Online Appendix for "Information Nudges and Self-Control"

Thomas Mariotti* Nikolaus Schweizer†
Nora Szech‡ Jonas von Wangenheim§

March 5, 2022

Proofs for Section 4

Proof of Corollary 1. We focus with no loss of generality on the case where (8) holds neither under $\overline{\mathbf{P}}$ nor under $\underline{\mathbf{P}}$. We need to show that $\underline{F}(\underline{t}^c) \geq \overline{F}(\bar{t}^c)$. For this we show that the unique $\hat{t}$ defined by $\underline{F}(\underline{t}^c) = \overline{F}(\hat{t})$ satisfies $\hat{t} \geq \bar{t}^c$. We have

$$\begin{aligned}
\int_{\hat{t}}^1 (\theta - t^a)\, \overline{\mathbf{P}}(\mathrm{d}\theta) &= \int_{\overline{F}^{-1}(\underline{F}(\underline{t}^c))}^1 (\theta - t^a)\, \overline{\mathbf{P}}(\mathrm{d}\theta) \\
&= \int_{\underline{F}(\underline{t}^c)}^1 [\overline{F}^{-1}(p) - t^a]\, \mathrm{d}p \\
&\geq \int_{\underline{F}(\underline{t}^c)}^1 [\underline{F}^{-1}(p) - t^a]\, \mathrm{d}p \\
&= \int_{\underline{t}^c}^1 (\theta - t^a)\, \underline{\mathbf{P}}(\mathrm{d}\theta) \\
&= 0 \\
&= \int_{\bar{t}^c}^1 (\theta - t^a)\, \overline{\mathbf{P}}(\mathrm{d}\theta),
\end{aligned}$$

where the inequality follows from Shaked and Shanthikumar (2007, Section 4.A.1), and the last two equalities follow from (9). If $\hat{t} \geq t^a$, then a fortiori $\hat{t} > \bar{t}^c$. Otherwise, $\max\{\bar{t}^c, \hat{t}\} < t^a$ implies that $\theta - t^a < 0$ for θ between $\bar{t}^c$ and $\hat{t}$, so that $\hat{t} \geq \bar{t}^c$ from the above inequality. Hence the result. ■

Proof of Corollary 2. For future reference, we more generally show the result for any right-truncation $\mathbf{P}_b \equiv \mathbf{P}[\cdot \,|\, \theta \leq b]$ of $\mathbf{P}$, with cdf F_b and density f_b over $[0, b]$, where $\frac{1}{\delta^2 C} < b \leq 1$. Corollary 2 corresponds to the special case $b = 1$. A first observation is that MHRP is preserved

*Toulouse School of Economics, CNRS, University of Toulouse Capitole, Toulouse, France, CEPR, and CESifo. Email: thomas.mariotti@tse-fr.eu.

†Department of Econometrics and Operations Research, Tilburg University, Tilburg, The Netherlands. E-mail: n.f.f.schweizer@uvt.nl.

‡Karlsruhe Institute of Technology (KIT), ECON Institute, Karlsruhe, Germany, Berlin Social Science Center (WZB), and CESifo. Email: nora.szech@kit.edu.

§Institute for Microeconomics, University of Bonn, Bonn, Germany. Email: jwangenheim@uni-bonn.de.

by right-truncation.

Lemma A.1 *If* $\mathbf{P}$ *satisfies MHRP, then so does* $\mathbf{P}_b$ *for any* $b \in (0,1)$.

Proof. For each $t \in [0,b)$, we have

$$r_b(t) \equiv \frac{f_b(t)}{1-F_b(t)} \propto \frac{f(t)}{F(b)-F(t)} = r(t)\,\frac{1-F(t)}{F(b)-F(t)},$$

so that $r_b(t)$ is the product of two strictly positive and strictly increasing functions of t. The result follows. ■

Now, fix some $b \in (\frac{1}{\delta^2 C}, 1)$ and, for each $\beta \in (\frac{1}{b\delta^2 C}, 1)$, define

$$\phi_b(\beta) \equiv \mathbf{E}_b\left[\theta \,|\, \theta \geq \frac{1+\beta\delta}{(1+\delta)\beta\delta^2 C}\right] - \frac{1}{\beta\delta^2 C}. \tag{A.4}$$

We show that there exists a unique solution β_b^u to $\phi_b(\beta) = 0$ and that $\phi_b(\beta) \geq 0$ if and only if $\beta \geq \beta_b^u$. This, in particular, implies Corollary 2, with $\beta^u \equiv \beta_1^u$. Because f is continuous, so is ϕ_b. Hence, by the intermediate value theorem, we only need to check that $\phi_b(\frac{1}{b\delta^2 C}) < 0$, that $\phi_b(1) > 0$, and that ϕ_b is strictly increasing. As for the first two statements, we have

$$\phi_b\left(\frac{1}{b\delta^2 C}\right) = \mathbf{E}_b\left[\theta \,|\, \theta \geq \frac{1+b\delta C}{(1+\delta)\delta C}\right] - b \quad \text{and} \quad \phi_b(1) = \mathbf{E}_b\left[\theta \,|\, \theta \geq \frac{1}{\delta^2 C}\right] - \frac{1}{\delta^2 C},$$

and the result follows from $b\delta^2 C > 1$ and the fact that $\mathbf{P}_b$ has full support over $[0,b]$. As for the third statement, notice that, letting $\xi \equiv \frac{1}{\beta\delta^2 C}$ and changing variables accordingly, it is equivalent to the claim that

$$\mathbf{E}_b\left[\theta \,|\, \theta \geq \frac{\xi + \frac{1}{\delta C}}{1+\delta}\right] - \xi$$

is strictly decreasing in $\xi \in (\frac{1}{\delta^2 C}, b)$. A classical result from reliability theory (see, e.g., Bryson and Siddiqui 1969) states that, for a distribution that satisfies MHRP, the mean residual life is strictly decreasing in the age. By Lemma A.1, such is the case of $\mathbf{P}_b$, and thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{E}_b[\theta \,|\, \theta \geq t] < 1$$

for all $b \in (0,1)$ and $t \in [0,b)$. It follows that

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\,\mathbf{E}_b\left[\theta \,|\, \theta \geq \frac{\xi + \frac{1}{\delta C}}{1+\delta}\right] < \frac{1}{1+\delta} < 1$$

for all $\xi \in (\frac{1}{\delta^2 C}, b)$. Hence the result. ■

Proof of Corollary 3. By (3)–(4) and (9), the probability of harmful consumption is

$$F(t^c) - F(t^h) = F(t^c) - F\left(\frac{\mathbf{E}[\theta \,|\, \theta \geq t^c] + \frac{1}{\delta C}}{1+\delta}\right).$$

As observed in the main text, t^c is strictly decreasing in $\beta \in (\frac{1}{\delta^2 C}, \beta^u)$. Hence it is sufficient to show that

$$H(t) \equiv F(t) - F\left(\frac{\mathbf{E}[\theta \,|\, \theta \geq t] + \frac{1}{\delta C}}{1+\delta}\right) \tag{A.5}$$

is strictly increasing in $t \in (t^u, 1)$, where

$$t^u \equiv \frac{1+\beta^u \delta}{(1+\delta)\beta^u \delta^2 C}.$$

Notice for future reference that, for each $t \in (t^u, 1)$,

$$t > \frac{\mathbf{E}[\theta \,|\, \theta \geq t] + \frac{1}{\delta C}}{1+\delta} \tag{A.6}$$

because, as β^u is the unique value of $\beta \in (\frac{1}{\delta^2 C}, 1)$ that achieves equality in (10), (A.6) becomes an equality at $t = t^u$ and because, as $\mathbf{P}$ satisfies MHRP, the mapping $t \mapsto (1+\delta)t - \mathbf{E}[\theta \,|\, \theta \geq t]$ is strictly increasing over $[0, 1)$. Then, for each $t \in (t^u, 1)$,

$$\begin{aligned} H'(t) &= f(t) - \frac{1}{1+\delta} f\left(\frac{\mathbf{E}[\theta \,|\, \theta \geq t] + \frac{1}{\delta C}}{1+\delta}\right) \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{E}[\theta \,|\, \theta \geq t] \\ &\geq f(t) - \frac{1}{1+\delta} f\left(\frac{\mathbf{E}[\theta \,|\, \theta \geq t] + \frac{1}{\delta C}}{1+\delta}\right) \\ &> 0, \end{aligned} \tag{A.7}$$

where the first inequality again follows from MHRP, and the second inequality follows from (11) and (A.6). Hence the result. ■

Corollary A.1 *If* $\mathbf{P}$ *satisfies MHRP and* f *is nonincreasing in a left-neighborhood of* $t = 1$ *or strictly positive at* $t = 1$*, and if*

$$f(1) < \frac{1}{2(1+\delta)} f\left(\frac{1 + \frac{1}{\delta C}}{1+\delta}\right), \tag{A.8}$$

then the probability $F(t^c) - F(t^h)$ *that harmful consumption takes place under the optimal IC mechanism is strictly increasing in* β *in a right-neighborhood of* $\beta = \frac{1}{\delta^2 C}$*.*

Proof. Defining H as in (A.5), we have

$$\frac{\mathrm{d}}{\mathrm{d}\beta} [F(t^c) - F(t^h)] > 0$$

in a strict right-neighborhood of $\beta = \frac{1}{\delta^2 C}$ if and only if $H' < 0$ in a strict left-neighborhood of $t = 1$ or, equivalently, by (A.7),

$$f(1) - \frac{1}{1+\delta} f\left(\frac{1 + \frac{1}{\delta C}}{1+\delta}\right) \liminf_{t \to 1^-} \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{E}[\theta \,|\, \theta \geq t] < 0. \tag{A.9}$$

We need to show that (A.8) implies (A.9) if $f(1) > 0$ or, if $f(1) = 0$, if f is nonincreasing in a left-neighborhood of $t = 1$.[1] That is, we need to show that, under these assumptions,

$$\liminf_{t \to 1^-} \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{E}[\theta \,|\, \theta \geq t] \geq \frac{1}{2}.$$

Suppose, by way of contradiction, that there exists a sequence $(t_n)_{n \in \mathbb{N}}$ in $(0, 1)$ converging to 1 such that, for some $\varepsilon > 0$,

$$\left. \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{E}[\theta \,|\, \theta \geq t] \right|_{t = t_n} < \frac{1 - \varepsilon}{2}$$

for all n. Then, because

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbf{E}[\theta \,|\, \theta \geq t] = \frac{f(t)}{1 - F(t)} \{\mathbf{E}[\theta \,|\, \theta \geq t] - t\},$$

we have

$$f(t_n) \left\{ \int_{t_n}^1 \theta f(\theta) \, \mathrm{d}\theta - t_n[1 - F(t_n)] \right\} - \frac{1 - \varepsilon}{2} [1 - F(t_n)]^2 < 0 \tag{A.10}$$

for all n. Consider then the function

$$I(t) \equiv f(t) \left\{ \int_t^1 \theta f(\theta) \, \mathrm{d}\theta - t[1 - F(t)] \right\} - \frac{1 - \varepsilon}{2} [1 - F(t)]^2.$$

