Notice that the limit is well defined because F^k_j is a nonnegative supermartingale with respect to j.[50] We will call this limit F^k_J to be consistent with the case in which J is finite.

Combining these observations with (9), we conclude that

$$\mathbb{E}_{i+1}\left[\sum_{j>i} \mathbb{1}_{\mathcal{A}_j}\beta_j\right] \leq \frac{2C}{\hat{F}^k}\left(F^k_{i+1} - \mathbb{E}_{i+1}[F^k_J]\right).$$

We have $\beta_j = \gamma_j F^1_j \lambda$. Moreover, Lemma 1 shows that $\gamma_j > 0$ only if $F^1_j \geq c/R$. This implies[51] that $\gamma_j \leq g\beta_j$ where $g = \frac{R}{\lambda c}$ and hence that

$$\mathbb{E}_{i+1}\left[\sum_{j\geq i+1} \mathbb{1}_{\mathcal{A}_j}\gamma_j\right] \leq \frac{2Cg}{\hat{F}^k}\left(F^k_{i+1} - \mathbb{E}_{i+1}[F^k_J]\right) \tag{10}$$

Let $\mathcal{Z}$ denote the event that at least one agent $j > i$ works and $\pi_{i+1}(m_i) = \Pr_{i+1}(\mathcal{A} \cap \mathcal{Z})$, i.e., the probability that $\mathcal{A}$ and $\mathcal{Z}$ both occur conditional on m^i_1. We have

$$\begin{aligned}
\pi_{i+1}(m_i) &= \Pr_{i+1}\left(\mathbb{1}_{\mathcal{A}}\sum_{j>i}\mathbb{1}_{j \text{ works}} \geq 1\right)\\
&\leq \mathbb{E}_{i+1}\left[\mathbb{1}_{\mathcal{A}}\sum_{j>i}\mathbb{1}_{j \text{ works}}\right]\\
&= \sum_{j>i}\mathbb{E}_{i+1}\left[\mathbb{1}_{\mathcal{A}}\mathbb{1}_{j \text{ works}}\right]\\
&\leq \sum_{j>i}\mathbb{E}_{i+1}\left[\mathbb{1}_{\mathcal{A}_j}\mathbb{1}_{j \text{ works}}\right]\\
&= \sum_{j>i}\mathbb{E}_{i+1}\left[\mathbb{E}_j\left[\mathbb{1}_{\mathcal{A}_j}\mathbb{1}_{j \text{ works}}\right]\right]\\
&= \sum_{j>i}\mathbb{E}_{i+1}\left[\mathbb{1}_{\mathcal{A}_j}\mathbb{E}_j\left[\mathbb{1}_{j \text{ works}}\right]\right]\\
&= \sum_{j>i}\mathbb{E}_{i+1}\left[\mathbb{1}_{\mathcal{A}_j}\gamma_j\right]
\end{aligned}$$

[50] The argument is similar to the one used to prove Lemma 4. For any fixed j, let $\bar{F}^{k+1}_l$ denote the probability assigned by $l \geq j$ to there remaining at least $k+1$ signals *at the beginning of round* j. The process $\{\bar{F}^{k+1}_l\}_{l\geq j}$ is a martingale in j, by the law of iterated expectations and the fact that that l's filtration grows finer as l increases. Moreover, $\bar{F}^{k+1}_l \geq F^{k+1}_l$ for all $l \geq j$, because the actual number of remaining signals only decreases over time. Therefore, we have $F^{k+1}_j = \bar{F}^{k+1}_j = E_j[\bar{F}^{k+1}_{j+1}] \geq E_j[F^{k+1}_{j+1}]$. This, together with the fact that F^{k+1}_l is uniformly bounded and measurable with respect to the information at the beginning round l, shows that it is a supermartingale.

[51] The inequality clearly holds if $\ {}_j = 0$.

The first equality comes from the fact that $\mathcal{Z}$ is identical to the event $\{\sum_{j>i} \mathbb{1}_{j \text{ works}} \geq 1\}$. The first inequality comes from the fact that $\mathbb{1}_{\mathcal{A}} \sum_{j>i} \mathbb{1}_{j \text{ works}}$ is nonnegative and integer valued, so that its expectation exceeds the probability that it is strictly positive. The second equality is an application of Tonelli's theorem. The second inequality comes from the fact that $\mathcal{A} \subset \mathcal{A}_j$ and, hence, $\mathbb{1}_{\mathcal{A}} \leq \mathbb{1}_{\mathcal{A}_j}$. The next equality comes from the law of iterated expectations and the next one comes from the fact that $\mathcal{A}_j$ is measurable with respect to the information at the beginning of round j. The last equality holds by definition of γ_j.

From (10) and Tonelli's theorem, this implies that

$$\pi_{i+1}(m_i) \leq \frac{2Cg}{\hat{F}^k} \left(F_{i+1}^k - \mathbb{E}_{i+1}[F_J^k]\right) \tag{11}$$

Let $F_i^{k,r}(m_i)$ denote the probability that there are at least k signals left conditional on i working and reporting m_i, and for any signal s_i let $F_i^{k,w}(s_i)$ denote the probability that there are at least k signals left conditional on i working and discovering s_i. $F_i^{k,w}(s_i)$ represents i's belief after discovering s_i, while $F_i^{k,r}(m_i)$ represents what $i+1$ would believe conditional on knowing that i worked and observing report m_i.

The following lemmas are proved in Appendices C.5 and C.6. Let $N_i = \{m_i : F_i^{k,r}(m_i) > (1+\eta)F_i^k\}$.

Lemma 5 *(i)* $\gamma_i(N_i) \leq \frac{1}{2\eta G^2}$, *(ii) for* $m_i \notin N_i$, $\pi_{i+1}(m_i) \leq \frac{2Cg}{\hat{F}^k}\left((1+\eta)F_i^k - \mathbb{E}_{i+1}[F_J^k]\right)$.

Let $T_i = \{m_i : F_{i+1}^{k+1}(m_i) \geq \sqrt{G}\varepsilon\}$.

Lemma 6 *(i)* $\Pr_{i+1}(\mathcal{A}) \geq 1 - 1/\sqrt{G}$ *for all* $m_i \notin T_i$, *(ii)* $\gamma_i(T_i) \leq 1/\sqrt{G}$.

Let V^w denote i's expected gross utility if he works and $V^w(m_i)$ denote his expected gross utility conditional on working and reporting m_i. We have

$$\begin{aligned} V^w &= \sum_{i \in M_i} \gamma_i(m_i) V^w(m_i) \\ &\leq (\gamma_i(N_i) + \gamma_i(T_i))R + \sum_{m_i \notin N_i \cup T_i} \gamma_i(m_i) V^w(m_i) \\ &\leq \left(\frac{1}{2\eta G^2} + \frac{1}{\sqrt{G}}\right) R + \sum_{m_i \notin N_i \cup T_i} \gamma_i(m_i) V^w(m_i) \end{aligned} \tag{12}$$

Moreover,

$$V^w(m_i) = \Pr(\mathcal{Z}|m_1^i) V^w(m_i|\mathcal{Z}) + \Pr(\mathcal{Z}^c|m_1^i) V^w(m_i|\mathcal{Z}^c). \tag{13}$$

Conditional on m_1^i, the event $\mathcal{Z}$ is independent of how m_i was produced (i.e., whether m_i was obtained by work or fabrication). Indeed, as long as no one works, the distribution of reports m_j made by agents following i depends only on m_1^i, not on the signals that remain to be discovered in the case. And as soon as someone works, then by definition $\mathcal{Z}$ has occurred. Thus, what triggers the event $\mathcal{Z}$ (whenever it occurs) is a sequence of uninformative (until $\mathcal{Z}$ occurs) reports m_j for agents following i, whose probability distribution is completely pinned down by m_1^i.

