Online Appendix to “Forgetful Consumers and Consumption Tracking” by Ying Bao, Peter Landry, and Mengze Shi

B. Supplementary Appendix

In this appendix, we analyze several variations and extensions of our model.

B.1. Alternate Formulations of Consumer Forgetfulness

In our primary analysis, we modeled consumer forgetfulness in a rather particular way. Namely, according to our model, a consumer who consumed at $t = 1$ will believe, upon her arrival at $t = 2$, that α is the probability that she did *not* consume at $t = 1$. Next, we will consider five potential alternate formulations of consumer forgetfulness.

B.1.1. Binary Forgetting

In our original model with $0 < \alpha < 1$, any consumer who consumes at $t = 1$ will “partially” forget her $t = 1$ consumption in that, at $t = 2$, she still believes it is possible that she consumed at $t = 1$. Now, however, we consider an alternate formulation in which a consumer who consumes at $t = 1$ either fully forgets or fully remembers her past consumption, with α now denoting the probability that she fully forgets. Formally, in this binary forgetting model, a consumer who consumes at $t = 1$ has an α probability of arriving at $t = 2$ with an incorrect belief that $\Pr[d_1 = 0] = 1$ and a $1-\alpha$ probability of arriving at $t = 2$ with the correct belief that $\Pr[d_1 = 0] = 0$. (By contrast, in our original model the consumer would always arrive at $t = 2$ with a belief that $\Pr[d_1 = 0] = \alpha$.)

Benchmark Analysis

Under this alternate formulation of consumer forgetfulness, and without consumption tracking, a consumer who did not consume at $t = 1$ will still consume at $t = 2$ with probability 1. That is, $D_1(\phi|d_1 = 0) = 1$ for all $\phi \geq 0$. If the consumer did consume at $t = 1$, then with probability α she will forget and thus consume as she would with $d_1 = 0$; with probability $1-\alpha$, however, she will remember her past consumption and consume only if her $t = 2$ valuation exceeds the penalty fee ϕ. Given v_2 is uniformly distributed between 0 and 1, we can then verify

$$D_2(\phi|d_1 = 1) = \alpha + (1-\alpha)\max\{1-\phi, 0\} = \begin{cases} \alpha, & \phi > 1, \\ 1-(1-\alpha)\phi, & \phi \leq 1. \end{cases} \tag{B-1}$$

Next, we can verify that a consumer’s $t = 1$ perception of her expected $t = 2$ utilities, conditional on each possible d_1, are:

$$\widetilde{u}_2(\phi|d_1 = 0) = \frac{1}{2}, \quad \widetilde{u}_2(\phi|d_1 = 1) = \widetilde{\alpha}\left(\frac{1}{2} - \phi\right) + \frac{(1-\widetilde{\alpha})(1-\phi)^2 \mathrm{I}[\phi<1]}{2}. \tag{B-2}$$

Using our above work, while noting $D_1 = \Pr[v_1 + \widetilde{u}_2(\phi|d_1 = 1) \geq \widetilde{u}_2(\phi|d_1 = 0)]$, we can compute:

$$D_1(\phi) = \begin{cases} 0, & \phi > \frac{1+\widetilde{\alpha}}{2\widetilde{\alpha}}, \\ \frac{1+\widetilde{\alpha}}{2} - \widetilde{\alpha}\phi, & 1 < \phi \leq \frac{1+\widetilde{\alpha}}{2\widetilde{\alpha}}, \\ 1 - \phi + \frac{(1-\widetilde{\alpha})\phi^2}{2}, & \phi < 1. \end{cases} \tag{B-3}$$

The probability of consuming in both periods, $D_{12} = D_1 \cdot D_2(\phi|d_1 = 1)$, is then:

$$D_{12}(\phi) = \begin{cases} 0, & \phi > \frac{1+\widetilde{\alpha}}{2\widetilde{\alpha}}, \\ \alpha\left(\frac{1+\widetilde{\alpha}}{2} - \widetilde{\alpha}\phi\right), & 1 < \phi \leq \frac{1+\widetilde{\alpha}}{2\widetilde{\alpha}}, \\ (1-(1-\alpha)\phi)\left(1 - \phi + \frac{(1-\widetilde{\alpha})\phi^2}{2}\right), & \phi < 1. \end{cases} \tag{B-4}$$

Using $p = \widetilde{V} = \frac{1+D_1^2}{2}$ and $\Pi = p + D_{12} \cdot \phi$, we can express the firm's profits as:

$$\Pi(\phi) = \begin{cases} \frac{1}{2}, & \phi > \frac{1+\widetilde{\alpha}}{2\widetilde{\alpha}}, \\ \frac{1}{2} + \frac{(1-\widetilde{\alpha}(2\phi-1))(1-\widetilde{\alpha}(2\phi-1)+4\alpha\phi)}{8}, & 1 \leq \phi \leq \frac{1+\widetilde{\alpha}}{2\widetilde{\alpha}}, \\ \frac{1}{2} + \frac{(2(1-\phi)+(1-\widetilde{\alpha})\phi^2)(2(1-\phi)(1+2\phi)+(1-\widetilde{\alpha}+4\alpha)\phi^2)}{8}, & \phi < 1 \end{cases} \tag{B-5}$$

From the above work, we can verify the optimal contract in the benchmark equilibrium with binary forgetting as follows.

Remark 4. *In the benchmark equilibrium with binary forgetting:*

(i) *If* $\widetilde{\alpha} \geq \frac{\alpha(8-\alpha-\sqrt{8(7-2\alpha)})}{8+\alpha^2}$, *then* $\phi^*_{NT} = 0$ *and* $p^*_{NT} = 1$*;*

(ii) *If* $\widetilde{\alpha} < \frac{\alpha(8-\alpha-\sqrt{8(7-2\alpha)})}{8+\alpha^2}$, *then* $\phi^*_{NT} = \frac{(1+\widetilde{\alpha})(\alpha-\widetilde{\alpha})}{2\widetilde{\alpha}(2\alpha-\widetilde{\alpha})} > 0$ *and* $p^*_{NT} = \frac{\alpha(1+\widetilde{\alpha})}{2(2\alpha-\widetilde{\alpha})} < 1.$

Thus, the firm optimally uses a positive penalty fee when consumers are sufficiently naive (in that $\widetilde{\alpha}$ is below some threshold that depends on α) and otherwise sets its penalty fee to zero.

Analysis with Consumption Tracking

Now suppose consumption tracking is available to consumers. A consumer who consumes at $t = 1$ but forgets her past consumption (as she would with probability α) will choose to use consumption tracking if $k < k_\tau(1, \phi)$; if she remembers her past consumption (as she would with probability $1-\alpha$), she will choose to use consumption tracking if $k < k_\tau(0, \phi)$. Note here that $k_\tau(\alpha, \phi)$ is still as given in (3) for our original model, except with α set to 0 or 1 because — at the time of the tracking decision — a consumer who completely forgets effectively has $\alpha = 0$ (in the original sense of α) while a consumer who completely remembers effectively has $\alpha = 1$ (again in the original sense of α). Crucially, however, (3) implies $k_\tau(1, \phi) = k_\tau(0, \phi) = 0$. This means that the consumer will never use consumption tracking, regardless of whether she remembers or forgets.

Intuitively, this makes sense because the consumer — in both cases — believes she already knows, with certainty, whether she consumed at $t = 1$. Furthermore, a consumer knows she will not use consumption tracking conditional on either forgetting or remembering at $t = 2$ because, even with $\widetilde{\alpha} \neq \alpha$, she still knows she will either end up completely forgetting or completely remembering her past consumption. Formally, this ensures $\widetilde{D}_\tau = D_\tau = 0$.

The fact that a consumer never uses or expects to use consumption tracking in the model with binary forgetting means that the availability of consumption tracking does enter the objective functions of consumers or the firm. It also means that the availability of consumption tracking cannot impact on market behaviors or outcomes. This leads to the following characterization of the equilibrium with consumption tracking.

Remark 5. *With consumption tracking, the equilibrium under binary forgetting will still be identical to the corresponding benchmark equilibrium without consumption tracking.*

In other words, the availability of consumption tracking is irrelevant (for all parameter values) under the presently considered variation of our model with binary forgetting. The optimal contract will therefore be given by $\phi^* = \phi^*_{NT}$ and $p^* = p^*_{NT}$, with ϕ^*_{NT} and p^*_{NT} as given in Remark 4.

B.1.2. Inferring Past Consumption from Past Valuation

As another potential formulation of consumer forgetfulness, we now suppose that the consumer cannot remember her $t = 1$ consumption d_1, but can remember her $t = 1$ valuation v_1 and forms beliefs of d_1 based her memory of v_1.

Benchmark Analysis

Under this formulation of consumer forgetfulness, a consumer's memory of v_1 will effectively allow her to remember d_1. This is because the consumer can simply infer d_1 based on whether or not v_1 exceeds the threshold value of v_1 for consumption (assuming the consumer also remembers this threshold). As a result, the consumer's $t = 2$ consumption behavior will be the same as a consumer in our original model with $\alpha = 0$.

Formally, the $t = 2$ consumption probability is now given by

$$D_2 = \max\{1 - d_1\phi, 0\}, \tag{B-6}$$

while the consumer's $t = 1$ perceptions of her $t = 2$ utilities, conditional on each possible d_1, are

$$\widetilde{u}_{2|d_1=0} = \frac{1}{2}, \qquad \widetilde{u}_{2|d_1=1} = \frac{(1-\phi)^2 \mathrm{I}[\phi<1]}{2}. \tag{B-7}$$

Note here that these perceptions are accurate since the consumer would correctly anticipate her ability to infer her $t = 1$ consumption.

Given v_1 is uniformly distributed between 0 and 1, we can compute the $t = 1$ consumption probability and the probability of consuming in both periods as

$$D_1 = \frac{1+(1-\phi)^2 \mathrm{I}[\phi<1]}{2}, \tag{B-8}$$

$$D_{12} = \frac{(1-\phi)(1+(1-\phi)^2)\mathrm{I}[\phi<1]}{2}. \tag{B-9}$$

Given $p = \widetilde{V} = \frac{1+(D_1)^2}{2}$ and $\Pi = p + D_{12}\phi$, we can then derive benchmark profits:

$$\Pi(\phi) = \begin{cases} 1 - \frac{(8(1-\phi)+3\phi^2)\phi^2}{8}, & \phi < 1, \\ \frac{5}{8}, & \phi \geq 1. \end{cases} \tag{B-10}$$

By inspection, we can see $\Pi(\phi) < 1$ for all $\phi > 0$. Thus, in the absence of consumption tracking, the zero-penalty equilibrium with $\phi^* = 0$ and $p^* = 1$ must arise when consumers can infer their $t = 1$ consumption from their memory of v_1.

Analysis with Consumption Tracking

Since consumers can, under this variation of the model, perfectly infer their $t = 1$ consumption, there will be no demand for consumption tracking. Formally, $k_\tau = \widetilde{k}_\tau = 0$, implying $D_\tau = \widetilde{D}_\tau = 0$. That is, consumers will never use nor expect to use consumption tracking regardless of the size of the penalty fee ϕ. As a result, the optimal contract will still be given by $\phi^* = \phi^*_{NT} = 0$ and $p^* = p^*_{NT} = 1$, while all aspects of equilibrium behavior will be unaffected by the availability of consumption tracking.

B.1.3. Inferring Past Consumption from $\widetilde{\alpha}$

As our next potential formulation of consumer forgetfulness, we now suppose that the consumer cannot remember her $t = 1$ consumption d_1 (or her $t = 1$ valuation v_1), but can infer her $t = 1$ consumption probability based on her perceived forgetfulness $\widetilde{\alpha}$ on which her $t = 1$ consumption decision was based.

Benchmark Analysis

Under this formulation of consumer forgetfulness, the consumer's belief, held at $t = 2$, of her $t = 1$ consumption is given by the unconditional $t = 1$ consumption probability D_1. As a result, in the absence of consumption tracking, the consumer consumes at $t = 2$ if and only if her $t = 2$ valuation exceeds the penalty fee weighted by her expectation of having consumed at $t = 1$, i.e.

if and only if $v_2 > D_1\phi$. Given v_2 is uniformly distributed between 0 and 1, the consumer's $t=2$ consumption probability is given by

$$D_2 = \max\{1 - D_1\phi, 0\}. \tag{B-11}$$

Next, we can express the consumer's expected $t=2$ utilities conditional on her $t=1$ choice as

$$\widetilde{u}_{2|d_1=0} = D_2\Big(1 - \frac{D_2}{2}\Big), \qquad \widetilde{u}_{2|d_1=1} = D_2\Big(1 - \frac{D_2}{2} - \phi\Big). \tag{B-12}$$

Given v_1 is uniformly distributed between 0 and 1, we can compute the $t=1$ consumption probability as:

$$D_1 = 1 - (1 - D_1\phi)\phi = \frac{1}{1+\phi}, \tag{B-13}$$

which implies $D_2 = \max\{1 - D_1\phi, 0\} = \frac{1}{1+\phi}$ from (B-11) as well as

$$D_{12} = D_1 D_2 = \frac{1}{(1+\phi)^2}. \tag{B-14}$$

Noting $\widetilde{V} = V = D_1(\mathrm{E}[v_1|d_1=1] + \widetilde{u}_{2|d_1=1}) + (1-D_1)\widetilde{u}_{2|d_1=0}$, we can then verify that a consumer's initial valuation of a subscription reduces to

$$\widetilde{V} = \frac{1}{1+\phi}. \tag{B-15}$$

Since the firm optimally sets $p = \widetilde{V}$, we can compute the firm's benchmark profits under this formulation of consumer forgetting as

$$\Pi = \frac{1}{1+\phi} + \frac{1}{(1+\phi)} \cdot \phi = \frac{1+2\phi}{(1+\phi)^2} = 1 - \Big(\frac{\phi}{1+\phi}\Big)^2. \tag{B-16}$$

By inspection, we can see $\Pi(\phi) < 1$ for all $\phi > 0$. Thus, in the absence of consumption tracking, the zero-penalty equilibrium with $\phi^* = 0$ and $p^* = 1$ must arise when consumers infer their $t=1$ consumption probability from $\widetilde{\alpha}$. Note, however, that $\widetilde{\alpha}$ does not appear in the above expressions, nor does it affect the optimal contract. This makes sense considering the true forgetfulness parameter α no longer matters if consumers base their $t=2$ consumption decisions on their unconditional $t=1$ consumption probability, D_1; since α has no meaning under this formulation of consumer forgetting, the consumer's initial perception of α, i.e. $\widetilde{\alpha}$, also becomes irrelevant.

Analysis with Consumption Tracking

With consumption tracking, a consumer's $t=1$ perception of her $t=2$ utilities conditional on not using and on using consumption tracking are given by

$$\widetilde{u}_{2|d_\tau=0} = D_2\Big(1 - \frac{D_2}{2} - D_1\phi\Big), \qquad \widetilde{u}_{2|d_\tau=1} = \begin{cases} D_1 + \dfrac{(1-D_1)(1-\phi)^2}{2} - k, & \phi \le 1, \\ D_1 - k, & \phi > 1, \end{cases} \tag{B-17}$$

which allows us to compute the threshold monitoring cost (i.e., the cost of using the consumption tracking technology to remember past consumption):

$$k_\tau = \begin{cases} D_1 + \dfrac{(1-D_1)(1-\phi)^2}{2} - D_2\left(1 - \frac{D_2}{2} - D_1\phi\right), & \phi \le 1, \\ D_1 - D_2\left(1 - \frac{D_2}{2} - D_1\phi\right), & \phi > 1 \end{cases} \tag{B-18}$$

If the consumer does use consumption tracking, $d_\tau = 1$, her $t = 1$ expectations of her $t = 2$ utilities will be given by

$$\widetilde{u}_{2|d_1=0} = \frac{1}{2} - k, \qquad \widetilde{u}_{2|d_1=1} = \frac{(1-\phi)^2 \mathrm{I}[\phi<1]}{2} - k. \tag{B-19}$$

In this case, $\widetilde{u}_{2|d_1=0} - \widetilde{u}_{2|d_1=1}$ will be the same as in the benchmark analysis, implying $D_1 = \frac{1}{1+\phi}$ still holds. Her valuation of a subscription will then be given by $\widetilde{V} = V = D_1\left(1 - \frac{D_1}{2} + \widetilde{u}_{2|d_1=1}\right) + (1-D_1)\widetilde{u}_{2|d_1=0}$, which reduces to

$$\widetilde{V}(\phi|d_\tau = 1) = \begin{cases} 1 - \frac{(2+2\phi-\phi^2)\phi}{2(1+\phi)^2} - k, & \phi \le 1, \\ 1 - \frac{1+\phi+\phi^2}{2(1+\phi)^2} - k, & \phi > 1. \end{cases} \tag{B-20}$$

Noting that $D_{12} = D_1 D_{2|d_1=1}$, $D_1 = \frac{1}{1+\phi}$, and $D_2 = \max\{1-\phi, 0\}$, we can verify that total profits $\Pi = \widetilde{V} + D_{12}\phi$ are given by:

$$\Pi(\phi|d_\tau = 1) = \begin{cases} 1 - \frac{(2+\phi)\phi^2}{2(1+\phi)^2} - k, & \phi \le 1, \\ 1 - \frac{1+\phi+\phi^2}{2(1+\phi)^2} - k, & \phi > 1 \end{cases} \tag{B-21}$$

Observe $\Pi < 1$ in both cases. Furthermore, the case with $d_\tau = 0$ is identical to the benchmark analysis without consumption tracking. This implies that, even when consumption tracking is available, the optimal contract still entails $\phi^* = 0$ and $p^* = 1$ under the present formulation of consumer forgetfulness in which consumers base their $t = 2$ consumption decisions on their unconditional $t = 1$ consumption probabilities through their memory of $\widetilde{\alpha}$. In other words, consumption tracking is (once again) irrelevant.