We clearly have $I(1) = 0$. We now show that, under the stated assumptions on f, I is strictly decreasing in a left-neighborhood of $t = 1$, which, given (A.10), yields the desired contradiction as the sequence $(t_n)_{n \in \mathbb{N}}$ converges to 1. As I is continuous, it is sufficient to show that its right upper Dini derivative $\mathrm{D}^+ I$ is strictly negative in a strict left-neighborhood of $t = 1$ (Giorgi and Komlósi 1992, Theorem 1.14). By continuity of f, the mapping $t \mapsto \int_t^1 \theta f(\theta) \, \mathrm{d}\theta - t[1 - F(t)]$ is continuously differentiable. A simple calculation then shows that, for each $t \in (0, 1)$,

$$\mathrm{D}^+ I(t) = [1 - F(t)](\mathrm{D}^+ f(t) \{\mathbf{E}[\theta \,|\, \theta \geq t] - t\} - \varepsilon f(t)).$$

Now, recall that, by assumption, f is strictly positive over $(0, 1)$. Thus, if $f(1) > 0$, then $\mathrm{D}^+ I$ is strictly negative in a strict left-neighborhood of $t = 1$ because the mean residual life $\mathbf{E}[\theta \,|\, \theta \geq t] - t$ converges to zero as t goes to 1; similarly, if $f(1) = 0$, then, because the mean residual life $\mathbf{E}[\theta \mid \theta \geq t] - t$ is strictly positive for all $t \in [0, 1)$, the same conclusion obtains if f is nonincreasing, so that its right upper Dini derivative $\mathrm{D}^+ f$ is nonpositive in a strict left-neighborhood of $t = 1$. Hence the result. ■

The somewhat convoluted condition (A.8) can be interpreted as follows. If initially $\beta \approx \frac{1}{\delta^2 C}$, then nearly all types consume under the optimal IC mechanism, that is, $t^c \approx 1$. If β increases

[1] Notice that, in the latter case, condition (A.8) is automatically satisfied.

by $\mathrm{d}\beta$, then t^c decreases by some amount $\mathrm{d}t^c$, so that a mass of types approximately equal to $f(1)\,\mathrm{d}t^c$ can be neutralized. At the same time,

$$t^h = \frac{\mathbf{E}[\theta\,|\,\theta \geq t^c] + \frac{1}{\delta C}}{1+\delta} \approx \frac{1 + \frac{1}{\delta C}}{1+\delta}$$

decreases by an amount

$$\mathrm{d}t^h \approx \frac{\mathrm{d}t^c}{1+\delta}\,\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{E}[\theta\,|\,\theta \geq t]\bigg|_{t=1^-} \geq \frac{\mathrm{d}t^c}{2(1+\delta)}$$

under the weak conditions we impose on f.[2] The mass of new types thus trapped in harmful consumption is bounded below by $\frac{1}{2(1+\delta)}\,f(t^h)\,\mathrm{d}t^c$, which exceeds the mass $f(1)\,\mathrm{d}t^c$ of neutralized types if condition (A.8) is satisfied. As a result, the probability of the harmful-consumption trap (t^h, t^c) locally increases in β. Notice that, because f is assumed to be strictly positive over $(0,1)$, condition (A.8) is satisfied as soon as $f(1) = 0$.

Proofs for Section 5

Proof of Proposition 2. For each $\beta_H \in (\beta_L, 1)$, we denote by $t^a(\beta_H)$, $t^h(\beta_H)$, $t^c(\beta_H)$, and $t^*(\beta_H) \equiv \max\{t^h(\beta_H), t^c(\beta_H)\}$ the relevant cutoffs defined in Sections 2–3. It follows from (3)–(4) that t^a and t^h are continuous. As for t^c and t^*, notice that, for each $\beta_H \in (\beta_L, 1)$, the assumption that $\mathbf{P}$ has a continuous density f allows us to rewrite (9) as

$$\frac{\int_{t^c(\beta_H)}^{1} \theta f(\theta)\,\mathrm{d}\theta}{1 - F(t^c(\beta_H))} = \frac{1}{\beta_H \delta^2 C}, \tag{A.11}$$

which implies, using again the assumption that f is continuous, that t^c and t^* are continuous as well. Now, for each $\beta_H \in (\beta_L, 1)$, define

$$\varphi_{t_L^*}(\beta_H) \equiv \mathbf{E}[\theta\,|\,t^*(\beta_H) \leq \theta < t_L^*] - t^a(\beta_H) = \frac{\int_{t^*(\beta_H)}^{t_L^*} \theta f(\theta)\,\mathrm{d}\theta}{F(t_L^*) - F(t^*(\beta_H))} - \frac{1}{\beta_H \delta^2 C}. \tag{A.12}$$

Because f and t^* are continuous, so is $\varphi_{t_L^*}$. Hence, by the intermediate value theorem, we only need to check that $\varphi_{t_L^*}(\beta_L^+) < 0$, that $\varphi_{t_L^*}(1) > 0$, and that $\varphi_{t_L^*}$ crosses zero only once. As for the first two statements, we have

$$\varphi_{t_L^*}(\beta_L^+) = t_L^* - t^a(\beta_L) \quad \text{and} \quad \varphi_{t_L^*}(1) = \mathbf{E}[\theta\,|\,t^*(1) \leq \theta < t_L^*] - t^a(1),$$

and the result follows from $t_L^* < t_L^a = t^a(\beta_L)$, $t^*(1) = t^a(1) = t^h(1) < t_L^h \leq t_L^*$, and the fact that $\mathbf{P}$ has full support over $[0,1]$. As for the third statement, we distinguish two cases.

Case 1 If $\beta_L < \beta_H < \beta^u$, with β^u defined as in Corollary 2, then the unconstrained-optimal

[2] The intuition for the factor $\frac{1}{2}$ is easy to grasp when $f(1) > 0$. Indeed, in that case, the distribution of θ conditional on $\theta \geq t$ is approximately uniform when t is close to 1 as f is continuous, and, hence, a marginal increase $\mathrm{d}t$ in t increases $\mathbf{E}[\theta\,|\,\theta \geq t]$ by approximately $\mathrm{d}[\frac{1}{2}(t+1)] = \frac{1}{2}\,\mathrm{d}t$.

mechanism for type H is not IC and, therefore, $t^*(\beta_H) = t^c(\beta_H) > t^h(\beta_H)$. In this case, from (A.11)–(A.12), we have

$$\varphi_{t_L^*}(\beta_H) = \frac{\int_{t^c(\beta_H)}^{t_L^*} \theta f(\theta)\, \mathrm{d}\theta}{F(t_L^*) - F(t^c(\beta_H))} - \frac{\int_{t^c(\beta_H)}^{1} \theta f(\theta)\, \mathrm{d}\theta}{1 - F(t^c(\beta_H))} < 0$$

as $t_L^* < 1$ and $\mathbf{P}$ has full support over $[0, 1]$. It follows that $\varphi_{t_L^*}$ cannot cross zero over (β_L, β^u), and thus that the desired threshold cannot belong to this interval.

Case 2 If $\beta_H \geq \max\{\beta_L, \beta^u\}$, then the unconstrained-optimal mechanism for type H is IC and, therefore, $t^*(\beta_H) = t^h(\beta_H)$. In this case, we have

$$\varphi_{t_L^*}(\beta_H) = \mathbf{E}\left[\theta \,|\, \frac{1 + \beta_H \delta}{(1 + \delta)\beta_H \delta^2 C} \leq \theta < t_L^*\right] - \frac{1}{\beta_H \delta^2 C} = \phi_{t_L^*}(\beta_H),$$

where $\phi_{t_L^*}(\beta_H)$ is as defined in (A.4) with $b = t_L^*$. As shown in the proof of Corollary 2, because $\mathbf{P}$ satisfies MHRP, $\phi_{t_L^*}$ is strictly increasing and vanishes at a single point $\beta^u_{t_L^*}$, which defines the desired threshold $\beta_H^{ni}(\beta_L)$. That $\beta_H^{ni}(\beta_L) > \beta^u$ was shown in Case 1. That $\beta_H^{ni}(\beta_L)$ is strictly increasing in β_L follows from the fact that $t_L^* = t^*(\beta_L)$ and, thus, $\phi_{t_L^*}$ are strictly decreasing in β_L. Hence the result. ■

Proof of Proposition 3. A useful preliminary observation is that, because the mechanism designer always prefers a higher abstinence rate than the DM, we can, in analogy with the proof of Proposition 1, neglect constraints (17) and (19) in our quest for an optimal IC joint mechanism. That is, the following result holds.

Lemma A.2 *Any solution to the relaxed problem*

$$\max \left\{ \sum_{i=L,H} p_i \beta_i \{t_i^h \mathbf{E}[\Pi_i(\theta)] - \mathbf{E}[\theta \Pi_i(\theta)]\} : \pi \text{ satisfies (16) and (18)} \right\} \tag{A.13}$$

is a solution to problem (20).

Proof. We show that any solution to (A.13) satisfies (17) and (19), and thus is a solution to (20). We accordingly distinguish two cases.

Case 1 Suppose, by way of contradiction, that a solution $(\pi_{LH}, \pi_L, \pi_\emptyset)$ to (A.13) violates (17). Then type L would prefer to abstain whenever the recommendation is L. Because the utility from consumption is weakly lower for the mechanism designer than for type L, the former prefers that type L abstain in this case, and a fortiori that type H abstain as $t_H^a < t_L^a$. Therefore, the joint mechanism $(\pi_{LH}, 0, \pi_\emptyset + \pi_L)$ would satisfy (16) and (18) and improve upon the solution to (A.13), a contradiction.

Case 2 Suppose, by way of contradiction, that a solution $(\pi_{LH}, \pi_L, \pi_\emptyset)$ to (A.13) violates

(19). Then type H would prefer to abstain whenever the recommendation is LH. Because the utility from consumption is weakly lower for the mechanism designer than for type H, the former prefers that type H abstain in this case. Therefore, the joint mechanism $(0, \pi_L + \pi_{LH}, \pi_\emptyset)$ would satisfy (16) and (18) and improve upon the solution to (A.13), once again a contradiction. The result follows. ■

Among all joint mechanisms $(\pi_{LH}, \pi_L, \pi_\emptyset)$ that issue the recommendation LH with some probability γ_{LH}, those such that

$$\pi_{LH}(\theta) = 1_{\{\theta < t_{\gamma_{LH}}\}}$$

for $t_{\gamma_{LH}} \equiv F^{-1}(\gamma_{LH})$ are the best for efficiency purposes as they minimize the expected harm from consumption for a given probability of joint consumption. The following lemma shows that they are also best at satisfying the incentive constraints (16) and (18), as they issue recommendations to abstain to higher-risk types than any other joint mechanism that satisfies the same constraints and issues the same recommendations with the same probabilities.

