \mu)(\mu - \overline{\theta}^l) = c$.*

Proof of Lemma 16: For $\underline{\theta}^t - \mu \geq \mu - \overline{\theta}^l$, Lemma 9 and Lemma 15 imply $\underline{\theta}^t \geq \overline{\theta}^0 > \underline{\theta}^0 \geq \overline{\theta}^l$. Since $\Theta_q^t(\mathcal{M}_q^+) \cup \Theta_q^t(\mathcal{M}_q^+) \cup \Theta_q^0 = [0,1]$, we have $\underline{\theta}^t = \overline{\theta}^0 > \underline{\theta}^0 = \overline{\theta}^l$.

For $\underline{\theta}^t - \mu < \mu - \overline{\theta}^l$, by Lemma 9 $\mu - \overline{\theta}^l > \underline{\theta}^t - \mu \geq \overline{\theta}^0 - \mu > 0$, so $\underline{\theta}^t = \overline{\theta}^0$ and $\overline{\theta}^l \geq \underline{\theta}^0$. Then Lemma 8 and Lemma 14 imply $(\underline{\theta}^t - \mu)(\mu - \overline{\theta}^l) \geq c \geq (\overline{\theta}^0 - \mu)(\mu - \underline{\theta}^0) = (\underline{\theta}^t - \mu)(\mu - \underline{\theta}^0)$, so $\overline{\theta}^l = \underline{\theta}^0$ and $(\underline{\theta}^t - \mu)(\mu - \overline{\theta}^l) = c$.

■