B.1.4. Memory States with Bayesian Updating

As our next potential formulation of consumer forgetfulness, we now consider a variant of our model in which a consumer may enter two possible memory states at $t = 2$: *recall* and *not recall*. Formally, we let $r \in \{Y, N\}$ denote whether a consumer recalls consuming at $t = 1$ upon their arrival at $t = 2$, where $r = Y$ corresponds to "yes" (as in the consumer does recall) and $r = N$ corresponds to "no" (the consumer does not recall). If the consumer consumes at $t = 1$, the probability of entering the "no" state is presumed to be α. If the consumer does not consume at $t = 1$, the probability of entering the "no" state is presumed to be 1. Formally,

$$\Pr[r = Y] = (1-\alpha)d_1, \qquad \Pr[r = N] = 1 - (1-\alpha)d_1.$$

Given a consumer's memory state, the consumer uses Bayes' Rule to form expectations on her probability of having consumed at $t = 1$. These expectations (conditional on the realized state) are:

$$\mathrm{E}[d_1|r=Y] = 1, \qquad \mathrm{E}[d_1|r=N] = \frac{\alpha D_1}{\alpha D_1 + 1 - D_1}.$$

Note, here that the consumer is certain she consumed at $t = 1$ when she enters memory state $r = Y$ because memory state $r = Y$ can only be reached following consumption at $t = 1$.

Benchmark Analysis

Without consumption tracking, a consumer's probability of consuming at $t = 2$ conditional on her memory state is then

$$\begin{aligned} D_{2|r=Y} &= \Pr[v_2 \geq \phi] = \max\{1-\phi, 0\}, \\ D_{2|r=N} &= \Pr\left[v_2 \geq \frac{\alpha D_1 \phi}{\alpha D_1 + 1 - D_1}\right] = \max\left\{1 - \frac{\alpha D_1 \phi}{\alpha D_1 + 1 - D_1}, 0\right\}. \end{aligned} \tag{B-22}$$

Noting $D_{2|d_1=0} = D_{2|r=N} =$ and $D_{2|d_1=1} = \alpha D_{2|r=N} + (1-\alpha) D_{2|r=Y}$, we can compute the $t = 2$ consumption probabilities conditional on the $t = 1$ choice as

$$\begin{aligned} D_{2|d_1=0} &= \max\left\{1 - \frac{\alpha D_1 \phi}{\alpha D_1 + 1 - D_1}, 0\right\}, \\ D_{2|d_1=1} &= \begin{cases} (1-\alpha)(1-\phi) + \frac{\alpha(1-D_1(1-\alpha(1-\phi))}{1-(1-\alpha)D_1}, & \phi \leq 1 \\ \frac{\alpha(1-D_1(1-\alpha(1-\phi))}{1-(1-\alpha)D_1}, & 1 < \phi < \frac{1-(1-\alpha)D_1}{\alpha D_1}, \\ 0, & \phi \geq \frac{1-(1-\alpha)D_1}{\alpha D_1}. \end{cases} \end{aligned} \tag{B-23}$$

Next, we would like to compute the consumer's $t = 1$ perception of her expected $t = 2$ utility conditional on her $t = 1$ choice (and $t = 2$ memory state). Using $\widetilde{u}_{2|d_1=0} = \Pr[d_2 = 1|r = N]\mathrm{E}[v_2|d_2 = 1, r = N]$, $\widetilde{u}_{2|d_1=1,r=Y} = \Pr[d_2 = 1|r = Y](\mathrm{E}[v_2|d_2 = 1, r = Y] - \phi)$, and $\widetilde{u}_{2|d_1=1,r=N} = \Pr[d_2 = 1|r = N](\mathrm{E}[v_2|d_2 = 1, r = N] - \phi)$, we can derive

$$\widetilde{u}_{2|d_1=0} = \begin{cases} \frac{1}{2} \cdot \left(1 - \left(\frac{\widetilde{\alpha} D_1 \phi}{\widetilde{\alpha} D_1 + 1 - D_1}\right)^2\right), & \phi < \frac{1-(1-\widetilde{\alpha})D_1}{\widetilde{\alpha} D_1}, \\ 0, & \phi \geq \frac{1-(1-\widetilde{\alpha})D_1}{\widetilde{\alpha} D_1}. \end{cases} \tag{B-24}$$

$$\widetilde{u}_{2|d_1=1,r=Y} = \begin{cases} \frac{(1-\phi)^2}{2}, & \phi < 1, \\ 0, & \phi \geq 1. \end{cases} \tag{B-25}$$

$$\widetilde{u}_{2|d_1=1,r=N} = \begin{cases} \frac{(1-D_1(1-\widetilde{\alpha}(1-\phi)))((1-(1-\widetilde{\alpha})D_1)(1-2\phi)+\widetilde{\alpha} D_1 \phi)}{2(1-(1-\widetilde{\alpha})D_1)^2}, & \phi < \frac{1-(1-\widetilde{\alpha})D_1}{\widetilde{\alpha} D_1}, \\ 0, & \phi \geq \frac{1-(1-\widetilde{\alpha})D_1}{\widetilde{\alpha} D_1}. \end{cases} \tag{B-26}$$

Noting $\widetilde{u}_{2|d_1=1} = (1-\widetilde{\alpha})\widetilde{u}_{2|d_1=1,r=Y} + \widetilde{\alpha}\widetilde{u}_{2|d_1=1,r=N}$, the latter two equations imply

$$\widetilde{u}_{2|d_1=1} = \begin{cases} \frac{(1-\widetilde{\alpha})(1-\phi)^2}{2} + \frac{\widetilde{\alpha}(1-D_1(1-\widetilde{\alpha}(1-\phi)))((1-(1-\widetilde{\alpha})D_1)(1-2\phi)+\widetilde{\alpha}D_1\phi)}{2(1-(1-\widetilde{\alpha})D_1)^2}, & \phi < 1 \\ \frac{\widetilde{\alpha}(1-D_1(1-\widetilde{\alpha}(1-\phi)))((1-(1-\widetilde{\alpha})D_1)(1-2\phi)+\widetilde{\alpha}D_1\phi)}{2(1-(1-\widetilde{\alpha})D_1)^2}, & 1 \le \phi < \frac{1-(1-\widetilde{\alpha})D_1}{\widetilde{\alpha}D_1}, \\ 0, & \phi \ge \frac{1-(1-\widetilde{\alpha})D_1}{\widetilde{\alpha}D_1}. \end{cases} \tag{B-27}$$

It is then readily verifiable that a consumer's willingness-to-pay for a subscription is given by

$$\widetilde{V} = \widetilde{u}_{2|d_1=0} + \frac{D_1^2}{2} = \frac{1}{2}\left(\left(1-\left(\frac{\widetilde{\alpha}D_1\phi}{\widetilde{\alpha}D_1+1-D_1}\right)^2\right)\cdot \mathrm{I}\left[\phi < \frac{1-(1-\widetilde{\alpha})D_1}{\widetilde{\alpha}D_1}\right] + D_1^2\right). \tag{B-28}$$

Remark 6. *In the memory state model without consumption tracking, the optimal contract is given by $\phi_{NT}^* = 0$ and $p_{NT}^* = 1$.*

To verify the optimality of $\phi_{NT}^* = 0$ and $p_{NT}^* = 1$, first suppose that the actual optimal contract entails $\phi_{NT}^* > 0$. To demonstrate that $\phi_{NT}^* > 0$ cannot in fact hold, we need to consider two cases.

Case 1. $\phi_{NT}^* > 1$

Noting $\Pi = \widetilde{V} + D_{12}\phi$, we can use our above work to verify the following upper bound on Π given $\phi > 1$:

$$\Pi \le \frac{1}{2}\left(1-\left(\frac{\widetilde{\alpha}D_1\phi}{1-(1-\widetilde{\alpha})D_1}\right)^2 + D_1^2\right) + D_1\phi\alpha\left(1-\frac{\alpha D_1\phi}{1-(1-\alpha)D_1}\right) \tag{B-29}$$

It is then readily verifiable that the right-side expression is decreasing in $\widetilde{\alpha}$. This ensures that Π is bounded above by the value of the expression when evaluated at $\widetilde{\alpha} = 0$. That is,

$$\Pi \le \frac{1}{2}\left(1+D_1^2\right) + D_1\phi\alpha\left(1-\frac{\alpha D_1\phi}{1-(1-\alpha)D_1}\right) \tag{B-30}$$

Letting $A(\phi) \equiv \frac{1}{2}\left(1+D_1^2\right) + D_1\phi\alpha\left(1-\frac{\alpha D_1\phi}{1-(1-\alpha)D_1}\right)$, we can then verify that $\phi = \phi^A \equiv \frac{1-(1-\alpha)D_1}{2D_1\alpha}$ solves $A'(\phi) = 0$ and $A''(\phi) = -\frac{2\alpha^2 D_1^2}{1-(1-\alpha)D_1} < 0$, while $\phi^A > 1$ if and only $D_1 < \frac{1}{1+\alpha}$. This implies that, given $\phi > 1$, $A(\phi)$ is maximized at $\phi = \phi^A$ if $D_1 \le \frac{1}{1+\alpha}$ and at $\phi = 1$ if $D_1 > \frac{1}{1+\alpha}$. For the case with $D_1 \le \frac{1}{1+\alpha}$, we can compute $A(\phi^A) = \frac{3+2D_1^2-(1-\alpha)D_1}{4}$. Since $\frac{\partial^2 A(\phi^A)}{\partial D_1^2} = 1 > 0$, $A(\phi^A|D_1)$ is maximized at either $D_1 = 0$ or $D_1 = \frac{1}{1+\alpha}$. Next, $A(\phi^A|D_1 = 0) = \frac{3}{4} < 1$ and $A\left(\phi^A|D_1 = \frac{1}{1+\alpha}\right) = 1 - \frac{\alpha}{2(1+\alpha)^2} < 1$. For the case with $D_1 \le \frac{1}{1+\alpha}$, we can compute $A(1) = \frac{1}{2}\left(1+D_1^2+\frac{2\alpha(1-D_1)D_1}{1-(1-\alpha)D_1}\right)$ and $\frac{\partial A(1|\alpha)}{\partial\alpha} = \frac{(1-D_1)^2 D_1}{(1-(1-\alpha)D_1)^2} > 0$, implying $A(1|\alpha)$ is maximized at $\alpha = 1$. Next, observe $A(1|\alpha = 1) = \frac{1+2D_1-D_1^2}{2}$ and $\frac{\partial A(1|\alpha=1,D_1)}{\partial D_1} = 1-D_1 \ge 0$. Thus, $A(1|\alpha = 1, D_1)$ is maximized at $D_1 = 1$, while $A(1|\alpha = 1, D_1 = 1) = 1$. Since Π is bounded above by $A(\phi)$, $\Pi > 1$ cannot hold with $\phi > 1$. Since $\Pi = 1$ with $\phi = 0$ and $p = 1$, the optimal contract cannot entail $\phi_{NT}^* > 1$.

Case 2. $\phi^*_{NT} \leq 1$

Noting $\Pi = \widetilde{V} + D_{12}\phi$, we can use our above work to verify the following upper bound on Π given $\phi \leq 1$:

$$\Pi \leq \frac{1}{2}\left(1 - \left(\frac{\widetilde{\alpha} D_1 \phi}{1-(1-\widetilde{\alpha})D_1}\right)^2 + D_1^2\right) + D_1\phi\left((1-\alpha)(1-\phi) + \alpha\left(1 - \frac{\alpha D_1 \phi}{1-(1-\alpha)D_1}\right)\right). \quad \text{(B-31)}$$

It is then readily verifiable that the right-side of (B-31) is decreasing in $\widetilde{\alpha}$. This ensures that Π is bounded above by the value of the expression when evaluated at $\widetilde{\alpha} = 0$. That is,

$$\Pi \leq \frac{1+D_1^2}{2} + D_1\phi\left((1-\alpha)(1-\phi) + \alpha\left(1 - \frac{\alpha D_1 \phi}{1-(1-\alpha)D_1}\right)\right). \quad \text{(B-32)}$$

Next, given $\phi < 1$, $\widetilde{\alpha} = 0$ implies $\widetilde{u}_{2|d_1=0} = \frac{1}{2}$ and $\widetilde{u}_{2|d_1=1} = \frac{(1-\phi^2)}{2}$, implying $D_1 = \Pr[v_1 > \widetilde{u}_{2|d_1=0} - \widetilde{u}_{2|d_1=1}] = 1 - \left(\frac{1}{2} - \frac{(1-\phi^2)}{2}\right) = \frac{1+(1-\phi)^2}{2}$. Substituting this expression into (B-32), we can see $\Pi \leq B(\phi) \equiv \frac{1}{2}\big(1 + \big(\frac{1+(1-\phi)^2}{2}\big)^2 + (1+(1-\phi)^2)\big(\phi(1-\phi) + \frac{\alpha\phi^2(1-(1-\phi)^2)}{1+\alpha-(1-\alpha)(1-\phi)^2}\big)\big)$. Next, $\frac{\partial B(\phi)}{\partial \alpha} = \frac{(1+(1-\phi)^2)\phi^2(1-(1-\phi)^2)^2}{2(1+\alpha-(1-\alpha)(1-\phi)^2)^2} > 0$, implying $B(\phi|\alpha)$ is maximized at $\alpha = 1$. In turn, $B(\phi|\alpha = 1) = (1 + (1-\phi)^2)\big(\phi(1-\phi) + \frac{\phi^2(1-(1-\phi)^2)}{2}\big)$, while $B'(\phi|\alpha = 1) = (1-\phi)(1-\phi(1-\phi))(2(1-\phi)^2 + \phi^2) > 0$ given $\phi < 1$. Therefore, $B(\phi|\alpha = 1) < B(1|\alpha = 1) = 1 - \frac{4\phi^2(1-\phi)^2(2+\phi^2)+\phi^4(3-2\phi^2)}{8} < 1$ for $0 < \phi < 1$. Thus, since $\Pi < B(1|\alpha = 1) < 1$ for $0 < \phi < 1$, yet $\Pi = 1$ with $\phi = 0$ and $p = 1$, the optimal contract cannot entail $0 < \phi^*_{NT} < 1$. Furthermore, with case 1, considered above, we can now see that $\phi^*_{NT} = 0$ and $p^*_{NT} = 1$ must hold in the benchmark equilibrium under the present formulation of consumer forgetfulness.