Lemma A.3 *For any joint mechanism $\pi \equiv (\pi_{LH}, \pi_L, \pi_\emptyset)$ that satisfies (16) and (18), there exists a joint mechanism $\tilde{\pi} \equiv (\tilde{\pi}_{LH}, \tilde{\pi}_L, \tilde{\pi}_\emptyset)$ that also satisfies (16) and (18), and such that*

$$\mathbf{E}[\tilde{\pi}_j(\theta)] = \mathbf{E}[\pi_j(\theta)], \quad j = LH, L, \emptyset, \tag{A.14}$$

$$\tilde{\pi}_{LH}(\theta) = 1_{\{\theta < t_{\gamma_{LH}}\}} \tag{A.15}$$

for $\gamma_{LH} \equiv \mathbf{E}[\pi_{LH}(\theta)]$ and $t_{\gamma_{LH}} \equiv F^{-1}(\gamma_{LH})$. Moreover, $\tilde{\pi}$ achieves a weakly higher value in (A.13) than π, and strictly so if π does not satisfy (A.15) on a $\mathbf{P}$-nonnull set.

Proof. For the proof of this lemma, it is convenient to work with an equivalent formulation of a joint mechanism, due to Aumann (1964). Let $\Omega \equiv [0, 1]$ be a sample space, endowed with Lebesgue measure $\boldsymbol{\lambda}$ over the Borel sets. It follows from Aumann (1964, Lemma F) that, for every joint mechanism $\pi \equiv (\pi_{LH}, \pi_L, \pi_\emptyset)$, there exists a measurable direct mechanism $x : \Theta \times \Omega \to \{LH, L, \emptyset\}$ issuing, for every type $\theta \in \Theta$ and for every element $\omega \in \Omega$, a recommendation for both types to consume (LH), for only type L to consume (L), or for both types to abstain ($\emptyset$), and such that

$$\pi_j(\theta) \equiv \boldsymbol{\lambda}[\{\omega \in \Omega : x(\theta, \omega) = j\}], \quad (\theta, j) \in \Theta \times \{LH, L, \emptyset\}. \tag{A.16}$$

Conversely, for any measurable direct persuasion mechanism $x : \Theta \times \Omega \to \{LH, L, \emptyset\}$, (A.16) uniquely defines a joint mechanism $\pi \equiv (\pi_{LH}, \pi_L, \pi_\emptyset)$.

Now, let $x : \Theta \times \Omega \to \{LH, L, \emptyset\}$ be the direct mechanism associated to π, and, for each

$j \in \{LH, L, \emptyset\}$, let

$$\gamma_j(t_{\gamma_{LH}}) \equiv \mathbf{P} \otimes \boldsymbol{\lambda}[\{(\theta, \omega) \in \Theta \times \Omega : x(\theta, \omega) = j \wedge \theta < t_{\gamma_{LH}}\}]$$

be the probability that x issues recommendation j and $\theta < t_{\gamma_{LH}}$. Define a new direct joint persuasion mechanism

$$\tilde{x}(\theta, \omega) \equiv \begin{cases} LH & \text{if} \quad \theta < t_{\gamma_{LH}}, \\ L & \text{if} \quad \theta \geq t_{\gamma_{LH}} \wedge \left(x(\theta, \omega) = L \vee \left(x(\theta, \omega) = LH \wedge \omega < \frac{\gamma_L(t_{\gamma_{LH}})}{\gamma_L(t_{\gamma_{LH}}) + \gamma_\emptyset(t_{\gamma_{LH}})}\right)\right), \\ \emptyset & \text{if} \quad \theta \geq t_{\gamma_{LH}} \wedge \left(x(\theta, \omega) = \emptyset \vee \left(x(\theta, \omega) = LH \wedge \omega \geq \frac{\gamma_L(t_{\gamma_{LH}})}{\gamma_L(t_{\gamma_{LH}}) + \gamma_\emptyset(t_{\gamma_{LH}})}\right)\right), \end{cases}$$

and let $\tilde{\pi} \equiv (\tilde{\pi}_{LH}, \tilde{\pi}_L, \tilde{\pi}_\emptyset)$ be the corresponding joint mechanism. The direct mechanism $\tilde{x}$ is constructed so that the recommendation probabilities are the same as under x, but consumption is recommended to both types if and only if $\theta < t_{\gamma_{LH}}$. Hence (A.14)–(A.15) hold by construction. Moreover, $\tilde{\pi}$ satisfies the incentive constraints (16) and (18), as it issues recommendations to abstain to higher-risk types than π. Finally, $\tilde{\pi}$ weakly improves efficiency upon π, as it induces the same expected consumption levels with a lower expected harm from consumption, and strictly so if π does not satisfy (A.15) on a $\mathbf{P}$-nonnull set. The result follows. ■

Lemma A.3 implies that any solution $\pi^{**} \equiv (\pi^{**}_{LH}, \pi^{**}_L, \pi^{**}_\emptyset)$ to (A.13) is such that, for some cutoff t^{**}_{LH}, we have (up to a $\mathbf{P}$-null set)

$$\pi^{**}_{LH}(\theta) = 1_{\{\theta < t^{**}_{LH}\}}.$$

For any such joint mechanism, type H consumes if and only if $\theta < t^{**}_{LH}$. Thus his consumption behavior is already fully determined. Hence, given an optimal cutoff t^{**}_{LH}, problem (A.13) reduces to finding a measurable function $\pi^{**}_L : [0, 1] \to [0, 1]$ that vanishes over $[0, t^{**}_{LH})$ and that solves the following problem:

$$\max \{t^h_L \mathbf{E}[\pi_L(\theta)] - \mathbf{E}[\theta \pi_L(\theta)] : \pi \text{ satisfies (16) and (18)}\}. \tag{A.17}$$

As in Section 3, the left-hand side of constraint (16) is not well-defined if $\pi_\emptyset = 0$ $\mathbf{P}$-almost surely over $[t^{**}_{LH}, 1)$, and similarly the left-hand side of constraint (18) is not well-defined if $\pi_L = 0$ $\mathbf{P}$-almost surely over $[t^{**}_{LH}, 1)$. To circumvent this problem, we again adopt the convention that the undefined constraint is emptily satisfied, which allows us to linearize the constraints (16) and (18). We start with an existence result.

Lemma A.4 *Problems (A.17), (A.13), and (20) have a solution.*

Proof. Our convention on the constraints (16) and (18) allows us to rewrite (A.17) as

$$\max \{t^h_L \mathbf{E}[\pi_L(\theta)] - \mathbf{E}[\theta \pi_L(\theta)] : \mathbf{E}[\theta[1 - \pi_L(\theta)] \,|\, \theta \geq t^{**}_{LH}] \geq t^a_L \mathbf{E}[1 - \pi_L(\theta) \,|\, \theta \geq t^{**}_{LH}] \text{ and } \mathbf{E}[\theta \pi_L(\theta)] \geq t^a_H \mathbf{E}[\pi_L(\theta)]\}, \tag{A.18}$$

where the maximum is taken over the set

$$S \equiv \{\pi_L \in L_\infty(\mathbf{P}) : \pi_L(\theta) \in [0,1] \text{ for all } \theta \in [0,1] \text{ and } \pi_L(\theta) = 0 \text{ for all } \theta \in [0, t_{LH}^{**})\}. \quad (A.19)$$

Notice that S is a closed subset of the unit ball $B_{L_\infty(\mathbf{P})}$ of $L_\infty(\mathbf{P})$ when the latter set is endowed with the weak* topology $\sigma(L_\infty(\mathbf{P}), L_1(\mathbf{P}))$, which we henceforth assume without further mention. By the Banach–Alaoglu compactness theorem (Aliprantis and Border 2006, Theorem 6.21), S is thus compact in that topology, and so is by duality the set S' of the functions in S that satisfy the constraints in (A.18); notice furthermore that S' is nonempty as it contains the function

$$\pi_L(\theta) = 1_{\{t_H^a \le \theta < t_L^a\}} 1_{\{\theta \ge t_{LH}^{**}\}}.$$

Because S' is a nonempty compact set and the objective function in (A.18) is continuous in π_L by duality, (A.18) and, hence, (A.17) have a solution. To complete the proof, observe that, by Lemma A.2, we only need to show that (A.13) has a solution. Treating t_{LH}^{**} as a parameter, Berge's maximum theorem (Aliprantis and Border 2006, Theorem 17.31) implies that the solutions to (A.17) as t_{LH}^{**} varies are described by an upper hemicontinuous correspondence $\varpi_L^{**} : [0,1] \twoheadrightarrow B_{L_\infty(\mathbf{P})}$ with nonempty compact values. Thus, by Lemma A.3, (A.13) reduces to maximizing a continuous function of $(t_{LH}^{**}, \pi_L^{**})$ over $\{(t_{LH}^{**}, \pi_L^{**}) : t_{LH}^{**} \in [0,1] \text{ and } \pi_L^{**} \in \varpi_L^{**}(t_{LH}^{**})\}$, which is a nonempty compact set by the closed graph theorem (Aliprantis and Border 2006, Theorem 17.11). The result follows. ■

We are now ready to characterize the solutions to (A.17).

Lemma A.5 *Under Assumption 2, problem (A.17) has a solution of the form (23).*

Proof. We distinguish two cases.

Case 1 If constraint (18) is slack at the optimum, then (A.17) reduces to finding an optimal mechanism for type L alone, as described in Section 3. Proposition 1 yields that this mechanism is given (up to a $\mathbf{P}$-null set) by

$$\Pi_L^{**}(\theta) = 1_{\{\theta < t_L^*\}},$$

so that

$$\pi_L^{**}(\theta) = 1_{\{t_{LH}^{**} \le \theta < t_L^*\}}.$$

Hence we must have $t_{LH}^{**} = t_H^*$ as (18) is slack. We thus fall back on the joint mechanism (14), which is IC if and only if condition (15) holds.

