Online Appendix to "Dynamic Disclosure with(out) Timestamps"

Aaron Kolb*   Beixi Zhou†

September 18, 2026

Abstract

This Online Appendix contains omitted proofs, discussion of the support restriction, and additional results on receiver welfare and multiple evidence arrivals.

1 Supporting material for Section 4

1.1 Motivation and discussion of support restriction

In this section, we provide a tremble-based motivation for the support restriction on off-path beliefs used in Lemma 1 and discuss the restriction's connection to neutral beliefs.

Our support restriction is related to, but substantively different from, earlier work on support restrictions in games with fixed types. These impose that the support of beliefs at later information sets be a subset of the support of beliefs at earlier information sets; see Madrigal, Tan, and da Costa Werlang (1987), who show an example where imposing a support restriction eliminates the (outcome-unique) Nash equilibrium. That notion of support restriction does not directly map into our setting, which has evolving private information: a sender without evidence can later acquire good or bad evidence. Hence, it would be inappropriate to force the support of beliefs to decrease over time in a set-inclusion sense. Our support restriction is narrower, and merely states that if the stock of undisclosed good evidence is empty at some date T, those histories cannot re-enter the support of later beliefs after an off-path disclosure. New (after T) evidence histories may still enter.

Our support restriction has a natural tremble-based motivation. Conditional on possessing undisclosed good evidence at a given date, the sender's incentives do not depend

*Kelley School of Business, Indiana University, kolba@iu.edu.

†Department of Economics, University of Pittsburgh, beixi.zhou@pitt.edu.

on the evidence's arrival date, since timestamps are unobserved and payoff-irrelevant. This motivates perturbations that bound the ratio of trembles for different ages.

Consider a two-period version of the game, where in each period $t = 0, 1$, evidence arrives with probability μ if it hasn't already arrived, and the disclosure decision is made after the sender observes the arrival or not. The state evolves in between periods, with probability λ_i from state i to state $1-i$. Suppose good evidence arriving in period 0 is disclosed with probability 1, and good evidence arriving in period 1 is not disclosed. Perturb this strategy so that good evidence in period 0 is disclosed with probability $1-\varepsilon_0$, and in period 1, any good evidence (whether from period 0 or period 1) is disclosed with probability ε_1. Then after disclosure of good evidence in period 1, the receiver's posterior probability that it arrived in period 0 is $\frac{p_0\mu\varepsilon_0\varepsilon_1}{p_0\mu\varepsilon_0\varepsilon_1+\phi_1(1-\mu)\mu\varepsilon_1}$, where $\phi_1 = p_0(1-\lambda_1) + (1-p_0)\lambda_0$ is the no-information belief about the state in period 1. Then canceling ε_1 and taking $\varepsilon_0 \to 0$ gives a limiting probability of 0 on evidence from time 0, in line with the support restriction.

The argument above imposes an age-independent tremble probability ε_1 in period 1; this is consistent with the payoff irrelevance of timestamps. However, the result goes through even for age-dependent trembles, provided there is a uniform bound on their ratio. To that end, let ε_1^O and ε_1^N be the probability of disclosing in period 1 for old (period 0) and new (period 1) evidence, respectively. Consider a sequence of trembles $(\varepsilon_0, \varepsilon_1^O, \varepsilon_1^N)$ (suppressing the index) and suppose there is a constant K such that $\frac{\varepsilon_1^O}{\varepsilon_1^N} \leq K$ along the sequence. Then the posterior probability that good evidence disclosed in period 1 is from period 0 is

$$\frac{p_0\mu\varepsilon_0\varepsilon_1^O}{p_0\mu\varepsilon_0\varepsilon_1^O + \phi_1(1-\mu)\mu\varepsilon_1^N} \leq \frac{p_0\mu\varepsilon_0 K}{p_0\mu\varepsilon_0 K + \phi_1(1-\mu)\mu} \to 0.$$

In other words, to justify a positive probability on time 0 arrival would require the ratio of period-1 trembles to favor time 0 with a ratio that becomes arbitrarily large.

Neutral beliefs and the support restriction Neutral beliefs imply the support restriction at all times when the probability that the sender possesses undisclosed good evidence is strictly positive. In the previous example, under the original (unperturbed) strategy, the neutral belief at time 1, just prior to the disclosure decision, assigns probability 0 to date 0 good evidence, but probability $\phi_1(1-\mu)\mu$ to date 1 good evidence. Neutral beliefs then assign probability $\frac{0}{0+\phi_1(1-\mu)\mu} = 0$ to date 0 good evidence conditional on the sender possessing good evidence just prior to the disclosure decision in period 1.

1.2 Proof of second part of Lemma 1

Toward a contradiction, suppose there is an equilibrium with some $T > 0$ such that $S_T = 0$ but $S_{T-} > 0$.[1] Since a positive mass of types disclose at T, disclosure at time T must be optimal. Fix $\delta \in (0, S_{T-})$. For all sufficiently small $\eta > 0$ and all $t < T$ sufficiently close to T, the probability of evidence arrival in $(T - \eta, T)$ conditional on no disclosure by time t is strictly less than δ. Hence, for each such t, there is a measure of at least $S_t - \delta$ of undisclosed good evidence that arrived before $T - \eta$, which is bounded below away from 0 for t sufficiently close to T. Since $\lambda_1 > 0$, the probability (conditional on no disclosure by t) of the following event is bounded away from 0: undisclosed good evidence arrived before $T - \eta$ and $\theta_t = 0$. This implies $G_{0,T-} > 0$. It follows that after disclosure of good evidence at time T, the posterior is $q_T = \frac{G_{1,T-}}{G_{0,T-}+G_{1,T-}} < 1$, and the payoff of such disclosure is $\Pi(q_T) := \frac{p^*}{r} + \frac{q_T - p^*}{r+\Lambda}$, where, as in the main body, $\Lambda := \lambda_0 + \lambda_1$. By disclosing at time $T + \varepsilon$, the support restriction implies that the receiver's belief is at least $p^* + (1-p^*)e^{-\Lambda\varepsilon}$. The continuation payoff at T of disclosing at $T + \varepsilon$ is therefore at least $e^{-r\varepsilon}\Pi(p^* + (1-p^*)e^{-\Lambda\varepsilon})$, which tends to $\Pi(1) > \Pi(q_T)$ as $\varepsilon \to 0$. This contradicts the optimality of disclosure at time T.

1.3 Proof of Lemma 16

If $T < \infty$, then $T < T^{\max}$ by definition of T. Since the right hand side of (35) contains a factor of p_t, the comparison theorem implies $p_t > 0$ for all $t \in [\underline{T}, T^{\max})$, and in particular, $p_T > 0$.

Note that for all $t \in [\underline{T}, T)$, we have $(p_t, q_t) \in R$ and therefore κ_t is well-defined and strictly positive. Moreover, S_t is well-defined and strictly positive. Hence, by construction, the ODEs for (G_1, G_0, p) are all satisfied.

The original ODE for p can be written

$$\dot{p}_t = -\lambda_1 p_t - S_t \kappa_t (q_t - p_t). \tag{OA.1}$$

Defining $\delta_t = q_t - p_t$ and subtracting this ODE for p from the indifference condition for q, we have (10). It follows that δ is strictly increasing on $[\underline{T}, T)$, so $\delta_T > 0$; in other words, $q_T > p_T$.

We now establish that $A_T > 0$. Since $A_t > 0$ for all $t < T$, we must have $A_T \geq 0$. Toward

[1] Note that S, G_0, and G_1 have left limits because they are conditional probabilities and the underlying event probabilities have left limits.

a contradiction, suppose $A_T = 0$. We claim that $\dot{A}_T > 0$. A_t is differentiated as

$$\begin{aligned}
\dot{A}_t &= (r-\mu)\dot{q}_t - 2\dot{p}_t(\lambda_1 + r) \\
&= (r-\mu)[rq_t - (\lambda_1 + r)p_t] \\
&\quad - 2(\lambda_1 + r)\left[p_t \frac{2(\lambda_1 + r)\mu p_t + \mu q_t(\lambda_1 + \mu - r) - (\lambda_1 + \mu)(\lambda_1 + r + \mu)}{\lambda_1 + r + \mu(1 - q_t)}\right].
\end{aligned}$$

Now if $A_t = 0$, then $p_t = \frac{r+\lambda_1+\mu+(r-\mu)q_t}{2(\lambda_1+r)}$. Substituting this into the expression for $\dot{A}_t$ and simplifying yields

$$\begin{aligned}
\dot{A}_t|_{A_t=0} &= C_0 + C_1 q_t, \qquad \text{where} \\
C_0 &:= \frac{1}{2}(2\lambda_1 + \mu - r)(\lambda_1 + r + \mu) \\
C_1 &:= \frac{1}{2}(r - \mu)(2\lambda_1 + \mu + r).
\end{aligned}$$

Since this is linear in q_t, it suffices to check the endpoints. Using $q_t = 1$ yields $C_0 + C_1 = \frac{1}{2}\lambda_1(3r + 2\lambda_1 + \mu) > 0$. To evaluate at a lower bound on q_t, recall that $A_T = 0$ implies

$$p_T = \frac{r + \lambda_1 + \mu + (r-\mu)q_T}{2(\lambda_1 + r)} < q_T,$$

as we have already shown $p_T < q_T$. This in turn yields $q_T > \underline{q} := \frac{\lambda_1+\mu+r}{2\lambda_1+\mu+r}$. Hence,

$$\begin{aligned}
C_0 + C_1\underline{q} &= \frac{1}{2}(2\lambda_1 + \mu - r)(\lambda_1 + r + \mu) + \frac{1}{2}(r - \mu)(\lambda_1 + \mu + r) \\
&= \lambda_1(\lambda_1 + r + \mu) > 0.
\end{aligned}$$

As $q_t \in (\underline{q}, 1]$ it follows that $C_0 + C_1 q_t$ lies between two positive values obtained at the extreme values $\underline{q}$ and 1, and is therefore positive. We conclude that $\dot{A}_T|_{A_T=0} > 0$, so A_t cannot intersect 0 from above at $t = T$. Thus, $A_T > 0$.

Having shown that $q_T > p_T > 0$ and $A_T > 0$, the only remaining possibility is $q_T = 1$.

1.4 Boundedness of $\overline{T}$ as $p_0 \uparrow 1$

As stated in Section 5.2, the start of transparency stays bounded even as $p_0 \uparrow 1$. First, note that $\underline{T}(p_0)$ is positive and continuous at $p_0 = 1$; we have $\underline{T}(p_0) = b^{-1}(p_0) \to b^{-1}(1) < \infty$ where b is defined in the proof of Lemma 13. Then we have $\delta_{\underline{T}(1)} = q_{\underline{T}(1)} - p_{\underline{T}(1)} > 0$, as $p_{\underline{T}(1)} = \phi_{\underline{T}(1)}$ which is strictly below the belief conditional on good evidence arriving at any $\tau \leq \underline{T}(1)$. Therefore, the argument of Lemma 17 gives a finite exit time $\overline{T}(1)$ where $q_{\overline{T}(1)} = 1$, and we have $\dot{q}_{\overline{T}(1)-} > 0$. By continuous dependence of the purging phase system on its initial conditions at time $\underline{T}(p_0)$ and the transversal exit at $\overline{T}(1)$, we have $\overline{T}(p_0) \to \overline{T}(1)$ as $p_0 \uparrow 1$.