Analysis with Consumption Tracking

With consumption tracking, a consumer will only potentially use consumption tracking if she arrives at $t = 2$ in memory state $r = N$ (since a consumer can perfectly infer her $t = 1$ consumption when $r = Y$). Using $\widetilde{u}_{2|d_\tau=0,r=N} = D_2\left(1 - \frac{D_2}{2} - \mathrm{E}[d_1|r = N]\phi\right)$ and $\widetilde{u}_{2|d_\tau=1,r=N} = \mathrm{E}[d_1|r = N]\widetilde{u}_{2|d_1=1} + (1 - \mathrm{E}[d_1|r = N])\widetilde{u}_{2|d_1=0} - k$, a consumer's perceived $t = 2$ utilities conditional on not using and on using consumption tracking in memory state $r = N$ are given by

$$\widetilde{u}_{2|d_\tau=0,r=N} = D_2\left(1 - \frac{D_2}{2} - \mathrm{E}[d_1|r = N]\phi\right) = \frac{1}{2}\big(1 - \frac{\alpha D_1 \phi}{1-(1-\alpha)D_1}\big)^2, \quad \text{(B-33)}$$

$$\widetilde{u}_{2|d_\tau=1} = \frac{1}{2}\left(1 - \frac{\alpha D_1 \phi(1-(1-\phi)^2 \mathrm{I}[\phi<1])}{1-(1-\alpha)D_1}\right) - k \quad \text{(B-34)}$$

This implies the threshold monitoring cost is given by:

$$k_\tau = \begin{cases} \frac{\alpha D_1}{2(1-(1-\alpha)D_1)}\left(1 - \phi + \frac{(1-D_1)\phi^2}{1-(1-\alpha)D_1}\right), & \phi \leq 1, \\ \frac{\alpha\phi D_1(1-D_1(1+\alpha(\phi-1)))}{2(1-(1-\alpha)D_1)^2}, & \phi > 1 \end{cases} \quad \text{(B-35)}$$

Remark 7. *In the memory state model with consumption tracking, the optimal contract is (still) given by* $\phi_{NT}^* = 0$ *and* $p^* = 1$.

To verify the optimality of $\phi_{NT}^* = 0$ and $p^* = 1$, first note that the potential equilibria with consumption tracking can be divided into four types: a planned use equilibrium (with $\widetilde{D}_\tau > 0$ and $D_\tau > 0$), an unplanned nonuse equilibrium (with $\widetilde{D}_\tau > 0 = D_\tau$), a planned nonuse equilibrium (with $\widetilde{D}_\tau = D_\tau = 0$), and an unplanned use equilibrium (with $\widetilde{D}_\tau = 0 < D_\tau$). Note, however, that the firm's profits in a planned nonuse equilibrium will take the same form as in the case without consumption tracking, and as we've already established, $\Pi > 1$ cannot hold under this scenario. Next, in an unplanned use equilibrium, it is readily verifiable that the applicable profit expression only differs from that given in our benchmark analysis through the difference in $D_{2|d_1=1}$ resulting from the consumer's unplanned use of consumption tracking, and that $D_{2|d_1=1}$ is lower (all else equal) with $d_\tau = 1$, ensuring profits in an unplanned use equilibrium cannot exceed the equilibrium profits without consumption tracking. Thus, $\Pi > 1$ cannot hold in this case either. This leaves us with the first two cases to consider.

Case 1. Planned Use of Consumption Tracking ($\widetilde{D}_\tau > 0$ and $D_\tau > 0$)

A tedious yet straightforward derivation of the firm's equilibrium profits (following the same steps used in prior derivations) under planned use of consumption tracking yields:

$$\Pi = \begin{cases} 1-(1-\phi)\phi^2 - \widetilde{\alpha}k - \frac{k(1-\widetilde{\alpha})(\phi^2 - k(1-\widetilde{\alpha}))}{2} - \frac{3\phi^4}{8}, & \phi \le 1,\ k \le \frac{1-(1-\phi)^2}{2(1-\widetilde{\alpha})}, \\ 1-k+\phi(1-\phi), & \phi \le 1,\ k > \frac{1-(1-\phi)^2}{2(1-\widetilde{\alpha})}, \\ \frac{1}{2}\left(1-2k+\left(\frac{1}{2}+(1-\widetilde{\alpha})k\right)^2\right), & \phi > 1,\ k \le \frac{1}{2(1-\widetilde{\alpha})}, \\ 1-k, & \phi > 1,\ k > \frac{1}{2(1-\widetilde{\alpha})}. \end{cases} \tag{B-36}$$

It is readily verifiable that the third and fourth expressions in (B-36) are decreasing in k and satisfy $\Pi \le 1$ at $k = 0$, implying $\Pi \le 1$ must hold for all k in these cases. It is also verifiable that the first expression is decreasing in $\widetilde{\alpha}$, while its value at $\widetilde{\alpha} = 0$ satisfies $\Pi = 1-(1-\phi)\phi^2 - \frac{k(\phi^2-k)}{2} - \frac{3\phi^4}{8} \le 1$ given its restriction $k \le \frac{1-(1-\phi)^2}{2(1-\widetilde{\alpha})}$, thus ensuring $\Pi \le 1$ for all $\widetilde{\alpha}$ in this case. Thus, we only need to consider the second case in (B-36). In turn, it is readily verifiable that the restriction $k \le \frac{1-(1-\phi)^2}{2(1-\widetilde{\alpha})}$ in this case implies $\widetilde{u}_{2|d_1=0} \le \widetilde{u}_{2|d_1=1}$, and thus $D_1 = 1$. However, $D_1 = 1$ implies $\widetilde{k}_\tau = k_\tau = 0$, thus precluding the possibility of planned use. Therefore, this case cannot apply in equilibrium. Thus, given our work above, a planned use equilibrium with $\phi > 0$ must entail $\Pi \le 1$. However, $\phi = 0$ and $p = 1$ yields $\Pi = 1$, thus ensuring that a planned use equilibrium cannot arise under the current formulation of consumer forgetting.

Case 2. Unplanned Nonuse of Consumption Tracking ($\widetilde{D}_\tau > 0 = D_\tau$)

A tedious yet straightforward derivation of the firm's equilibrium profits (again following the same steps used in prior derivations) under unplanned nonuse of consumption tracking yields:

$$\Pi = \begin{cases} \frac{1-2k}{2} + \frac{\phi(1-(1-\phi)^2+2(1-\widetilde{\alpha})k)(\alpha\phi(1-(1-\phi)^2-2(1-\widetilde{\alpha})k))}{2(1-(1-\phi)^2-2(1-\widetilde{\alpha})(1-\alpha)k+\alpha(1+(1-\phi)^2))} \\ \quad + \frac{(1+(1-\phi)^2+2(1-\widetilde{\alpha})k)^2}{8} + \frac{\phi(1-\phi)(1-(1-\phi)^2+2(1-\widetilde{\alpha})k)}{2}, & \phi < 1,\ k \le \frac{1-(1-\phi)^2}{2(1-\widetilde{\alpha})}; \\ 1-k+\phi(1-\phi), & \phi < 1,\ k > \frac{1-(1-\phi)^2}{2(1-\widetilde{\alpha})}; \\ \frac{1-2k}{2} - \frac{\alpha\phi(1+2(1-\widetilde{\alpha})k)(1+2(1-\widetilde{\alpha})k)}{2(1-2(1-\widetilde{\alpha})(1-\alpha)k+\alpha)} \\ \quad + \frac{(1+2(1-\widetilde{\alpha})k)^2}{8} + \frac{\alpha\phi(1+2(1-\widetilde{\alpha})k)}{2}, & 1 \le \phi < \frac{1-2(1-\widetilde{\alpha})(1-\alpha)k+\alpha}{\widetilde{\alpha}(1+2(1-\widetilde{\alpha})k)},\ k \le \frac{1}{2(1-\widetilde{\alpha})}; \\ \frac{1-2k}{2} + \frac{(1+2(1-\widetilde{\alpha})k)^2}{8}, & \phi \ge \max\{1, \frac{1-2(1-\widetilde{\alpha})(1-\alpha)k+\alpha}{\widetilde{\alpha}(1+2(1-\widetilde{\alpha})k)}\},\ k \le \frac{1}{2(1-\widetilde{\alpha})}; \\ 1-k, & \phi \ge 1,\ k > \frac{1}{2(1-\widetilde{\alpha})}. \end{cases} \tag{B-37}$$

It is then readily verifiable that the first expression in (B-37) is decreasing in $\widetilde{\alpha}$ and satisfies $\Pi = 1-k-\phi^2+2\phi^3-\frac{15\phi^4}{8}+\phi^5-\frac{\phi^6}{4} < 1$ given $k > 0$ and $0 < \phi < 1$. Thus, $\Pi < 1$ must hold in this case. It is similarly verifiable that the second expression in (B-37) also yields $\Pi < 1$ given $\phi < 1$ and $k > \frac{1-(1-\phi)^2}{2(1-\widetilde{\alpha})}$, along with $0 < \widetilde{\alpha} < 1$. Next, it is verifiable that the third and fourth expressions in (B-37) are both decreasing in k, with $\Pi = \frac{5}{8} < 1$ when evaluated at $k = 0$ given $\phi < \frac{1-2(1-\widetilde{\alpha})(1-\alpha)k+\alpha}{\widetilde{\alpha}(1+2(1-\widetilde{\alpha})k)}$ in the third expression and $\Pi = \frac{5}{8} + \frac{\alpha\phi(1-\alpha(\phi-1))}{2(1+\alpha)}$ when evaluated at $k = 0$ given $\phi \ge \frac{1-2(1-\widetilde{\alpha})(1-\alpha)k+\alpha}{\widetilde{\alpha}(1+2(1-\widetilde{\alpha})k)}$ in the fourth expression. Meanwhile, $\Pi = \frac{5}{8} + \frac{\alpha\phi(1-\alpha(\phi-1))}{2(1+\alpha)}$ is maximized at $\phi = \frac{1+\alpha}{2\alpha}$, at which $\Pi = \frac{6+\alpha}{8} < 1$. Thus, $\Pi < 1$ must always hold in the third and fourth expressions of (B-37). By inspection, $\Pi = 1-k \le 1$ in the last case as well. Thus, $\Pi \le 1$ must hold for all cases in (B-37). This, along with our prior work, ensures that $\phi^* = 0$ and $p^* = 1$ indeed holds (giving $\Pi = 1$) under the present formulation of consumer forgetting.

B.1.5. Symmetric Forgetting

In our primary model, consumers may be forgetful of past consumption but *not* past abstinence. Thus, when $t = 2$ arrives, a consumer who did not consume at $t = 1$ perfectly remembers her past abstinence, or put differently, she is *certain* that she did not consume at $t = 1$. As discussed in Section 2 (see footnote 4 and surrounding discussion), this "asymmetric forgetting" formulation is compatible with empirical evidence that individuals are much more likely to forget an event that did occur than to falsely recall an event that did not occur. That said, it is still interesting to consider an alternate formulation. In this appendix, we consider such a formulation whereby consumers are *symmetrically* forgetful of past consumption as well as past abstinence. Namely, a consumer who did not consume at $t = 1$ may now be forgetful of past abstinence in that, when $t = 2$ arrives, she believes she did not consume at $t = 1$ with probability $1-\alpha$ and that she did consume at $t = 1$ with probability α. Thus, regardless of whether the consumer consumed ($d_1 = 1$) or abstained ($d_1 = 0$) at $t = 1$, she assigns probability α to the incorrect belief and probability $1-\alpha$ to the correct belief.

As we recall, in our original model a consumer's memory was entirely uninformative if $\alpha = 1$. That is, with $\alpha = 1$, her perceived likelihood of having consumed at $t = 1$ was the same (in this case zero) regardless of whether she consumed or abstained. With symmetric forgetting, however, a consumer's memory of past consumption is entirely uninformative if $\alpha = \frac{1}{2}$. That is, with $\alpha = \frac{1}{2}$, her perceived likelihood of having consumed at $t = 1$ is the same (in this case $\frac{1}{2}$) regardless of whether she consumed or abstained. Meanwhile, $\alpha > \frac{1}{2}$ would represent "negative memory," in which a consumer's belief regarding her past consumption is *negatively* correlated with her true consumption. To avoid such nonsensical cases, we restrict $\alpha \leq \frac{1}{2}$ in our analysis with symmetric forgetting.

We now provide a more detailed characterization of consumers' consumption decisions in the model with symmetric forgetting. To start, in the symmetric forgetting model without consumption tracking, a consumer's probability of consuming at $t = 2$ conditional on $d_1 = 0$ is

$$D_2(\phi|d_1 = 0) = \Pr[v_2 \geq \alpha\phi] = \max\{1-\alpha\phi, 0\}. \tag{B-38}$$

If $d_1 = 1$, then at $t = 2$ the consumer believes there is $1-\alpha$ probability that $d_1 = 1$, implying the expected penalty fee from consuming at $t = 2$ is $(1-\alpha)\phi$. Since v_2 is uniformly distributed between 0 and 1, the consumer's probability of consuming at $t = 2$ conditional on $d_1 = 1$ is

$$D_2(\phi|d_1 = 1) = \Pr[v_2 \geq (1-\alpha)\phi] = \max\{1-(1-\alpha)\phi, 0\}. \tag{B-39}$$

Next, we would like to compute the consumer's $t = 1$ perception of her expected $t = 2$ utility conditional on her $t = 1$ choice. If $d_1 = 0$, the consumer knows she will be free to consume at $t = 2$ without risk of incurring the penalty. This implies the following $t = 1$ expectation of her $t = 2$ utility, conditional on $d_1 = 0$:

$$\widetilde{u}_2(\phi|d_1 = 0) = \Pr[v_2 \geq \widetilde{\alpha}\phi] \cdot \mathrm{E}[v_2|v_2 \geq \widetilde{\alpha}\phi] = \tfrac{1-\widetilde{\alpha}^2\phi^2}{2} \cdot \mathrm{I}\big[\phi < \tfrac{1}{\widetilde{\alpha}}\big]. \tag{B-40}$$

If $d_1 = 1$, the consumer expects to consume at $t = 2$ (and thus incur the penalty fee) if and only if $v_2 \geq (1-\widetilde{\alpha})\phi$. Thus, her $t = 1$ expectation of her $t = 2$ utility conditional on $d_1 = 1$ is:

$$\widetilde{u}_2(\phi|d_1 = 1) = \Pr[v_2 \geq (1-\widetilde{\alpha})\phi] \cdot \mathrm{E}[v_2-\phi|v_2 \geq (1-\widetilde{\alpha})\phi] = \tfrac{(1-\phi)^2-\widetilde{\alpha}^2\phi^2}{2} \cdot \mathrm{I}\big[\phi < \tfrac{1}{1-\widetilde{\alpha}}\big]. \tag{B-41}$$

Using (B-38), (B-39), (B-40), and (B-41) while noting $D_1(\phi) = \Pr[v_1 + \widetilde{u}_2(\phi|d_1 = 1) \geq \widetilde{u}_2(\phi|d_1 = 0)]$ and $D_2(\phi) = D_1(\phi) \cdot D_2(\phi|d_1 = 1) + (1 - D_1(\phi)) \cdot D_2(\phi|d_1 = 0)$, we can now derive the unconditional probabilities of consuming at each $t = 1, 2$:

$$D_1(\phi) = \begin{cases} \frac{1+(1-\phi)^2}{2}, & \phi \leq \frac{1}{1-\widetilde{\alpha}}, \\ \frac{1+\widetilde{\alpha}^2\phi^2}{2}, & \frac{1}{1-\widetilde{\alpha}} < \phi \leq \frac{1}{\widetilde{\alpha}}, \\ 1, & \phi > \frac{1}{\widetilde{\alpha}}, \end{cases} \tag{B-42}$$

$$D_2(\phi) = \begin{cases} 1 - \alpha\phi - \frac{(1+(1-\phi)^2)(1-2\alpha)\phi}{2}, & \phi \leq \frac{1}{1-\widetilde{\alpha}}, \\ 1 - \alpha\phi - \frac{(1+\widetilde{\alpha}^2\phi^2)(1-2\alpha)\phi}{2}, & \frac{1}{1-\widetilde{\alpha}} < \phi \leq \frac{1}{1-\alpha}, \\ \frac{(1-\alpha\phi)(1-\widetilde{\alpha}^2\phi^2)}{2}, & \frac{1}{1-\alpha} < \phi \leq \frac{1}{\alpha}, \\ 0, & \phi > \frac{1}{\alpha}. \end{cases} \tag{B-43}$$

Next, we can express a consumer's maximum $t = 0$ willingness-to-pay for a subscription as $\widetilde{V}(\phi) = \widetilde{u}_2(\phi|d_1 = 0) + D_1(\phi) \cdot (\mathrm{E}[v_1|v_1 > \Delta] - \Delta)$ given $\Delta \equiv \widetilde{u}_2(\phi|d_1 = 0) - \widetilde{u}_2(\phi|d_1 = 1) \geq 0$. Since v_1 is uniformly distributed between 0 and 1, we can verify that a consumer's willingness-to-pay reduces to:

$$\widetilde{V}(\phi) = \frac{\max\{1-\widetilde{\alpha}^2\phi^2, 0\} + D_1(\phi)^2}{2} = \begin{cases} \frac{1}{2}\left(1 - \widetilde{\alpha}^2\phi^2 + \left(\frac{1+(1-\phi)^2}{2}\right)^2\right), & \phi \leq \frac{1}{1-\widetilde{\alpha}}, \\ \frac{1}{2}\left(1 - \widetilde{\alpha}^2\phi^2 + \left(\frac{1+\widetilde{\alpha}^2\phi^2}{2}\right)^2\right), & \frac{1}{1-\widetilde{\alpha}} < \phi \leq \frac{1}{\widetilde{\alpha}}, \\ \frac{1}{2}, & \phi > \frac{1}{\widetilde{\alpha}}. \end{cases} \tag{B-44}$$