Case 2 If constraint (18) is binding at the optimum, that is, according to Case 1, if condition (15) does not hold, then

$$\frac{\mathbf{E}[\theta\pi_L(\theta)]}{\mathbf{E}[\pi_L(\theta)]} = t_H^a. \tag{A.20}$$

Plugging (A.20) into the objective function in (A.17), the problem becomes[3]

$$\max\{(t_L^h - t_H^a)\mathbf{E}[\pi_L(\theta)] : \pi \text{ satisfies (16) and (A.20)}\}. \tag{A.21}$$

Our convention on the constraints (16) and (18) allows us to replace expectations in (A.21) by integrals, yielding the equivalent problem

$$\max\left\{(t_L^h - t_H^a)\int_{t_{LH}^{**}}^1 \pi_L(\theta)f(\theta)\,\mathrm{d}\theta : \int_{t_{LH}^{**}}^1 \theta[1-\pi_L(\theta)]f(\theta)\,\mathrm{d}\theta \geq t_L^a \int_{t_{LH}^{**}}^1 [1-\pi_L(\theta)]f(\theta)\,\mathrm{d}\theta \right.$$
$$\left. \text{and } \int_{t_{LH}^{**}}^1 \theta\pi_L(\theta)f(\theta)\,\mathrm{d}\theta = t_H^a \int_{t_{LH}^{**}}^1 \pi_L(\theta)f(\theta)\,\mathrm{d}\theta\right\},$$

where the maximum is taken over the set S defined by (A.19). Because S is convex and the objective function as well as the constraints are affine in π_L, this equivalent problem is convex. Therefore, by the Kuhn–Tucker theorem (Clarke 2013, Theorem 9.4), for any solution π_L^{**} to this problem, which by construction is a solution to (A.21) and (A.17), there exists a vector of Lagrange multipliers $(\eta^{**}, \lambda^{**}, \mu^{**})$ such that the following properties are satisfied:

1. *Nontriviality*:

$$(\eta^{**}, \lambda^{**}, \mu^{**}) \neq (0, 0, 0). \tag{A.22}$$

2. *Positivity*:

$$\eta^{**} \in \{0, 1\} \quad \text{and} \quad \lambda^{**} \in \mathbb{R}_+. \tag{A.23}$$

3. *Lagrangian maximization*:

$$\pi_L^{**} \in \arg\max\left\{\int_{t_{LH}^{**}}^1 h^{**}(\theta)\pi_L(\theta)f(\theta)\,\mathrm{d}\theta : \pi_L \in S\right\}, \tag{A.24}$$

where h^* is the affine function defined by

$$h^{**}(\theta) \equiv \eta^{**}(t_L^h - t_H^a) + \lambda^{**}t_L^a + \mu^{**}t_H^a - (\lambda^{**} + \mu^{**})\theta.$$

4. *Complementary slackness*:

$$\lambda^{**}\left\{\int_{t_{LH}^{**}}^1 \theta[1-\pi_L^{**}(\theta)]f(\theta)\,\mathrm{d}\theta - t_L^a\int_{t_{LH}^{**}}^1 [1-\pi_L^{**}(\theta)]f(\theta)\,\mathrm{d}\theta\right\} = 0. \tag{A.25}$$

[3]We keep the multiplicative constant $t_L^h - t_H^a$, which is strictly positive under Assumption 2, in order to make Lemma A.5 relevant when this assumption does not hold, as in Proposition 4.

5 *Equality constraint*:

$$\int_{t_{LH}^{**}}^{1} \theta \pi_L^{**}(\theta) f(\theta) \, \mathrm{d}\theta = t_H^a \int_{t_{LH}^{**}}^{1} \pi_L^{**}(\theta) f(\theta) \, \mathrm{d}\theta. \tag{A.26}$$

We distinguish four subcases.

Subcase 2.1 If $h^{**}(\theta) > 0$ for all $\theta \in (t_{LH}^{**}, 1)$, then the objective function in (A.24) is uniquely (up to a $\mathbf{P}$-null set) maximized over S by

$$\pi_L^{**}(\theta) = 1_{\{\theta \geq t_{LH}^{**}\}},$$

which corresponds to a cutoff $t_L^{**} = 1$ in (23). Notice that (A.25) is automatically satisfied and that (A.26) becomes

$$\mathbf{E}[\theta \,|\, \theta \geq t_{LH}^{**}] = t_H^a.$$

Hence we must have $t_{LH}^{**} = t_H^c = t_H^*$. That is, type L always consumes and type H is facing his individually optimal mechanism.

Subcase 2.2 If $h^{**}(\theta) < 0$ for all $\theta \in (t_{LH}^{**}, 1)$, then the objective function in (A.24) is uniquely (up to a $\mathbf{P}$-null set) maximized over S by

$$\pi_L^{**}(\theta) = 0,$$

which corresponds to a cutoff $t_L^{**} = t_{LH}^{**}$ in (23). Notice that (A.26) is automatically satisfied, and that (A.25) becomes

$$\lambda^{**}\{\mathbf{E}[\theta \,|\, \theta \geq t_{LH}^{**}] - t_L^a\} = 0.$$

Hence we must have $t_{LH}^{**} = t_L^c$ if $\lambda^{**} > 0$.

Subcase 2.3 Suppose that h^{**} changes sign over $(t_{LH}^{**}, 1)$—so that, in particular, $\lambda^{**} + \mu^{**} \neq 0$—at the cutoff

$$t_L^{**} \equiv \frac{\eta^{**}(t_L^h - t_H^a) + \lambda^{**} t_L^a + \mu^{**} t_H^a}{\lambda^{**} + \mu^{**}}.$$

We claim that $\lambda^{**} + \mu^{**} > 0$. Indeed, if $\lambda^{**} + \mu^{**} < 0$, then the objective function in (A.24) is uniquely (up to a $\mathbf{P}$-null set) maximized over S by

$$\pi_L^{**}(\theta) = 1_{\{\theta \geq t_L^{**}\}}, \tag{A.27}$$

so that

$$\pi_\emptyset^{**}(\theta) = 1_{\{t_{LH}^{**} \leq \theta < t_L^{**}\}}. \tag{A.28}$$

Now, given (A.28), (16) requires

$$\mathbf{E}[\theta \,|\, t^{**}_{LH} \leq \theta < t^{**}_{L}] \geq t^a_L. \tag{A.29}$$

However, we know from Lemma A.2 that any solution to (A.13) and, hence, to (A.17) and (A.21), is also a solution to (20). In particular, given (A.27), (17) requires

$$\mathbf{E}[\theta \,|\, \theta \geq t^{**}_{L}] \leq t^a_L,$$

in contradiction with (A.29) as $\mathbf{P}$ has full support over $[0, 1]$. Thus $\lambda^{**} + \mu^{**} > 0$, as claimed, and the objective function in (A.24) is uniquely (up to a $\mathbf{P}$-null set) maximized over S by

$$\pi^{**}_L(\theta) = 1_{\{t^{**}_{LH} \leq \theta < t^{**}_{L}\}},$$

once again in line with (23).

Subcase 2.4 Suppose finally that h^{**} is identically zero over $(t^{**}_{LH}, 1)$—so that, in particular, $\lambda^{**} + \mu^{**} = 0$. Then

$$\eta^{**}(t^h_L - t^a_H) + \lambda^{**}(t^a_L - t^a_H) = 0.$$

Because $t^a_L > t^a_H$, we have $\eta^{**} = 1$ by (A.23); otherwise, by (A.23) again, $\eta^{**} = \lambda^{**} = \mu^{**} = 0$, which violates (A.22). Applying (A.23) yet again, we obtain $t^a_H \geq t^h_L$, with equality if and only if $\lambda^{**} = 0$. Hence this subcase cannot arise under Assumption 2. The result follows. ■

Proposition 3 is then an immediate consequence of Lemma A.5. Hence the result. ■

Proof of Lemma 3. We solve (A.13) for the optimal cutoffs $(t^{**}_{LH}, t^{**}_{L})$—the existence of which we established in Proposition 3—under the assumption that the individually optimal mechanisms with cutoffs t^*_H and t^*_L are not simultaneously implementable, that is, (15) does not hold. We first claim that we can restrict attention to cutoffs (t_{LH}, t_L) such that $t_L \geq t^*_L$. To prove this claim, we distinguish two cases. If $t^*_L > t^h_L$, then (16) is satisfied if and only if $t_L \geq t^*_L$. If $t^*_L = t^h_L$, then, for any given t_{LH}, any cutoff $t_L < t^h_L$ would induce an inefficiently high rate of abstinence for type L and would tighten (18) compared to $t_L = t^h_L$; hence an optimal cutoff t_L must satisfy $t_L \geq t^h_L$, which is IC as $t^h_L = t^*_L$. The claim follows. Replacing expectations in (A.13) by integrals then yields the equivalent problem

$$\max \left\{ p_L \beta_L \int_0^{t_L} (t^h_L - \theta) f(\theta)\, \mathrm{d}\theta + p_H \beta_H \int_0^{t_{LH}} (t^h_H - \theta) f(\theta)\, \mathrm{d}\theta \right\}, \tag{A.30}$$

subject to the constraints

$$\int_{t_{LH}}^{t_L} (\theta - t^a_H) f(\theta)\, \mathrm{d}\theta \geq 0, \tag{A.31}$$

$$t_L - t^*_L \geq 0, \tag{A.32}$$

$$1 - t_L \geq 0. \tag{A.33}$$

The objective function in (A.30) is continuous in (t_{LH}, t_L) and the feasible set defined by $(t_{LH}, t_L) \in [0,1]^2$ and (A.31)–(A.33) is nonempty and compact. Hence problem (A.30)–(A.33) has a solution (t_{LH}^{**}, t_L^{**}). The proof consists of four steps.

Step 1 We first show that $t_L^{**} > t_H^a > t_{LH}^{**} \geq t_H^h$ in any solution (t_{LH}^{**}, t_L^{**}) to (A.30)–(A.33). That $t_L^{**} > t_H^a$ follows from our preliminary observation that $t_L \geq t_L^h$ along with Assumption 2. As for t_{LH}^{**}, suppose, by way of contradiction, that $t_{LH}^{**} \geq t_H^a$. Because $t_L^{**} > t_H^a$, we have

$$\int_{t_H^a}^{t_L^{**}} (\theta - t_H^a) f(\theta)\, \mathrm{d}\theta > 0.$$

Hence lowering t_{LH}^{**} to a value $t_H^a - \varepsilon$ for some small $\varepsilon > 0$ would preserve (A.31) and strictly increase the objective in (A.30), a contradiction. Thus $t_H^a > t_{LH}^{**}$, as claimed. The proof that $t_{LH}^{**} \geq t_H^h$ is similar, observing that the left-hand side of (A.31) is strictly increasing in $t_{LH} \in [0, t_H^a]$ and the objective function in (A.30) is strictly increasing in $t_{LH} \in [0, t_H^h]$.

Step 2 We next verify that the constraints (A.31)–(A.33) satisfy the Mangasarian–Fromovitz qualification conditions at (t_{LH}^{**}, t_L^{**}) (Mangasarian 1969, 11.3.5). Letting g be the mapping defined by the left-hand sides of the binding constraints at (t_{LH}^{**}, t_L^{**}), we must prove that

$$\nabla g(t_{LH}^{**}, t_L^{**}) z^T > 0$$

has a solution $z \in \mathbb{R}^2$, where $\nabla g(t_{LH}^{**}, t_L^{**})$ is the Jacobian matrix of g at (t_{LH}^{**}, t_L^{**}). This is obvious if (A.31) is not binding. If (A.31) is binding, then the first line of $\nabla g(t_{LH}^{**}, t_L^{**})$ is

$$\mathrm{D}g_1(t_{LH}^{**}, t_L^{**}) \equiv \big((t_H^a - t_{LH}^{**}) f(t_{LH}^{**}) \quad (t_L^* - t_H^a) f(t_L^{**})\big).$$

We shall exploit the fact that f is strictly positive over $(0,1)$. Notice first that, because $t_H^a > t_{LH}^{**} \geq t_H^h$ by Step 1, we always have $(t_H^a - t_{LH}^{**}) f(t_{LH}^{**}) > 0$. If only (A.31) is binding, then $1 > t_L^{**} > t_H^a$ by Step 1, so that $(t_L^* - t_H^a) f(t_L^{**}) > 0$ and

$$\nabla g(t_{LH}^{**}, t_L^{**}) = \mathrm{D}g_1(t_{LH}^{**}, t_L^{**}).$$

We can then take any $z \in \mathbb{R}^2_{++}$. Next, if (A.31) and (A.32) are binding, then $t_L^{**} = t_L^*$, so that $(t_L^* - t_H^a) f(t_L^{**}) > 0$ and

$$\nabla g(t_{LH}^{**}, t_L^{**}) = \begin{pmatrix} (t_H^a - t_{LH}^{**}) f(t_{LH}^{**}) & (t_L^* - t_H^a) f(t_L^{**}) \\ 0 & 1 \end{pmatrix}.$$

We can then take any $z \in \mathbb{R}^2_{++}$. Finally, if (A.31) and (A.33) are binding, then it is optimal to have $t_{LH}^{**} = t_H^*$ by Proposition 1, and

$$\nabla g(t_{LH}^{**}, t_L^{**}) = \begin{pmatrix} (t_H^a - t_{LH}^{**}) f(t_{LH}^{**}) & (t_L^* - t_H^a) f(t_L^{**}) \\ 0 & -1 \end{pmatrix}.$$

We can then take $z = (1, \varepsilon)$ for some small enough $\varepsilon < 0$.