1.5 Purging phase for moderate p_0 case

For all $t \in (0, \overline{T})$, (p, q) must satisfy (35) and (36), and must obey initial conditions $(p_0, q_0) = (p_0, 1)$. Note that $A(p_0, 1) = 2r + \lambda_1 - 2p_0(\lambda_1 + r) \geq 0$ because $p_0 \leq \overline{p}$. As we already showed, $\dot{A}_T|_{A_T=0} > 0$, so $A_t > 0$ for all $t \in (0, \overline{T})$; moreover, $A_0 > 0$ if and only if $p_0 < \overline{p}$. By standard arguments, there exists a unique local solution to the system in (p, q). Since $q_0 = 1$, the initial condition is not in R, but $p_0 > \underline{p}$ implies $\dot{q}_{0+} < 0$, so we have $0 < p_t < q_t < 1$ and $A_t > 0$ for all sufficiently small t; in other words, the solution immediately enters R. A repeat of the arguments from the high-p_0 case establishes that the solution exits R at some finite time $\overline{T}$ through $q_{\overline{T}} = 1$. $S_t = S(p_t, q_t)$ and $\kappa_t = \kappa(p_t, q_t)$ are then well defined and strictly positive by (32) and (33) for all $t \in (0, \overline{T})$. We have $S_0 = S(p_0, 1) = 0$, while $\kappa(p_0, 1)$ is undefined; if $p_0 < \overline{p}$, we have $\lim_{t \to 0} \kappa(p_t, q_t) = +\infty$ and if $p_0 = \overline{p}$, $\lim_{t \to 0} \kappa(p_t, q_t)$ is finite and positive. The variables G_0 and G_1 are recovered by $G_{1,t} = S_t q_t$ and $G_{0,t} = S_t - G_{1,t}$; note that their initial conditions $(G_{1,0}, G_{0,0}) = (0, 0)$ are satisfied. By construction, (7)-(9), (19), and (20) are satisfied at all $t > 0$.

To recover the CDF of the strategy, for all $0 < t \leq s < \overline{T}$, we have $H(s; G, \tau, t) = 1 - e^{-\int_t^s \kappa_u \, du}$, and for $s \geq \overline{T}$ $H(s; G, \tau, t) = 1$. For $t = 0$, the only possible arrival date is $\tau = 0$. We specify $H(s; G, 0, 0) = 1$ for all $s \geq 0$, which does not violate the atomless property of Definition 1 since evidence arrival at time 0 is a zero-measure event.

1.6 Proof of Proposition 4

The proof has two main steps. First, we prove the following corollary to Lemma 19, which establishes optimality of non-disclosure of bad evidence; it is stronger than needed for our purposes because it allows for arbitrary good evidence disclosure strategies, whereas our construction will involve a particular (three-phase) good evidence disclosure strategy. Second, we use a continuity argument to extend the good evidence disclosure strategy of Theorem 2.

For the first step, we use the following intermediate result.

Lemma OA.1. *Fix an arbitrary strategy H such that bad evidence is never disclosed. Let H^0 denote the strategy under which good evidence is disclosed immediately and bad evidence is never disclosed. Let p and p^0 denote the no-disclosure belief paths under H and H^0, respectively. We claim that for all $t \geq 0$, $p_t \geq p_t^0$.*

Proof. The proof uses similar logic to that of Proposition 2. Fix $t \geq 0$. For strategy H, define N as the ex ante probability that no evidence is disclosed by time t. For each $s \in [0, t]$, let ν_s^G denote the probability that good evidence arriving at time s is disclosed by time t. Let $\nu_s^{G,0} \equiv 1$ and N^0 have analogous definitions to ν_s^G and N, but under strategy H^0 instead of H.

We now apply the law of iterated expectations to H and H^0. Even though timestamps are not observed by the receiver, we can condition on the true evidence-arrival time in the law-of-iterated-expectations decomposition. Thus,

$$\phi_t = Np_t + \int_0^t \mu\phi_s e^{-\mu s}\nu_s^G q_t^{G,s}\, ds \qquad \text{and} \qquad \phi_t = N^0 p_t^0 + \int_0^t \mu\phi_s e^{-\mu s}\nu_s^{G,0} q_t^{G,s}\, ds,$$

where we have used that bad evidence is not disclosed under either H or H^0. Note that because $q_t^{G,s} \geq \phi_t$ at all times and $N, N^0 > 0$, we have $\phi_t \geq p_t, p_t^0$ at all times.

Equating the right hand sides and subtracting p_t^0 throughout yields

$$\begin{aligned} N(p_t - p_t^0) + \int_0^t \mu\phi_s e^{-\mu s}\nu_s^G (q_t^{G,s} - p_t^0)\, ds &= 0 + \int_0^t \mu\phi_s e^{-\mu s}\nu_s^{G,0}(q_t^{G,s} - p_t^0)\, ds \\ \implies N(p_t - p_t^0) &= \int_0^t \mu\phi_s e^{-\mu s}(\nu_s^{G,0} - \nu_s^G)(q_t^{G,s} - p_t^0)\, ds. \end{aligned}$$

Since $N \geq e^{-\mu t} > 0$ and $1 = \nu_s^{G,0} \geq \nu_s^G$ for all s and $q_t^{G,s} \geq \phi_t \geq p_t^0$, we conclude that $p_t \geq p_t^0$. $\square$

Corollary OA.1. *For sufficiently small $\lambda_0 > 0$, if the receiver conjectures that bad evidence is never disclosed and after off-path disclosure of bad evidence the receiver updates to* 0*, then it is optimal for the sender to never disclose bad evidence, independent of the good-evidence disclosure strategy.*

Proof of Corollary OA.1. From the proof of Lemma 19, under the given off-path beliefs, delaying disclosure of bad evidence is locally incentive compatible if $p_t \geq a(\lambda_0) := \frac{\lambda_0}{\lambda_0+\lambda_1+r}$ for all $t \geq 0$. Hence, it suffices to show that if $\lambda_0 > 0$ is sufficiently small, this inequality holds at all times for all possible good-evidence disclosure strategies. Given Lemma OA.1, we only need to show that $p_t^0 \geq a(\lambda_0)$ for all $t \geq 0$.

First, fix $\varepsilon_0 > 0$. We show that there exists $T_0 > 0$ such that for all $\lambda_0 \in (0, \varepsilon_0)$ and $t \geq T_0$, $\bar{q}_t^B$ cannot cross $a(\lambda_0)$ from above, where $\bar{q}_t^B$ is calculated under the strategy of immediate good disclosure and no bad disclosure. From (57), $\dot{\bar{q}}_t^B|_{\bar{q}_t^B = a(\lambda_0)} = \lambda_0(1-\bar{q}_t^B) - \lambda_1\bar{q}_t^B - \bar{q}_t^B\frac{\dot{B}_t^{\mathrm{un}}}{B_t^{\mathrm{un}}} = a(\lambda_0)\left(r - \frac{\dot{B}_t^{\mathrm{un}}}{B_t^{\mathrm{un}}}\right)$. Now define $\bar{\phi} := \max\{p_0, \frac{\varepsilon_0}{\varepsilon_0+\lambda_1}\} < 1$. Then for all $\lambda_0 \in (0, \varepsilon_0)$ and all t, we have $\phi_t \leq \max\{p_0, \frac{\lambda_0}{\lambda_0+\lambda_1}\} \leq \bar{\phi}$. Thus, from (50),

$$\frac{\dot{B}_t^{\mathrm{un}}}{B_t^{\mathrm{un}}} = \frac{\mu e^{-\mu t}(1-\phi_t)}{\int_0^t \mu e^{-\mu s}(1-\phi_s)\,ds} \leq \frac{\mu e^{-\mu t}}{(1-\bar{\phi})(1-e^{-\mu t})}. \tag{OA.2}$$

The right hand side of (OA.2) is independent of λ_0 and tends to 0 as $t \to \infty$. We conclude that for all $t > T_0$, $\bar{q}_t^B$ cannot cross $a(\lambda_0)$ from above.

Next, we show that if $T \geq T_0$ is sufficiently large and $\varepsilon_1 \leq \varepsilon_0$ sufficiently small, $\bar{q}_T^B > a(\lambda_0)$ for all $\lambda_0 \in (0, \varepsilon_1)$. This follows from

$$\begin{aligned}
\frac{\bar{q}_T^B}{a(\lambda_0)} &= \frac{\frac{\lambda_0}{\Lambda}\int_0^T \mu e^{-\mu s}(1-\phi_s)(1-e^{-\Lambda(T-s)})\,ds}{\frac{\lambda_0}{\Lambda+r}\int_0^T \mu e^{-\mu s}(1-\phi_s)\,ds} \\
&\xrightarrow{\lambda_0 \to 0} \frac{\lambda_1+r}{\lambda_1}\frac{\int_0^T \mu e^{-\mu s}(1-p_0 e^{-\lambda_1 s})(1-e^{-\lambda_1(T-s)})\,ds}{\int_0^T \mu e^{-\mu s}(1-p_0 e^{-\lambda_1 s})\,ds} \\
&\xrightarrow{T\to\infty} \frac{\lambda_1+r}{\lambda_1} > 1.
\end{aligned}$$

The claim then follows by choosing $T \geq T_0$ and then choosing ε_1 accordingly. Since we have ruled out crossing from above, and also $p_t^0 \geq \bar{q}_t^B$ by (54), this establishes that $p_t^0 \geq \bar{q}_t^B \geq a(\lambda_0)$ for all $t \geq T$ and $\lambda_0 \in (0, \varepsilon_1)$.

Finally, we prove there exists $\varepsilon_2 \leq \varepsilon_1$ such that for all $\lambda_0 \in (0, \varepsilon_2)$, $p_t^0 > a(\lambda_0)$ on $[0, T]$. Recall from the proof of Lemma 19 that p_t^0 is bounded below by the solution $\tilde{p}$ to

$$\dot{p}_t = \lambda_0(1-p_t) - \lambda_1 p_t - \mu p_t(1-p_t), \tag{OA.3}$$

which by the comparison theorem is bounded below by the solution $p_t^\flat = p_0 e^{-(\lambda_1+\mu)t}$ to

$$\dot{p}_t = -\lambda_1 p_t - \mu p_t. \tag{OA.4}$$

For sufficiently small $\varepsilon_2 > 0$, we have $p_T^\flat > a(\varepsilon_2)$. It follows that for all $\lambda_0 \in (0, \varepsilon_2)$ and all $t \in [0, T]$,

$$p_t \geq p_t^0 \geq p_t^\flat \geq p_T^\flat > a(\varepsilon_2) > a(\lambda_0). \tag{OA.5}$$

We have therefore established that $p_t \geq a(\lambda_0)$ for all $t \geq 0$ and all $\lambda_0 \in (0, \varepsilon_2)$.