The probability of consuming in both periods, i.e. $D_{12}(\phi) = D_1(\phi) \cdot D_2(\phi|d_1 = 1)$ — and thus, the probability of accruing the penalty fee — is then:

$$D_{12}(\phi) = \begin{cases} \frac{(1+(1-\phi)^2)(1-(1-\alpha)\phi)}{2}, & \phi \leq \frac{1}{1-\widetilde{\alpha}}, \\ \frac{(1+\widetilde{\alpha}^2\phi^2)(1-(1-\alpha)\phi)}{2}, & \frac{1}{1-\widetilde{\alpha}} < \phi \leq \frac{1}{1-\alpha}, \\ 0, & \phi > \frac{1}{1-\alpha}. \end{cases} \tag{B-45}$$

By substituting in $\widetilde{\alpha}$ for each of α in (B-42), (B-43), and (B-45), while applying $\max\{1-(1-\widetilde{\alpha})\phi, 0\} = 0$ given $\phi \geq \frac{1}{1-\widetilde{\alpha}}$, we can also derive the perceived consumption probabilities:

$$\widetilde{D}_1(\phi) = \begin{cases} \frac{1+(1-\phi)^2}{2}, & \phi \leq \frac{1}{1-\widetilde{\alpha}}, \\ \frac{1+\widetilde{\alpha}^2\phi^2}{2}, & \frac{1}{1-\widetilde{\alpha}} < \phi \leq \frac{1}{\widetilde{\alpha}}, \\ 1, & \phi > \frac{1}{\widetilde{\alpha}}. \end{cases} \tag{B-46}$$

$$\widetilde{D}_2(\phi) = \begin{cases} 1 - \widetilde{\alpha}\phi - \frac{(1+(1-\phi)^2)(1-2\widetilde{\alpha})\phi}{2}, & \phi \leq \frac{1}{1-\widetilde{\alpha}}, \\ \frac{(1-\widetilde{\alpha}\phi)(1-\widetilde{\alpha}^2\phi^2)}{2}, & \frac{1}{1-\widetilde{\alpha}} < \phi \leq \frac{1}{\widetilde{\alpha}}, \\ 0, & \phi > \frac{1}{\widetilde{\alpha}}. \end{cases} \tag{B-47}$$

$$\widetilde{D}_{12}(\phi)=\begin{cases}\frac{(1+(1-\phi)^2)(1-(1-\widetilde{\alpha})\phi)}{2}, & \phi\leq\frac{1}{1-\widetilde{\alpha}},\\ 0, & \phi>\frac{1}{1-\widetilde{\alpha}}.\end{cases} \tag{B-48}$$

With consumption tracking in the symmetric forgetting model, if $d_1=0$ a consumer may now believe that there is a risk of accruing the penalty fee if she consumes at $t=2$. Thus, there may now be an incentive to incur the monitoring cost k to use consumption tracking even if she did not consume at $t=1$. If she does not use consumption tracking following abstinence, her probability of consuming at $t=2$ is still $D_2(\phi|d_1=0,d_\tau=0)=\max\{1-\alpha\phi,0\}$, as in (B-38), while $\widetilde{u}_2(\phi|d_1=0,d_\tau=0)=\frac{1-\widetilde{\alpha}^2\phi^2}{2}\cdot\mathrm{I}\left[\phi<\frac{1}{\widetilde{\alpha}}\right]$, as in (B-40). If she does use consumption tracking following abstinence, her probability of consuming at $t=2$ is then $D_2(\phi|d_1=0,d_\tau=1)=1$, while $\widetilde{u}_2(\phi|d_1=0,d_\tau=1)=\mathrm{E}[v_2]-k=\frac{1}{2}-k$. If $d_1=1$ and $d_\tau=0$, the consumer's probability of consuming at $t=2$ is still $D_2(\phi|d_1=1,d_\tau=0)=\max\{1-(1-\alpha)\phi,0\}$, as in (B-39), and $\widetilde{u}_2(\phi|d_1=1,d_\tau=0)=\frac{(1-\phi)^2-\widetilde{\alpha}^2\phi^2}{2}\cdot\mathrm{I}\left[\phi<\frac{1}{1-\widetilde{\alpha}}\right]$, as in (B-41). If $d_\tau=1$, she will know $d_1=1$, in which case her probability of consuming at $t=2$ becomes $D_2(\phi|d_1=1,d_\tau=1)=\max\{1-\phi,0\}$ and $\widetilde{u}_2(\phi|d_1=1,d_\tau=1)=\frac{(1-\phi)^2}{2}\cdot\mathrm{I}[\phi<1]-k$.

Next, the probability of consumption in the second period following a decision to use consumption tracking is simply $D_2(\phi|d_\tau=1)=\max\{1-d_1\cdot\phi,0\}$. In turn, the perceived $t=2$ utility given the consumer expects to use consumption tracking would then be

$$\widetilde{u}_2(\phi|d_\tau=1)=\frac{(\max\{1-d_1\cdot\phi,0\})^2}{2}-k. \tag{B-49}$$

Drawing on our work thus far, the following result summarizes four consumption profiles that can arise in the symmetric forgetting model without consumption tracking.

Proposition 3. *For a consumer who subscribes to the service in the symmetric forgetting model without consumption tracking:*

(i) [future abstinence-based avoidance] *if* $\phi\geq\frac{1}{\alpha}$, *then* $D_1=\min\left\{\frac{1+\widetilde{\alpha}^2\phi^2}{2},1\right\}$, $D_2=0$, *and* $D_{12}=\widetilde{D}_{12}=0$*;*

(ii) [memory-based avoidance] *if* $\frac{1}{\alpha}>\phi\geq\frac{1}{1-\alpha}$, *then* $D_1=\frac{1+\widetilde{\alpha}^2\phi^2}{2}$, $D_2=\frac{(1-\alpha\phi)(1-\widetilde{\alpha}^2\phi^2)}{2}$, *and* $D_{12}=\widetilde{D}_{12}=0$*;*

(iii) [unintentional accrual] *if* $\frac{1}{1-\alpha}>\phi\geq\frac{1}{1-\widetilde{\alpha}}$, *then* $D_1=\frac{1+\widetilde{\alpha}^2\phi^2}{2}$, $D_2=1-\alpha\phi-\frac{(1+\widetilde{\alpha}^2\phi^2)(1-2\alpha)\phi}{2}$, $D_{12}=\frac{(1+\widetilde{\alpha}^2\phi^2)(1-(1-\alpha)\phi)}{2}>0$, *and* $\widetilde{D}_{12}=0$*;*

(iv) [intentional accrual] *if* $\phi<\frac{1}{1-\widetilde{\alpha}}$, *then* $D_1=\frac{1+(1-\phi)^2}{2}$, $D_2=1-\alpha\phi-\frac{(1+(1-\phi)^2)(1-2\alpha)\phi}{2}$, $D_{12}=\frac{(1+(1-\phi)^2)(1-(1-\alpha)\phi)}{2}>0$, *and* $\widetilde{D}_{12}=\frac{(1+(1-\phi)^2)(1-(1-\widetilde{\alpha})\phi)}{2}>0.$

Proof of Proposition 3.

Part (i). The expressions for D_1, D_2, D_{12}, and $\widetilde{D}_{12}$ given $\phi \geq \frac{1}{\alpha}$ follow from (B-42), (B-43), (B-45), and (B-48). Note here that the first-period consumption probability reduces to $D_1 = \min\left\{\frac{1+\widetilde{\alpha}^2\phi^2}{2}, 1\right\}$ in light of the fact that $\phi \geq \frac{1}{\alpha}$ implies $\phi \geq \frac{1}{1-\widetilde{\alpha}}$ given $0 \leq \widetilde{\alpha} \leq \alpha \leq \frac{1}{2}$ and that $\frac{1+\widetilde{\alpha}^2\phi^2}{2} > 1$ holds if and only if $\phi > \frac{1}{\widetilde{\alpha}}$.

Part (ii). The expressions for D_1, D_2, D_{12}, and $\widetilde{D}_{12}$ given $\frac{1}{1-\alpha} \leq \phi < \frac{1}{\alpha}$ again follow from (B-42), (B-43), (B-45), and (B-48), while also noting that $\frac{1}{1-\alpha} \leq \phi < \frac{1}{\alpha}$ implies $\frac{1}{1-\widetilde{\alpha}} \leq \phi < \frac{1}{\widetilde{\alpha}}$ since $\frac{1}{1-\widetilde{\alpha}} \leq \frac{1}{1-\alpha} \leq \frac{1}{\alpha} \leq \frac{1}{\widetilde{\alpha}}$ given $0 \leq \widetilde{\alpha} \leq \alpha \leq \frac{1}{2}$.

Part (iii). The expressions for D_1, D_2, D_{12}, and $\widetilde{D}_{12}$ given $\frac{1}{1-\widetilde{\alpha}} \leq \phi < \frac{1}{1-\alpha}$ again follow from (B-42), (B-43), (B-45), and (B-48), while also noting that $\phi < \frac{1}{1-\alpha}$ implies $\phi < \frac{1}{\widetilde{\alpha}}$ since $\frac{1}{1-\alpha} \leq \frac{1}{\widetilde{\alpha}}$ given $0 \leq \widetilde{\alpha} \leq \alpha \leq \frac{1}{2}$.

Part (iv). The expressions for D_1, D_2, D_{12}, and $\widetilde{D}_{12}$ given $\phi < \frac{1}{1-\widetilde{\alpha}}$ again follow from (B-42), (B-43), (B-45), and (B-48). □

By comparing Proposition 3 to Lemma 1, we can see that the symmetric forgetting model may give rise to qualitatively distinct consumption profiles in comparison to our original model with asymmetric forgetting. Part (i), for instance, describes a consumption profile characterized by *future* abstinence-based avoidance of the penalty fee. Here, the penalty fee is sufficiently large to deter the consumer from consuming in the second period *even* if the consumer did not consume in the first period (and thus would not be at risk of incurring the penalty fee). As a result of such abstinence, the consumer is therefore always able to avoid the penalty fee, though here it is abstinence in the second period that guarantees penalty avoidance as opposed to the abstinence-based avoidance profile from our original model (Lemma 1, part ii), in which the consumer always abstains in the first period to avoid incurring the penalty fee. Interestingly, the assured abstinence in the second period may or may not be anticipated. If ϕ is sufficiently large ($\phi \geq \frac{1}{\widetilde{\alpha}}$), the consumer correctly anticipates that she will abstain in the second period regardless of her first-period consumption choice and thus always consumes in the first period knowing it is her only opportunity to attain value from the service. Otherwise (i.e. if $\frac{1}{\widetilde{\alpha}} > \phi > \frac{1}{\alpha}$), the consumer incorrectly believes that if she abstains in the first period then she might still consume — and thus attain value from the service — in the second period. As a result, the consumer might make the mistake of abstaining in the first period (as well as the second period), in the false hope of attaining a higher value from consumption in the second period.

Next, part (ii) describes a consumption profile characterized by memory-based avoidance of the penalty fee. Here, the consumer will have a sufficiently strong memory of first period consumption to ensure that she does not mistakenly consume in both periods. Unlike the case of future abstinence-based avoidance in part (i), however, now the consumer might consume in the second period if she abstains in the first period. That said, the consumer may still unnecessarily abstain in the second period following first-period abstinence, based on a misplaced concern that she might have consumed in the first period. The possibility of abstinence in *both* periods distinguishes the present case from the memory-based avoidance consumption profile seen in our original analysis with asymmetric forgetting (Lemma 1, part i). The present case also entails a higher probability of consumption in the first period. In turn, the consumption profiles characterized by unintentional and intentional accrual of the penalty fee described in parts (iii) and (iv) of Proposition 3 likewise feature higher probabilities of first-period consumption compared to the analogous consumption profiles from our original analysis (Lemma 1, part iii and iv).

In addition to its influence on consumers' consumption decisions, symmetric forgetting can also alter consumers' decisions of whether or not to use consumption tracking, if available. These latter effects are apparent in the following lemma, which characterizes a consumer's tracking decision in the model with symmetric forgetting.

Lemma 3. *In the symmetric forgetting model, a consumer chooses to use consumption tracking ($d_\tau = 1$) if and only if $k < k_\tau(\alpha, \phi | d_1)$, where:*

$$k_\tau(\alpha,\phi|d_1) = \begin{cases} \frac{\alpha(1-\alpha)\phi^2}{2}, & \phi \leq 1, \\ \frac{\alpha((1-\alpha)\phi^2-(\phi-1)^2)}{2}, & 1 < \phi < \frac{1}{\alpha},\ d_1 = 0, \\ \frac{1-\alpha}{2}, & \phi \geq \frac{1}{\alpha},\ d_1 = 0, \\ \frac{(1-\alpha)(\alpha\phi^2-(\phi-1)^2)}{2}, & 1 < \phi < \frac{1}{1-\alpha},\ d_1 = 1, \\ \frac{\alpha}{2}, & \phi \geq \frac{1}{1-\alpha},\ d_1 = 1. \end{cases}$$

Proof of Lemma 3.

A consumer with $d_1 = 1$ has the same $t = 2$ beliefs in the symmetric forgetting model as in the original model with asymmetric forgetting. In light of this, the expression for $k_\tau(\alpha, \phi | d_1 = 1)$ in Lemma 3 is equivalent to, and can be derived in the same way (see proof of Lemma 2) as the expression for $k_\tau(\alpha, \phi)$ in (3). Suppose Consumer A has $d_1 = 0$ and $\alpha = \alpha_A$, while Consumer B has $d_1 = 1$ and $\alpha = \alpha_B$. Then, in the symmetric forgetting model, Consumer A's $t = 2$ beliefs are equivalent to Consumer B's $t = 2$ beliefs if and only if $\alpha_A = 1 - \alpha_B$. The expression for $k_\tau(\alpha, \phi | d_1 = 0)$ in Lemma 3 can then be confirmed by substituting $1 - \alpha$ for α in the expression for $k_\tau(\alpha, \phi | d_1 = 1)$. $\square$

As in our original analysis (see Lemma 2), a consumer still chooses to use consumption tracking if the monitoring cost k is below some threshold that depends on the consumer's forgetfulness α and the penalty fee ϕ. Note, however, that the threshold in Lemma 3 now also depends on the first period consumption choice d_1, as the consumer may now use consumption tracking even if she did not consume in the first period. Naturally, a consumer in the original asymmetric forgetting model would, due to her perfect recollection of past abstinence, only consider tracking her first period consumption if she did, in fact, consume in the first period. With symmetric forgetting, however, she may now choose to use consumption tracking even if she abstained in the first period. In such cases, the consumer is worried she might have consumed in the first period, while choosing to use consumption tracking assures her that she can consume in the second period without penalty.

The next result characterizes the market equilibrium — which is much simpler than before — in the model with symmetric forgetting.

Proposition 4. *In the symmetric forgetting model, the zero-penalty benchmark equilibrium necessarily arises in that $\phi^* = 0$ and $p^* = 1$ regardless of α, $\widetilde{\alpha}$, and k; in this equilibrium, $D_{12} = \widetilde{D}_{12} = 1$ and $D_\tau = 0$, while $\Pi = 1$ and $CS = 0$.*

Proof of Proposition 4

To prove Proposition 4, we will show that an equilibrium with $\phi^* > 0$ cannot exist. In this proof, we will often make use of the fact that the expression for $k_\tau(\alpha, \phi | d_1 = 0)$ in Lemma 3 given $0 \leq \widetilde{\alpha} \leq \alpha \leq \frac{1}{2}$ implies: $\widetilde{k}_{\tau|d_1=1} \leq \widetilde{k}_{\tau|d_1=0}$, $\widetilde{k}_{\tau|d_1=1} \leq k_{\tau|d_1=1}$, and $\widetilde{k}_{\tau|d_1=0} \leq k_{\tau|d_1=0}$.