Step 3 According to Step 2, constraints (A.31)–(A.33) are qualified at any solution (t_{LH}^{**}, t_L^{**}) to (A.30)–(A.33). Therefore, by the Kuhn–Tucker necessary optimality conditions for nonconvex optimization problems (Mangasarian 1969, 11.3.6), there exists a vector of Lagrange multipliers $(\zeta^{**}, \nu^{**}, \chi^{**})$ such that the following properties are satisfied:

1. *Positivity*:

$$(\zeta^{**}, \nu^{**}, \chi^{**}) \in \mathbb{R}_+^3. \tag{A.34}$$

2. *First-order conditions*:

$$p_L\beta_L(t_L^h - t_L^{**})f(t_L^{**}) + \zeta^{**}(t_L^{**} - t_H^a)f(t_L^{**}) + \nu^{**} - \chi^{**} = 0, \tag{A.35}$$

$$p_H\beta_H(t_H^h - t_{LH}^{**})f(t_{LH}^{**}) - \zeta^{**}(t_{LH}^{**} - t_H^a)f(t_{LH}^{**}) = 0. \tag{A.36}$$

3. *Complementary slackness*:

$$\zeta^{**}\int_{t_{LH}^{**}}^{t_L^{**}} (\theta - t_H^a)f(\theta)\,\mathrm{d}\theta = 0, \tag{A.37}$$

$$\nu^{**}(t_L^{**} - t_L^*) = 0, \tag{A.38}$$

$$\chi^{**}(1 - t_L^{**}) = 0. \tag{A.39}$$

We distinguish three cases.

Case 1 Suppose first that (A.32) is binding, so that $t_L^{**} = t_L^*$ and $\chi^{**} = 0$ by (A.39), and suppose further, by way of contradiction, that $\zeta^{**} = 0$. Then, by (A.36) along with the fact that $f(t_{LH}^{**}) > 0$ as $t_H^a > t_{LH}^{**} \geq t_H^h$ by Step 1 and f is strictly positive over $(0,1)$, we must have $t_{LH}^{**} = t_H^h \leq t_H^*$. Therefore, using the assumption that the individually optimal mechanisms with cutoffs t_H^* and t_L^* are not simultaneously implementable, we obtain

$$\mathbf{E}[\theta\,|\,t_{LH}^{**} \leq \theta < t_L^*] \leq \mathbf{E}[\theta\,|\,t_H^* \leq \theta < t_L^*] < t_H^a.$$

But then (A.31) is violated at (t_{LH}^{**}, t_L^*), a contradiction. Hence, by (A.34), $\zeta^{**} > 0$, so that, by (A.37), (A.31) must be binding at (t_{LH}^{**}, t_L^*). That is, t_{LH}^{**} must satisfy

$$\int_{t_{LH}^{**}}^{t_L^*} (\theta - t_H^a)f(\theta)\,\mathrm{d}\theta = 0. \tag{A.40}$$

Because f is strictly positive over $(0,1)$, we have $f(t_L^*) > 0$; moreover, as argued above, $f(t_{LH}^{**}) > 0$. Because $\chi^{**} = 0 \leq \nu^{**}$ by (A.34), the first-order conditions (A.35)–(A.36) rewrite as

$$p_L\beta_L(t_L^h - t_L^*) + \zeta^{**}(t_L^* - t_H^a) \leq 0, \tag{A.41}$$

$$p_H\beta_H(t_H^h - t_{LH}^{**}) - \zeta^{**}(t_{LH}^{**} - t_H^a) = 0. \tag{A.42}$$

Because $\zeta^{**} > 0$ and $t_L^* \geq t_L^h > t_H^a$, (A.41) implies $t_L^* > t_L^h$. Hence the bracketed terms in (A.41) are different from zero. Moreover, because the bracketed terms in (A.42) cannot simultaneously be zero, none of them can be zero. Because $t_H^h \leq t_{LH}^{**} < t_H^a$ by Step 1, we can thus divide (A.42) by (A.41), which yields

$$\frac{p_H \beta_H}{p_L \beta_L} \frac{t_{LH}^{**} - t_H^h}{t_L^* - t_L^h} \leq \frac{t_H^a - t_{LH}^{**}}{t_L^* - t_H^a}. \tag{A.43}$$

Case 2 Suppose next that (A.33) is binding, so that $t_L^{**} = 1$ and $\nu^{**} = 0$ by (A.38). By Proposition 1, it is then optimal to have $t_{LH}^{**} = t_H^*$. Because f is strictly positive over $(0, 1)$, we have $f(t_H^*) > 0$. The first-order condition (A.36) then rewrites as

$$p_H \beta_H (t_H^h - t_H^*) - \zeta^{**}(t_H^* - t_H^a) = 0, \tag{A.44}$$

so that $t_H^* > t_H^h$ if and only if $\zeta^{**} > 0$. If $f(1) > 0$, then, because $\chi^{**} \geq 0 = \nu^{**}$ by (A.34), we can also simplify (A.35) to obtain

$$p_L \beta_L (t_L^h - 1) + \zeta^{**}(1 - t_H^a) \geq 0. \tag{A.45}$$

The argument leading to (A.45) is a bit more involved if $f(1) = 0$. In that case, it follows from (A.35) and $\nu^{**} = 0$ that $\chi^{**} = 0$ as well. Hence the relevant part of the Lagrangian, to be maximized with respect to t_L, can be written as

$$\int_{t_H^*}^{t_L} [p_L \beta_L (t_L^h - \theta) + \zeta^{**}(\theta - t_H^a)] f(\theta)\, \mathrm{d}\theta,$$

which, as f is strictly positive over $(0, 1)$, is maximum for $t_L = 1$ only if (A.45) holds. By (A.34) and (A.45), $\zeta^{**} > 0$, so that, by (A.37), (A.31) must be binding at $(t_H^*, 1)$. That is, t_H^* must satisfy

$$\int_{t_H^*}^{1} (\theta - t_H^a) f(\theta)\, \mathrm{d}\theta = 0, \tag{A.46}$$

which generically implies that $t_H^* > t_H^h$, so that the unconstrained-optimal mechanism for type H is not IC. The terms $t_H^* - t_H^a$ and $1 - t_H^a$ in (A.44)–(A.45) are by construction different from zero. We can thus divide (A.44) by (A.45), which yields

$$\frac{p_H \beta_H}{p_L \beta_L} \frac{t_H^* - t_H^h}{1 - t_L^h} \geq \frac{t_H^a - t_H^*}{1 - t_H^a}. \tag{A.47}$$

Case 3 Suppose finally that (A.32)–(A.33) are not binding, so that $\nu^{**} = \chi^{**} = 0$ by (A.38)–(A.39). As f is strictly positive over $(0, 1)$, we have $f(t_L^{**}) > 0$ and, as argued in Case 1, $f(t_{LH}^{**}) > 0$. The first-order conditions (A.35)–(A.36) then rewrite as

$$p_L \beta_L (t_L^h - t_L^{**}) + \zeta^{**}(t_L^{**} - t_H^a) = 0, \tag{A.48}$$

$$p_H\beta_H(t_H^h - t_{LH}^{**}) - \zeta^{**}(t_{LH}^{**} - t_H^a) = 0. \tag{A.49}$$

We must have $\zeta^{**} > 0$ and, hence, by (A.37), (A.31) must be binding, for, otherwise, by (A.48)–(A.49), we would have $t_{LH}^{**} = t_H^h$ and $t_L^{**} = t_L^h$, so that the individually unconstrained-optimal mechanisms for types H and L would be simultaneously implementable, a contradiction. That is, (t_{LH}^{**}, t_L^{**}) must satisfy

$$\int_{t_{LH}^{**}}^{t_L^{**}} (\theta - t_H^a) f(\theta)\,\mathrm{d}\theta = 0. \tag{A.50}$$

Because $t_L^{**} > t_{LH}^{**}$ by Step 1, it follows that the bracketed terms on the left-hand sides of (A.48)–(A.49) cannot be zero. Dividing yields

$$\frac{p_H\beta_H}{p_L\beta_L}\,\frac{t_{LH}^{**} - t_H^h}{t_L^{**} - t_L^h} = \frac{t_H^a - t_{LH}^{**}}{t_L^{**} - t_H^a}. \tag{A.51}$$

Step 4 To complete the proof, we only need to delineate the circumstances under which each of the cases discussed in Step 3 arises. In each case, (A.31) is binding, see (A.40), (A.46), and (A.50). Let accordingly

$$\mathcal{T}_L \equiv \{t_L \geq t_L^* : \text{there exists } t_H \leq t_L \text{ such that } \mathbf{E}[\theta \,|\, t_H \leq \theta < t_L] = t_H^a\}. \tag{A.52}$$

Because $t_L^* > t_H^a$ and $\mathbf{E}[\theta \mid t_H^* \leq \theta < t_L^*] < t_H^a$ as the individually optimal mechanisms with cutoffs t_H^* and t_L^* are not simultaneously implementable, $t_L^* \in \mathcal{T}_L$. Because $\mathbf{E}[\theta \,|\, t_H \leq \theta < t_L]$ is strictly increasing in t_H and t_L, $\mathcal{T}_L$ is thus an interval $[t_L^*, \sup \mathcal{T}_L]$, and there exists a unique strictly decreasing function $\hat{t}_{LH} : \mathcal{T}_L \to [0, t_H^a)$ implicitly defined by

$$\mathbf{E}[\theta \,|\, \hat{t}_{LH}(t_L) \leq \theta < t_L] = t_H^a \tag{A.53}$$

for all $t_L \in \mathcal{T}_L$. By (A.40), (A.46), and (A.50), given t_L^{**}, t_{LH}^{**} is uniquely pinned down by

$$t_{LH}^{**} = \hat{t}_{LH}(t_L^{**}). \tag{A.54}$$

As f is strictly positive over $(0, 1)$, a straightforward application of the implicit function theorem implies that $\hat{t}_{LH}$ is differentiable over the interior of $\mathcal{T}_L$, with

$$\hat{t}'_{LH}(t_L) = -\frac{f(t_L)}{f(\hat{t}_{LH}(t_L))}\,\frac{t_L - \mathbf{E}[\theta \,|\, \hat{t}_{LH}(t_L) \leq \theta < t_L]}{\mathbf{E}[\theta \,|\, \hat{t}_{LH}(t_L) \leq \theta < t_L] - \hat{t}_{LH}(t_L)} < 0. \tag{A.55}$$