$\square$

For the second step, we now characterize the disclosure of good evidence and prove its optimality. During the stockpiling and purging phases, the laws of motion are

$$\dot{G}_{1,t} = G_{0,t}\lambda_0 - G_{1,t}\lambda_1 + (p_t - G_{1,t} - B_{1,t})\mu - G_{1,t}\kappa_t(1 - G_{1,t} - G_{0,t}) \tag{OA.6}$$

$$\dot{G}_{0,t} = -G_{0,t}\lambda_0 + G_{1,t}\lambda_1 - G_{0,t}\kappa_t(1 - G_{1,t} - G_{0,t}) \tag{OA.7}$$

$$\dot{B}_{1,t} = B_{0,t}\lambda_0 - B_{1,t}\lambda_1 + B_{1,t}\kappa_t(G_{0,t} + G_{1,t}) \tag{OA.8}$$

$$\dot{B}_{0,t} = -B_{0,t}\lambda_0 + B_{1,t}\lambda_1 + (1 - p_t - G_{0,t} - B_{0,t})\mu + B_{0,t}\kappa_t(G_{0,t} + G_{1,t}) \tag{OA.9}$$

$$\dot{p}_t = \lambda_0(1 - p_t) - \lambda_1 p_t - G_{1,t}\kappa_t + p_t(G_{0,t} + G_{1,t})\kappa_t. \tag{OA.10}$$

The local delay IC condition becomes

$$\dot{q}^G_t \geq (\lambda_0 + \lambda_1 + r)(q^G_t - p_t) + (\lambda_0 + \lambda_1)(p^* - q^G_t) \tag{OA.11}$$

$$= rq^G_t - (\lambda_0 + \lambda_1 + r)p_t + \lambda_0. \tag{OA.12}$$

Define $c(t;\lambda_0) = \dot{q}^G_t - [rq^G_t - (\lambda_0 + \lambda_1 + r)p_t + \lambda_0]$, generalizing $c(t)$ from the $\lambda_0 = 0$ case.

By adapting arguments from the proof of Lemma 13, allowing for $\lambda_0 > 0$, we obtain generalized cutoffs

$$\overline{p} := \frac{\lambda_0 + r + \frac{\lambda_1}{2}}{\lambda_0 + \lambda_1 + r}$$

$$\underline{p} := \frac{\lambda_0 + r}{\lambda_0 + \lambda_1 + r}.$$

For the rest of the proof, all references to $\underline{p}$ and $\overline{p}$ or $\underline{p}(\lambda_0)$ and $\overline{p}(\lambda_0)$ refer to these updated expressions. Note that $\overline{p}(\lambda_0)$ and $\underline{p}(\lambda_0)$ are strictly increasing and continuous in λ_0. It follows that for a fixed prior p_0, for sufficiently small $\lambda_0 \geq 0$, the classification below does not change with λ_0. Specifically, if $p_0 = \overline{p}(0)$, then for all sufficiently small $\lambda_0 > 0$, we have $p_0 \in (\underline{p}(\lambda_0), \overline{p}(\lambda_0))$, and therefore p_0 lies in the moderate prior region in either case. And if $p_0 = \underline{p}(0)$, then for all sufficiently small $\lambda_0 > 0$ we have $p_0 < \underline{p}(\lambda_0)$, and p_0 lies in the low prior region in either case. All other priors in $(0, 1)$ of course remain in their respective regions for all sufficiently small λ_0.

We now characterize the unique equilibrium within the class for each case (i) $p_0 \leq \underline{p}$, (ii) $p_0 > \overline{p}$, and (iii) $p_0 \in (\underline{p}, \overline{p}]$.

For $p_0 \leq \underline{p}$, the equilibrium must start in the transparency phase, so $q^G_t \equiv 1$, $\dot{q}^G_t \equiv 0$. The local IC condition for immediate disclosure to be optimal at all times then reduces to $0 \leq r \cdot 1 - (\Lambda + r)p_t + \lambda_0$, which is equivalent to $p_t \leq \underline{p}$. Now by assumption $p_0 \leq \underline{p}$, and it is easy to see from the ODE $\dot{p}_t = -\lambda_1 p_t + \lambda_0(1 - p_t) - \mu(p_t - B_{1,t})(1 - p_t)$ that whenever $p_t = \underline{p}$, $\dot{p}_t < 0$, so $p_t < \underline{p}$ for all $t > 0$.

For $p_0 > \overline{p}$, the equilibrium must start in the stockpiling phase. To characterize this phase, we argue that for sufficiently small λ_0, there exists a $\underline{T}(\lambda_0) > 0$ such that (OA.11) is slack on $[0, \underline{T}(\lambda_0))$ and binds at $\underline{T}(\lambda_0)$. We have an IVP consisting of (OA.6)-(OA.10) with $\kappa \equiv 0$ and the initial conditions (which are independent of λ_0). This system has the form $\dot{z} = F(z_t; \lambda_0)$ for $z = (G_1, G_0, B_1, B_0, p)$, where F is of class C^1. Now the unique solution $z(t, 0)$ for $\lambda_0 = 0$ exists on $[0, \underline{T}]$, and for some $\varepsilon > 0$, it can be extended locally to $[0, \underline{T} + \varepsilon]$, a compact interval. By standard continuous dependence results (e.g., Teschl (2012, Theorem 2.11)) for all $\lambda_0 > 0$ in a neighborhood of 0, a unique solution $z(t, \lambda_0)$ to the system exists on $[0, \underline{T} + \varepsilon)$ and is of class C^1. Hence, $c(t, \lambda_0)$ is of class C^1. At $\lambda_0 = 0$, we have $c(\underline{T}; 0) = 0$, $c(t; 0) > 0$ for $t \in [0, \underline{T})$, and $\frac{\partial}{\partial t} c(\underline{T}; 0) < 0$ by the proof of Theorem 2. By the implicit function theorem, there exists $\underline{T}(\lambda_0)$ solving $c(\underline{T}(\lambda_0); \lambda_0) = 0$ and $\underline{T}(\lambda_0)$ is C^1. Therefore, $c_t(\underline{T}(\lambda_0); \lambda_0) < 0$ for all λ_0 sufficiently small. Fix any $\delta \in (0, \underline{T}(0))$. Since $c(t; 0) > 0$ on the compact interval $[0, \underline{T}(0) - \delta]$ and $c(t; \lambda_0)$ converges uniformly to $c(t; 0)$ on $[0, \underline{T}(0) - \delta]$, we have $c(t; \lambda_0) > 0$ on $[0, \underline{T}(0) - \delta]$ for all sufficiently small λ_0. Finally, by choosing δ sufficiently small and λ_0 sufficiently small, we can also ensure that $\underline{T}(\lambda_0) \in (\underline{T}(0) - \delta, \underline{T}(0) + \delta)$, and is the only solution of $c(\underline{T}(\lambda_0); \lambda_0) = 0$. Fixing this δ, for all sufficiently small λ_0, we then have $c(t; \lambda_0) > 0$ for all $t \in [0, \underline{T}(\lambda_0))$; i.e., $\underline{T}(\lambda_0)$ is the first solution to $c(t; \lambda_0) = 0$, and must be the start of the purging phase. Continuous dependence of the solution on λ_0 also implies that the values $(p_{\underline{T}(\lambda_0)}, q^G_{\underline{T}(\lambda_0)}, B_{1,\underline{T}(\lambda_0)}, B_{0,\underline{T}(\lambda_0)})$ are C^1 in λ_0.

Next, we characterize the purging phase. We solve for $S(p, q^G, B_1)$ and $\kappa(p, q^G, B_1, B_0)$ using similar steps to Theorem 2. Suppressing superscript G in q^G, we have the indifference equation

$$\dot{q}_t = r q_t - (\lambda_0 + \lambda_1 + r) p_t + \lambda_0 \tag{OA.13}$$

and the belief-updating laws of motion

$$\dot{S}_t = (p_t - S_t q_t - B_{1,t})\mu - S_t \kappa_t (1 - S_t) \tag{OA.14}$$

$$\dot{q}_t = -\lambda_1 q_t + \lambda_0 (1 - q_t) + (1 - q_t)\mu \frac{p_t - B_{1,t} - S_t q_t}{S_t}. \tag{OA.15}$$

Combining (OA.13) and (OA.15) yields

$$r q_t - (\lambda_0 + \lambda_1 + r) p_t + \lambda_0 = -\lambda_1 q_t + \lambda_0 (1 - q_t) + (1 - q_t)\mu \frac{p_t - B_{1,t} - S_t q_t}{S_t}$$

$$\implies S_t = S(p_t, q_t, B_{1,t}) := \frac{\mu (p_t - B_{1,t})(1 - q_t)}{(q_t - p_t)(\Lambda + r) + \mu q_t (1 - q_t)}.$$

By differentiating this expression, we obtain a second expression for $\dot{S}_t$; equating it to (OA.14) allows us to solve for κ in the form

$$
\begin{aligned}
\kappa_t &= \kappa(p_t, q_t, B_{1,t}, B_{0,t}) \\
&= \frac{((\Lambda + r)(q_t - p_t) + \mu q_t(1 - q_t))N(p_t, q_t, B_{1,t}, B_{0,t})}{(\Lambda + r)(p_t - B_{1,t})(q_t - p_t)(1 - q_t)[\mu B_{1,t}(1 - q_t) + (q_t - p_t)(\Lambda + r + \mu(1 - q_t))]},
\end{aligned}
$$

where

$$
\begin{aligned}
N(p_t, q_t, B_{1,t}, B_{0,t}) =& 2(r + \Lambda)^2 p_t^3 - (r + \Lambda)p_t^2 \left(r + 2\lambda_0 + \lambda_1 + \mu + (3r + 2\Lambda - \mu)q_t\right) \\
& - \lambda_0(1 - B_{0,t})(1 - q_t)q_t[r + \Lambda + \mu(1 - q_t)] \\
& + p_t\Big(\lambda_0(r + \Lambda + \mu) - \lambda_0(r + \Lambda)B_{0,t}(1 - q_t) \\
& \quad + \left((r + \Lambda)^2 - \lambda_0\mu + \lambda_1\mu + r\mu\right) q_t \\
& \quad + \left(r^2 + r(\Lambda - \mu) - \lambda_1\mu\right) q_t^2\Big) \\
& - B_{1,t}\Big(\lambda_0\mu + 2(r + \Lambda)^2 p_t^2 \\
& \quad + \left((r + \Lambda)^2 - 2\lambda_0\mu + \lambda_1\mu + r\mu\right) q_t \\
& \quad + \left(r^2 + \lambda_0^2 + r(2\lambda_0 + \lambda_1 - \mu) - \lambda_1\mu + \lambda_0(\lambda_1 + 2\mu)\right) q_t^2 \\
& \quad - \lambda_0\mu q_t^3 - (r + \Lambda)p_t \left(r + \Lambda + \mu + (3r + 3\lambda_0 + 2\lambda_1 - \mu)q_t\right)\Big).
\end{aligned}
$$