Case 1: $k < \widetilde{k}_{\tau|d_1=1}$

Suppose $k < \widetilde{k}_{\tau|d_1=1}$ in equilibrium with $\phi^* > 0$. This implies $\widetilde{d}_{\tau|d_1=0} = \widetilde{d}_{\tau|d_1=1} = d_{\tau|d_1=1} = 1$. Using (B-49) with $D_1 = \Pr[v_1 > \widetilde{u}_2(\phi|d_1 = 0) - \widetilde{u}_2(\phi|d_1 = 1)]$, while noting that a consumer's maximum $t = 0$ willingness-to-pay for a subscription would be given by $\widetilde{V}(\phi) = \widetilde{u}_2(\phi|d_1 = 0) + D_1(\phi) \cdot (\mathrm{E}[v_1|v_1 > \Delta] - \Delta)$ for $\Delta \equiv \widetilde{u}_2(\phi|d_1 = 0) - \widetilde{u}_2(\phi|d_1 = 1) \geq 0$ and with v_1 uniformly distributed between 0 and 1, we can verify that in this case a consumer's willingness-to-pay and probability of incurring the penalty fee would be given by:

$$\widetilde{V}(\phi) = \frac{5}{8} - k + \frac{(1-\phi)^2(2+(1-\phi)^2)\mathrm{I}[\phi<1]}{8}, \tag{B-50}$$

$$D_{12}(\phi) = \frac{(1-\phi)(1+(1-\phi)^2)\mathrm{I}[\phi<1]}{2}. \tag{B-51}$$

We can express the firm's maximum expected profits from a given consumer as

$$\Pi(\phi) = \begin{cases} 1-k-\phi^2\big(1-\phi+\frac{3\phi^2}{8}\big), & \phi < 1, \\ \frac{5}{8}-k, & \phi \geq 1, \end{cases} \tag{B-52}$$

which uses $\Pi(\phi) = p + D_{12}(\phi)\cdot\phi$ along with $p^*(\phi) = V(\phi)$. We can then see $\Pi < 1$ must hold. Since the firm can always earn $\Pi = 1$ by setting $\phi = 0$ (with $p = 1$), the firm cannot earn higher profits than would be attained with $\phi = 0$ given $k < \widetilde{k}_{\tau|d_1=1}$.

Case 2: $\widetilde{k}_{\tau|d_1=1} < k < \widetilde{k}_{\tau|d_1=0}$ *and* $k < k_{\tau|d_1=1}$

Suppose $\widetilde{k}_{\tau|d_1=1} < k < \widetilde{k}_{\tau|d_1=0}$ and $k < k_{\tau|d_1=1}$ in equilibrium with $\phi^* > 0$. This implies $\widetilde{d}_{\tau|d_1=0} = 1$, $\widetilde{d}_{\tau|d_1=1} = 0$, and $d_{\tau|d_1=1} = 1$. Using (B-49) with $D_1 = \Pr[v_1 > \widetilde{u}_2(\phi|d_1=0) - \widetilde{u}_2(\phi|d_1=1)]$, while noting that a consumer's maximum $t = 0$ willingness-to-pay for a subscription would be given by $\widetilde{V}(\phi) = \widetilde{u}_2(\phi|d_1=0) + D_1(\phi)\cdot(\mathrm{E}[v_1|v_1>\Delta]-\Delta)$ for $\Delta \equiv \widetilde{u}_2(\phi|d_1=0) - \widetilde{u}_2(\phi|d_1=1) \geq 0$ and with v_1 uniformly distributed between 0 and 1, we can verify that in this case a consumer's willingness-to-pay and probability of incurring the penalty fee would be given by:

$$\widetilde{V}(\phi) = \begin{cases} 1+\frac{k^2}{2} - \frac{\phi(2-(1-\widetilde{\alpha}^2)\phi)(4(1+k)-\phi(2-(1-\widetilde{\alpha}^2)\phi))}{8}, & \phi < \frac{1}{1-\widetilde{\alpha}}, \\ \frac{5}{8} - \frac{k(1-k)}{2}, & \phi \geq \frac{1}{1-\widetilde{\alpha}}, \end{cases} \tag{B-53}$$

$$D_{12}(\phi) = 0. \tag{B-54}$$

Note, we do not need to consider $\phi < 1$ in the above expressions since $\phi < 1$ would imply $\widetilde{k}_{\tau|d_1=0} = \widetilde{k}_{\tau|d_1=1}$, thus precluding $\widetilde{k}_{\tau|d_1=1} < k < \widetilde{k}_{\tau|d_1=0}$.

Next, we can express the firm's maximum expected profits as

$$\Pi(\phi) = \begin{cases} 1+\frac{k^2}{2} - \frac{\phi(2-(1-\widetilde{\alpha}^2)\phi)(4(1+k)-\phi(2-(1-\widetilde{\alpha}^2)\phi))}{8}, & \phi < \frac{1}{1-\widetilde{\alpha}}, \\ \frac{5}{8} - \frac{k(1-k)}{2}, & \phi \geq \frac{1}{1-\widetilde{\alpha}}, \end{cases} \tag{B-55}$$

which uses $\Pi(\phi) = p + D_{12}(\phi)\cdot\phi$ along with $p^*(\phi) = V(\phi)$.

For $\phi < \frac{1}{1-\widetilde{\alpha}}$, it is then readily verifiable that $\Pi(\phi) > 1$ requires $k > \phi\big(1+\frac{2-(1-\widetilde{\alpha}^2)\phi}{2}\big) + \sqrt{2\phi\big(1+\frac{2-(1-\widetilde{\alpha}^2)\phi}{2}\big)}$ and that this threshold is increasing in ϕ for $\phi < \frac{1}{1-\widetilde{\alpha}^2}$ and decreasing in ϕ for $\phi > \frac{1}{1-\widetilde{\alpha}^2}$. Furthermore, the threshold is equal to $\frac{3}{2}$ if $\phi = 1$ and equal to $\frac{1+\widetilde{\alpha}^2}{2} + \sqrt{1+\widetilde{\alpha}^2} > \frac{3}{2}$ if $\phi = \frac{1}{1-\widetilde{\alpha}}$. Thus, $k > \frac{3}{2}$ must hold for $\Pi > 1$ in this case. However, $k < k_{\tau|d_1=1}$ with $1 < \phi < \frac{1}{1-\widetilde{\alpha}}$ requires $k < \frac{(1-\alpha)(\alpha\phi^2-(\phi-1)^2)}{2} < \frac{\alpha}{2} < \frac{1}{4} < \frac{3}{2}$. Thus, $\Pi(\phi) > 1$ cannot hold in this case with $1 < \phi < \frac{1}{1-\widetilde{\alpha}}$. Lastly, by inspection, we can see that $\Pi(\phi) = \frac{5}{8} - \frac{k(1-k)}{2} < \frac{5}{8} < 1$ in this case given $\phi > \frac{1}{1-\widetilde{\alpha}}$. Thus, $\Pi(\phi) > 1$ cannot hold in this case with $\phi > \frac{1}{1-\widetilde{\alpha}}$.

Collectively, the above work shows that $\Pi < 1$ must hold. Since the firm can always earn $\Pi = 1$ by setting $\phi = 0$ (with $p = 1$), the firm cannot earn higher profits than would be attained with $\phi = 0$ given $\widetilde{k}_{\tau|d_1=1} < k < \widetilde{k}_{\tau|d_1=0}$ and $k < k_{\tau|d_1=1}$.

Case 3: $\widetilde{k}_{\tau|d_1=1} < k < \widetilde{k}_{\tau|d_1=0}$ *and* $k > k_{\tau|d_1=1}$

Suppose $\widetilde{k}_{\tau|d_1=1} < k < \widetilde{k}_{\tau|d_1=0}$ and $k > k_{\tau|d_1=1}$ in equilibrium with $\phi^* > 0$. This implies $\widetilde{d}_{\tau|d_1=0} = 1$, $\widetilde{d}_{\tau|d_1=1} = 0$, and $d_{\tau|d_1=1} = 0$. Using (B-49) with $D_1 = \Pr[v_1 > \widetilde{u}_2(\phi|d_1 = 0) - \widetilde{u}_2(\phi|d_1 = 1)]$, while noting that a consumer's maximum $t = 0$ willingness-to-pay for a subscription would be $\widetilde{V}(\phi) = \widetilde{u}_2(\phi|d_1 = 0) + D_1(\phi) \cdot (\mathrm{E}[v_1|v_1 > \Delta] - \Delta)$ for $\Delta \equiv \widetilde{u}_2(\phi|d_1 = 0) - \widetilde{u}_2(\phi|d_1 = 1) \geq 0$ and with v_1 uniformly distributed between 0 and 1, we can verify that in this case a consumer's willingness-to-pay would be given by (B-53) with

$$D_{12}(\phi) = \begin{cases} (1-(1-\alpha)\phi)\left(1+k-\phi+\frac{(1-\widetilde{\alpha}^2)\phi^2}{2}\right), & \phi < \frac{1}{1-\widetilde{\alpha}}, \\ (1-(1-\alpha)\phi)\left(\frac{1}{2}+k\right), & \frac{1}{1-\widetilde{\alpha}} < \phi < \frac{1}{1-\alpha}, \\ 0, & \phi \geq \frac{1}{1-\alpha}. \end{cases} \tag{B-56}$$

Next, we can express the firm's maximum expected profits as

$$\Pi(\phi) = \begin{cases} 1 - \frac{\widetilde{\alpha}^2\phi^2}{2} + \frac{(2k+(1-\widetilde{\alpha}^2)\phi^2)^2}{8} - \frac{(1-\alpha)\phi^2(2(1-\phi+k)+(1-\widetilde{\alpha}^2)\phi^2)}{2}, & \phi < \frac{1}{1-\widetilde{\alpha}}, \\ \frac{5}{8} - \frac{k(1-k)}{2} + (1-(1-\alpha)\phi)\left(\frac{1}{2}+k\right)\phi, & \frac{1}{1-\widetilde{\alpha}} \leq \phi < \frac{1}{1-\alpha}, \\ \frac{5}{8} - \frac{k(1-k)}{2}, & \phi \geq \frac{1}{1-\alpha}, \end{cases} \tag{B-57}$$

which uses $\Pi(\phi) = p + D_{12}(\phi) \cdot \phi$ along with $p^*(\phi) = V(\phi)$.

For $\phi < \frac{1}{1-\widetilde{\alpha}}$, it is readily verifiable that $\Pi(\phi) > 1$ requires $k > \frac{(1+\widetilde{\alpha}^2-2\alpha)\phi^2}{2} + \phi\sqrt{1+\widetilde{\alpha}^2-2\alpha+(1-(1-\alpha)\phi)^2}$. However, $\widetilde{k}_{\tau|d_1=1} < k < \widetilde{k}_{\tau|d_1=0}$ cannot hold with $\phi < 1$ because $\phi < 1$ implies $\widetilde{k}_{\tau|d_1=0} = \widetilde{k}_{\tau|d_1=1}$. In turn, $k < \widetilde{k}_{\tau|d_1=0}$ with $1 < \phi < \frac{1}{1-\widetilde{\alpha}}$ requires $k < \frac{\widetilde{\alpha}((1-\widetilde{\alpha})\phi^2-(\phi-1)^2)}{2}$, which is less than this threshold. Thus, $\Pi(\phi) > 1$ cannot hold in this case with $\phi < \frac{1}{1-\widetilde{\alpha}}$.

For $\frac{1}{1-\widetilde{\alpha}} < \phi < \frac{1}{1-\alpha}$, it is readily verifiable that $\Pi(\phi) > 1$ requires $k > \frac{3}{2} - 2\phi(1-(1-\alpha)\phi)$. We can demonstrate that this threshold is higher than $\frac{\alpha}{2}$. To see this, note that $\frac{\partial}{\partial\alpha}\left(\frac{3}{2} - 2\phi(1-(1-\alpha)\phi) - \frac{\alpha}{2}\right) = -\frac{1}{2} - 2\phi < 0$, implying this difference, i.e. $\frac{3}{2} - 2\phi(1-(1-\alpha)\phi) - \frac{\alpha}{2}$, is minimized at $\alpha = \frac{1}{2}$ given $\alpha \leq \frac{1}{2}$, in which case the difference reduces to $\frac{5}{4} - 2\phi + \phi^2 = \frac{1}{4} + (\phi-1)^2 > 0$. Thus, $\Pi(\phi) > 1$ requires $k > \frac{3}{2} - 2\phi(1-(1-\alpha)\phi) > \frac{\alpha}{2}$, but $k < \widetilde{k}_{\tau|d_1=0}$ requires $k < \frac{\alpha}{2}$. As a result, $\Pi(\phi) > 1$ cannot hold in this case with $\frac{1}{1-\widetilde{\alpha}} < \phi < \frac{1}{1-\alpha}$. Lastly, by inspection, we can see $\Pi(\phi) = \frac{5}{8} - \frac{k(1-k)}{2} < \frac{5}{8} < 1$ in this case given $\phi > \frac{1}{1-\alpha}$. Thus, $\Pi(\phi) > 1$ cannot hold in this case with $\phi > \frac{1}{1-\alpha}$.

Collectively, the above work shows that $\Pi < 1$ must hold in case 3. Since the firm can always earn $\Pi = 1$ by setting $\phi = 0$ (with $p = 1$), the firm cannot earn higher profits than would be attained with $\phi = 0$ given $\widetilde{k}_{\tau|d_1=1} < k < \widetilde{k}_{\tau|d_1=0}$ and $k > k_{\tau|d_1=1}$.

Case 4: $k > \widetilde{k}_{\tau|d_1=1}$

We can break this case down into subcases depending on whether $k < k_{\tau|d_1=1}$ or $k \geq k_{\tau|d_1=1}$. Either way, $\widetilde{V}(\phi)$ will be the same as its benchmark value, while D_{12} will be less than or equal to its benchmark value. Thus, the firm cannot increase its profits from trackers relative to the no-tracking benchmark if $k > \widetilde{k}_{\tau|d_1=1}$. In conjunction with our prior work, we see that it is never possible for the firm to increase its profits relative to the benchmark without consumption tracking. It then follows that $\phi^* = 0$ and $p^* = 1$ must hold in the symmetric forgetting model, while $\phi^* = 0$ and $p^* = 1$ imply $D_{12} = \widetilde{D}_{12} = 1$, $D_\tau = 0$, $\Pi = 1$, and $CS = 0$. □

Thus, with symmetric forgetting, it is never optimal for the firm to use a positive penalty fee. Note that this does not depend on the costs or availability of consumption tracking. In this way, the model with symmetric forgetting is not helpful for understanding the empirical use of penalty fees. Furthermore, since consumption tracking is irrelevant in the absence of penalty fees, the result also suggests that the model with symmetric forgetting — as with the other, previously considered alternate formulations of consumer forgetfulness — is not helpful for understanding recent advances in consumption tracking technologies. With that said, these empirical limitations are not too concerning in light of the fact that they are based on a symmetric forgetting formulation that lacks empirical support (as discussed earlier). Rest assured, and as illustrated by our prior analysis, our original model with asymmetric forgetting does offer a useful lens for understanding the incidence of penalty fees as well as the recent proliferation of (and demand for) technologies that help consumers track their consumption as a means to avoid penalty fees.

Lastly, we may also consider a model in which consumers are heterogeneous with regards to the symmetry of their forgetfulness. In particular, suppose that a λ share of consumers are asymmetrically forgetful, as in our original model, while the remaining $1-\lambda$ share are symmetrically forgetful, as in the model variant considered above. In this case, it will once again be possible that the firm optimally sets a positive penalty fee and will also once again be possible that the availability of consumption tracking leads to a different equilibrium than the corresponding benchmark equilibrium. In other words, our model can have some symmetrically forgetful consumers while retaining its ability to explain the use of positive penalty fees and predict meaningful effects of consumption tracking on market behavior.

These possibilities (i.e., that the firm optimally sets a positive penalty fee and that the availability of consumption tracking can lead to a different equilibrium) are established in the following example.

Example 3. *In the model where consumers are heterogeneous in whether they are symmetrically or asymmetrically forgetful, as described above, suppose* $k = \frac{13}{100}$ *and that* $\lambda = \frac{9}{10}$ *is the share of asymmetrically forgetful consumers. Also suppose that these asymmetrically forgetful consumers have* $\alpha = \frac{9}{10}$ *and* $\widetilde{\alpha} = \frac{25}{28}$*. Then the availability of consumption tracking compels the firm to impose a penalty fee that would not otherwise exist,* $\phi^* > \phi^*_{NT} = 0$*, regardless of the values of* α *and* $\widetilde{\alpha}$ *for the* $1-\lambda$ *share of consumers who are symmetrically forgetful.*

To verify that the availability of consumption tracking would compel the firm to impose a penalty fee that would not otherwise exist in Example 3, it helps to note that this example uses the same parameter values from Example 1, with the exception of λ. As we saw in Example 1, $\phi^*_{NT} = 0$ was optimal in the benchmark equilibrium when all consumers were asymmetrically forgetful. From Proposition 4, we know that $\phi^*_{NT} = 0$ would also be optimal in the benchmark equilibrium if all consumers were symmetrically forgetful (for any α and $\widetilde{\alpha}$). In the current setup with both types of consumers, setting $\phi^*_{NT} = 0$ (with $p^*_{NT} = 1$) thus maximizes profits within each segment and must therefore still be optimal in Example 3, leading to $\Pi_{NT} = 1$.