While (A.54) holds in each of Cases 1, 2, and 3, these cases differ as to whether (A.43), (A.47), or (A.51) holds. Defining accordingly

$$\kappa(t_L) \equiv \frac{p_H\beta_H}{p_L\beta_L}\,\frac{\hat{t}_{LH}(t_L) - t_H^h}{t_L - t_L^h} - \frac{t_H^a - \hat{t}_{LH}(t_L)}{t_L - t_H^a}, \tag{A.56}$$

we have $\kappa(t_L^*) \leq 0$, $\kappa(1) \geq 0$, and $\kappa(t_L^{**}) = 0$ in Cases 1, 2, and 3, respectively. To conclude, we only need to show that these cases are mutually exclusive. For this, we only need to show that κ crosses zero only once, from above. Indeed, if $\kappa(t_L) = 0$, then

$$\begin{aligned}
\kappa'(t_L) &= \frac{p_H \beta_H}{p_L \beta_L}\left[\frac{\hat{t}'_{LH}(t_L)}{t_L - t^h_L} - \frac{\hat{t}_{LH}(t_L) - t^h_H}{(t_L - t^h_L)^2}\right] + \frac{\hat{t}'_{LH}(t_L)}{t_L - t^a_H} + \frac{t^a_H - \hat{t}_{LH}(t_L)}{(t_L - t^a_H)^2} \\
&< -\frac{p_H \beta_H}{p_L \beta_L}\,\frac{\hat{t}_{LH}(t_L) - t^h_H}{(t_L - t^h_L)^2} + \frac{t^a_H - \hat{t}_{LH}(t_L)}{(t_L - t^a_H)^2} \\
&= \frac{[t^a_H - \hat{t}_{LH}(t_L)](t^a_H - t^h_L)}{(t_L - t^h_L)(t_L - t^a_H)^2} \\
&< 0, \qquad \text{(A.57)}
\end{aligned}$$

where the first inequality follows from (A.55), the second equality follows from (A.56) along with $\kappa(t_L) = 0$, and the second inequality follows from Assumption 2. Thus Case 1 occurs if and only if $\kappa(t^*_L) \leq 0$, so that $\kappa(t_L) < 0$ for all $t_L > t^*_L$, Case 2 occurs if and only if $\kappa(1) \geq 0$, so that $\kappa(t_L) > 0$ for all $t_L < 1$, and Case 3 occurs if and only if $\kappa(t^*_L) > 0$ and $\kappa(1) < 0$, so that $\kappa(t_L)$ changes sign from positive to negative only at $t_L = t^{**}_L$. The result follows. ■

Proof of Corollary 4. The proof consists of three steps.

Step 1 Consider first the boundary $\underline{p}$, starting with the case $t^*_L > t^h_L$. Define the function $\hat{t}_{LH}$ as in (A.53). By Assumption 2, $t^*_L > t^a_H$, and, by construction, $\hat{t}_{LH}(t^*_L) < t^a_H$. Moreover, because the individually optimal mechanisms with cutoffs t^*_H and t^*_L are not simultaneously implementable, $\hat{t}_{LH}(t^*_L) > t^*_H$ and thus $\hat{t}_{LH}(t^*_L) > t^h_H$. Hence

$$\frac{\beta_H}{\beta_L}\,\frac{\hat{t}_{LH}(t^*_L) - t^h_H}{t^*_L - t^h_L} > 0 \ \text{ and } \ \frac{t^a_H - \hat{t}_{LH}(t^*_L)}{t^*_L - t^a_H} > 0.$$

As $p \mapsto \frac{p}{1-p}$ is a strictly increasing continuous mapping from $(0, 1)$ to $(0, \infty)$, there exists a unique $\underline{p} \in (0, 1)$ such that

$$\frac{\underline{p}\beta_H}{(1 - \underline{p})\beta_L}\,\frac{\hat{t}_{LH}(t^*_L) - t^h_H}{t^*_L - t^h_L} = \frac{t^a_H - \hat{t}_{LH}(t^*_L)}{t^*_L - t^a_H},$$

so that

$$\frac{p_H \beta_H}{p_L \beta_L}\,\frac{\hat{t}_{LH}(t^*_L) - t^h_H}{t^*_L - t^h_L} \leq \frac{t^a_H - \hat{t}_{LH}(t^*_L)}{t^*_L - t^a_H}$$

if and only if $p_H \in [0, \underline{p}]$. Defining κ as in (A.56), we thus have $\kappa(t^*_L) \leq 0$ for any such p_H. It then follows from Step 4 of the proof of Lemma 3 that $(t^{**}_{LH}, t^{**}_L) = (\hat{t}_{LH}(t^*_L), t^*_L)$. We have thus proven that, if $t^*_L > t^h_L$, there exists $\underline{p} \in (0, 1)$ such that, for all $p_H \in (0, \underline{p}]$, type L faces his individually optimal mechanism. To complete the proof, we only need to check that if $t^*_L = t^h_L$ and type L faces his individually optimal mechanism, so that $t^{**}_L = t^*_L = t^h_L$, then it must be that $p_H = 0$, in which case we can set $\underline{p} \equiv 0$ by convention. Indeed, from (A.41) in Case 1 of the proof of Lemma 3, if we impose the constraint (A.31), which is relevant only if $p_H > 0$, then $\zeta^{**} > 0$, and $t^{**}_L = t^*_L$ implies $t^*_L > t^h_L$. Thus $t^{**}_L = t^*_L = t^h_L$ implies $p_H = 0$, as desired.

Step 2 Consider next the boundary $\overline{p}$, starting with the case $t^*_H > t^h_H$. Then

$$\frac{\beta_H}{\beta_L}\,\frac{t^*_H - t^h_H}{1-t^h_L} > 0 \quad\text{and}\quad \frac{t^a_H - t^*_H}{1-t^a_H} > 0.$$

As $p \mapsto \frac{p}{1-p}$ is a strictly increasing continuous mapping from $(0,1)$ to $(0,\infty)$, there exists a unique $\overline{p} \in (0,1)$ such that

$$\frac{\overline{p}\beta_H}{(1-\overline{p})\beta_L}\,\frac{t^*_H - t^h_H}{1-t^h_L} = \frac{t^a_H - t^*_H}{1-t^a_H},$$

so that

$$\frac{p_H\beta_H}{p_L\beta_L}\,\frac{t^*_H - t^h_H}{1-t^h_L} \geq \frac{t^a_H - t^*_H}{1-t^a_H}$$

if and only if $p_H \in [\overline{p}, 1]$. Defining κ as in (A.56), we thus have $\kappa(1) \geq 0$ for any such p_H. It then follows from Step 4 of the proof of Lemma 3 that $(t^{**}_{LH}, t^{**}_L) = (t^*_H, 1)$. We have thus proven that, if $t^*_H > t^h_H$, there exists $\overline{p} \in (0,1)$ such that, for all $p_H \in [\overline{p}, 1)$, type H faces his individually optimal mechanism. To complete the proof, we only need to check that if $t^*_H = t^h_H$ and type H faces his individually optimal mechanism, so that $t^{**}_{LH} = t^*_H = t^h_H$, then it must be that $p_H = 1$, in which case we can set $\overline{p} \equiv 1$ by convention. Indeed, from (A.44) in Case 2 of the proof of Lemma 3, $t^*_H = t^h_H$ implies $\zeta^{**} = 0$. Because $t^{**}_{LH} = t^*_H$ implies $t^{**}_L = 1$, (A.45) implies $p_L = 0$, as desired.

Step 3 According to Steps 1–2,

$$\frac{p_H\beta_H}{p_L\beta_L}\,\frac{\hat{t}_{LH}(t^*_L) - t^h_H}{t^*_L - t^h_L} > \frac{t^a_H - \hat{t}_{LH}(t^*_L)}{t^*_L - t^a_H} \quad\text{and}\quad \frac{p_H\beta_H}{p_L\beta_L}\,\frac{t^*_H - t^h_H}{1-t^h_L} < \frac{t^a_H - t^*_H}{1-t^a_H}$$

if and only if $p_H \in (\underline{p}, \overline{p})$. Defining κ as in (A.56), we thus have

$$\kappa(p_H, t^{**}_L) = \frac{p_H\beta_H}{(1-p_H)\beta_L}\,\frac{\hat{t}_{LH}(t^{**}_L) - t^h_H}{t^{**}_L - t^h_L} - \frac{t^a_H - \hat{t}_{LH}(t^{**}_L)}{t^{**}_L - t^a_H} = 0 \tag{A.58}$$

for any such p_H, where we make the dependence of κ on p_H explicit. It then follows from Step 4 of the proof of Lemma 3 that (t^{**}_{LH}, t^{**}_L) is the unique solution to (26). Let us accordingly denote by $\hat{t}_L(p_H)$ the unique solution to (A.58). We clearly have $\mathrm{D}_{p_H}\kappa(p_H, t_L) > 0$ and, from (A.57), $\mathrm{D}_{t_L}\kappa(p_H, t_L) < 0$ if $\kappa(p_H, t_L) = 0$. A straightforward application of the implicit function theorem then implies that $\hat{t}_L$ is differentiable over $(\underline{p}, \overline{p})$, with $\hat{t}'_L > 0$. Summarizing, because, for each $p_H \in (\underline{p}, \overline{p})$,

$$(t^{**}_{LH}, t^{**}_L) = (\hat{t}_{LH}(\hat{t}_L(p_H)), \hat{t}_L(p_H)),$$

where $\hat{t}_{LH}$ is strictly decreasing over the interval $\mathcal{T}_L$ by (A.55), the probabilities $F(\hat{t}_{LH}(\hat{t}_L(p_H)))$ and $F(\hat{t}_L(p_H))$ that type H and type L consume, respectively, are strictly decreasing and strictly increasing in $p_H \in (\underline{p}, \overline{p})$, respectively. Hence the result. ■

Proof of Proposition 4. By Proposition 3, if Assumption 2 holds, then there exists an optimal IC joint mechanism of the form (27) with $\bar{t}_L^{**} = 1$. Suppose then that Assumption 2 does not hold. The result is immediate if we are in Case 1 or Subcases 2.1–2.3 of Lemma A.5; note, incidentally, that we can be in Subcase 2.2, which, according to Lemma 3, cannot arise under Assumption 2. There remains to consider Subcase 2.4 of Lemma A.5, in which the affine function h^{**} is identically zero over $(t_{LH}^{**}, 1)$.