Substituting $S_t = S(p_t, q_t, B_{1,t})$ and $\kappa_t = \kappa(p_t, q_t, B_{1,t}, B_{0,t})$ into (OA.6)-(OA.10) and shifting time so that $t = 0$ represents the start of the purging phase, we obtain an autonomous initial value problem in $z_t = (p_{t+\underline{T}}, q^G_{t+\underline{T}}, B_{1,t+\underline{T}}, B_{0,t+\underline{T}})$ of the form

$$
\dot{z}_t = F(z_t; \lambda_0) \tag{OA.16}
$$

with initial condition $z_0(\lambda_0) = (p_{\underline{T}}, q^G_{\underline{T}}, B_{1,\underline{T}}, B_{0,\underline{T}})$ given by the end of the stockpiling phase. The right-hand sides of the ODEs are ratios of polynomials in (z, λ). After simplifications which cancel $(1 - q_t)/(1 - q_t)$, the denominators are all nonzero if z lies in the set

$$
\begin{aligned}
U(\lambda_0) := \{z \in \mathbb{R}^4 : & \, q > p > B_1, \\
& \mu B_1(1 - q) + (q - p)(\Lambda + r + \mu(1 - q)) > 0, \\
& (q - p)(\Lambda + r) + \mu q(1 - q) > 0\}.
\end{aligned}
$$

Thus, we define the open set

$$\mathcal{U} := \{(z, \lambda_0) \in \mathbb{R}^4 \times \mathbb{R} : z \in U(\lambda_0)\}$$

and F is C^1 on $\mathcal{U}$. Note that, given the solutions (G_1, G_0, p, κ) to the purging phase system in the proof of Theorem 2 for $\lambda_0 = 0$, and letting $\overline{T}(0)$ denote the end of the purging interval there, B_0 can be calculated from its reduced ODE; these variables together with $B_1 \equiv 0$ constitute a solution $z(t, 0)$ on $[0, \overline{T}(0) - \underline{T}(0)]$ to (OA.6)-(OA.10) when $\lambda_0 = 0$. Now, at $\lambda_0 = 0$, it is easy to verify that $(z(0, 0), 0) \in \mathcal{U}$, and that $(z(t, 0), 0)$ stays inside $\mathcal{U}$ on $[0, \overline{T}(0) - \underline{T}(0)]$. Since the end point $(z(\overline{T}(0) - \underline{T}(0), 0), 0)$ is in $\mathcal{U}$, the solution can be extended locally to $[0, \overline{T}(0) - \underline{T}(0) + \varepsilon]$ for some $\varepsilon > 0$. By continuous dependence, for all λ_0 sufficiently small, $(z_0(\lambda_0), \lambda_0) \in \mathcal{U}$ and a unique solution $z_t(\lambda_0)$ to the IVP exists on $[0, \overline{T}(0) - \underline{T}(0) + \varepsilon]$ and $(z_t(\lambda_0), \lambda_0)$ stays inside $\mathcal{U}$ and $z_t(\lambda_0)$ is C^1 jointly in (t, λ_0). Moreover, it is easy to check that $q_0(0) < 1$, and by definition, $\overline{T}(0) - \underline{T}(0)$ is the first time that $q_t(0)$ hits 1. Also, the proof of Theorem 2 shows that the crossing is transversal: $\dot{q}_{\overline{T}(0)-\underline{T}(0)}(0) > 0$. Hence, by a repeat of the argument used for $\underline{T}$ at the end of the stockpiling phase, for sufficiently small $\lambda_0 > 0$, there exists $\overline{T}(\lambda_0)$ such that $q_{\overline{T}(\lambda_0)-\underline{T}(\lambda_0)}(\lambda_0) = 1$ and $\overline{T}(\lambda_0) - \underline{T}(\lambda_0)$ is the first time $q_t(\lambda_0)$ hits 1; i.e., $q_t(\lambda_0) < 1$ for all $t \in [0, \overline{T}(\lambda_0) - \underline{T}(\lambda_0))$.

We now recover S and κ. For S, the denominator is already nonzero by the definition of $\mathcal{U}$, so S is well-defined. For κ, the denominator is nonzero by the definition of $\mathcal{U}$ and the fact that $q_t(\lambda_0) < 1$ for all $t \in [0, \overline{T}(\lambda_0) - \underline{T}(\lambda_0))$. We must also verify that κ is strictly positive. At $\lambda_0 = 0$, recall that the baseline solution satisfies $\kappa_t(0) > 0$ throughout the purging phase; in fact, the proof of Theorem 2 shows that the numerator is strictly positive on $[0, \overline{T}(0) - \underline{T}(0)]$ for $(z(0), 0)$; by continuity in t, it remains strictly positive on $[0, \overline{T}(0) - \underline{T}(0) + \varepsilon]$ for sufficiently small $\varepsilon > 0$ (chosen sufficiently small to still satisfy its earlier definition). Also, the numerator is polynomial, and therefore continuous, in $(p, q, B_1, B_0, \lambda_0)$ on $\mathcal{U}$. It follows that for λ_0 sufficiently small, the numerator in $\kappa(p_t, q_t, B_{1,t}, B_{0,t})$ is strictly positive on $[0, \overline{T}(0) - \underline{T}(0) + \varepsilon]$. By our definition of $\mathcal{U}$, for sufficiently small λ_0, the denominator terms other than the factor $1 - q_t$ are also strictly positive on $[0, \overline{T}(0) - \underline{T}(0) + \varepsilon]$. Therefore, κ_t stays well defined and strictly positive on $[0, \overline{T}(\lambda_0) - \underline{T}(\lambda_0))$.

Having characterized the purging phase, we can easily establish optimality of the transparency phase. By continuity of solutions and $p_{\overline{T}(0)-\underline{T}(0)}(0) < \underline{p}(0) = \frac{\lambda_0 + r}{\lambda_0 + \lambda_1 + r}\big|_{\lambda_0=0}$, we have $p_{\overline{T}(\lambda_0)-\underline{T}(\lambda_0)}(\lambda_0) < \underline{p}(\lambda_0)$ for all sufficiently small λ_0. Since p cannot cross $\underline{p}$ from below in the transparency phase (as argued already when starting from $p_0 \leq \underline{p}$), the same inequality holds at all future times. It follows that immediate disclosure of good evidence is optimal at all future times given that the receiver conjectures it.

Global optimality follows from local optimality exactly as in Theorem 2.

Finally, consider the moderate prior case, $p_0 \in (\underline{p}, \overline{p}]$. We have $(z_0, \lambda_0) = (p_0, 1, 0, 0, \lambda_0) \in \mathcal{U}$. The equilibrium must begin in the purging phase, and the system in z obtained above applies in the interior of this phase. At time 0, we have $q_0 = 1$, but as in the moderate priors part of the proof of Theorem 2, a unique local solution exists, $(p_t, q_t, B_{1,t}, B_{0,t}, \lambda_0)$ is in $\mathcal{U}$, and $q_t < 1$ for all sufficiently small $t > 0$. To verify the latter claim, note that at time 0,

$$\dot{q}_0(\lambda_0) = r - (\Lambda + r)p_0 + \lambda_0 = (\Lambda + r)(\underline{p} - p_0) < 0,$$

so $q_t(\lambda_0) < 1$ for all sufficiently small $t > 0$. At $\lambda_0 = 0$ the solution coincides with the moderate prior purging solution from the proof of Theorem 2. The same continuous dependence and transversal crossing arguments as in the high prior case establish a solution on $[0, \overline{T}(\lambda_0))$ for all sufficiently small λ_0, with $\overline{T}(\lambda_0)$ continuously differentiable, and with $\overline{T}(\lambda_0)$ being the first time after time 0 that $q_t = 1$. The arguments that recover S and κ and establish the positivity of κ_t on $(0, \overline{T}(\lambda_0))$ apply unchanged. The remaining arguments for the transparency and overall optimality of the sender's strategy are also unchanged.

Proof of Lemma 15. In the stockpiling phase, we have

$$\dot{p}_t = -\lambda_1 p_t \tag{OA.17}$$

$$\dot{q}_t = -\lambda_1 q_t + (1 - q_t)\mu \frac{p_t - S_t q_t}{S_t} \tag{OA.18}$$

$$\dot{S}_t = (p_t - S_t q_t)\mu, \tag{OA.19}$$

with initial conditions $(p_0, q_0, S_0) = (p_0, 1, 0)$. That p_t is strictly positive for all $t \in [0, \underline{T}]$ follows from the comparison theorem and is also evident from the closed-form solution.

It is useful to show that $S_t > 0$ until $\underline{T}$. We have $\dot{S}_0 > 0$, so $S_t > 0$ for all sufficiently small t. Now write (OA.19) as $\dot{S}_t = f(t, S_t)$, and define $y_t \equiv 0$. Then $S_t > y_t$ and $\dot{S}_t - f(t, S_t) = 0 > -p_t\mu = \dot{y}_t - f(t, y_t)$. By the comparison theorem, $S_t > y_t = 0$ for all $t > 0$.

The comparison theorem also implies $q_t > 0$ for all $t \in [0, \underline{T}]$; moreover, $q_t < 1$ for all $t \in (0, \underline{T}]$, since $\dot{q}_0 = -\frac{\lambda_1}{2} < 0$ and the RHS of the q-ODE is negative whenever $q_t = 1$.

Also note that $N_t := p_t - S_t q_t = p_t - G_{1,t}$, which has ODE $\dot{N}_t = -(\lambda_1 + \mu)N_t$. Since $N_0 = p_0 - G_{1,0} = p_0 \in (0, 1)$, we have $N_t = p_t - G_{1,t} > 0$ for all $t \in [0, \underline{T})$.

Define $\delta_t := q_t - p_t$, where $\delta_0 = 1 - p_0 > 0$. We have

$$\dot{\delta}_t = -\lambda_1 \delta_t + (1 - q_t)\mu \frac{N_t}{S_t}. \tag{OA.20}$$

The inequalities just established imply the second term is positive for all $t \in (0, \underline{T}]$. Hence, by the comparison theorem, we have $\delta_t > 0$ for all $t \in [0, \underline{T}]$.