To verify that $\phi^* > 0$ must hold in Example 3, suppose the firm uses the contract that was optimal in Example 1, i.e. with $p = \frac{5}{8}$ and $\phi = 2$. In this case, the minimum profit attained from symmetrically forgetful consumers is simply zero (which would be realized if these consumers choose not to subscribe). Given the equilibrium profits in Example 3, we can then see that the expected profit from all consumers must satisfy $\Pi \geq \frac{9}{10} \cdot \frac{23{,}209}{20{,}000} + \frac{1}{10} \cdot 0 = \frac{208{,}881}{200{,}000} > 1 = \Pi_{NT}$ with $p = \frac{5}{8}$ and $\phi = 2$. Since setting $\phi = \phi^*_{NT} = 0$ and $p = p^*_{NT} = 1$ would once again lead to $\Pi = \Pi_{NT} = 1$, which is less than the lower bound on the total expected profit with $p = \frac{5}{8}$ and $\phi = 2$, the optimal penalty fee when consumption tracking is available cannot be zero, ensuring $\phi^* > 0$. Thus, the availability of consumption tracking does compel the firm to impose a penalty fee that would not otherwise exist. (Note here that this conclusion holds even if $\phi^* \neq 2$ and, once again, does not depend on the levels of forgetfulness and sophistication among symmetrically forgetful consumers.)

B.2. Alternate Assumptions on the Timing of Decisions to Use Consumption Tracking

As we recall, consumers with access to consumption tracking in our model could decide whether to use the technology upon their arrival at $t = 2$. As discussed in Section 5, however, this assumption is only suited to a particular class of consumption tracking technologies (such as mobile banking apps) on which our study is focused. For other consumption tracking technologies, such as automated text alerts requiring consumers to "opt-in" well before receiving any potential alerts or technologies which require consumers to manually input their past consumption, it may be more natural to model consumers' potential decisions to use consumption tracking as occurring prior to $t = 2$. Here, we consider three such variations of our model: one in which consumers' decisions to use consumption tracking occur at $t = 0$ (i.e., at the same time as the $t = 0$ subscription decision), one in which consumers' decisions to use consumption tracking occur at the beginning of $t = 1$ (i.e., immediately before learning v_1), and one in which consumers' decisions to use consumption tracking occur at the end of $t = 1$ (i.e., after choosing d_1).

B.2.1. Decisions to Use Consumption Tracking at $t = 0$

We now analyze a variation of our model in which consumers with access to consumption tracking decide whether to use consumption tracking at $t = 0$ instead of at $t = 2$. To start, we may note that the timing of any hypothetical decisions to use consumption tracking is irrelevant when consumption tracking is not actually available. Thus, our benchmark analysis is unaffected by changing the timing of consumers' decisions to use consumption tracking. For this reason, we can proceed directly to our analysis with consumption tracking.

With decisions to use consumption tracking at $t = 0$, a consumer decides whether to use consumption tracking in conjunction with her subscription decision. As before, the firm will always set $p^* = \widetilde{V}(\phi^*)$ to ensure the consumer subscribes in equilibrium. In turn, the consumer's willingness-to-pay for a subscription, $\widetilde{V}(\phi^*)$ can be separated into two cases. If the consumer perceives that it is optimal to use consumption tracking given the size of the penalty fee, i.e. if $\widetilde{V}(\phi^*) = \widetilde{V}(\phi^*|d_\tau = 1) > \widetilde{V}(\phi^*|d_\tau = 0)$, then she will in fact choose to use consumption tracking while her $t = 1$ and $t = 2$ consumption decisions will be identical to those of a consumer with perfect memory (i.e. with $\alpha = 0$) in the absence of consumption tracking; however, her willingness-to-pay will be reduced by k relative to the consumer with perfect memory who does need to use consumption tracking. If the consumer perceives that it is optimal not to use consumption tracking, i.e. if $\widetilde{V}(\phi^*) = \widetilde{V}(\phi^*|d_\tau = 0) \geq \widetilde{V}(\phi^*|d_\tau = 1)$, then she will in fact choose not to use consumption tracking while her $t = 1$ and $t = 2$ consumption decisions will be the same as in the benchmark model without consumption tracking.

Remark 8. *In the model with decisions to use consumption tracking at $t = 0$, the firm's optimal contract with consumption tracking is the same as its optimal contract in the benchmark equilibrium without consumption tracking: $\phi^* = \phi^*_{NT}$ and $p^* = p^*_{NT}$. Furthermore, the equilibrium will be identical to the benchmark equilibrium without consumption tracking.*

To verify the irrelevance of consumption tracking with decisions to use consumption tracking at $t = 0$, as asserted in Remark 8, let $\phi_0^* = \arg\max_\phi \{\Pi(\phi) : d_\tau = 0\}$ denote the optimal penalty fee (with corresponding optimal price $p_0^* = \widetilde{V}(\phi_0^*)$) conditional on the penalty fee being at a level for which the consumer chooses not to use consumption tracking. Similarly, let $\phi_1^* = \arg\max_\phi \{\Pi(\phi) : d_\tau = 1\}$ denote the optimal penalty fee (with $p_1^* = \widetilde{V}(\phi_1^*)$) conditional on the penalty fee being at a level for which the consumer chooses not to use consumption tracking.

Next, from our benchmark analysis (Proposition 1), we know that if $\phi_{NT}^* = 0$, a consumer will consume in both periods; if $\phi_{NT}^* > 0$, a consumer consumes with probability $\frac{1}{2}$ at $t = 1$ and initially believes she will consume at $t = 2$ if and only if she did not consume at $t = 1$. Importantly, for both $\phi_{NT}^* = 0$ and $\phi_{NT}^* > 0$, the consumer believes she will make the same choices, for the given value of ϕ_{NT}^*, that would be made by a consumer with $\alpha = 0$. As a result, she believes she will receive the same total value (in expectation) from a subscription that she would receive with $\alpha = 0$. Therefore, if the firm offers its benchmark equilibrium contract, the consumer would not be willing to incur any positive monitoring cost to use consumption tracking as a means of "endowing" her with $\alpha = 0$, while her willingness-to-pay for a subscription would be the same as before. This implies $\Pi(\phi_{NT}^*) = \Pi_{NT}(\phi_{NT}^*)$, and since ϕ_{NT}^* maximized profits in the absence of consumption tracking, $\phi_0^* = \phi_{NT}^*$ must hold.

Since the use of consumption tracking allows a consumer to enter the $t = 1, 2$ consumption decisions with $\alpha = 0$, her valuation of a subscription conditional on her use of consumption tracking — and hence, the optimal price for a given penalty fee — will be reduced by the monitoring cost k relative to the case in which she truly has $\alpha = 0$ and consumption tracking is not available. Furthermore, from Proposition 1 we can see $\phi_{NT}^* = 0$ when $\alpha = 0$, in which case $\Pi_{NT} = 1$. Thus, given the consumer optimally tracks her consumption at ϕ_1^*, $\Pi(\phi_1^*) = \Pi_{NT}(\phi_1^* | \alpha = 0, d_\tau = 1) - k \leq \Pi_{NT}(0 | \alpha = 0, d_\tau = 1) - k = 1 - k$. Next, since $\Pi_{NT}(\phi_{NT}^*) \geq 1 > 1 - k$, and recalling $\phi_0^* = \phi_{NT}^*$ with $\Pi(\phi_0^*) = \Pi(\phi_{NT}^*) = \Pi_{NT}(\phi_{NT}^*)$, we can be assured that $\Pi(\phi_0^*) > \Pi(\phi_1^*)$. In turn, this guarantees $\phi^* = \phi_0^* = \phi_{NT}^*$ along with $p^* = p_{NT}^*$. Moreover, since the consumer does not track (or expect to track) her consumption, it follows that this equilibrium will be entirely identical to the corresponding benchmark equilibrium without consumption tracking.

Thus, as we can see, the potential availability of consumption tracking becomes irrelevant if consumers' decisions to use consumption tracking occur at $t = 0$ instead of at $t = 2$. Intuitively, consumption tracking becomes irrelevant in this case because, at the time of the $t = 0$ tracking decision, a consumer in the benchmark equilibrium already expects to receive just as much value from the subscription as the value the she would receive with perfect memory of her past consumption. As a result, the consumer has no incentive to use consumption tracking as a means to help her remember her past consumption. Moreover, the firm has no incentive to adjust its contract from the benchmark equilibrium given the lack of consumer demand for consumption tracking.

Of note, with consumers' decisions to use consumption tracking at $t = 0$, the availability of consumption tracking may again be relevant if we allow consumers to be heterogeneous in $\widetilde{\alpha}$. To formally consider this possibility, suppose a λ share of consumers have $\widetilde{\alpha} = \widetilde{\alpha}_L$ and a $1-\lambda$ share of consumers have $\widetilde{\alpha} = \widetilde{\alpha}_H$, with $0 < \lambda < 1$ and $\widetilde{\alpha}_L < \widetilde{\alpha}_H \geq \alpha$. (Note here that λ now parameterizes a different type of heterogeneity than the type considered in Appendix B.1.5.) We now consider a very simple numerical example, in which the equilibrium with consumption tracking differs from the equilibrium without consumption tracking:

Example 4. *Let $\alpha = \widetilde{\alpha}_H = \frac{9}{10}$, $\widetilde{\alpha}_L = 0$, $k = 0$, and $\lambda = \frac{9}{10}$. (Note this means that consumers in the smaller, size $1-\lambda = \frac{1}{10}$ segment with $\widetilde{\alpha} = \widetilde{\alpha}_H$ are fully sophisticated, consumers in the larger, size $\lambda = \frac{9}{10}$ segment are naive, and the use of consumption tracking is costless). Then, assuming equilibria that yield higher profits are selected over equilibria that yield lower profits, the optimal contract with consumption tracking will entail $p^* = \frac{5}{8}$ and $\phi^* = 5$. In this equilibrium, all consumers subscribe to the service, while consumers with $\widetilde{\alpha} = \widetilde{\alpha}_H = \frac{9}{10}$ (but not consumers with $\widetilde{\alpha} = \widetilde{\alpha}_L = 0$) choose to use consumption tracking.*

From Proposition 1, we can verify that the equilibrium contract in Example 4 is the same as the contract in the benchmark equilibrium for the case in which $\lambda = 1$ instead of $\lambda = \frac{9}{10}$. From our prior analysis, we can verify that this benchmark equilibrium would have entailed $\Pi = \frac{5}{8} + \frac{1}{2}\big(1 - \big(1 - \frac{9}{10}\big)5\big)5 = \frac{15}{8}$, implying this would also be the expected profit for a consumer with $\widetilde{\alpha} = \widetilde{\alpha}_L = 0$ and $\lambda = \frac{9}{10}$ when consumption tracking is available in light of Remark 8. Meanwhile, the expected value of a subscription for a fully sophisticated consumer would be $\widetilde{V}(5|\alpha = 0, d_\tau = 1) = p^* - k = \frac{5}{8} = p^*$ if the consumer uses consumption tracking (noting that a consumer for whom consumption tracking endows them with a perfect memory would expect to derive as much gross value from the service as a naive consumer with $\widetilde{\alpha}_L = \widetilde{\alpha} = 0$) and $\widetilde{V}(5|\widetilde{\alpha} = \frac{9}{10}, d_\tau = 0) = \frac{1}{2} < p^*$ if the consumer does not use consumption tracking (since this consumer would always set $d_1 = 0$ and $d_2 = 1$). Thus, a fully sophisticated consumer is willing to pay for a subscription at p^* and will necessarily choose to use consumption tracking as means of avoiding the penalty fee (with certainty). The expected profit from a fully sophisticated consumer will then simply be equal to the subscription price, $p^* = \frac{5}{8}$. With our above work, the total expected profit in this equilibrium will then be $\Pi = \frac{9}{10} \cdot \frac{15}{8} + \frac{1}{10} \cdot \frac{5}{8} = \frac{7}{4}$. Importantly, this equilibrium must be different than the corresponding benchmark equilibrium without consumption tracking because fully sophisticated consumers will actually use consumption tracking to avoid penalty fees — a behavior that would not be possible had consumption tracking not been available.

We can readily verify that this equilibrium is indeed optimal. In particular, first suppose $1 \leq \phi^* \neq 5$. In this case, the profit per naive consumer must be less than with $\phi = 5$ because we know that $\phi = 5$ maximizes expected profits when all consumers are naive (with or without consumption tracking). As in the case with $\phi = 5$, a fully sophisticated consumer will still be willing to subscribe to the service while using consumption tracking as a means of avoiding the penalty fee, provided $p \leq \frac{5}{8}$. Thus the expected profit per fully sophisticated consumer is bounded above by $\frac{5}{8}$. Since expected profits are lower in both segments (and strictly lower in one segment), the total expected profit with ϕ^* must be less than with $\phi = 5$, precluding the possibility that $1 \leq \phi^* \neq 5$. If instead we suppose $\phi^* \leq 1$, it is then verifiable from our prior work that the maximum expected profit per naive consumer (i.e., that attained with p^* set equal to naive consumers' valuation of the subscription) would be given by $1 - \frac{(4-4\phi-3\phi^2)\phi^2}{10}$, which is decreasing for $\phi < \frac{\sqrt{33}-3}{6} \approx .457$ and increasing for $\phi > \frac{\sqrt{33}-3}{6} \approx .457$, and is thus maximized at $\phi = 0$ or $\phi = 1$ given $\phi \leq 1$. Inserting these values into our prior expression we see that the maximum profit per naive consumer in this case is attained with $\phi = 1$, and is equal to $1 - \frac{(4-4(1)-3(1^2))(1^2)}{10} = \frac{43}{40}$. From our prior analysis, we also know that the zero-penalty benchmark equilibrium arises with fully sophisticated consumer, yielding a profit of $\Pi = 1$, thus implying the maximum expected profit from a fully sophisticated consumer given $\phi \leq 1$ is 1. With our above work, we can see that the maximum total expected profit satisfies $\Pi \leq \frac{9}{10} \cdot \frac{43}{40} + \frac{1}{10} \cdot 1 = \frac{427}{400} < \frac{7}{4}$. This ensures that $\phi^* \leq 1$ also cannot be optimal, thus confirming that $\phi^* = 5$, with $p^* = \frac{5}{8}$, is indeed optimal in Example 4.

B.2.2. Decisions to Use Consumption Tracking at the Beginning of $t = 1$

Next, we consider the case in which consumers' potential decisions to use consumption tracking are made upon their arrival at $t = 1$. As in the case with decisions to use consumption tracking at $t = 0$, we can proceed directly to our analysis with consumption tracking because our benchmark analysis is the same regardless of the timing of consumers' decisions to use consumption tracking.

Remark 9. *In the model with decisions to use consumption tracking at the beginning of $t = 1$, the firm's optimal contract with consumption tracking is the same as its optimal contract in the benchmark equilibrium without consumption tracking: $\phi^* = \phi^*_{NT}$ and $p^* = p^*_{NT}$. Furthermore, the equilibrium will be identical to the benchmark equilibrium without consumption tracking.*

As it turns out, the model with decisions to use consumption tracking at the beginning of $t = 1$ is quite similar to the model with decisions to use consumption tracking at $t = 0$. Of particular importance, in the benchmark equilibrium a consumer will still believe upon arrival at $t = 1$ —

and not just at $t = 0$, as noted above — that she will act identically to a consumer with $\alpha = 0$; as a result, the consumer has zero demand for consumption tracking. This means the firm can earn the same profits as before, i.e. $\Pi(\phi^*_{NT}) = \Pi_{NT} \geq 1$, as long as it maintains its benchmark equilibrium contract. Moreover, any contract for which a consumer would choose to use consumption tracking will, as in the case with decisions to use consumption tracking at $t = 0$, still be bounded above by $1 - k < 1 < \Pi(\phi^*_{NT})$. It therefore follows, as in the prior case with decisions to use consumption tracking at $t = 0$, that the benchmark equilibrium contract remains optimal when consumption tracking is available and with decisions to use consumption tracking at the beginning of $t = 1$, resulting in identical market behavior and outcomes.