From the previous lemmas, we have $\Pr(\mathcal{A} \cap \mathcal{Z}|m_1^i) \leq \frac{2Cg}{\hat{F}^k}\left((1+\eta)F_i^k - \mathbb{E}_{i+1}[F_J^k]\right)$ for all $m_i \notin N_i$ and $\Pr(\mathcal{A}^c|m_1^i) \leq 1/\sqrt{G}$ for all $m_i \notin T_i$. Letting $\hat{M}_i = M_i \setminus (N_i \cup T_i)$, this implies that

$$\Pr(\mathcal{Z}|m_1^i) = \Pr(\mathcal{Z} \cap \mathcal{A}|m_1^i) + \Pr(\mathcal{Z} \cap \mathcal{A}^c|m_1^i) \leq \frac{2Cg}{\hat{F}^k}\left(F_i^k(1+\eta) - \mathbb{E}_{i+1}[F_J^k]\right) + 1/\sqrt{G} \quad (14)$$

for all $m_i \in \hat{M}_i$.

Conditional on $\mathcal{A}$, $F_j^{k+1} \leq G\varepsilon$ for all $j \geq i$. By definition of J, all continuation equilibria until round J included are informative, which implies that $F_j^k \geq \mathcal{F}^k(F_j^{k+1})$ for all $j \leq J$. Since $\mathcal{F}^k(\cdot)$ is nonincreasing, this implies that $F_j^k \geq \mathcal{F}^k(G\varepsilon) = \hat{F}^k$ for all $j \leq J$.

We thus have for $m_i \in \hat{M}_i$

$$\begin{aligned}
\mathbb{E}_{i+1}F_J^k &= \Pr_{i+1}(\mathcal{A})\ \mathbb{E}_{i+1}[F_J^k|\mathcal{A}] + \Pr_{i+1}(\mathcal{A}^c)\ \mathbb{E}_{i+1}[F_J^k|\mathcal{A}^c] \\
&\geq \Pr_{i+1}(\mathcal{A})\mathbb{E}_{i+1}[F_J^k|\mathcal{A}] \\
&\geq (1 - 1/\sqrt{G})\hat{F}^k.
\end{aligned}$$

By construction, $F_i^k \leq \bar{F}^k + \hat{F}^k\eta$ and $\hat{F}^k \geq \bar{F}^k - \hat{F}^k\eta$. Therefore, $F_i^k - \hat{F}^k \leq (\bar{F}^k + \hat{F}^k\eta) - (\bar{F}^k - \hat{F}^k\eta) = 2\eta\hat{F}^k$. Letting $B = 8Cg$, (14) then implies (for $\eta \leq 1$, which we assume) that for all m_i in $\hat{M}_i$

$$\Pr(\mathcal{Z}|m_1^i) \leq B\eta + \frac{B}{2\sqrt{G}} + 1/\sqrt{G}. \quad (15)$$

For each m_i, let $V_i^f(m_i|\mathcal{Z}^c)$ denote i's expected gross utility if he sends message m_i conditional on no $j > i$ working and $V^{f,*}$ denote the maximizer of $V_i^f(m_i|\mathcal{Z}^c)$ over all messages $m_i \in \hat{M}_i$. Notice that i's expected gross utility conditional on m_i and no $j > i$ working does not depend on whether i worked or shirked: either way, the subsequent reports $\{m_j\}_{j>i}$ are independent of the signals that remain to be discovered. Therefore, i's conditional expected gross utilities satisfy $V^w(m_i|\mathcal{Z}^c) = V_i^f(m_i|\mathcal{Z}^c)$.

Combining these observations with (13) and (15), we obtain

$$V^w(m_i) \leq \left(B\eta + \frac{B}{2\sqrt{G}} + \frac{1}{\sqrt{G}}\right) R + V^{f,*},$$

for $m_i \in \hat{M}_i$. Combining this with (12) yields

$$V^w \leq \left(\frac{1}{2\eta G^2} + \frac{1}{\sqrt{G}}\right) R + \left(B\eta + \frac{B}{2\sqrt{G}} + \frac{1}{\sqrt{G}}\right) R + V^{f,*}.$$

i's utility from working thus satisfies

$$U^w \leq \left(\frac{1}{2\eta G^2} + \frac{1}{\sqrt{G}} + B\eta + \frac{B}{2\sqrt{G}} + \frac{1}{\sqrt{G}}\right) R + V^{f,*} - c. \tag{16}$$

If i sends a message $m_i^* \in \hat{M}_i$ that achieves $V^{f,*}$, his utility U^f satisfies

$$\begin{aligned} U^f &\geq \Pr(\mathcal{Z}|m_1^{i-1}, m_i^*) \times 0 + \Pr(\mathcal{Z}^c|m_1^{i-1}, m_i^*) V^{f,*} \\ &\geq \left(1 - B\eta - \frac{B}{2\sqrt{G}} - \frac{1}{\sqrt{G}}\right) V^{f,*} \\ &\geq V^{f,*} - \left(B\eta - \frac{B}{2\sqrt{G}} - \frac{1}{\sqrt{G}}\right) R \end{aligned}$$

where 0 is used as a lower bound on i's realized gross utility in the first inequality.

Therefore, working is strictly suboptimal if

$$\left(\frac{1}{2\eta G^2} + \frac{1}{\sqrt{G}} + B\eta + \frac{B}{2\sqrt{G}} + \frac{1}{\sqrt{G}}\right) R + V^{f,*} - c < V^{f*} - \left(B\eta + \frac{B}{2\sqrt{G}} + \frac{1}{\sqrt{G}}\right) R$$

or

$$\frac{c}{R} > \frac{1}{2\eta G^2} + \frac{3}{\sqrt{G}} + 2B\eta + \frac{B}{\sqrt{G}}. \tag{17}$$

This inequality is always satisfied as long as η is small enough and ηG is large enough. In particular, since $B = 8Cg$, $g = R/\lambda c$, and C can be taken to equal $2R/c$ if $\lambda = 1$ and $2R/(c\lambda(1-\lambda))$ as noted in Proposition 2, the inequality is satisfied if $\eta = 1/\sqrt{G}$ and $\sqrt{G} = 128R^3/c^3$ for $\lambda = 1$ and $\sqrt{G} = 128R^3/(c^3\lambda^2(1-\lambda))$ if $\lambda < 1$ (recalling that $R > c$), in which case each term on the right-hand side of (17) is less than $c/4R$ with some strict inequalities. ■

B Proof of Theorem 2

Suppose first that $\lambda = 1$ and $\rho = 1$, which means that any agent who works finds a signal with probability 1. We construct compensation functions for which the strategy profile proposed constitutes an equilibrium. Under this strategy profile, as long as p_i lies in $(\underline{p}, \bar{p})$ all agents work and truthfully report their signal about ω. Moreover, given the symmetric signal structure, p_i depends only on the number of "H" and "L" signals as long as all agents $j < i$ work with probability 1. Therefore, the set of equilibrium posteriors forms a grid $\{q^k\}$ containing $\hat{p}$ and containing a single point on each side of $(\underline{p}, \bar{p})$. Let $q^0 \leq \underline{p} < q^1, \ldots, \hat{p}, \ldots, q^N < \bar{p} < q^{N+1}$ denote this grid. Along the candidate equilibrium, the belief p_i evolves on this grid until it hits either q^0 or q^{N+1}, after which the investigation stops.