To prove that $A_{\underline{T}} > 0$, recall from earlier in the proof that $c'(\underline{T}) < 0$. Hence, $c(\underline{T})$ crosses 0 from above at $\underline{T}$. Now $c(\underline{T}) = 0$ implies

$$0 = \dot{q}_{\underline{T}} - rq_{\underline{T}} + (\lambda_1 + r)p_{\underline{T}}, \tag{OA.21}$$

and combining this with (OA.18) at $t = \underline{T}$ yields $S_{\underline{T}} = S(p_{\underline{T}}, q_{\underline{T}})$. Now when $c(t) = 0$, some algebra shows that $c'(t) = -\frac{(\lambda_1+r)(q_t-p_t)}{1-q_t} A_t$. Since $c(\underline{T}) = 0$ by construction and $c'(\underline{T}) < 0$, it follows that $A_{\underline{T}} > 0$. □

2 Supporting results for Section 4.2

Proof of Lemma 18. Let π_N, π_G, and π_B denote the probabilities of no disclosure, good evidence disclosure, and bad evidence disclosure, respectively, on $(t, t+s]$ conditional on no disclosure on $[0, t]$. Recall that $\phi_s(p_t)$ denotes the belief at time $t + s$ based on the belief p_t at time t and no information other than Markov switching dynamics in $(t, t + s]$. By the law of iterated expectations,

$$\phi_s(p_t) = \pi_N p_{t+s} + \pi_B \, \mathbb{E}[\theta_{t+s} | \sigma^B \in (t, t+s]] + \pi_G \, \mathbb{E}[\theta_{t+s} | \sigma^G \in (t, t+s]].$$

We claim that $\mathbb{E}[\theta_{t+s} | \sigma^B \in (t, t+s]] \leq \phi_s(\bar{q}_t^B) \leq \phi_s(p_t)$. To derive the first inequality, write

$$\begin{aligned}\mathbb{E}[\theta_{t+s} | \sigma^B \in (t, t+s]] = {} & \Pr(\tau \leq t | \sigma^B \in (t, t+s])\mathbb{E}[\theta_{t+s} | \tau \leq t, \sigma^B \in (t, t+s]] \\ & + \Pr(\tau > t | \sigma^B \in (t, t+s])\mathbb{E}[\theta_{t+s} | \tau > t, \sigma^B \in (t, t+s]].\end{aligned}$$

For the first term, since bad evidence disclosure is age-independent, we have $\mathbb{E}[\theta_{t+s} | \tau \leq t, \sigma^B \in (t, t+s]] = \phi_s(\bar{q}_t^B)$. For the second term, if evidence arrives at some time $u \in (t, t+s]$, then $\theta_u = 0$, and the conditional belief about θ_{t+s} is $\phi_{t+s-u}(0) \leq \phi_s(0) \leq \phi_s(\bar{q}_t^B)$. Averaging over u, we have $\mathbb{E}[\theta_{t+s} | \tau > t, \sigma^B \in (t, t+s]] \leq \phi_s(\bar{q}_t^B)$ and overall $\mathbb{E}[\theta_{t+s} | \sigma^B \in (t, t+s]] \leq \phi_s(\bar{q}_t^B) \leq \phi_s(p_t)$, where the last inequality is due to $\bar{q}_t^B \leq p_t$.

Also, $\mathbb{E}[\theta_{t+s} | \sigma^G \in (t, t+s]] \leq 1$. Rearranging the equation and using these inequalities,

$$\begin{aligned} & \pi_N p_{t+s} \geq \phi_s(p_t) - \pi_B \phi_s(p_t) - \pi_G \\ \iff \; & \pi_N (p_{t+s} - p_t) \geq (\pi_N + \pi_G)(\phi_s(p_t) - p_t) - \pi_G(1 - p_t).\end{aligned}$$

Now $\pi_N > 0$ and we have

$$\begin{aligned} p_{t+s} - p_t &\geq \left(1 + \frac{\pi_G}{\pi_N}\right)(\phi_s(p_t) - p_t) - \frac{\pi_G}{\pi_N}(1 - p_t) \\ &= (\phi_s(p_t) - p_t) - \frac{\pi_G}{\pi_N}(1 - \phi_s(p_t)). \end{aligned} \tag{OA.22}$$

Expanding the first term yields

$$\begin{aligned} \phi_s(p_t) - p_t = p^* - (p^* - p_t)e^{-(\lambda_0+\lambda_1)s} - p_t &= (p^* - p_t)(1 - e^{-(\lambda_0+\lambda_1)s}) \\ \geq -(1-p^*)(1 - e^{-(\lambda_0+\lambda_1)s}) &= -\frac{\lambda_1}{\lambda_0+\lambda_1}(1 - e^{-(\lambda_0+\lambda_1)s}) \geq -\lambda_1 s. \end{aligned}$$

Let M be the probability that no evidence has arrived in $[0,t]$ conditional on no disclosure in $[0,t]$. We have $\pi_G \leq M(1-e^{-\mu s})$ and $\pi_N \geq Me^{-\mu s}$, and therefore $\frac{\pi_G}{\pi_N} \leq \frac{1-e^{-\mu s}}{e^{-\mu s}} = e^{\mu s} - 1$. Hence, returning to (OA.22),

$$p_{t+s} - p_t \geq -\lambda_1 s - (e^{\mu s} - 1).$$

Since we can partition $[t, t+s]$ into N intervals of width s/N and apply the same bound on each interval, summing, and taking N arbitrarily large, we have

$$p_{t+s} - p_t \geq -(\lambda_1 + \mu)s.$$

To derive the implication $p_t \geq p_{t-}$, fix $u > 0$, set $t = u - s$, and take $s \to 0$. Then the previous result and existence of a left limit imply $p_u \geq p_{u-s} - (\lambda_1 + \mu)s \to p_{u-}$. $\square$

The following lemma states a useful upper bound on p_t. Intuitively, for p_t to be close to 1, it must be that the sender possesses very recent evidence with very high probability, which is impossible.

Lemma OA.2. *Fix any $\delta > 0$. Then in all equilibria, p_t is uniformly bounded away from 1: $p_t \leq p(\delta) := e^{-\mu\delta}\max\{p_0, p^* + (1-p^*)e^{-\Lambda\delta}\} + (1 - e^{-\mu\delta}) \cdot 1 < 1$.*

Proof. Suppose there has been no disclosure up to time t. By definition, the belief is p_t. Now fix any $\delta > 0$ (allowing for $\delta > t$). There are four possibilities: (i) no evidence has arrived by t, (ii) good evidence arrived by $t - \delta$, (iii) good evidence arrived in $(t-\delta, t]$, and (iv) bad evidence arrived. In case (i), the updated belief is $\phi_t \leq \max\{p_0, p^*\}$. In case (ii), it is at most $\phi_\delta(1) = p^* + (1-p^*)e^{-\Lambda\delta}$. In case (iii), it is at most 1. In case (iv), it is at most $\phi_t(0) = p^*(1 - e^{-\Lambda t}) \leq p^*$. Let A be the probability of (iii) conditional on no disclosure up to time t. Then we have

$$p_t \leq (1-A)\max\{p_0, p^* + (1-p^*)e^{-\Lambda\delta}\} + A \cdot 1. \tag{OA.23}$$

Also, by Bayes' rule we have

$$\begin{aligned}
A &= \frac{\Pr(\tau^G \in (t-\delta,t]), \sigma > t)}{\Pr(\sigma > t)} \\
&\leq \frac{\Pr(\tau \in (t-\delta,t]), \sigma > t)}{\Pr(\sigma > t)} \\
&= \frac{\Pr(\tau \in (t-\delta,t]), \sigma > t)}{\Pr(\tau \in (t-\delta,t], \sigma > t) + \Pr(\tau \leq t-\delta, \sigma > t) + \Pr(\tau > t)} \\
&\leq \frac{\Pr(\tau \in (t-\delta,t]), \sigma > t)}{\Pr(\tau \in (t-\delta,t], \sigma > t) + \Pr(\tau > t)} \\
&\leq \frac{e^{-\mu(t-\delta)} - e^{-\mu t}}{e^{-\mu(t-\delta)} - e^{-\mu t} + e^{-\mu t}} \\
&= 1 - e^{-\mu\delta},
\end{aligned}$$

where the second to last line uses that $\Pr(\tau \in (t-\delta,t]) = e^{-\mu\max\{t-\delta,0\}} - e^{-\mu t} \leq e^{-\mu(t-\delta)} - e^{-\mu t}$ for all $\delta > 0$, including $\delta > t$. Using this in the inequality for p_t yields the uniform upper bound in the lemma. □

Proof of Lemma 20. The fact that $\overline{\Delta}(r) \to \infty$ can be seen in (59), where r only enters the term $(\Lambda + r)$, and so on any finite interval, the inequality holds for all sufficiently large r.

Next, we show that $\lim_{r\to\infty} r^2\Gamma(\overline{\Delta}(r)) = 0$. We have

$$r^2\Gamma(\overline{\Delta}(r)) = r^2\int_0^{\varepsilon} e^{-rs}y_s\,ds + r^2\int_{\varepsilon}^{\overline{\Delta}(r)} e^{-rs}y_s\,ds - r^2(1-e^{-r\overline{\Delta}(r)})\frac{p^* - \bar{q}^B_{\overline{\Delta}(r)}}{\Lambda + r}. \tag{OA.24}$$

On the right hand side of (OA.24), the second term clearly tends to 0 as $r \to \infty$. The first term Taylor-expands as

$$\begin{aligned}
&r^2\int_0^{\varepsilon} e^{-rs}[y_0 + s\dot{y}_0 + \frac{s^2}{2}\ddot{y}_0 + o(s^2)]\,ds \\
&= \dot{y}_0[1 - e^{-r\varepsilon}(1 + r\varepsilon)] + \frac{\ddot{y}_0}{2r}[2 - e^{-r\varepsilon}(2 + r\varepsilon(2 + r\varepsilon))] + r^2\int_0^{\varepsilon} e^{-rs}o(s^2)\,ds \\
&\to \dot{y}_0 + 0 + 0 = \dot{y}_0.
\end{aligned}$$

The last term in (OA.24) has the same limit as $r(p^* - \bar{q}^B_{\overline{\Delta}(r)})$ as $r \to \infty$ because $\frac{r}{\Lambda+r}(1 - e^{-r\overline{\Delta}(r)}) \to 1$ (due to $\overline{\Delta}(r) \to \infty$). Because of the stationary prior, $p^* = \phi_{\overline{\Delta}(r)}$, and then since local IC (60) binds at $\overline{\Delta}(r)$, we have

$$r(p^* - \bar{q}^B_{\overline{\Delta}(r)}) = \frac{r}{\Lambda + r}\bar{q}^B_{\overline{\Delta}(r)} \frac{\mu(1-p^*)(N^{\text{un}}_{\overline{\Delta}(r)} + B^{\text{un}}_{\overline{\Delta}(r)})}{B^{\text{un}}_{\overline{\Delta}(r)}} \to \mu p^*(1-p^*) = \dot{y}_0, \tag{OA.25}$$

where we have used that $N^{\text{un}}_{\overline{\Delta}(r)} = e^{-\mu\overline{\Delta}(r)}$, that $\overline{\Delta}(r) \to \infty$, and the limits derived in the proof of Lemma 2. Thus, the last term in (OA.24) tends to $\dot{y}_0$, so $\lim_{r\to\infty} r^2\Gamma(\overline{\Delta}(r)) = \dot{y}_0 + 0 - \dot{y}_0 = 0$. □