The fact that consumers and the firm effectively face equivalent decisions, resulting in identical behavior, in the model variants with decisions to use consumption tracking at $t = 0$ or at the beginning of $t = 1$ implies that other aspects of the former variant would carry over to the latter variant. This includes the previously-established feature of the former variant by which consumption tracking can once again matter in equilibrium if we allow heterogeneity in $\widetilde{\alpha}$. In fact, the equilibrium with consumption tracking in Example 4 would — and the fact that it cannot arise without consumption tracking — also apply in the present case with consumers' decisions to use consumption tracking at the beginning of $t = 1$ (with all logic from the discussion and mathematical derivations that follow Example 4 carrying over to this setting).

B.2.3. Decisions to Use Consumption Tracking at the End of $t = 1$

Lastly, we consider the case in which consumers' potential decisions to use consumption tracking are made at the end of $t = 1$, immediately after their $t = 1$ consumption decision. Once again, here we can proceed directly to our analysis with consumption tracking because our benchmark analysis does not depend on the timing of consumers' potential decisions to use consumption tracking.

With decisions to use consumption tracking at the end of $t = 1$, a consumer's choice of whether to use consumption tracking now depends on her perceived forgetfulness $\widetilde{\alpha}$ instead of her true forgetfulness α. That is, the consumer now tracks her consumption if $k < k_\tau(\widetilde{\alpha}, \phi)$ but not if $k > k_\tau(\widetilde{\alpha}, \phi)$. Note that this means the consumer's actual use of consumption tracking must now align with her prior expected use of consumption tracking.

Proposition 5. *In the model with decisions to use consumption tracking at the end of $t = 1$, the equilibrium with consumption tracking will entail planned nonuse ($\widetilde{D}_\tau = D_\tau = 0$). The firm's optimal contract in this equilibrium is given by $\phi^* = \phi^*_{NT}$ and $p^* = p^*_{NT}$ if $k \geq$*

$k_\tau(\widetilde{\alpha},\phi^*_{NT})$ and otherwise given by $\phi^* = \widetilde{\phi}_\tau \cdot \mathrm{I}\left[\alpha > 1 - \frac{((1-\widetilde{\alpha}^2)\widetilde{\phi}_\tau)^2 - 4\widetilde{\alpha}^2}{8(1-\widetilde{\phi}_\tau)+4(1-\widetilde{\alpha}^2)\widetilde{\phi}_\tau^2}\right] < \phi^*_{NT}$ and $p^* = \frac{1}{2} + \frac{(1+((1-\phi^*)^2-(\widetilde{\alpha}\phi^*)^2)\mathrm{I}[\phi^*<(1-\widetilde{\alpha})^{-1}])^2}{8}$.

Proof of Proposition 5

With decisions to use consumption tracking at the end of $t = 1$, $D_1(\phi)$ is still given by (A-13), while $D_2(\phi|d_1 = 0)$, $D_2(\phi|d_1 = 1, d_\tau = 0)$, and $D_2(\phi|d_1 = 1, d_\tau = 1)$ are still given by (A-1), (A-2), and (A-12), respectively. Given these relationships, it is still true that the equilibrium with consumption tracking cannot entail $D_\tau > 0$, as established in the proof of Proposition 2. Now, however, we must have $D_\tau = \widetilde{D}_\tau$ since the consumer's expected and actual thresholds for consumption tracking are both $k_\tau(\widetilde{\alpha},\phi)$. Thus, the equilibrium with consumption tracking must entail $D_\tau = \widetilde{D}_\tau = 0$. As a result, the firm's equilibrium profits at $\phi = \phi^*$ will take the form given in (A-14). Since $\phi = \phi^*_{NT}$ maximizes profits as given in (A-14), $\phi^* = \phi^*_{NT}$ (along with $p^* = p^*_{NT}$) will hold as long as $k \geq k_\tau(\widetilde{\alpha},\phi^*_{NT})$. Furthermore, as established in the proof of Proposition 1, $\Pi(\phi)$ as given in (A-14) is increasing in ϕ for all $\phi < \phi^*_{NT}$ such that $\Pi(\phi) > 1$. Thus, if $k < k_\tau(\widetilde{\alpha},\phi^*_{NT})$ and $\Pi(\phi^*) > 1$, it must be the case that $\phi^* = \max\{\phi : k \leq k_\tau(\widetilde{\alpha},\phi)\} = \widetilde{\phi}_\tau$. Meanwhile, we can use (A-14) to confirm that $\Pi(\widetilde{\phi}_\tau) > 1$ in this case if and only if $\alpha > 1 - \frac{((1-\widetilde{\alpha}^2)\widetilde{\phi}_\tau)^2 - 4\widetilde{\alpha}^2}{8(1-\widetilde{\phi}_\tau)+4(1-\widetilde{\alpha}^2)\widetilde{\phi}_\tau^2}$. This ensures that $\phi^* = \widetilde{\phi}_\tau \cdot \mathrm{I}\left[\alpha > 1 - \frac{((1-\widetilde{\alpha}^2)\widetilde{\phi}_\tau)^2 - 4\widetilde{\alpha}^2}{8(1-\widetilde{\phi}_\tau)+4(1-\widetilde{\alpha}^2)\widetilde{\phi}_\tau^2}\right] < \phi^*_{NT}$ if $k < k_\tau(\widetilde{\alpha},\phi^*_{NT})$. The expressions for p^* in each case then follow from $p^* = \widetilde{V}(\phi^*) = \frac{1+D_1^2(\phi^*)}{2}$ with D_1 as given in (A-5) for the case (that must apply in this equilibrium) with $D_\tau = \widetilde{D}_\tau = 0$. □

As Proposition 5 demonstrates, only a planned nonuse equilibrium can arise when consumers' decisions to use consumption tracking occur at the end of $t = 1$. Here, it makes sense that an unplanned equilibrium is no longer possible because the consumer's actual use of consumption tracking must now align with her prior expected use of consumption tracking (as both are now based on $\widetilde{\alpha}$). Naturally, without the possibility of an unplanned equilibrium, the potential counter-intuitive effects of consumption tracking in an unplanned equilibrium (Corollary 1) are no longer possible. More specifically, with decisions to use consumption tracking at the end of $t = 1$, the availability of consumption tracking can only lead to a decrease in the equilibrium penalty fee ($\phi^* \leq \phi^*_{NT}$), a decrease in the firm's equilibrium profits ($\Pi \leq \Pi_{NT}$), and an increase in consumer surplus ($CS \geq CS_{NT}$).

B.3. Other Model Variations and Extensions

B.3.1. Time-invariant valuations

As noted in the main text, we follow the precedent of Grubb (2015) in assuming that consumers' valuations v_1 and v_2 are temporally independent. To help us better understand the role of this assumption in our analysis, this appendix considers an alternate version of our model that lacks temporally-independent valuations. In particular, and as a simple benchmark for comparison, we now consider an opposite formulation in which consumers' valuations are the same in both periods as opposed to being independent.

Formally, we now impose $v_1 = v_2$, while maintaining that this (now common) valuation is drawn from a uniform distribution between 0 and 1. Note that the realization of v_1 is now fully informative of v_2, in contrast to our original specification in which the realization of v_1 has no bearing on v_2. Thus, with time-invariant valuations a consumer essentially learns v_2 at $t = 1$ instead of at $t = 2$ (as in our original model).

The following lemma characterizes consumer decision-making for a given penalty fee when consumers have time-invariant valuations.

Lemma 4. *For a consumer who subscribes to the service in the model with time-invariant valuations, $d_1 = \mathrm{I}[v_1 > \phi]$, $d_2 = 1$, and $d_\tau = 0$. Thus, $D_1 = D_{12} = \widetilde{D}_{12} = \max\{1-\phi, 0\}$ and $D_2 = \widetilde{D}_2 = 1$.*

Proof of Lemma 4

As in our original model, with time-invariant valuations, a consumer who subscribes to the service will always choose $d_2 = 1$ and $d_\tau = 0$ in cases with $d_1 = 0$; if $d_1 = 1$, the consumer will choose both $d_2 = 1$ and $d_\tau = 0$ if $v_2 > \phi$. If $v_1 = v_2 > \phi$, it then follows that choosing $d_1 = 1$ yields a total expected payoff of $v_1 + v_2 - \phi$ while choosing $d_1 = 0$ yields a total expected payoff of $v_2 < v_1 + v_2 - \phi$, implying the consumer will choose $d_1 = 1$ — followed by $d_\tau = 0$ and $d_2 = 1$ — if $v_1 > \phi$. If $v_1 \leq \phi$, choosing $d_1 = 1$ yields a total expected payoff of $v_1 - (\phi - v_2)\mathrm{I}[v_2 > (1-\alpha)\phi](1-d_\tau) - k \cdot d_\tau$ while choosing $d_1 = 0$ yields a total expected payoff of $v_2 = v_1 \geq v_1 - (\phi - v_2)\mathrm{I}[v_2 > (1-\alpha)\phi](1-d_\tau) - k \cdot d_\tau$, implying the consumer will choose $d_1 = 0$ — again followed by $d_\tau = 0$ and $d_2 = 1$ — if $v_1 \leq \phi$. Given v_1 is uniformly distributed between 0 and 1, it follows that $D_1 = \max\{1-\phi, 0\}$. Since $d_2 = 1$ given $d_1 = 0$ and $d_2 = 1$ given $d_1 = 1$ and $v_1 > \phi$, while $v_1 > \phi$ must hold given $d_1 = 1$ since $d_1 = 1$ itself only holds if $v_1 > \phi$, we can then confirm $D_2 = \widetilde{D}_2 = 1$. Lastly, $D_1 = \max\{1-\phi, 0\}$ and $D_2 = \widetilde{D}_2 = 1$, ensuring $D_{12} = \widetilde{D}_{12} = D_1 = \max\{1-\phi, 0\}$. □

From Lemma 4, we can see that, with time-invariant valuations, a consumer who subscribes to the service *always* consumes at $t = 2$. Anticipating this, the consumer consumes at $t = 1$ if and only if her valuation $v_1 = v_2$ is higher than the penalty fee ϕ. Intuitively, she knows she will consume at $t = 2$ regardless of her $t = 1$ consumption choice, which means consuming at $t = 1$ will be worth it if and only if the value the consumer derives from consumption exceeds the penalty fee that will inevitably be incurred following a choice to consume at $t = 1$. Note here that the consumer's consumption choices do not depend on her memory parameters α and $\widetilde{\alpha}$.

Lemma 4 also reveals that a consumer will never use consumption tracking in the model with time-invariant valuations; in effect, consumption tracking is irrelevant in this version of the model. As the next result shows, penalty fees likewise become irrelevant when we look at the full equilibrium under time-invariant valuations.

Proposition 6. *In the model with time-invariant valuations, the zero-penalty equilibrium necessarily arises in that* $\phi^* = 0$ *and* $p = 1$.

Proof of Proposition 6

Using Lemma 4, we can derive a consumer's willingness-to-pay for a subscription as $\widetilde{V}(\phi) = \widetilde{D}_1 \cdot \mathrm{E}[v_1|d_1 = 1] + \widetilde{D}_2 \cdot \mathrm{E}[v_2|d_2 = 1] - k \cdot \widetilde{D}_\tau - \widetilde{D}_{12} \cdot \phi = \max\{1-\phi, 0\} \cdot \frac{\phi+1}{2} + 1 \cdot \frac{1}{2} - k \cdot 0 - \max\{1-\phi, 0\} \cdot \phi = \frac{1+\max\{1-\phi,0\}^2}{2}$, which used $\widetilde{D}_1 = D_1 = \max\{1-\phi, 0\}$, $\widetilde{D}_2 = D_2 = 1$, and $\widetilde{D}_\tau = D_\tau = 0$. Again using Lemma 4 while setting $p = \widetilde{V}$, we can then derive the firm's maximum expected profit for a given penalty fee as $\Pi(\phi) = \widetilde{V}(\phi) + D_{12} \cdot \phi = \frac{1+\max\{1-\phi,0\}^2}{2} + \max\{1-\phi, 0\} \cdot \phi = 1 - \frac{\min\{1,\phi^2\}}{2}$. We can then observe $\Pi(\phi) = 1 - \frac{\min\{1,\phi^2\}}{2} < 1$ for all $\phi > 0$ and $\Pi(0) = 1$. This ensures $\phi^* = 0$ with $p^* = \widetilde{V}(0) = \frac{1+\max\{1-0,0\}^2}{2} = 1$. □

To understand Proposition 6's implication that the firm would never set a positive penalty fee under time-invariant valuations, recall from Lemma 4 that a consumer accurately perceives her consumption choices in this version of the model; that is, $\widetilde{D}_1 = D_1$, $\widetilde{D}_2 = D_2$, and most notably, $\widetilde{D}_{12} = D_{12}$, which means a consumer has accurate expectations regarding her likelihood of incurring the penalty fee. Due to the consumer's accurate perceptions, any potential revenues from penalty fees to the firm will be offset by an equal-sized reduction in a consumer's willingness to pay for a subscription. In turn, a consumer's willingness to pay for a subscription with a positive penalty fee will be further diminished by the fact that the penalty fee might deter the consumer from consuming in the first period, thus preventing the consumer from realizing the full value of the service in both periods. As a result, the firm's expected profits when using a positive penalty fee must be less than the expected profits without a penalty fee.

B.3.2. Heterogeneous monitoring costs

Up until now, we have assumed that all consumers face the same monitoring cost k. In this appendix, we consider the possibility that that consumers face heterogeneous costs of using consumption tracking. In particular, we now suppose that a λ share of consumers face a monitoring cost of $k = k^{\ell}$ while the remaining $1-\lambda$ share of consumers face a monitoring cost of $k = k^h$, with $0 \leq k^{\ell} < k^h$. Thus, the size $1-\lambda$ segment may be understood as being comprised of consumers who face a relatively high monitoring cost in comparison to consumers in the size λ segment. (Note that λ now represents a different form of heterogeneity than the forms considered in Appendix B.1.5 and Appendix B.2.1.)

Naturally, potential heterogeneity in monitoring costs is irrelevant if consumers do not have access to consumption tracking. Thus, the benchmark analysis will be the same as before. With consumption tracking, however, new predictions may emerge. That said, the equilibria with heterogeneous monitoring costs may often still resemble, in several key regards, the equilibria from our original analysis with consumption tracking.

Remark 10. *In the model with heterogeneous monitoring costs:*

- *if λ is sufficiently large, the firm's optimal contract will be the same as the optimal contract given in Proposition 2 with $k = k^{\ell}$;*
- *if λ is sufficiently small, the firm's optimal contract will be the same as the optimal contract given in Proposition 2 with $k = k^h$.*

These observations are not surprising since they are trivially true in the limiting cases as $\lambda \to 0$ and $\lambda \to 1$. In any event, Remark 10 implies that "traditional" planned and unplanned nonuse equilibria can still arise, with the caveat that these designations may only reflect the perspectives of low-k consumers with sufficiently large λ and of high-k consumers with sufficiently small λ.

In the case that the equilibrium entails planned nonuse from the perspective of low-k consumers with λ sufficiently large, the equilibrium will in fact be identical to the corresponding equilibrium with a homogeneous monitoring cost and $k = k^{\ell}$. This is because high-k consumers (and not just low-k consumers) will also not use or expect to use consumption tracking in this equilibrium, implying high-k consumers will behave identically to low-k consumers. If, however, the equilibrium entails *un*planned nonuse from the perspective of low-k consumers, high-k consumers may not be willing to pay the firm's subscription price, leading these consumers to decide against subscribing to the service. Thus, with large λ, the firm may not serve high-k consumers in an unplanned nonuse equilibrium (where nonuse is "unplanned" from the perspective of low-k consumers). This possibility is illustrated in the following example, which is a slight modification of our original Example 1:[23]

[23] Since Example 5 is equivalent to Example 1 as $\lambda \to 1$, ϕ^* and p^* are not affected, while Π, CS, and $\widetilde{D}_\tau$ are simply scaled by the share of low-k consumers (λ), provided high-k consumers do not subscribe. Indeed, we can verify that high-k consumers do not subscribe since $\widetilde{V}(2|k = k^h) = \frac{1}{2}\big(1 + \big(\frac{1}{2} - k\big)^2 \mathrm{I}\big[k < k_\tau\big(\frac{25}{28}, 2\big) = \frac{27}{196}\big]\big) < \frac{1}{2}\big(1 + \big(\frac{1}{2} - \frac{13}{100}\big)^2 = \frac{11{,}369}{20{,}000} = p^*$ given $k^h > k^{\ell}$.