Case 1 We first assume that $t_H^a > t_L^h$. Then, by arguments already invoked, $\lambda^{**} > 0$ and, by (A.25), any solution π_L^{**} to (A.17) must satisfy (A.26) and

$$\int_{t_{LH}^{**}}^1 \theta[1 - \pi_L^{**}(\theta)]f(\theta)\,\mathrm{d}\theta = t_L^a \int_{t_{LH}^{**}}^1 [1 - \pi_L^{**}(\theta)]f(\theta)\,\mathrm{d}\theta. \tag{A.59}$$

Notice that, because Lemma A.4 guarantees that a solution π_L^{**} to (A.17) exists, there exists a solution to (A.26) and (A.59). Conversely, because h^{**} is identically zero over $(t_{LH}^{**}, 1)$, any solution to (A.26) and (A.59) is a solution to the maximization condition (A.24) and, hence, to (A.17) as this is a convex problem and $\eta^{**} > 0$ (Clarke 2013, Exercise 9.7). Let us then fix a solution π_L^{**} to (A.26) and (A.59). We focus with no loss of generality on the case where π_L^{**} is not equal to 0 or to 1, $\mathbf{P}$-almost surely over $(t_{LH}^{**}, 1)$; otherwise, we are back to Subcases 2.1 or 2.2 of Lemma A.5. That is, we focus on the case where both constraints (16) and (18) in (A.13) are well-defined and binding. In particular, we must have

$$t_{LH}^{**} < t_H^a < \mathbf{E}[\theta \,|\, \theta \geq t_{LH}^{**}] < t_L^a. \tag{A.60}$$

Summing (A.26) and (A.59) and rearranging, we obtain that any solution to (A.26) and (A.59) satisfies

$$\int_{t_{LH}^{**}}^1 [1 - \pi_L^{**}(\theta)]f(\theta)\,\mathrm{d}\theta = \rho \equiv \frac{\mathbf{E}[\theta \,|\, \theta \geq t_{LH}^{**}] - t_H^a}{t_L^a - t_H^a}[1 - F(t_{LH}^{**})] < 1 - F(t_{LH}^{**}). \tag{A.61}$$

We claim that, in line with (27), there exists a solution to (A.26) and (A.59) of the form

$$\pi_L^{**}(\theta) = 1_{\{t_{LH}^{**} \leq \theta < \underline{t}_L^{**}\}} + 1_{\{\theta \geq \bar{t}_L^{**}\}}$$

for some cutoffs $\bar{t}_L^{**} > \underline{t}_L^{**} > t_{LH}^{**}$. To prove this claim, we show that the system in $(\underline{t}, \bar{t})$

$$\int_{\underline{t}}^{\bar{t}} \theta f(\theta)\,\mathrm{d}\theta = t_L^a[F(\bar{t}) - F(\underline{t})] \tag{A.62}$$

$$\int_{t_{LH}^{**}}^{\underline{t}} \theta f(\theta)\,\mathrm{d}\theta + \int_{\bar{t}}^1 \theta f(\theta)\,\mathrm{d}\theta = t_H^a[F(\underline{t}) - F(t_{LH}^{**}) + 1 - F(\bar{t})], \tag{A.63}$$

has a unique solution. As above, summing (A.62)–(A.63) yields

$$F(\bar{t}) - F(\underline{t}) = \rho, \tag{A.64}$$

and, hence, (A.62) rewrites as

$$\psi(\underline{t}) \equiv \frac{\int_{\underline{t}}^{F^{-1}(F(\underline{t})+\rho)} \theta f(\theta)\,\mathrm{d}\theta}{\rho} = \mathbf{E}[\theta\,|\,\underline{t} \leq \theta < F^{-1}(F(\underline{t})+\rho)] = t_L^a,$$

which we must solve for $\underline{t} \in (t_{LH}^{**}, F^{-1}(1-\rho)]$. By the intermediate value theorem, we only need to check that $\psi(t_{LH}^{**}) < t_L^a$, that ψ is strictly increasing over $(t_{LH}^{**}, F^{-1}(1-\rho)]$, and that $\psi(F^{-1}(1-\rho)) \geq t_L^a$. The first statement follows from

$$\psi(t_{LH}^{**}) = \mathbf{E}[\theta\,|\,t_{LH}^{**} \leq \theta < F^{-1}(F(t_{LH}^{**})+\rho)] < \mathbf{E}[\theta\,|\,\theta \geq t_{LH}^{**}] < t_L^a,$$

where the first inequality follows from the fact that $F(t_{LH}^{**}) + \rho < 1$ by (A.64) and that $\mathbf{P}$ has full support over $[0,1]$, and the second inequality follows from (A.60). The second statement follows from a straightforward computation,

$$\psi'(\underline{t}) = \frac{f(\underline{t})[F^{-1}(F(\underline{t})+\rho) - \underline{t}]}{\rho} > 0.$$

The third statement amounts to

$$\frac{\int_{F^{-1}(1-\rho)}^{1} \theta f(\theta)\,\mathrm{d}\theta}{\rho} \geq t_L^a. \tag{A.65}$$

But we know that there exists a solution to (A.26) and (A.59), which satisfies

$$t_L^a = \frac{\int_{t_{LH}^{**}}^{1} \theta[1-\pi_L^{**}(\theta)]f(\theta)\,\mathrm{d}\theta}{\int_{t_{LH}^{**}}^{1} [1-\pi_L^{**}(\theta)]f(\theta)\,\mathrm{d}\theta} = \frac{\int_{t_{LH}^{**}}^{1} \theta[1-\pi_L^{**}(\theta)]f(\theta)\,\mathrm{d}\theta}{\rho}$$

by (A.61), and clearly

$$\int_{F^{-1}(1-\rho)}^{1} \theta f(\theta)\,\mathrm{d}\theta = \max\left\{\int_{t_{LH}^{**}}^{1} \theta[1-\pi_L(\theta)]f(\theta)\,\mathrm{d}\theta : \int_{t_{LH}^{**}}^{1} [1-\pi_L(\theta)]f(\theta)\,\mathrm{d}\theta = \rho\right\},$$

which yields the desired inequality (A.65). The claim follows. In case (A.65) holds as an equality, we have $\bar{t}_L^{**} = 1$, and π_L^{**} has the same form as in Subcase 2.3 of Lemma A.5.

Case 2 The proof for the limiting case $t_H^a = t_L^h$ or, equivalently, $\beta_H = \beta_H^{no}(\beta_L)$, relies on a simple continuity argument. From the proof of Lemma A.4, for each $\beta_H \geq \beta_H^{no}(\beta_L)$, any solution to (A.13) for β_H can be represented by a pair $(t_{LH}^{**}(\beta_H), \pi_L^{**}(\beta_H)) \in [0,1] \times B_{L_\infty(\mathbf{P})}$. Consider a strictly decreasing sequence $(\beta_{H,n})_{n\in\mathbb{N}}$ converging to $\beta_H^{no}(\beta_L)$. By Berge maximum theorem (Aliprantis and Border 2006, Theorem 17.31) along with the fact that $B_{L_\infty(\mathbf{P})}$ is metrizable as $L_1(\mathbf{P})$ is separable (Aliprantis and Border 2006, Theorems 6.30 and 13.16), any sequence $((t_{LH}^{**}(\beta_{H,n}), \pi_L^{**}(\beta_{H,n})))_{n\in\mathbb{N}}$ of solutions to (A.13) for each term of the sequence $(\beta_{H,n})_{n\in\mathbb{N}}$ has a subsequence that converges in $[0,1] \times B_{L_\infty(\mathbf{P})}$ to a solution $(t_{LH}^{**}(\beta_H^{no}(\beta_L)), \pi_L^{**}(\beta_H^{no}(\beta_L)))$ to (A.13) for $\beta_H^{no}(\beta_L)$. We can with no loss of generality assume that this sequence converges. For each $n \in \mathbb{N}$, we have $\beta_{H,n} > \beta_H^{no}(\beta_L)$ and, hence,

$$\pi_L^{**}(\beta_{H,n})(\theta) = 1_{\{t_{LH}^{**}(\beta_{H,n}) \leq \theta < t_L^{**}(\beta_{H,n})\}} \tag{A.66}$$

by Subcases 2.1–2.3 of the proof of Lemma A.5. Therefore,

$$
\begin{aligned}
\int \pi_L^{**}(\beta_H^{no}(\beta_L))(\theta)\,\mathbf{P}(\mathrm{d}\theta) &= \lim_{n\to\infty} \int \pi_L^{**}(\beta_{H,n})(\theta)\,\mathbf{P}(\mathrm{d}\theta) \\
&= \lim_{n\to\infty} F(t_L^{**}(\beta_{H,n})) - F(t_{LH}^{**}(\beta_{H,n})) \\
&= \lim_{n\to\infty} F(t_L^{**}(\beta_{H,n})) - F(t_{LH}^{**}(\beta_H^{no}(\beta_L))), \qquad \text{(A.67)}
\end{aligned}
$$

where the first equality follows from the fact that the sequence $(\pi_L^{**}(\beta_{H,n}))_{n\in\mathbb{N}}$ converges in $B_{L_\infty(\mathbf{P})}$ to $\pi_L^{**}(\beta_H^{no}(\beta_L))$, using the definition of the weak* topology $\sigma(L_\infty(\mathbf{P}), L_1(\mathbf{P}))$, the second equality follows from (A.66), and the third inequality follows from the fact that the sequence $(t_{LH}^{**}(\beta_{H,n}))_{n\in\mathbb{N}}$ converges to $t_{LH}^{**}(\beta_H^{no}(\beta_L))$ in $[0,1]$ and that F is continuous as $\mathbf{P}$ is nonatomic. Because F is strictly increasing as $\mathbf{P}$ has full support, (A.67) implies that the sequence $(t_L^{**}(\beta_{H,n}))_{n\in\mathbb{N}}$ converges to some limit t_∞. To complete the proof, notice that, for any Borel subset A of $[0,1]$,

$$
\begin{aligned}
\int_A \pi_L^{**}(\beta_H^{no}(\beta_L))(\theta)\,\mathbf{P}(\mathrm{d}\theta) &= \lim_{n\to\infty} \int_A \pi_L^{**}(\beta_{H,n})(\theta)\,\mathbf{P}(\mathrm{d}\theta) \\
&= \lim_{n\to\infty} \mathbf{P}[A \cap (t_{LH}^{**}(\beta_{H,n}), t_L^{**}(\beta_{H,n})]], \qquad \text{(A.68)}
\end{aligned}
$$

using again the definition of the weak* topology $\sigma(L_\infty(\mathbf{P}), L_1(\mathbf{P}))$ along with (A.66). Finally, we can substitute $A = (t_{LH}^{**}(\beta_H^{no}(\beta_L)), t_\infty]$ and $A = (t_\infty, 1]$ in (A.68) and use the fact that the sequence $((t_{LH}^{**}(\beta_{H,n}), t_L^{**}(\beta_{H,n})))_{n\in\mathbb{N}}$ converges to $(t_{LH}^{**}(\beta_H^{no}(\beta_L)), t_\infty)$ to conclude that in fact $t_\infty = t_L^{**}(\beta_H^{no}(\beta_L))$ and

$$
\pi_L^{**}(\beta_H^{no}(\beta_L))(\theta) = 1_{\{t_{LH}^{**}(\beta_H^{no}(\beta_L)) \le \theta < t_L^{**}(\beta_H^{no}(\beta_L))\}}
$$

up to a $\mathbf{P}$-null set. Hence the result. ■

Proof of Proposition 5. The proof consists in transforming our model into one to which the results of Kolotilin et al. (2017) can be adapted, and then to apply those results to characterize the optimal mechanism.