Let J denote the last investigator who works: we have $p_J \in \{q^1, q^N\}$ and $p_{J+1} \in \{q^0, q^{N+1}\}$. Also let $\tilde{p} = p_{J+1}$ denote the value of the belief when learning stops under the candidate equilibrium.

We construct utility functions in which an investigator's compensation depends only on his report and on the posterior $\tilde{p}$.

For any i such that $p_i = q^k \in (\underline{p}, \bar{p})$, if i reports "H", he gets a reward $R_H^k \geq 0$ if $\tilde{p} = q^{N+1}$ and a punishment $P_L^k \leq 0$ if $\tilde{p} = q^0$. If i reports "L", he gets $R_L^k \geq 0$ if $\tilde{p} = q^0$ and $P_L^k \leq 0$ if $\tilde{p} = q^{N+1}$.

For any p, q on the grid, let $\pi(p, q)$ denote the probability that the belief sequence ends with $\tilde{p} = q^{N+1}$, i.e., exits $(\underline{p}, \bar{p})$ through $\bar{p}$, from the perspective of an agent who assigns probability p to ω, but the prior used by investigators is $p_0 = q$. That is, $\pi(p, q)$ is the probability that an individual with prior p assigns to the sequence p_i converging to q^{N+1} in equilibrium given that the public belief, which serves as the state variable for the equilibrium, starts at q.

If i sends report "H" starting from prior $p_i = q^k$, he assigns a probability $\pi(q^k, q^{k+1})$ to the public belief converging to q^{N+1}. If i works and receives report "H''", his belief about the continuation equilibrium is $\pi(q^{k+1}, q^{k+1})$. Similarly, if i sends "L'', his belief is $\pi(q^k, q^{k-1})$ whereas if he works and reports "L'' his belief is $\pi(q^{k-1}, q^{k-1})$. It is straightforward to verify the inequalities

$$\pi(q^{k+1}, q^{k+1}) > \pi(q^k, q^{k+1}) \tag{18}$$

and

$$\pi(q^{k-1}, q^{k-1}) < \pi(q^k, q^{k-1}), \tag{19}$$

for all $k \in [2, N-1]$. The strictness of the inequalities comes from the fact that conditional on the true state ω, the dynamic of $\{p_j\}_{j \geq i+1}$ starting any given value of p_{i+1} is strictly increasing in ω in FOSD, as is easily checked. Therefore, the probability of hitting q^{N+1} before q^0 is strictly increasing in the belief p_i that the state is high.

For $k = 1$, the investigation stops if i reports "L" so (19) holds as an equality, but (18) is still strict, because this report triggers further investigation. The reverse is true for $k = N$: (18) only holds as an equality while (19) is strict.

If i shirks, his maximal utility is

$$\max\{\pi(q^k, q^{k+1})R_H^k + (1 - \pi(q^k, q^{k+1}))P_H^k; \pi(q^k, q^{k-1})P_L^k + (1 - \pi(q^k, q^{k-1}))R_L^k\}. \tag{20}$$

The left argument is i's expected payo if he sends "H", and the right one is his payo if he sends "L". Since i can send either message at no cost, his best payo from fabrication is the maximum of these two terms. If i works, he gets

$$z^k[\pi(q^{k+1}, q^{k+1})R_H^k + (1 - \pi(q^k, q^{k+1}))P_H^k] + (1 - z^k)[\pi(q^{k-1}, q^{k-1})P_L^k + (1 - \pi(q^{k-1},^{k-1}))R_L^k] \tag{21}$$

where z^k is the probability of receiving signal "H" given belief q^k, and is equal to $z^k = \Pr(``H''|q^k) = q^k\pi + (1 - q^k) \times (1 - \pi)$.

Working is optimal for i if (21) exceeds (20) by at least c.

This condition is obtained as follows: set $P_H^k = P_L^k = -Q$ where Q is a strictly positive constant, and let $R_H^k = Q\frac{1-\pi(q^k,q^{k+1})}{\pi(q^k,q^{k+1})}$ and $R_L^k = Q\frac{\pi(q^k,q^{k-1})}{1-\pi(q^k,q^{k-1})}$. This guarantees that i's expected payo from fabrication is zero, regardless of the outcome. From (18) and (19), his payo from working is of order Q and thus exceeds c, for Q high enough.[52]

If $k = 1$ or N, there is one signal that i can send after working which yields a payo of order Q, while the other signal yields 0. The signal associated with a positive payo arises with a probability that is bounded away from 0, since p_i lies in $(\underline{p}, \bar{p})$.

Moreover this scheme is feasible as long as the maximal reward R and and punishment P respectively exceed $\sup\{R_\theta^k : \theta \in \{L, H\}, k \in \{1, \ldots, N\}\}$ and Q.

The proof easily generalizes to $\lambda < 1$ and $\rho < 1$. See Appendix C.7.

[52]To see this, let $\bar{\pi}$ denote a strictly positive lower bound on all inequalities (18) and (19) over all k's whenever they hold strictly. Then, the gain from working is of order $Q\bar{\pi}$.

This implies that

$$\Delta_i^{\emptyset}(S'') - \Delta_i(S'') = -\frac{\lambda f_i^0 \Delta_i(S'')}{(1-f_i^0)(1-\lambda)+f_i^0} \tag{22}$$

for any $S'' \neq \emptyset$, and

$$\Delta_i^{\emptyset}(\emptyset) - \Delta_i(\emptyset) = \frac{\lambda(1-f_i^0)\Delta_i(\emptyset)}{(1-f_i^0)(1-\lambda)+f_i^0}. \tag{23}$$

Let $V_i(m_i, S'')$ denote i's expected gross utility conditional on i producing evidence m_i and on $S_{i+1} = S''$. Notice that $m_1^i = (m_1^{i-1}, m_i)$ and S_{i+1} completely determine the distribution of reports $\{m_j\}_{j>i}$. Therefore, $V_i(m_i, S'')$ is the same regardless of whether i has worked or shirked. Agent i's expected gross utility conditional on (i) working, (ii) finding no signal, and (iii) producing message m_i, is

$$V_i^w(\emptyset, m_i) = \sum_{S'' \in \mathcal{S}} V_i(m_i, S'')\Delta_i^{\emptyset}(S''),$$

whereas his expected gross utility if i shirks and sends message m_i is

$$V_i^s(m_i) = \sum_{S'' \in \mathcal{S}} V_i(m_i, S'')\Delta_i(S'')$$

because i has learned nothing from shirking and thus holds the same belief as his prior belief at the beginning of round i. Combining these expressions, we get

$$V_i^w(\emptyset, m_i) - V_i^s(m_i) = \sum_{S'' \in \mathcal{S}} V_i(m_i, S'')(\Delta_i^{\emptyset}(S'') - \Delta_i(S'')). \tag{24}$$