Proof of Lemma 21. For $X(r)$, we split the integral into $[0, \varepsilon)$ and $[\varepsilon, \overline{\Delta}(r))$ and repeat the Taylor expansion from the proof of Lemma 20:

$$\int_0^{\overline{\Delta}(r)} e^{-rs} y_s \, ds = \int_0^{\varepsilon} e^{-rs}[y_0 + \dot{y}_0 s + \frac{1}{2}\ddot{y}_0 s^2 + o(s^2)] \, ds + \int_{\varepsilon}^{\overline{\Delta}(r)} e^{-rs} y_s \, ds. \tag{OA.26}$$

We have $r^3 \int_{\varepsilon}^{\overline{\Delta}(r)} e^{-rs} y_s \, ds \to 0$ as $r \to \infty$, so $X(r)$ has the same limit as

$$r^3 \left(\int_0^{\varepsilon} e^{-rs}[y_0 + \dot{y}_0 s + \frac{1}{2}\ddot{y}_0 s^2 + o(s^2)] \, ds - \frac{\dot{y}_0}{r^2} \right). \tag{OA.27}$$

Recall that $y_0 = 0$ and $\dot{y}_0 = -\dot{p}_0$. Hence (OA.27) simplifies to

$$r^3 \dot{y}_0 \left(\int_0^{\varepsilon} e^{-rs} s \, ds - \frac{1}{r^2} \right) + r^3 \int_0^{\varepsilon} e^{-rs} \frac{1}{2}\ddot{y}_0 s^2 \, ds + r^3 \int_0^{\varepsilon} e^{-rs} o(s^2) \, ds. \tag{OA.28}$$

We address each term in (OA.28) individually. The first term reduces to

$$r^3 \dot{y}_0 \left(\frac{1}{r^2}(1 - (1 + r\varepsilon)e^{-r\varepsilon}) - \frac{1}{r^2} \right) = -\dot{y}_0 r(1 + r\varepsilon)e^{-r\varepsilon} \xrightarrow{r\to\infty} 0. \tag{OA.29}$$

The second term in (OA.28) gives

$$\frac{1}{2}\ddot{y}_0[2 - (2 + 2r\varepsilon + r^2\varepsilon^2)e^{-r\varepsilon}] \to \ddot{y}_0. \tag{OA.30}$$

Finally, the third limit can be made arbitrarily small by taking ε sufficiently small. We conclude that $\lim_{r\to\infty} X(r) = \ddot{y}_0 = \dot{y}_0(\mu(2p^* - 1) - \Lambda)$.

For $Y(r)$, rearrange it to

$$r^3(1 - e^{-r\overline{\Delta}(r)}) \left(\frac{p^* - \bar{q}^B_{\overline{\Delta}(r)}}{\Lambda + r} - \frac{\dot{y}_0}{r^2} \right) - r^3 e^{-r\overline{\Delta}(r)} \frac{\dot{y}_0}{r^2}.$$

The final term tends to 0 and $1 - e^{-r\overline{\Delta}(r)} \to 1$, leaving us to calculate the limit of

$$
\begin{aligned}
r^3\left(\frac{p^* - \bar{q}^B_{\overline{\Delta}(r)}}{\Lambda + r} - \frac{\dot{y}_0}{r^2}\right) &= \dot{y}_0 r^3 \left(\frac{\bar{q}^B_{\overline{\Delta}(r)}(N^{\text{un}}_{\overline{\Delta}(r)} + B^{\text{un}}_{\overline{\Delta}(r)})}{(\Lambda + r)^2 B^{\text{un}}_{\overline{\Delta}(r)} p^*} - \frac{1}{r^2}\right) \\
&= \dot{y}_0 r \frac{r^2 \bar{q}^B_{\overline{\Delta}(r)}(N^{\text{un}}_{\overline{\Delta}(r)} + B^{\text{un}}_{\overline{\Delta}(r)}) - (\Lambda + r)^2 B^{\text{un}}_{\overline{\Delta}(r)} p^*}{(\Lambda + r)^2 B^{\text{un}}_{\overline{\Delta}(r)} p^*} \\
&= \dot{y}_0 \left[\frac{r^3(\bar{q}^B_{\overline{\Delta}(r)}(N^{\text{un}}_{\overline{\Delta}(r)} + B^{\text{un}}_{\overline{\Delta}(r)}) - B^{\text{un}}_{\overline{\Delta}(r)} p^*)}{(\Lambda + r)^2 B^{\text{un}}_{\overline{\Delta}(r)} p^*} - \frac{2\Lambda r^2}{(\Lambda + r)^2} - \frac{r\Lambda^2}{(\Lambda + r)^2}\right]
\end{aligned} \tag{OA.31}
$$

where the first equality uses the expression for $r(p^* - \bar{q}^B_{\overline{\Delta}(r)})$ in (OA.25) and that $\mu(1 - p^*) = \frac{\dot{y}_0}{p^*}$, the second simply combines fractions and simplifies, and the third splits by degree of r. Now the second and third terms in the square brackets in (OA.31) contribute $-2\Lambda\dot{y}_0 - 0$ to the limit as $r \to \infty$. The first term in (OA.31) contributes

$$
\dot{y}_0 \frac{r^3 B^{\text{un}}_{\overline{\Delta}(r)}(\bar{q}^B_{\overline{\Delta}(r)} - p^*)}{(\Lambda + r)^2 B^{\text{un}}_{\overline{\Delta}(r)} p^*} + \dot{y}_0 \frac{r^3 \bar{q}^B_{\overline{\Delta}(r)} N^{\text{un}}_{\overline{\Delta}(r)}}{(\Lambda + r)^2 B^{\text{un}}_{\overline{\Delta}(r)} p^*}. \tag{OA.32}
$$

Since $r(\bar{q}^B_{\overline{\Delta}(r)} - p^*) \to -\dot{y}_0$ by Lemma 20, the first term of (OA.32) tends to $-\frac{(\dot{y}_0)^2}{p^*}$. The second term of (OA.32) has the same limit as $\dot{y}_0 r \frac{N^{\text{un}}_{\overline{\Delta}(r)}}{B^{\text{un}}_{\overline{\Delta}(r)}}$, because $\bar{q}^B_{\overline{\Delta}(r)} \to p^*$. We now calculate the limit of

$$
r\frac{N^{\text{un}}_{\overline{\Delta}(r)}}{B^{\text{un}}_{\overline{\Delta}(r)}} = r(p^* - \bar{q}^B_{\overline{\Delta}(r)}) \frac{\frac{N^{\text{un}}_{\overline{\Delta}(r)}}{B^{\text{un}}_{\overline{\Delta}(r)}}}{p^* - \bar{q}^B_{\overline{\Delta}(r)}},
$$

where again $r(p^* - \bar{q}^B_{\overline{\Delta}(r)}) \to \dot{y}_0$. Since r only enters the remaining term via $\overline{\Delta}(r)$, we are only left with calculating

$$
L := \lim_{t \to \infty} \frac{N^{\text{un}}_t}{B^{\text{un}}_t (p^* - \bar{q}^B_t)}, \tag{OA.33}
$$

for which we turn to closed forms. For $\Lambda \neq \mu$, under the stationary prior, we have

$$
\bar{q}^B_t = p^* \left(1 + \frac{\mu}{\Lambda - \mu} \frac{e^{-\Lambda t} - e^{-\mu t}}{1 - e^{-\mu t}}\right) \qquad \text{and} \qquad \frac{N^{\text{un}}_t}{B^{\text{un}}_t} = \frac{e^{-\mu t}}{(1 - p^*)(1 - e^{-\mu t})}.
$$

Thus

$$\frac{N_t^{\text{un}}}{B_t^{\text{un}}(p^* - \bar{q}_t^B)} = \frac{e^{-\mu t}}{-(1-p^*)(1-e^{-\mu t})p^* \frac{\mu}{\Lambda-\mu} \frac{e^{-\Lambda t} - e^{-\mu t}}{1-e^{-\mu t}}} = -\frac{\Lambda - \mu}{\dot{y}_0(e^{-(\Lambda-\mu)t} - 1)}. \qquad \text{(OA.34)}$$

If $\Lambda > \mu$, we have $L = \frac{\Lambda-\mu}{\dot{y}_0}$. If $\Lambda < \mu$, $L = 0$. For the knife-edge case $\Lambda = \mu$, we have $\frac{N_t^{\text{un}}}{B_t^{\text{un}}(p^* - \bar{q}_t^B)} = \frac{1}{\dot{y}_0 t}$, so again $L = 0$. Hence, in general, $L = \frac{\max\{\Lambda-\mu,0\}}{\dot{y}_0}$.

Retracing our steps, we have $\lim_{r\to\infty} Y(r) = -\frac{(\dot{y}_0)^2}{p^*} + (\dot{y}_0)^2 \frac{\max\{\Lambda-\mu,0\}}{\dot{y}_0} - 2\Lambda\dot{y}_0$, where

$$\begin{aligned}\lim_{r\to\infty}(X(r) - Y(r)) &= \dot{y}_0(\mu(2p^*-1) - \Lambda) - \left[-\frac{(\dot{y}_0)^2}{p^*} + \dot{y}_0 \max\{\Lambda-\mu, 0\} - 2\Lambda\dot{y}_0\right] \\ &= \dot{y}_0\left[\mu p^* + \Lambda - \max\{\Lambda-\mu,0\}\right] > 0.\end{aligned}$$

□

3 Proof of Proposition 6

For arbitrary $p_0 \in (0,1)$, suppose a FIFO equilibrium with $T > 0$ exists. We derive necessary and sufficient conditions for its existence and show that when it exists, it is unique, and D has a closed-form expression.

Since D is continuous and strictly increasing on $(0,T)$ with $D(t) > t$, it has an inverse on this interval $b = D^{-1}$ which is also strictly increasing and which satisfies $b(t) < t$ for all $t \in (0,T)$.