First, because the date-0 and date-1 selves share the same β and the same information about θ, the optimal consumption decisions x_0 and x_1 must coincide in our model, $x_0 = x_1 = x$. Second, we may identify the *receiver* as each self of the DM, with decision utility

$$
u \equiv x(1 - \beta\delta^2 C\theta) = x\left(\frac{t^a - \theta}{t^a}\right),
$$

and the *sender* as the mechanism designer, with intertemporal utility

$$
v \equiv x[1 + \beta\delta - (1+\delta)\beta\delta^2 C\theta] = -x(1-\beta)\delta + (1+\delta)u.
$$

Because a high action is good for high types in Kolotilin et al. (2017) and is denoted by a, we let $a \equiv 1 - x$ stand for abstention, and rewrite these utilities as

$$u \equiv a\left(\frac{\theta - t^a}{t^a}\right) + \frac{t^a - \theta}{t^a}$$

and

$$v \equiv a(1-\beta)\delta + (1+\delta)u - (1-\beta)\delta.$$

Any summand that does not depend on a does not enter any of the relevant optimization problems, and thus may be dropped without loss of generality. Similarly, the multiplicative factor $\frac{1}{t^a}$ does not affect the receiver's optimization problem, and can be omitted from its objective function u. However, as the sender does not know t^a, we cannot omit the multiplicative factor $\frac{1}{t^a}$ from her objective function v without implicitly changing the weight she puts on each private type.[4] We instead rescale v with a multiplicative factor $1 + \delta$. With a slight abuse of notation, we continue to denote the resulting transformed utilities by u and v, and obtain

$$u(a, \theta, t^a) \equiv a(\theta - t^a) \quad \text{and} \quad v(a, \theta, t^a) \equiv a\rho(t^a) + \frac{u(a, \theta, t^a)}{t^a}, \tag{A.69}$$

where

$$\rho(t^a) \equiv \frac{\delta}{1+\delta}(1-\beta) = \frac{\delta}{1+\delta}\left(1 - \frac{t_0}{t^a}\right).$$

These final expressions depend on β only through t^a. Now, the only differences with the persuasion problem studied in Kolotilin et al. (2017) are that t^a is distributed on $[t_0, 1]$ rather than on $[0, 1]$, and that there is an additional factor $\rho(t^a)$ in the first summand of $v(a, \theta, t^a)$. The following result is a modification of their Lemma 2, accounting for these changes.

Lemma A.6 *For every IC mechanism x, let $U^x(t^a)$ and $V^x(t^a)$ be the receiver's and the sender's expected interim utilities induced by x, respectively. Then the sender's expected utility is given by*

$$\int_{t_0}^{1} V^x(t^a)\,\mathrm{d}H(t^a) = \int_{t_0}^{1} U^x(t^a)J(t^a)\,\mathrm{d}t^a, \tag{A.70}$$

where $J(t^a) \equiv (\rho h)'(t^a) + \frac{h(t^a)}{t^a}$.

Proof. Define, for each $t^a \in [t_0, 1]$,

$$\tilde{v}(a, \theta, t^a) \equiv \frac{v(a, \theta, t^a)}{\rho(t^a)} = a + \frac{u(a, \theta, t^a)}{\rho(t^a)t^a}, \quad \tilde{h}(t^a) \equiv \frac{\rho(t^a)h(t^a)}{Z}, \quad \text{and} \quad \tilde{\rho}(t^a) \equiv \frac{1}{\rho(t^a)t^a},$$

where

$$Z \equiv \int_{t_0}^{1} h(t^a)\rho(t^a)\,\mathrm{d}t^a$$

[4] Indeed, Lemma A.6 below accounts for this problem by appropriately rescaling the density h.

is a normalizing constant that makes $\tilde{h}$ a density over $[t_0, 1]$. The reparameterized model written in terms of u, $\tilde{v}$, $\tilde{h}$ and $\tilde{\rho}$ is such that we can apply Kolotilin et al. (2017, Lemma 2) to obtain

$$\int_{t_0}^{1} V^x(t^a)\,\mathrm{d}H(t^a) = Z\int_{t_0}^{1} \tilde{V}^x(t^a)\tilde{h}(t^a)\,\mathrm{d}t^a = Z\tilde{h}(t_0)\mathbf{E}[\theta] + \int_{t_0}^{1} U^x(t^a)Z\tilde{J}(t^a)\,\mathrm{d}t^a, \tag{A.71}$$

where $\tilde{J}(t^a) \equiv \tilde{h}'(t^a) + \tilde{\rho}(t^a)\tilde{h}(t^a)$ for all $t^a \in [t_0, 1]$. The first term on the right-hand side of (A.71) is zero as ρ and, hence, $\tilde{h}$ vanish at $t_0 = \frac{1}{\delta^2 C}$, while the second term can be rewritten using

$$Z\tilde{I}(t^a) = Z\left[\tilde{h}'(t^a) + \frac{\tilde{h}(t^a)}{\rho(t^a)t^a}\right] = (h\rho)'(t^a) + \frac{h(t^a)}{t^a},$$

which implies (A.70). The result follows. ■

To derive an optimal mechanism, it is key to understand the sign pattern of the function J defined in Lemma A.6. The following result provides sufficient conditions for J to be either always nonnegative or to cross zero at most once, from above.

Lemma A.7 *Suppose that h is log-concave with $h'(t_0) > 0$. Then, either J is nonnegative over $[t_0, 1]$, or there exists $t_1 \in (t_0, 1)$ such that $J(t_1) = 0$ and $J(t) \gtrless 0$ if $t \lessgtr t_1$. The latter case obtains if $h(1) = 0$ and $h'(1) < 0$.*

Proof. A direct computation yields

$$J(t^a) = h(t^a)\left[\frac{1}{t^a} + \frac{1}{(1+\delta)\delta C(t^a)^2} + \frac{\delta}{1+\delta}\left(1 - \frac{t_0}{t^a}\right)\frac{h'(t^a)}{h(t^a)}\right]. \tag{A.72}$$

Clearly, if $h'(t_0) > 0$, then $J > 0$ in an open right-neighborhood of t_0. Similarly, if $h(1) = 0$ and $h'(1) < 0$, then $J(1) < 0$. In the latter case, the existence of a $t_1 \in (t_0, 1)$ such that $J(t_1) = 0$ follows from the intermediate value theorem. To conclude the proof, it thus remains to show that there can be at most one $t_1 \in (t_0, 1)$ such that $J(t_1) = 0$. We need to ensure that the term in square brackets in (A.72) crosses zero at most once, from above. Rearranging the condition that this term equal zero yields

$$-\frac{h'(t^a)}{h(t^a)} = \frac{(1+\delta)t^a + \frac{1}{\delta C}}{\delta t^a(t^a - t_0)}. \tag{A.73}$$

Because h is log-concave, the left hand-side of (A.73) is nondecreasing. Thus, because the right-hand side of (A.73) is strictly decreasing, we obtain that (A.73) can have at most one solution over $(t_0, 1)$. The result follows. ■

By Kolotilin et al. (2017, Theorem 1), the receiver's utility profile U^x is implementable by a mechanism x if and only if U^x is a convex function that lies between his utility profiles from full information and no information. With no information, the receiver abstains if and only if

$\mathbf{E}[\theta] \geq t^a$, in which case he obtains (abstention) utility $\mathbf{E}[\theta] - t^a$. His utility profile from no information is thus given by $\underline{U}(t^a) \equiv \max\{\mathbf{E}[\theta] - t^a, 0\}$. With full information, the receiver abstains if and only if $\theta \geq t^a$, in which case he obtains (abstention) utility $\theta - t^a$. His utility profile from full information is thus given by $\overline{U}(t^a) \equiv [1 - F(t^a)]\mathbf{E}[\theta - t^a \,|\, \theta \geq t^a]$. Clearly, $\underline{U}$ and $\overline{U}$ are nonincreasing convex functions that satisfy $\underline{U} \leq \overline{U}$, $\underline{U}(t_0) = \overline{U}(t_0)$, and $\underline{U}(1) = \overline{U}(1)$. Designing the optimal mechanism boils down to finding a convex function U^x that maximizes the right-hand side of (A.70) subject to the constraint $\underline{U} \leq U^x \leq \overline{U}$.

To complete the proof, we follow Kolotin et al. (2017, Section 4.2), in particular the discussion below their Theorem 2 and their Example 1, which treats the case where J switches signs at most once, from positive to negative. They argue that the optimal utility profile U^x must be piecewise linear wherever it does not coincide with the upper bound $\overline{U}$. If J is everywhere nonnegative, the optimal U^x satisfies $U^x = \overline{U}$, so that a full-information signal is optimal. If J switches sign once, from positive to negative, they show that there exists $\tilde{t} \in (t_0, 1]$ such that the optimal U^x satisfies $U^x(t^a) = \overline{U}(t^a)$ for $t^a \leq \tilde{t}$. For $t^a > \tilde{t}$, U_1^x continues as a tangent, that is, it is piecewise linear with slope $\overline{U}'(\tilde{t})$ until it hits $\underline{U}$, after which it coincides with $\underline{U}$. This choice of U^x corresponds to a mechanism that fully discloses each $\theta < \tilde{t}$, whereas all $\theta \geq \tilde{t}$ are pooled into a red warning. As a result, DMs with high self-control, $t^a \leq \tilde{t}$, consume whenever it is optimal for them to do so. DMs with intermediate self-control, $t^a \in (\tilde{t}, \mathbf{E}[\theta \,|\, \theta \geq \tilde{t}]]$, abstain from consuming whenever they observe a red warning, which leaves them to abstain more often than under full information. Finally, DMs with $t^a > \mathbf{E}[\theta \,|\, \theta \geq \tilde{t}]$ have so little self-control that they always consume and obtain (abstention) utility 0. Hence the result. ∎

References

[1] Aliprantis, C.D., and K.C. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Berlin, Heidelberg, New York: Springer.

[2] Aumann, R.J. (1964): "Mixed and Behavior Strategies in Infinite Extensive Games," in *Advances in Game Theory*, Annals of Mathematics Study, Vol. 52, ed. by M. Dresher, L.S. Shapley, and A.W. Tucker. Princeton: Princeton University Press, 627–650.

[3] Bryson, M.C., and M.M. Siddiqui (1969): "Some Criteria for Aging," *Journal of the American Statistical Association*, 64(328), 1472–1483.

[4] Clarke, F. (2013): *Functional Analysis, Calculus of Variations and Optimal Control*, London: Springer-Verlag.

[5] Giorgi, G., and S. Komlósi (1992): "Dini Derivatives in Optimization–Part I," *Decisions in*

Economics and Finance, 15(1), 3–30.

[6] Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2017): "Persuasion of a Privately Informed Receiver," *Econometrica*, 85(6), 1949–1964.

[7] Mangasarian, O.L. (1969): *Nonlinear Programming*, New York: McGraw-Hill.

[8] Shaked, M., and J.G. Shanthikumar (2007): *Stochastic Orders*, New York: Springer.