Since $V_i(m_i, S'') \in [0, R]$ for all m_i and S'', combining (24) with (22) and (23) yields

$$V_i^w(\emptyset, m_i) - V_i^s(m_i) \leq \frac{R\Delta_i(\emptyset)\lambda(1-f_i^0)}{(1-f_i^0)(1-\lambda)+f_i^0}.$$

Since $\lambda < 1$, the denominator is bounded below by $1-\lambda$. Since $\Delta_i(\emptyset) = f_i^0$, the numerator is bounded above by Rf_i^0. This yields

$$V_i^w(\emptyset, m_i) \leq V_i^s(m_i) + f_i^0 \frac{R}{1-\lambda} \leq V_i^* + f_i^0 \frac{R}{1-\lambda},$$

which proves the lemma. Intuitively, this results means that if f_i^0 is negligible relative to $(1-\lambda)$, then i's expected gross utility after working and finding nothing cannot be much higher than if i had shirked, because finding nothing in this case merely reveals that i was unlucky and otherwise conveys little else information.

C.3 Proof of Lemma 3

For each $m_i \in M$, let $V_i^w(m_i)$ denote i's expected gross utility conditional on working and sending message m_i and M_i^- denote the set of messages m_i after which no $j > i$ ever works, so that $M = M_i^+ \cup M_i^-$ and $M_i^+ \cap M_i^- = \emptyset$. Letting $\gamma_i(\tilde{M}_i)$ denote the probability that i sends a message in $\tilde{M}_i$ conditional on working and on m_1^{i-1}, we have:

$$V_i^w = \sum_{m_i \in M_i^-} \gamma_i(m_i) V_i^w(m_i) + \sum_{m_i \in M_i^+} \gamma_i(m_i) V_i^w(m_i). \tag{25}$$

For the first term, note that i's expected utility conditional on reporting m_i and on no $j > i$ ever producing real evidence does not depend on whether i worked or shirked: either way, the distribution of the reports $\{m_j\}_{j>i}$ is independent of the set of signals that remain in the case. Letting, as in the previous lemma, $V_i^s(m_i)$ denote i's expected gross utility conditional on shirking and sending message m_i, we thus have $V_i^w(m_i) = V_i^s(m_i)$ for all $m_i \in M_i^-$. Since $V_i^* = \max_{m_i \in M} V_i^s(m_i)$, the first term in (25) is bounded above by $\gamma_i(M_i^-)V_i^*$.

For the second term, we have $\gamma_i(m_i) = d_i(m_i) + g_i(m_i)$ and

$$\gamma_i(m_i) V_i^w(m_i) \leq d_i(m_i) V_i^w(\emptyset, m_i) + g_i(m_i) R,$$

where we used the fact that i's expected gross utility conditional on working, finding a signal, and reporting m_i is bounded by R.

Combining these observations yields

$$V_i^w \leq \gamma_i(M_i^-) V_i^* + g_i(M_i^+) R + \sum_{m_i \in M_i^+} d_i(m_i) V_i^w(\emptyset, m_i). \tag{26}$$

If $\lambda < 1$, Lemma 2 implies that $V_i^w(\emptyset, m_i) \leq V_i^* + \frac{f_i^0 R}{1-\lambda}$. Summing over all $m_i \in M_i^+$, we get

$$\sum_{m_i \in M_i^+} d_i(m_i) V_i^w(\emptyset, m_i) \leq d_i(M_i^+) V_i^* + d_i(M_i^+) \frac{f_i^0 R}{1-\lambda}. \tag{27}$$

Since $\gamma_i(M_i^-) + d_i(M_i^+) \leq \gamma_i(M_i^-) + \gamma_i(M_i^+) = 1$, combining (26) and (27) proves the lemma when $\lambda < 1$.

If $\lambda = 1$, using in (26) the fact that $V_i^w(\emptyset, m_i)$ is bounded above by R directly proves the lemma.

- $\delta_i(\tilde{M}_i)$: probability that i sends a message in $\tilde{M}_i$ conditional on working *and* finding no signal;[53]

- $\Phi(m_i)$ is the probability that (i) i works, (ii) i discovers a signal, (iii) i sends report m_i, and (iv) there remain at least k signals at the beginning of round $i+1$.

Let $p_i(m_i)$ denote the probability that i produces report m_i conditional on m_1^{i-1}: $p_i(m_i)$ is the denominator of (29). Rearranging (29) and simplifying, we have

$$F_i^k(\beta_i(m_i) + \gamma_i(1-F_i^1)\lambda\delta_i(m_i)) = (F_i^k - F_i^{k+1}(m_i))p_i(m_i) + \Phi(m_i) \tag{30}$$

Since $\gamma_i(1-F_i^1)\lambda\delta_i(m_i) \geq 0$, summing the previous equation over $m_i \in M_i^+$ yields

$$F_i^k\beta_i(M_i^+) \leq \mathbb{E}_i\left[\left(F_i^k - F_{i+1}^k(m_i)\right)\mathbb{1}_{m_i\in M_i^+}\right] + \sum_{m_i\in M_i^+}\Phi(m_i). \tag{31}$$

Since $\Phi(m_i) = \mathbb{E}_i\left[\mathbb{1}_{|S_i|\geq k+1}\ \mathbb{1}_{\text{i works, discovers a signal, and reports } m_i}\right]$, we have

$$\begin{aligned}\sum_{m_i\in M_i^+}\Phi(m_i) &\leq \sum_{m_i\in M_i^+}\mathbb{E}_i\left[\mathbb{1}_{|S_i|\geq k+1}\ \mathbb{1}_{\text{i works and reports } m_i}\right]\\ &= \mathbb{E}_i\left[\mathbb{1}_{|S_i|\geq k+1}\ \mathbb{1}_{\text{i works and reports } m_i\in M_i^+}\right]\\ &\leq \mathbb{E}_i\left[\mathbb{1}_{|S_i|\geq k+1}\ \mathbb{1}_{\text{i works}}\right]\\ &= F_i^{k+1}\gamma_i,\end{aligned} \tag{32}$$

noting, for the last equality, that the event that i works, which has probability γ_i, depends only on m_1^{i-1} and is thus independent of the event $\{|S_i| \geq k+1\}$ conditional on m_1^{i-1}.

Combining this with (31) yields

$$F_i^k\beta(M_i^+) \leq \mathbb{E}_i\left[(F_i^k - F_{i+1}^k(m_i))1_{m_i\in M_i^+}\right] + F_i^{k+1}\gamma_i. \tag{33}$$

Since $g_i(M_i^+) \geq c/2R$, we have $\beta(M_i^+) = \gamma_i g_i(M_i^+) \geq \gamma_i c/2R$. Inequality (33) then yields

$$\beta(M_i^+) \leq \frac{1}{F_i^k - 2R/cF_i^{k+1}}\mathbb{E}_i[(F_i^k - F_{i+1}^k(m_i))1_{m_i\in M_i^+}]. \tag{34}$$

Combining this with (28) yields (3) for $C(1) = 2R/c$.[54] Since $C(\lambda) \geq C(1)$ for all $\lambda \in (0,1]$, the monotonicity noted at the beginning of the proof yields the desired conclusion for $C(\lambda)$.