We claim that local indifference must hold at all times in $[0,T)$. Let $F(t)$ denote the payoff for a sender with time-0 good evidence from disclosing at t. Since each type $t > 0$ could disclose at any $s \in [t, D(t))$ but is willing to wait until $D(t) > t$, and timestamps are payoff irrelevant for the sender, we have $F(D(t)) \geq F(s)$. But at each time $s \in [t, D(t))$, there is a type $b(s) < s$ who chooses to disclose at s instead of $D(t)$; therefore $F(s) \geq F(D(t))$ for all $s \in [t, D(t))$. Hence, F is constant on $(t, D(t))$. Since these intervals cover $(0,T)$, there is indifference over the whole interval, and this extends to time 0 by continuity. In particular, $\int_0^t e^{-rs} p_s\, ds + e^{-rt}\frac{q_t}{\lambda_1+r}$ is constant (and hence differentiable). This implies q is differentiable and the familiar local indifference condition must hold on $(0,T)$:

$$\dot{q}_t = rq_t - (\lambda_1 + r)p_t.$$

Given $b(t)$, we calculate p_t and q_t as follows. We have

$$p_t = \frac{\Pr(\theta_t = 1, \sigma > t)}{\Pr(\sigma > t)} = \frac{\Pr(\theta_t = 1)\Pr(\sigma > t | \theta_t = 1)}{\Pr(\sigma > t)}.$$

Now $\Pr(\theta_t = 1) = \phi_t = p_0 e^{-\lambda_1 t}$. Also, $\theta_t = 1$ implies that $\theta_s = 1$ for all $s \leq t$, and given this, the probability of no disclosure by t is the probability that no news arrived by $b(t)$: $e^{-\mu b(t)}$. Hence $\Pr(\theta_t = 1, \sigma > t) = p_0 e^{-(\lambda_1 t + \mu b(t))}$. In the denominator,

$$\Pr(\sigma > t) = 1 - \Pr(\sigma \leq t) = 1 - \int_0^{b(t)} \mu e^{-\mu s} \phi_s \, ds = 1 - \frac{\mu p_0}{\lambda_1 + \mu}(1 - e^{-(\mu+\lambda_1)b(t)}), \quad \text{(OA.35)}$$

and therefore

$$p_t = \frac{p_0 e^{-(\lambda_1 t + \mu b(t))}}{1 - \frac{\mu p_0}{\lambda_1 + \mu}(1 - e^{-(\mu+\lambda_1)b(t)})}.$$

From the FIFO definition that for all $t \in [0, T)$, we have

$$q_t = e^{-\lambda_1 (t - b(t))}, \quad \text{(OA.36)}$$

reflecting that $\theta_{b(t)} = 1$ and that the probability that the state is good decays due to Markov switching until t. Since q is differentiable on $(0, T)$, $b(t) = \frac{\log q(t)}{\lambda_1} + t$ is also differentiable on $(0, T)$. This also implies

$$\dot{q}_t = -\lambda_1 (1 - b'(t)) e^{-\lambda_1 (t - b(t))}. \quad \text{(OA.37)}$$

Using these expressions, and multiplying through by $e^{\lambda_1 (t - b(t))}$, the indifference condition becomes

$$\begin{aligned}
-\lambda_1(1 - b'(t)) &= r - (\lambda_1 + r) \frac{p_0 e^{-(\mu+\lambda_1)b(t)}}{1 - \frac{\mu p_0}{\lambda_1+\mu}\left(1 - e^{-(\mu+\lambda_1)b(t)}\right)} \\
\iff \qquad b'(t) &= \frac{\lambda_1 + r}{\lambda_1}\left[1 - \frac{p_0 e^{-(\mu+\lambda_1)b(t)}}{1 - \frac{\mu p_0}{\lambda_1+\mu}\left(1 - e^{-(\mu+\lambda_1)b(t)}\right)}\right] \\
&= \frac{\lambda_1 + r}{\lambda_1}\left[\frac{(\lambda_1 + \mu(1 - p_0))e^{(\mu+\lambda_1)b(t)} - \lambda_1 p_0}{(\lambda_1 + \mu(1 - p_0))e^{(\mu+\lambda_1)b(t)} + \mu p_0}\right].
\end{aligned}$$

Changing variables with $t = D(s)$, we have $b'(t) = \frac{1}{D'(s)}$ and $b(t) = s$, and thus

$$D'(s) = \frac{\lambda_1}{\lambda_1 + r}\left[\frac{(\lambda_1 + \mu(1 - p_0))e^{(\mu+\lambda_1)s} + \mu p_0}{(\lambda_1 + \mu(1 - p_0))e^{(\mu+\lambda_1)s} - \lambda_1 p_0}\right].$$

Integrating and using $D(0) = 0$ yields the closed-form expression

$$D(t) = \frac{1}{\lambda_1 + r}\left[-\mu t + \log\left(\frac{e^{(\lambda_1+\mu)t}(\lambda_1 + \mu(1-p_0)) - \lambda_1 p_0}{(\lambda_1+\mu)(1-p_0)}\right)\right].$$

It is easy to verify that $D(0) = 0$. Moreover, D is strictly concave and has $\lim_{t\to\infty} D'(t) = \frac{\lambda_1}{\lambda_1+r} < 1$, so there exists $T > 0$ with $D(T) = T$ if and only if $D'(0) > 1$, and when such T exists, it is unique; we denote it by T^F. This condition is equivalent to $p_0 > \frac{r}{\lambda_1+r} = \underline{p}$.

To verify the equilibrium, recall that it is strictly optimal not to disclose bad evidence. For good evidence, note that by construction the sender is indifferent over disclosure at all times in $[0, T^F]$. Thus, we only need to show that immediate disclosure is optimal for all $t \geq T^F$. This is equivalent to $p_t \leq \underline{p}$. Now since $D'(T^F-) < 1$ by the strict concavity and single-crossing, we have $b'(T^F-) > 1$, and then (OA.37) yields $0 < \dot{q}_{T^F-} = r(1) - (\lambda_1 + r)p_{T^F}$ by local indifference; this is equivalent to $p_{T^F} < \underline{p}$. We also have $\dot{p}_t < 0$ for all $t \geq T^F$ under transparency, so $p_t < \underline{p}$ holds for all $t \geq T^F$ as desired.

Let T^F denote the start of the transparency phase in the FIFO equilibrium, and recall that $\overline{T}$ denotes the start of the transparency phase in the three-phase equilibrium. We now turn to the comparison of T^F and $\overline{T}$.

In either equilibrium, before transparency starts, define

$$g_t := \dot{q}_t - [rq_t - (\lambda_1 + r)p_t], \tag{OA.38}$$

which is the difference between the left and right hand sides of the local IC condition. Note that (i) in the three-phase equilibrium $g_t > 0$ for $t \in [0, \underline{T})$ and $g_t = 0$ for $t \in [\underline{T}, \overline{T})$ and (ii) in the FIFO equilibrium $g_t = 0$ for all $t \in [0, T^F)$. By definition, we have $\dot{q}_t = [rq_t - (\lambda_1 + r)p_t] + g_t$.

In either case, let ρ_t be the hazard rate of good evidence disclosure from the receiver's perspective, conditional on no disclosure before t. Then by Bayes' rule,

$$\dot{p}_t = -\lambda_1 p_t - \rho_t(q_t - p_t). \tag{OA.39}$$

Subtracting this from the previous equation for $\dot{q}_t$, and using $\delta := q - p$ as in the proof of Theorem 2, we have

$$\dot{\delta}_t = \dot{q}_t - \dot{p}_t = (r + \rho_t)\delta_t + g_t, \tag{OA.40}$$

which generalizes (10). This holds until the start of transparency in either equilibrium. Now using $N_t := \Pr(\sigma > t)$, we have $\dot{N}_t = -\rho_t N_t$. Hence,

$$\frac{d}{dt}(N_t\delta_t) = -\rho_t N_t \delta_t + N_t \dot{\delta}_t = N_t[r\delta_t + g_t].$$

This implies

$$\frac{d}{dt}(e^{-rt}N_t\delta_t) = -re^{-rt}N_t\delta_t + e^{-rt}N_t[r\delta_t + g_t] = e^{-rt}N_t g_t. \tag{OA.41}$$

Letting $Y_t := e^{-rt}N_t\delta_t$, and recalling the properties of g_t noted above, we have (i) $\dot{Y}_t > 0$ for $t \in [0, \underline{T})$ and $\dot{Y}_t = 0$ for $t \in [\underline{T}, \overline{T})$ in the three-phase equilibrium, and (ii) $\dot{Y}_t = 0$ for all $t \in [0, T^F)$ in the FIFO equilibrium.

In the FIFO equilibrium, therefore, we have $Y_0 = 1 - p_0 = Y_{T^F} = e^{-rT^F}N_{T^F}(1 - p_{T^F})$ since $q_{T^F} = 1$. In the three-phase equilibrium, we have $1 - p_0 \le Y_{\underline{T}} = Y_{\overline{T}} = e^{-r\overline{T}}N_{\overline{T}}(1 - p_{\overline{T}})$, with equality if and only if $\underline{T} = 0$, i.e. $p_0 \in (\underline{p}, \overline{p}]$, where again, we have used $q_{\overline{T}} = 1$.

We now argue that there is a function $G(T)$, independent of the equilibrium, such that in the FIFO equilibrium $Y_{T^F} = G(T^F)$ and in the three-phase equilibrium $Y_{\overline{T}} = G(\overline{T})$. In either equilibrium, let T denote the start of transparency. Since all good evidence that has arrived before T has been disclosed before T, the same calculation as in (OA.35) implies $N_T = 1 - \frac{\mu p_0}{\lambda_1 + \mu}(1 - e^{-(\mu+\lambda_1)T})$, which depends only on time. Also, $N_T p_T = \Pr(\theta_T = 1, \sigma > T) = \Pr(\theta_T = 1)\Pr(\sigma > T|\theta_T = 1) = p_0 e^{-(\lambda_1+\mu)T}$. The last equality uses that the bad state is absorbing: $\theta_T = 1$ implies $\theta_t = 1$ for all $t \in [0, T)$, and conditional on this, disclosure occurs by T if and only if evidence arrives by T. Therefore, $\Pr(\sigma > T|\theta_T = 1) = \Pr(\tau > T) = e^{-\mu T}$. Hence, define

$$\begin{aligned} G(T) &:= e^{-rT}N_T(1 - p_T) = e^{-rT}(N_T - N_T p_T) \\ &\qquad\qquad = e^{-rT}\left[\frac{\lambda_1 + \mu(1 - p_0)}{\lambda_1 + \mu} - \frac{p_0\lambda_1}{\lambda_1 + \mu}e^{-(\lambda_1+\mu)T}\right], \end{aligned}$$

which has derivative

$$G'(T) = e^{-rT}\left[-r\frac{\lambda_1 + \mu(1 - p_0)}{\lambda_1 + \mu} + \lambda_1 p_0 \frac{\lambda_1 + \mu + r}{\lambda_1 + \mu}e^{-(\lambda_1+\mu)T}\right].$$

By construction, in each respective equilibrium, $Y_{T^F} = G(T^F)$ and $Y_{\overline{T}} = G(\overline{T})$, and $G(0) = 1 - p_0 = G(T^F)$. It is also straightforward to show that whenever $p_0 > \underline{p}$, we have (i) $G'(0) > 0$, (ii) G' is eventually negative, and (iii) G' changes sign exactly once on $(0, \infty)$. Hence $G(T') > 1 - p_0$ for all $T' \in (0, T^F)$ and $G(T') < 1 - p_0$ for all $T' > T^F$. We have already shown that when $p_0 \in (\underline{p}, \overline{p}]$, we have $Y_{T^F} = Y_{\overline{T}}$, so $G(T^F) = G(\overline{T})$ and therefore $T^F = \overline{T}$.