[53] Note that $\delta_i(\tilde{M}_i) \geq d_i(\tilde{M}_i)$, where $d_i(\tilde{M}_i)$ was defined before Lemma 3.

[54] Note that the proposition's assumption that $F_i^k - C(\lambda)F_i^{k+1} > 0$ implies that $F_i^k - \frac{2R}{c}F_i^{k+1} > 0$ since $C(\lambda) \geq C(1)$ regardless of λ.

Case 2: $g_i(M_i^+) < \frac{c}{2R}$**.** We prove that γ_i is bounded above by the right-hand side of (3) for $C = C(\lambda)$. Since $\gamma_i \geq \beta_i$, this will yield the desired conclusion.

Intuitively, in Case 2 the probability of discovering a signal that, together with i's equilibrium message strategy, triggers subsequent work is too low to incentivize i to work. The only way of incentivizing i to work is therefore for him to signal by his message that he found nothing through work. For this to happen, the probability f_i^0 that there remains no evidence must be high enough. We will use this fact to obtain a bound on γ_i.

From Lemma 3, if $g_i(M_i^+) < c/2R$, i's utility from working is bounded above by

$$U_i^w = V_i^w - c \leq V_i^* + d_i(M_i^+)\frac{f_i^0 R}{1-\lambda} - \frac{c}{2}$$

if $\lambda < 1$, and by

$$U_i^w \leq V_i^* + d_i(M_i^+)R - \frac{c}{2}$$

if $\lambda = 1$. Therefore, working is optimal only if $d_i(M_i^+)f_i^0 \geq c(1-\lambda)/2R$ when $\lambda < 1$ and only if $d_i(M_i^+) \geq c/2R$ when $\lambda = 1$.

Summing (30) over M_i^+ and using (32) and $f_i^0 = 1 - F_i^1$ yields

$$F_i^k \beta_i(M_i^+) + \gamma_i F_i^k \delta_i(M_i^+) f_i^0 \lambda \leq \mathbb{E}_i\left[(F_i^k - F_{i+1}^k(m_i))1_{m_i \in M_i^+}\right] + F_i^{k+1}\gamma_i. \tag{35}$$

For $\lambda < 1$, we have $d_i(M_i^+)f_i^0 \geq c(1-\lambda)/2R$. Since $\delta_i(m_i) \geq d_i(m_i)$ for all m_i (by definition of these variables) and $F_i^k \beta_i(M_i^+) \geq 0$, (35) implies that

$$\gamma_i \leq \frac{\mathbb{E}_i\left[(F_i^k - F_{i+1}^k(m_i))1_{m_i \in M_i^+}\right]}{F_i^k c\lambda(1-\lambda)/2R - F_i^{k+1}}.$$

Multiplying the numerator and denominator by $C(\lambda)$ yields the result.

For $\lambda = 1$, we have $d_i(m_i) = \delta_i(m_i)f_i^0$ for all m_i and, hence, $\delta_i(M_i^+)f_i^0\lambda = d_i(M_i^+)$, which is greater than $c/2R$ as noted earlier. Therefore, (35) implies that

$$\gamma_i \leq \frac{\mathbb{E}_i\left[(F_i^k - F_{i+1}^k(m_i))1_{m_i \in M_i^+}\right]}{F_i^k (c/2R) - F_i^{k+1}}.$$

Multiplying the numerator and the denominator by $C(1)$ yields the result. ■

C.5 Proof of Lemma 5

(i) Let $S_i' = \{s_i : F_i^{k,w}(s_i) > F_i^k\}$, and, for each s_i, let $\gamma_i'(s_i)$ (resp. $\beta_i'(s_i)$) denote the probability that i discovers s_i given that he works (resp., the probability that i works and discovers s_i). Also let $\gamma_i'(s_i|k+1)$ (resp. $\beta_i'(s_i|k+1)$) denote the same probabilities conditional on $|S_i| \geq k+1$.

We have for all s_i

$$F_i^{k,w}(s_i) = \frac{F_i^{k+1}\gamma_i'(s_i|k+1)}{\gamma_i'(s_i)} = \frac{F_i^{k+1}\beta_i'(s_i|k+1)}{\beta_i'(s_i)}$$

where the second equality comes from $\beta_i'(s_i) = \gamma_i\gamma_i'(s_i)$ and $\beta_i'(s_i|k+1) = \gamma_i\gamma_i'(s_i|k+1)$. Therefore, $F_i^{k,w}(s_i) > F_i^k$ only if $\beta_i'(s_i) < \beta_i'(s_i|k+1)F_i^{k+1}/F_i^k$.

We have $\sum_{s_i \in S_i'} F_i^{k+1}\beta_i'(s_i|k+1) \leq \gamma_i F_i^{k+1}$. Therefore, the probability $\beta_i'(S_i')$ that i works and finds a signal in S_i' satisfies

$$\beta_i'(S_i') = \sum_{s_i \in S_i'} \beta_i'(s_i) \leq \gamma_i \frac{F_i^{k+1}}{F_i^k}.$$

Since $\beta_i'(S_i') = \gamma_i\gamma_i'(S_i')$, we have

$$\gamma_i'(S_i') \leq \frac{F_i^{k+1}}{F_i^k} \leq \frac{\varepsilon}{\hat{F}^k},$$

since $F_i^{k+1} \leq \varepsilon$ and $F_i^k \geq \hat{F}^k$. From (4), the right-hand side is bounded above by $\frac{\hat{F}^k}{2G^2}$.

For any m_i, let $q(m_i)$ denote the probability, conditional on i working and sending report m_i, that i has discovered a signal $s_i \in S_i'$, and let $\sigma(s_i|m_i)$ denote the probability that i discovered s_i given that he worked and reported m_i. We also let $s_i = \emptyset$ denote the event that i did not find anything, $\sigma(\emptyset|m_i)$ denote the probability that i found nothing given that he worked and reported m_i, $F_i^{k,w}(\emptyset)$ denote the probability that there at least k signals conditional on i working and finding nothing. We have

$$\begin{aligned} F_i^{k,r}(m_i) &= \sum_{s_i \in S_i} \sigma(s_i|m_i)F_i^{k,w}(s_i) \\ &= \sum_{s_i \in S_i'} \sigma(s_i|m_i)F_i^{k,w}(s_i) + \sigma(\emptyset|m_i)F_i^{k,w}(\emptyset) + \sum_{s_i \neq \emptyset, s_i \in S_i \setminus S_i'} \sigma(s_i|m_i)F_i^{k,w}(s_i) \end{aligned}$$

By construction, $F_i^{k,w}(s_i) \leq F_i^k$ for all s_i in the last term. Moreover, we have $F_i^{k,w}(\emptyset) \leq F_i^k$, as is easily checked:[55] intuitively, finding nothing always increases the probability that there

[55] Formally, for $k \geq 1$, we have $F_i^{k,w}(\emptyset) = \frac{F_i^k(1-\lambda)}{(1-F_i^1)+F_i^1(1-\lambda)} = F_i^k \frac{1-\lambda}{(1-F_i^1)+F_i^1(1-\lambda)} \leq F_i^k$.

are no signals remaining to be found. Finally, the first term is bounded above by $\sigma(S'_i|m_i) = q(m_i)$. Therefore,

$$F_i^{k,r}(m_i) \geq (1+\eta)F_i^k \Rightarrow q(m_i) + (1-q(m_i))F_i^k \geq (1+\eta)F_i^k$$
$$\Rightarrow q(m_i) \geq \eta F_i^k.$$

To conclude, note that

$$\sum_{m_i} \gamma_i(m_i) q(m_i) = \Pr(s_i \in S'_i | m_1^{i-1}, i \text{ works}) \leq \frac{\hat{F}^k}{2G^2}$$

The left-hand side is bounded below by $\gamma(N_i)\eta F_i^k$. Since $F_i^k \geq \hat{F}^k$, this implies that

$$\gamma(N_i) \leq \frac{1}{2\eta G^2}.$$

(ii) $F_{i+1}^k(m_i)$ is a convex combination[56] of F_i^k and $F_i^{k,r}(m_i)$. This implies that $F_{i+1}^k(m_i) \leq (1+\eta)F_i^k$ for all $m_i \notin N_i$. From (11), this further implies that

$$\pi_{i+1}(m_i) \leq \frac{2Cg}{\hat{F}^k}\left((1+\eta)F_i^k - \mathbb{E}_{i+1}[F_J^k]\right)$$

for all $m_i \notin N_i$.

C.6 Proof of Lemma 6

(i) If $m_i \notin T_i$, we have $F_{i+1}^{k+1} \leq \sqrt{G}\varepsilon$. Using this inequality in Lemma 4 instead of $F_i^{k+1} \leq \varepsilon$ and repeating its argument applied to round $i+1$, we conclude that $i+1$ assigns probability at least $1 - 1/\sqrt{G}$ to $\mathcal{A}$ whenever $m_i \notin T_i$.

(ii) Let $F_i^{k+1,r}(m_i)$ denote the probability that there are at least $k+1$ signals left at the beginning of round $i+1$ conditional on i working and reporting m_i. F_{i+1}^{k+1} is a convex combination of F_i^{k+1} and $F_i^{k+1,r}(m_i)$. This, together with the fact that $F_i^{k+1} \leq \varepsilon$ and the

[56] $F_{i+1}^k(m_i)$ is the probability that $i+1$ assigns to there being at least k signals left upon observing m_i. If $i+1$ knew that i didn't work and simply sent message m_i, this belief should be F_i^k since m_i conveys no additional information. And if $i+1$ knew that i produced m_i through working and then reporting m_i, his updated belief should be $F_i^{k,r}(m_i)$. Since $i+1$ doesn't observe i's action, in general $F_{i+1}^k(m_i)$ is a convex combination of these two posteriors, where the weights corresponds to the probability assigned by $i+1$ to i fabricating or working conditional on observing m_i. This fact is straightforward to check using Bayesian updating.

definition of T_i, shows that $m_i \in T_i$ only if $F_i^{k+1,r}(m_i) \geq \sqrt{G}\varepsilon$. Let T_i' denote the set of messages m_i for which the last inequality holds. As noted, $T_i \subset T_i'$.

Since i's prior probability that there are at least $k+1$ signals left is $F_i^{k+1} \leq \varepsilon$, the law of iterated expectations implies that the probability $\bar{F}_i^{k+1,r}(m_i)$ that there were at least $k+1$ signals left at the beginning of round i conditional on i working and finding m_i must satisfy

$$\mathbb{E}_i[\bar{F}^{k+1,r}|\text{working}] = \sum_{i \in M_i} \gamma_i(m_i)\bar{F}_i^{k+1,r}(m_i) = F_i^{k+1} \leq \varepsilon.$$

Using Markov's inequality, this implies that $\Pr(m_i : \bar{F}_i^{k+1,r}(m_i) \geq \sqrt{G}\varepsilon |\ i \text{ works}) \leq \frac{\varepsilon}{\sqrt{G}\varepsilon} = 1/\sqrt{G}$. Since also $\bar{F}_i^{k+1,r}(m_i) \geq F_i^{k+1,r}(m_i)$, we get $\gamma_i(T_i') \leq 1/\sqrt{G}$. Since $T_i \subset T_i'$, this shows that $\gamma_i(T_i) \leq 1/\sqrt{G}$.

C.7 Proof of Theorem 2: General case

The argument of Section B extends easily when ρ and/or λ are less than 1.

With $\rho < 1$ and $\lambda = 1$, the informative equilibrium is identical to the one described in Proposition 1 except that learning stops as soon as an agent fails to report evidence, in which case he gets a zero compensation. By construction of the equilibrium in the proof above, shirking and reporting that no evidence was found has the same value as fabricating any other message and can thus be deterred. Since a working agent may find nothing, or the learning process may be interrupted before the belief process exits $(\underline{p}, \bar{p})$, in which case the working agent receives 0, the rewards and punishments must be scaled up by $1/\pi_\rho(q^k)$, where $\pi_\rho(q^k)$ is the probability that the belief process exits $(\underline{p}, \bar{p})$ in equilibrium, given the current belief q^k, so that the expected compensation of a working agent still exceeds the cost c of working.

If $\lambda < 1$ and $\rho = 1$, a working agent may fail to find evidence even when there surely exists some. In this case, we assume once more that the compensation is zero, which deters shirking and reporting the empty message, and scale up all rewards and punishments by $1/\lambda$ to incentive the agent to work, as in the previous paragraph. The belief process will surely exit $(\underline{p}, \bar{p})$ since the amount of evidence is unlimited (only individual agents may be unlucky and find nothing with probability $1 - \lambda$).

The case in which both λ and ρ are less than 1 is a convex combination of the previous cases and addressed accordingly.

D Proof of Theorem 3

Without loss of generality, we assume once more that agents' gross utility functions $\{V_i\}_{i\in\mathbb{N}}$ all take values in $[0, R]$.

For any round i, consider the event $\mathcal{Q}_i$ that all past investigators who may have failed to discover signals, given their strategy and reporting history m_1^{i-1}, have indeed failed to discover signals. Conditional on $\mathcal{Q}_i$, the number q_i of signals that have been uncovered until round i is equal to the number of past witnesses plus the number of past investigators whose messages reveal that they have surely discovered signals given m_1^{i-1}. Put differently, q_i is the number of signals that have been surely discovered by round i. Let $\hat{F}_i^k$ denote the probability that $|S_i| \geq k$ conditional on $\mathcal{Q}_i$ and q_i.

We observe that (i) q_i is nondecreasing along any equilibrium path and is strictly increasing whenever a witness arrives, (ii) given the construction of S, the distribution of S_i conditional on m_1^{i-1} and $\mathcal{Q}_i$ is only a function of q_i, and (iii) $\hat{F}_1^k = F_1^k$ for all k.

We will prove that there exist strictly positive thresholds $\{\underline{F}^k\}_{k\geq 1}$ such that an informative continuation equilibrium exists in round i only if $\hat{F}_i^k \geq \underline{F}^k$ for all $k \geq 1$. Applied to $i = 1$, this result implies Theorem 3. The proof uses the following lemma.