We also have already shown that when $p_0 > \overline{p}$, we have $G(\overline{T}) = Y_{\overline{T}} > 1 - p_0 = Y_{T^F} = G(T^F)$, and therefore $\overline{T} < T^F$.

4 Receiver welfare

Here, we formalize the receiver welfare comparison in Section 5.1 involving age-independent bad evidence disclosure. Assume the (representative) receiver has the quadratic loss preferences specified in Section 5.1. We prove the following:

Proposition OA.1. *Assume $\lambda_0 > 0$. For sufficiently large $r > 0$, any equilibrium in the no-timestamps environment with immediate disclosure of good evidence and age-independent disclosure of bad evidence gives the receiver a lower expected payoff than the unique equilibrium in the timestamps environment.*

Proof. In any environment and equilibrium, the receiver's expected loss is

$$\begin{aligned}\mathbb{E}\int_0^\infty e^{-\rho t}(\pi_t - \theta_t)^2\,dt &= \int_0^\infty e^{-\rho t}\mathbb{E}[(\pi_t - \theta_t)^2]\,dt \\ &= \int_0^\infty e^{-\rho t}(\mathbb{E}[\theta_t^2] - \mathbb{E}[\pi_t^2])\,dt.\end{aligned}$$

To distinguish between environments, let π^{TS} denote the belief process in the unique timestamps equilibrium.

Let $T > 0$ be arbitrary. By Proposition 5, if r is sufficiently large, the receiver learns only from good evidence disclosure on $[0, T]$. Thus, for each T, define $\pi^{\text{NTS},T}$ as follows. For $t \in [0, T]$, define $\pi_t^{\text{NTS,T}}$ as the belief given immediate good evidence disclosure and no bad evidence disclosure on $[0, T]$. For $t > T$, define $\pi_t^{\text{NTS,T}} = \theta_t$, i.e. assuming perfect knowledge of the state. For sufficiently large r, without timestamps, the receiver's expected loss in any equilibrium in the given class is bounded below by $\int_0^\infty e^{-\rho t}(\mathbb{E}[\theta_t^2] - \mathbb{E}[(\pi_t^{\text{NTS,T}})^2])\,dt = \int_0^T e^{-\rho t}(\mathbb{E}[\theta_t^2] - \mathbb{E}[(\pi_t^{\text{NTS,T}})^2])\,dt$.

If $T > a^*(0)$, then for all $t \in (a^*(0), T])$, there is a positive probability of disclosure of bad evidence in the timestamps equilibrium, so the receiver's information in the unique timestamps equilibrium Blackwell dominates that under the strategy of immediate good evidence disclosure and no bad evidence disclosure. Moreover, $\pi_t^{\text{TS}} = \pi_t^{\text{NTS,T}}$ for all $t \in [0, a^*(0))$. Hence, comparing losses on $[0, T]$, we have

$$H(T) := \int_0^T e^{-\rho t}(\mathbb{E}[\theta_t^2] - \mathbb{E}[(\pi_t^{\text{NTS,T}})^2])\,dt - \int_0^T e^{-\rho t}(\mathbb{E}[\theta_t^2] - \mathbb{E}[(\pi_t^{\text{TS}})^2])\,dt > 0$$

and $H(T)$ is nondecreasing on $(a^*(0), \infty)$.

Meanwhile, $\int_T^\infty e^{-\rho t}(\mathbb{E}[\theta_t^2] - \mathbb{E}[(\pi_t^{\mathrm{TS}})^2])\, dt \leq \frac{e^{-\rho T}}{4\rho}$, since the quadratic loss is at most $\frac{1}{4}$. Hence, fixing any $T_0 > a^*(0)$, by choosing T sufficiently large that $\frac{e^{-\rho T}}{4\rho} < H(T_0)$, we have $\frac{e^{-\rho T}}{4\rho} < H(T)$ for all $T \geq T_0$, and thus

$$\int_0^\infty e^{-\rho t}(\mathbb{E}[\theta_t^2] - \mathbb{E}[(\pi_t^{\mathrm{TS}})^2])\, dt < \int_0^\infty e^{-\rho t}(\mathbb{E}[\theta_t^2] - \mathbb{E}[(\pi_t^{\mathrm{NTS,T}})^2])\, dt$$

for all $T > T_0$. By taking r sufficiently large for the purpose of Proposition 5, the result follows. □

5 Multiple evidence arrivals

In this section, we show that Proposition 3 extends to an environment with unlimited evidence arrivals. Suppose each piece of evidence continues to arrive at rate μ, and for simplicity, assume $\lambda_0 = 0$. The definitions of strategy, sequential rationality, belief system, and equilibrium can be extended to this setting.

Below, we show that there exists $\underline{p} < 1$ such that if $p_0 > \underline{p}$, then in any equilibrium there exists $t > 0$ such that there is a positive probability the sender possesses undisclosed good evidence. Due to $\lambda_0 = 0$, bad evidence is not disclosed.

Suppose there exists an equilibrium with no such t. Then almost surely, the sender discloses good evidence immediately. When the sender delays (going off the equilibrium path), his private belief diverges from the receiver's belief. Let $F(\pi, \tilde{\pi})$ denote the sender's value function immediately after disclosure of the first piece of good evidence when the receiver's belief is π, the sender's private belief is $\tilde{\pi}$, and the sender does not possess any remaining evidence. Note that although future pieces of evidence can arrive, under the properties that good evidence is disclosed immediately almost surely and bad evidence is never disclosed, $F(\pi, \tilde{\pi})$ is pinned down uniquely.

A useful identity is that for each $\hat{\pi} \in [0, 1]$

$$F(1, \hat{\pi}) = \hat{\pi} F(1, 1) + (1 - \hat{\pi}) F(1, 0). \tag{OA.42}$$

This follows from the law of iterated expectations: the sender's expectation of future receiver beliefs equals the sender's expectation, averaging over the current state, of those beliefs conditional on the current state being good or bad.

To establish a contradiction, first suppose the sender gets good evidence at time 0. Disclosing at time 0, the sender's expected payoff is simply

$$F(1,1) = \mathbb{E}\left[\int_0^\infty e^{-rt}\pi_t\,dt|\theta_0 = 1\right] = \mathbb{E}\left[\int_0^\infty e^{-rt}\theta_t\,dt|\theta_0 = 1\right] = \frac{1}{\lambda_1 + r} \tag{OA.43}$$

by the law of iterated expectations.

Next suppose the sender plans to disclose at time ε. For all ε except on a set of measure zero, the receiver believes the evidence disclosed at ε is fresh and updates to 1, but the sender's private belief and continuation play depend on the private history in $(0,\varepsilon)$. In particular, in $(0,\varepsilon)$, the sender can (i) get no additional evidence, (ii) get exactly one more piece of good evidence, (iii) get at least one piece of bad evidence or two or more pieces of good evidence. In case (i), the sender's private belief is just $e^{-\lambda_1\varepsilon}$ accounting for the possibility of unobserved switching since time 0. Let us specify that in this case, after disclosing at ε, the sender reverts to the equilibrium strategy of immediate good evidence disclosure. The continuation payoff is therefore $F(1, e^{-\lambda_1\varepsilon})$. In case (ii), suppose the sender never (for the remainder of the game) discloses the piece of evidence that arrives in $(0,\varepsilon)$, but discloses good evidence arriving after ε immediately. This yields a continuation payoff of $F(1, e^{-\lambda_1(\varepsilon-\tau)})$ where $\tau \in (0,\varepsilon)$ is the time of the second arrival, because the sender's private belief only decays since the most recent arrival. Moreover, $F(1, e^{-\lambda_1(\varepsilon-\tau)}) = F(1, e^{-\lambda_1\varepsilon}) + O(\varepsilon)$. Since the probability of case (ii) is also $O(\varepsilon)$, replacing this case with the payoff from case (i) is an approximation within $O(\varepsilon^2)$. Case (iii) has probability $O(\varepsilon^2)$ since it requires multiple arrivals and/or a Markov switch in the state before an arrival. Overall, the expected payoff of delaying until time ε (and disclosing immediately any good evidence that arrives after time ε) is

$$p_0\varepsilon + (1 - r\varepsilon)F(1, e^{-\lambda_1\varepsilon}) + o(\varepsilon). \tag{OA.44}$$

By the identity (OA.42)

$$F(1, e^{-\lambda_1\varepsilon}) = e^{-\lambda_1\varepsilon}F(1,1) + (1 - e^{-\lambda_1\varepsilon})F(1,0).$$

Hence, (OA.44) equals

$$p_0\varepsilon + (1 - r\varepsilon)[e^{-\lambda_1\varepsilon}F(1,1) + (1 - e^{-\lambda_1\varepsilon})F(1,0)] + o(\varepsilon). \tag{OA.45}$$

The payoff gain from delay is

$$\varepsilon[p_0 - (\lambda_1 + r)F(1,1) + \lambda_1 F(1,0)] + o(\varepsilon),$$

obtained by subtracting (OA.43) from (OA.45). Delay is therefore optimal if ε is sufficiently small and

$$p_0 > \underline{p} := (\lambda_1 + r)F(1,1) - \lambda_1 F(1,0).$$

Since $F(1,1) = \frac{1}{\lambda_1 + r}$ and $F(1,0)$ is positive, $\underline{p} < 1$. We conclude by continuity. Specifically, there exists $\varepsilon > 0$ such that for all sufficiently small $t > 0$, if the sender gets the first piece of evidence at time t, delay until ε is optimal. This contradicts the fact that almost surely the sender discloses good evidence immediately.

Note that in our baseline model with one piece of evidence, the receiver's belief after disclosure depends only on the Markov switching dynamics and is therefore independent of the sender's private belief; in this case, $F(1,0) = F(1,1)$, so the equation reduces to $p_0 > rF(1,1) = \frac{r}{\lambda_1 + r}$, reproducing the cutoff in Proposition 3.

References

Madrigal, Vicente, Tommy CC Tan, and Sérgio Ribeiro da Costa Werlang. 1987. "Support restrictions and sequential equilibria." *Journal of Economic Theory* 43 (2):329–334.

Teschl, Gerald. 2012. *Ordinary differential equations and dynamical systems*, vol. 140. American Mathematical Soc.