Lemma 7 *Under Assumption 1, the following inequalities hold: (i) $F_i^k \leq \hat{F}_i^k$ for all $i, k \geq 1$ and (ii) Path by path, $\hat{F}_j^k \leq \hat{F}_i^k$ for all $j \geq i$ and $k \geq 1$.*

Proof. Part (i): Let r_i denote the number of signals discovered by round i. We have $r_i \geq q_i$ and

$$\begin{aligned}
F_i^k &= \sum_{r\in\{q_i,\dots,i-1\}} \Pr(r_i = r \mid m_1^{i-1}) \Pr(\tilde{K} \geq r + k \mid \tilde{K} \geq r) \\
&\leq \sum_{r\in\{q_i,\dots,i-1\}} \Pr(r_i = r \mid m_1^{i-1}) \Pr(\tilde{K} \geq q_i + k \mid \tilde{K} \geq q_i) \\
&= \sum_{r\in\{q_i,\dots,i-1\}} \Pr(r_i = r \mid m_1^{i-1}) \hat{F}_i^k \\
&= \hat{F}_i^k.
\end{aligned}$$

The first equality comes from the independence of $\tilde{K}$ from S^∞: r_i is a sufficient statistic for $\tilde{K}$ given all the information produced before round i, and only to the extent that it reveals that $\tilde{K} \geq r_i$. The inequality comes from the increasing hazard rate condition, which implies that for any $k \geq 0$, $\Pr(\tilde{K} \geq k + q \mid \tilde{K} \geq q)$ is non-increasing in q.[57] The second equality is due to the equality $\hat{F}_i^k = \Pr(\tilde{K} \geq k + q_i \mid \tilde{K} \geq q_i)$, which again comes from the independence of $\tilde{K}$ and S^∞: the only relevant information about $\tilde{K}$ conditional on $\mathcal{Q}_i$ is the number of signals q_i discovered by round i.

Part (ii) For any $j \geq i$, we have $q_j \geq q_i$. As explained in the proof of Part (i), we have for all $k \geq 1$

$$\begin{aligned}\hat{F}_j^k &= \Pr(\tilde{K} \geq k + q_j \mid \tilde{K} \geq q_j) \\ &\leq \Pr(\tilde{K} \geq k + q_i \mid \tilde{K} \geq q_i) \\ &= \hat{F}_i^k,\end{aligned}$$

where the inequality comes from the increasing hazard rate property (see Footnote 57). ■

The existence of thresholds $\{\underline{F}^k\}_{k \geq 1}$ in Theorem 3 is proved by induction on k. We start with the base case $k = 1$ and then prove the induction step.

D.1 Proof for $k = 1$

Suppose that $\hat{F}_i^1 < c/2R$. Combining the two parts of Lemma 7, this implies that $F_j^1 < c/2R$ for all $j \geq i$. Lemma 1 still applies: no investigator $j \geq i$ works because the probability that he finds something is too small to justify the cost of e ort, given the maximal reward R.

We now show that if $\hat{F}_i^1$ lies below another threshold, smaller than $c/2R$, witnesses provide no informative message, either.

If i is a witness, we will use the following notation:

- β_i: probability that i produces an informative message given m_1^{i-1};
- M_i^+: set of messages m_i that are followed by an informative continuation equilibrium;
- $\Pr_i(M_i+)$: probability that i produces a message in M_i^+ given m_1^{i-1};
- $\gamma_i(m_i)$: probability that i sends m_i given m_1^{i-1};

[57] See, e.g., Barlow et al. (1963, p. 379). In brief, a random variable X with distribution F has the increasing hazard rate property if and only if the survival distribution $\bar{F} = 1 - F$ is log-concave. This property implies, as is easily checked, that for any $p, q \geq 0$, $\Pr(X \geq p + q \mid X \geq p) = \bar{F}(p+q)/\bar{F}(p)$ is decreasing in p.

- $\gamma_i(m_i|s_i)$: probability that i sends m_i after observing s_i.

LEMMA 8 *Consider $L \geq 2$ pairwise independent random variables $\{Y_\ell\}$ with non-atomic distributions over $\mathbb{R}$ and densities f_ℓ that are bounded above by $\bar{f}$. For any $\varepsilon \geq 0$, let E_ε^L denote the event that $\exists \ell, \ell' \leq L$ such that $|Y_\ell - Y_{\ell'}| \leq \varepsilon$. Then*

$$\Pr(E_\varepsilon^L) \leq L(L-1)\bar{f}\varepsilon.$$

Proof. The result is proved by induction on $L \geq 2$. For $L = 2$, we have

$$\Pr(|Y - Y'| \leq \varepsilon) = \int_{\mathbb{R}} f_Y(x) F_{Y'}[x - \varepsilon, x + \varepsilon] dx \leq 2\varepsilon \bar{f} \int_{\mathbb{R}} f_Y(x) dx = 2\varepsilon \bar{f}.$$

Now suppose that the claim holds for $L-1$. Notice that the event E_ε^L is the union of L events: the event E_ε^{L-1} concerning the first $L-1$ random variables, and, for each $\ell \leq L-1$, the event $E^{\ell,L}$ that the L^{th} random variable lies within ε of the ℓ^{th} random variable. Therefore,

$$\begin{aligned}
\Pr(E_\varepsilon^L) &\leq \Pr(E_\varepsilon^{L-1}) + \sum_{\ell \leq L-1} \Pr(|Y_\ell - Y_L| \leq \varepsilon) \\
&\leq (L-1)(L-2)\bar{f}\varepsilon + (L-1) \times 2\varepsilon\bar{f} \\
&= L(L-1)\bar{f}\varepsilon,
\end{aligned}$$

where the second inequality comes from the induction hypothesis and the fact that $\Pr(|Y_\ell - Y_L| \leq \varepsilon) \leq 2\bar{f}\varepsilon$, as shown in the first step of the induction for $L = 2$. ■

LEMMA 9 *There is a threshold $\underline{F}^1 \in (0, c/2R)$ such if i a witness, then $\beta_i > 0$ only if $\hat{F}_i^1 \geq \underline{F}^1$.*

Proof. Recall that if i is a witness, his message m_i is informative (in equilibrium) if it is statistically dependent of S conditional on m_1^{i-1}. Say that i's message is *ϵ-informative* if whenever i has preference ϵ there exist two signals $s_i \neq s_i'$ such that the equilibrium distributions of m_i conditional on i getting signals s_i versus s_i' are different across these two signals.

The following observation is straightforward to prove.

OBSERVATION 1 *i's message is informative if and only if the set of preference shocks ϵ for which i's message is ϵ-informative has positive probability.*

For any equilibrium and history up to some round i in which i is a witness, let ν_i denote the probability that i's preference shock ϵ_i is such that i's message is ϵ_i-informative.

For any $F < c/2R$, let $\nu(F)$ denote the supremum of ν_i over all witness rounds i of all equilibria such that $\hat{F}_i^1 \leq F$. We will show that $\nu(F) = 0$ for all F below some strictly positive threshold.

Consider such an equilibrium. For any witness round i and message m_i, let $z(m_i)$ denote the probability that at least some $j > i$ produces an informative message following message m_i.

i's expected utility if he receives signal s_i and sends message m_i is given by

$$U_i(m_i; s_i) = z(m_i)\mathbb{E}_i[V_i(m) \mid s_i, m_1^i] + (1 - z(m_i))\underline{V}_i(m_i) + \epsilon_i(m_i) \tag{36}$$

where

$$\underline{V}_i(m_i) = \mathbb{E}_i[V_i(m) \mid m_1^i, \text{no } j > i \text{ produces an informative message}].$$

Notice that $\underline{V}_i(m_i)$ does not depend on the signal s_i since this signal is payo irrelevant whenever no $j > i$ produces an informative message.

Given ϵ_i, i sends an informative signal only if there exist $m_i \neq m_i'$ and signals $s_i \neq s_i'$ such that

$$U_i(m_i; s_i) \geq U_i(m_i'; s_i)$$

and

$$U_i(m_i'; s_i') \geq U_i(m_i; s_i')$$

From (36) and the fact that $V_i(m) \in [0, R]$, this is possible only if:

$$|\epsilon_i(m_i) + \underline{V}_i(m_i) - \epsilon_i(m_i') + \underline{V}_i(m_i')| \leq R(z(m_i) + z(m_i')).$$

The random variables $Y_\ell = \epsilon_i(m_\ell) + \underline{V}_i(m_\ell)$ satisfy the assumptions of Lemma 8. Letting $|M|$ denote the cardinality of the message space M, we thus have

$$\Pr(i \text{ sends an informative message}) \leq |M|^2 \bar{f} R(z(m_i) + z(m_i')) \tag{37}$$

Since no investigator $j \geq i$ works when $\hat{F}_i^1 < c/2R$, we have for any message m_i:

$$z(m_i) \leq \sum_{j \geq 1} \Pr(\text{there are } j \text{ witnesses in the sequence after round } i, \text{ given message } m_i) j\nu(F). \tag{38}$$

Indeed, by definition of $\nu(F)$ and the fact that $\hat{F}_j^1 \leq F$ for all $j \geq i$, a witness provides an informative signal with probability at most $\nu(F)$. Thus $j\nu(F)$ is an upper bound on the probability that at least one witness provides an informative signal given there are j such witnesses.

The probability that at least j witnesses come after round i is bounded above by F^j, where j is an exponent (not a superscript): To show this for $j = 1$, notice that by Lemma 7, $\hat{F}_{i+1}^1 \leq F$, and the probability that there is at least one witness after round i is bounded above by the probability F_{i+1}^1 that there is at least one more signal, which is less than $\hat{F}_{i+1}^1$ again by Lemma 7. For $j = 2$, note that conditional on the first witness arriving, Lemma 7 implies that the probability that a second witness arrives is again by bounded by F since the probability that there remains another signal is bounded by F, and a witness can arise only if such a signal exists. By induction, this shows that the probability of having at least j witnesses and, hence, the probability of having exactly j witnesses, are bounded above by F^j. Combining this with (38) and using the standard formula $\sum_{j\geq 1} jx^j = x/(1-x)^2$ for all $x \in (0,1)$, we get

$$z(m_i) \leq \sum_{j\geq 1} F^j j\nu(F) = \frac{F\nu(F)}{(1-F)^2},$$

Combining this with (37), we obtain

$$\Pr(i \text{ sends an informative message} \mid \hat{F}_i^1 \leq F) \leq 2|M|^2 \bar{f} R\nu(F)\frac{F}{(1-F)^2}. \tag{39}$$

Taking the supremum of the left-hand side over all witness rounds i and equilibria such that $\hat{F}_i^1 \leq F$, we obtain

$$\nu(F) \leq \frac{2\nu(F)|M|^2\bar{f}RF}{(1-F)^2}.$$

For $2F/(1-F)^2 \leq 1/|M|^2\bar{f}R$, this relation is possible only if $\nu(F) = 0$, because the function on the right-hand side is a contraction of $\nu(F)$. The function $F/(1-F)^2$ is increasing on $[0,1)$ and starts at zero. Therefore, we conclude that there exists a threshold $\underline{F}^1 > 0$ such that $\nu(F) = 0$ for all $F \leq \underline{F}^1$. ■

D.2 Induction Step

Suppose that there exist strictly positive thresholds $\{\underline{F}^{k'}\}_{k'\in\{1,\ldots,k\}}$ such that a continuation equilibrium starting at round i is informative only if $\hat{F}_i^{k'} \geq \underline{F}^{k'}$ for all $k' \leq k$. We will show that a similar condition holds for $k+1$. The proof works by contradiction: we will suppose that for all $\varepsilon \in (0,1)$, there exists an informative continuation equilibrium such that $\hat{F}_i^{k+1} \leq \varepsilon$ and obtain an impossibility for ε small enough.

Consider any $\varepsilon < \underline{F}^k \times \underline{F}^1$ and any informative continuation equilibrium starting at round i such that $\hat{F}_i^{k+1} \leq \varepsilon$. Then, all continuation equilibria are uninformative as soon as some witness $j \geq i$ arrives. To see this, suppose that some witness arrives at round $j \geq i$. We have for any message m_j sent by this witness:

$$\begin{aligned}
\hat{F}_{j+1}^k(m_j) &= \Pr(\tilde{K} \geq k + q_j + 1 \mid \tilde{K} \geq q_j + 1) \\
&= \frac{\Pr(\tilde{K} \geq k + q_j + 1)}{\Pr(\tilde{K} \geq q_j + 1)} \\
&\leq \frac{\Pr(\tilde{K} \geq k + q_i + 1)}{\Pr(\tilde{K} \geq q_i + 1)} \\
&= \frac{\Pr(\tilde{K} \geq k + q_i + 1)}{\Pr(\tilde{K} \geq q_i)} \times \frac{\Pr(\tilde{K} \geq q_i)}{\Pr(\tilde{K} \geq q_i + 1)} \\
&= \frac{\hat{F}_i^{k+1}}{\hat{F}_i^1}
\end{aligned}$$

where the inequality comes for the monotone hazard rate assumption (see Footnote 57) and the fact that $q_j \geq q_i$. Since the continuation equilibrium from round i is informative, we must have $\hat{F}_i^1 \geq \underline{F}^1$. Therefore, $\hat{F}_i^{k+1} \leq \varepsilon < \underline{F}^k \underline{F}^1$, which implies that

$$\hat{F}_{j+1}^k < \underline{F}^k$$

and shows by induction hypothesis that all continuation equilibria are uninformative from round $j+1$ onwards.

Moreover, the first witness, j, knowing that continuation equilibria are uninformative regardless of his message, has no incentive to send an informative message.[58]

Therefore, if $\hat{F}_i^{k+1} \leq \varepsilon$, the only agents who may send informative messages are the investigators arriving between round i and the arrival of the first witness. The situation is therefore almost identical to the setting of Theorem 1, in the absence of witnesses, except that any sequential learning activity is interrupted at the apparition of the first witness. We know from Theorem 1 that such equilibria can be informative only if F_i^{k+1} exceeds the $(k+1)^{th}$-threshold given by Theorem 1, which we denote here $\tilde{F}^{k+1}$. Since $F_i^{k+1} \leq \hat{F}_i^{k+1}$ by Part (i) of

[58]Formally, the argument is similar to the proof of Lemma 9 except that here $z(m_i) = 0$ regardless of the message.

Lemma 7, we conclude, letting $\underline{F}^{k+1} = \min\{\tilde{F}_i^{k+1}, \underline{F}^k \underline{F}^1\}$, that no informative continuation equilibrium exists in round i if $\hat{F}_i^{k+1} \leq \underline{F}^{k+1}$. This concludes the induction step. ■