Example 5. *Let $\alpha = \frac{9}{10}$, $\widetilde{\alpha} = \frac{25}{28}$, and $k^{\ell} = \frac{13}{100}$, with sufficiently large $\lambda > 0$ and any $k^h > k^{\ell}$. Then $\phi^* = \phi_\tau(\alpha, k^{\ell}) = 2$ and $p^* = \frac{11{,}369}{20{,}000}$. In this equilibrium, only consumers with $k = k^{\ell}$ subscribe to the service, while $\Pi = \frac{23{,}209\lambda}{20{,}000}$, $CS = -\frac{3{,}663\lambda}{10{,}000}$, and $D_\tau = 0 < \widetilde{D}_\tau = \frac{37\lambda}{100}$.*

As for the case in which the equilibrium entails planned nonuse from the perspective of high-k consumers with λ sufficiently small, this equilibrium may differ from the corresponding equilibrium with homogeneous monitoring costs and $k = k^h$ (even though the optimal contract is the same). As one possibility, low-k consumers may not be willing to pay the firm's subscription price, leaving these consumers unserved in equilibrium. As another possibility, low-k consumers may still subscribe, but will now choose to track their consumption in equilibrium. Thus, we can now attain an equilibrium in which some consumers do make use of consumption tracking even with the firm optimally adjusting its penalty fee and subscription price relative to their benchmark levels. Moreover, this possibility, i.e. that low-k consumers may choose to track their consumption in equilibrium, can also exist in the case where the equilibrium entails *un*planned nonuse from the perspective of high-k consumers.

B.3.3. Premium Subscriptions with Heterogeneous monitoring costs

Our prior consideration of the model with heterogeneous monitoring costs presumed that the firm could still only offer one contract in equilibrium. Under this extension, however, different consumers may differ in their willingness-to-pay for a subscription. As a result, the firm could conceivably earn a higher profit if it offered two contracts instead of one. Building on this idea, we now consider the possibility that the firm can offer two subscriptions: a "premium" subscription with no penalty fee and a "basic" subscription with a positive penalty fee.

As we find, there are many equilibria in which it remains optimal for the firm to offer a single contract to all consumers. For instance, if k^{ℓ} is sufficiently large, then consumption tracking will be prohibitively costly for consumers with $k = k^{\ell}$ as well as for consumers with $k = k^h > k^{\ell}$. As a result, the benchmark equilibrium contract with $\phi = \phi^*_{NT}$ and $p = p^*_{NT}$ will maximize profits within each consumer segment, implying it is optimal for the firm to offer its benchmark equilibrium contract to all consumers (and thus not exercise its option to offer two contracts). Naturally, this scenario could either entail a single "premium" subscription if $\phi^*_{NT} = 0$ or a single "basic" subscription if $\phi^*_{NT} > 0$. It is also optimal for the firm to offer a single contract if k^h is sufficiently small. This is because, with sufficiently small k^h (implying $k^{\ell} < k^h$ is also sufficiently small), the firm would maximize its profits within each consumer segment by setting its penalty fee to zero (see Remark 3 and surrounding discussion). In other words, a single "premium" subscription is optimal with sufficiently small k^h.

With all that said, there are scenarios in which it is optimal for the firm to offer two subscriptions. Our next remark highlights one of these scenarios.

Remark 11. *In the model with heterogeneous monitoring costs, suppose k^h is sufficiently large, $k^\ell > 0$ is sufficiently small, and $\phi^*_{NT} > 0$. It is then optimal for the firm to offer a basic subscription with $\phi = \phi^*_{NT}$ and $p = p^*_{NT}$ as well as a premium subscription with $\phi = 0$ and $p = 1$. In equilibrium, consumers with $k = k^h$ choose the basic subscription and consumers with $k = k^\ell$ choose the premium subscription.*

Proof of Remark 11.

With sufficiently large k^h, Proposition 2 implies that the firm maximizes profits among consumers with $k = k^h$ by setting $\phi = \phi^*_{NT}$ and $p = p^*_{NT}$, while consumers with $k = k^\ell < k_\tau(\widetilde{\alpha}, \phi^*_{NT})$ must have $\widetilde{V}(\phi^*_{NT}) < p^*_{NT}$ since consumers with $k = k^h$ expect to follow memory-based avoidance in this equilibrium, identical to the expectations that would be held by a consumer with perfect memory (α), implying $\widetilde{V}(\phi^*_{NT}|k = k^\ell) < \widetilde{V}(\phi^*_{NT}|k = 0) = \widetilde{V}(\phi^*_{NT}|\alpha = 0) = \widetilde{V}(\phi^*_{NT}|k = k^h) = p^*_{NT}$. From Proposition 2, it is also readily verifiable that $\alpha < \min\left\{\widehat{\alpha}, 1 - \frac{(1-\widetilde{\alpha}^2)^2(\min\{\phi_\tau,\widetilde{\phi}_\tau\})^2-4\widetilde{\alpha}^2}{8(1-\min\{\phi_\tau,\widetilde{\phi}_\tau\})+4(1-\widetilde{\alpha}^2)(\min\{\phi_\tau,\widetilde{\phi}_\tau\})^2}\right\}$ must hold for sufficiently small k, ensuring the firm maximizes profits among consumers with $k = k^\ell$ by setting $\phi = 0$ and $p = 1$. As a result, the firm maximizes its profits by offering a basic subscription with $\phi = \phi^*_{NT}$ and $p = p^*_{NT}$ as well as a premium subscription with $\phi = 0$ and $p = 1$. Note here that consumers with $k = k^h$ are indifferent between the basic and premium subscriptions, but it is assumed that higher-profit equilibria are selected over lower-profit equilibria, ensuring consumers with $k = k^h$ still choose the basic subscription. $\square$

While the above scenario describes a case in which consumers with $k = k^\ell$ are served with the premium subscription, the following example shows how it may also be optimal to provide the premium subscription to consumers with $k = k^h$.

Example 6. *Let $\alpha = \frac{9}{10}$, $\widetilde{\alpha} = \frac{25}{28}$, and $k^\ell = \frac{13}{100}$, with sufficiently large $\lambda > 0$ and any $k^h > k^\ell$. Then it is optimal for the firm to offer a basic subscription with $\phi = \phi_\tau(\alpha, k^\ell) = 2$ and $p = \frac{11{,}369}{20{,}000}$ as well as a premium subscription with $\phi = 0$ and $p = 1$. In this equilibrium, consumers with k^ℓ choose the basic subscription and consumers with $k = k^h$ choose the premium subscription, while $\Pi = \frac{20{,}000+3{,}209\lambda}{20{,}000}$, $CS = -\frac{7{,}326\lambda}{20{,}000}$, and $D_\tau = 0 < \widetilde{D}_\tau = \frac{37}{100}$.*

Example 6 considers the same scenario as Example 5, except now the firm can offer more than one subscription. Here, the firm still offers the same basic subscription to low-k consumers. Now, however, the firm also offers a premium (no fee) subscription. In doing so, the firm is able to serve high-k consumers while increasing its total profits (since $\Pi = \frac{20{,}000+3{,}209\lambda}{20{,}000} > \frac{23{,}209\lambda}{20{,}000}$ for $\lambda < 1$).

Thus, we have seen how offering an additional premium contract (with no penalty fee) can allow the firm to increase its total profits, both as a means of serving consumers with $k = k^\ell$ and as a means of serving consumers with $k = k^h$. In both cases, consumers served by the premium contract may have otherwise been excluded from the equilibrium if they comprised a sufficiently small share of consumers (such that the firm would optimally focus on exclusively serving the other segment). As we noted earlier, however, there are many also cases in which both segments are optimally served by the same (single) contract.

B.3.4. Account Freezing

In this appendix, we explore the possibility that a consumer can have her account "frozen" before incurring a penalty fee. Formally, we now consider a variation of our model in which a consumer is prevented from choosing $d_2 = 1$ — in effect, having her account frozen — in cases where $d_1 = 1$ and $\phi > 0$.[24] The following result describes the potential consumption profiles that can arise in the model with this account freezing feature.

Proposition 7. *For a consumer who subscribes to the service in the model with account freezing:*

(i) [unrestrained use] *if* $\phi = 0$*, then* $D_1 = D_2 = D_{12} = \widetilde{D}_{12} = 1$*;*

(ii) [avoidance] *if* $\phi > 0$*, then* $D_1 = \frac{1}{2}$*,* $D_2 = \frac{1}{2}$*, and* $D_{12} = \widetilde{D}_{12} = 0$*.*

Proof Proposition 7.

Part (i). If $\phi = 0$, account freezing is inactive. It then follows from our original analysis without account freezing in Appendix A.1 that $D_1 = D_2 = D_{12} = \widetilde{D}_{12} = 1$ in this case.

Part (ii). If $\phi > 0$, account freezing is active. This is equivalent to taking $\phi > 0$ to be sufficiently large in our original benchmark model without account freezing (or consumption tracking), or more specifically, taking $\phi > 0$ to be large enough to ensure that the consumer never consumes and never expects to consume at $t = 2$ conditional on having consumed at $t = 1$. This corresponds to part (i) of Lemma 1, which yields $D_1 = D_2 = \frac{1}{2}$ and $D_{12} = \widetilde{D}_{12} = 0$, as desired. □

Proposition 7 is fairly straightforward. From part (i), if there is no penalty fee, account freezing is irrelevant and the consumer is free to consume in both periods. From part (ii), with a positive penalty fee account freezing would take effect following consumption in the first period. Anticipating this, a consumer consumes in the first period if and only if her $t = 1$ valuation is above average (i.e. at least $\frac{1}{2}$). Naturally, if she consumes at $t = 1$, her account is frozen, preventing her

[24] Since account freezing would not be relevant for a subscription that provides two units of consumption without penalty, here we assume that an account can only be frozen if $\phi > 0$.

from consuming at $t = 2$; however, if she abstains at $t = 1$ she is then free to consume at $t = 2$ without penalty.

The next result describes the market equilibrium with account freezing:

Proposition 8. *In the model with account freezing, the zero-penalty benchmark equilibrium necessarily arises in that $\phi^* = 0$ and $p^* = 1$ regardless of α, $\widetilde{\alpha}$, and k; in this equilibrium, $D_{12} = \widetilde{D}_{12} = 1$ and $D_\tau = 0$, while $\Pi = 1$ and $CS = 0$.*

Proof of Proposition 8.

Suppose $\phi > 0$. The consumer would consume at $t = 1$ if and only if $v_1 > \frac{1}{2}$ while consuming at $t = 2$ if and only if $d_1 = 0$. The consumer's valuation of a subscription would then be $\widetilde{V}(\phi|\phi > 0) = \Pr[v_1 > \frac{1}{2}] \cdot \mathrm{E}[v_1|v_1 > \frac{1}{2}] + (1 - \Pr[v_1 > \frac{1}{2}]) \cdot \mathrm{E}[v_2] = \frac{1}{2} \cdot 34 + \frac{1}{2} \cdot \frac{1}{2} = \frac{5}{8}$. The firm's maximum profits from setting $p = V$ would then be $\Pi = V + D_{12} \cdot \phi = \frac{5}{8} + 0 \cdot \phi = \frac{5}{8}$ since $D_{12} = 0$ given $\phi > 0$ from part (ii) of Proposition 7. Since $\phi = 0$ and $p = \widetilde{V}(0) = 1$ (noting $\widetilde{V}(0) = \mathrm{E}[v_1] + \mathrm{E}[v_2] = \frac{1}{2} + \frac{1}{2} = 1$ given $\phi = 0$) imply $\Pi = 1 > \frac{5}{8}$, $\phi^* = 0$ and $p^* = 1$ must be optimal with account freezing. $\square$

Proposition 8 shows that with an account freezing feature it is never optimal for the firm to use a positive penalty fee. To understand why, recall that with a positive penalty fee a consumer would only consume *once* (and only expect to consume once), and thus never incur the penalty fee (Proposition 7, part i). Without a penalty fee, however, the consumer consumes (and expects to consume) in both periods. As a result, the consumer is willing to pay more for a subscription in the absence of a penalty fee, and since the firm would not collect any penalty fees even with a positive penalty fee, the firm will attain higher profits (through a higher subscription price) if it does not use a positive penalty fee.

While our preceding analysis described the implications of an exogenously-imposed account freezing feature, it is also worth considering whether (or under what circumstances) such a feature may be implemented in the first place. To do this, we now suppose that the potential implementation of an account freezing feature is endogenously determined. In particular, we now model a decision of whether or not to implement account freezing, which is made prior to all other decisions in the model. As for the question of *who* gets to make this decision, we consider three possibilities: (i) the firm, which seeks to maximize its expected profits, Π; (ii) consumers, who seek to maximize their *subjective* expectation of consumer surplus, i.e. CS except calculated based on $\widetilde{\alpha}$ instead of α; and (iii) a "social planner," who may be thought of as a benevolent regulator and seems to maximize total welfare, given as the sum of Π and CS.

Since the zero-penalty equilibrium that arises with account freezing (Proposition 8) also often arises in the original model without account freezing, the decision-maker (whether the firm, a consumer, or the social planner) will often be indifferent regarding the potential implementation of account freezing. For this reason, our analysis of endogenous account freezing focuses on the less trivial cases where the decision-maker is *not* indifferent.

Proposition 9. *In the model with endogenously-implemented account freezing, in cases where the decision-maker is not indifferent:*

(i) *if the firm gets to decide, it will never implement account freezing;*

(ii) *if consumers get to decide, they will never implement account freezing;*

(iii) *if the social planner gets to decide, it will always implement account freezing.*

Proof of Proposition 9.

Part (i). With account freezing, $\Pi = 1$ must always hold from Proposition 8. Without account freezing, $\Pi = 1$ is always attainable by setting $\phi = 0$ and $p = 1$; therefore, $\Pi \geq 1$ without account freezing. The result then follows by comparing Π with and without account freezing.

Part (ii). With account freezing, $CS = 0$ must always hold from Proposition 8. In addition, $\widetilde{CS} = CS = 0$ must hold since consumers have accurate initial perceptions of their future behavior (i.e. $\widetilde{D_1} = D_1 = \widetilde{D}_2 = D_2 = 1$ and $\widetilde{D}_\tau = D_\tau = 0$) in this (zero-penalty) equilibrium with account freezing. Without account freezing, consumers would never expect $\widetilde{CS} < 0$ because $\widetilde{CS} = CS = 0$ can always be realized by simply choosing not to subscribe. Thus, $\widetilde{CS} \geq 0$ without account freezing. The result then follows by comparing $\widetilde{CS}$ with and without account freezing.

Part (iii). With account freezing, expected total welfare is given by $\Pi + CS = 1 + 0 = 1$, as implied by Proposition 8. Noting that payments of p and ϕ only amount to transfers that do not affect total welfare, without account freezing total welfare can be expressed as $\Pi + CS = D_1 \cdot \mathrm{E}[v_1|d_1 = 1] + D_2 \cdot \mathrm{E}[v_2|d_2 = 1] - k \cdot D_\tau \leq D_1 \cdot \mathrm{E}[v_1|d_1 = 1] + D_2 \cdot \mathrm{E}[v_2|d_2 = 1] \leq \mathrm{E}[v_1|d_1 = 1] + \mathrm{E}[v_2|d_2 = 1] = \frac{1}{2} + \frac{1}{2} = 1$. Thus, total welfare must satisfy $\Pi + CS \leq 1$ without account freezing. The result then follows by comparing $\Pi + CS$ with and without account freezing. $\square$

As Proposition 9 demonstrates, the potential endogenous implementation of an account freezing feature hinges on *who* gets to make this decision. From part (i), the firm will never choose to implement account freezing (in cases where it is not indifferent). This makes sense from the standpoint that account freezing guarantees the zero-penalty benchmark equilibrium with $\Pi = 1$. Since it is always possible for the firm to attain $\Pi = 1$ by setting its penalty fee to zero, any other equilibria that might arise without account freezing — and would thus be prevented by the implementation of account freezing — would entail higher profits.

From part (ii) of Proposition 9, consumers would also never choose to implement account freezing (again in cases where they are not indifferent). Intuitively, account freezing guarantees the realization of the zero-penalty equilibrium, in which consumers attain — and expect to attain — zero surplus as the firm sets its subscription price equal to a consumer's willingness-to-pay. Meanwhile, a consumer who does not subscribe to the service also attains — and expects to attain — zero surplus. Thus, in other equilibria that can arise without account freezing, if a consumer chooses to subscribe while expecting nonzero surplus then it must be the case that she expects *positive* surplus.[25] As a result, a consumer would never expect to benefit from — and may sometimes expect to be hurt by — the implementation of an account freezing feature.

Even though the firm and consumers would never choose to implement an account freezing feature, part (iii) of Proposition 9 implies that an account freezing feature can only *improve* total welfare — and would thus be implemented by a social planner (or regulator) who seeks to maximize total welfare. To understand why, recall that the implementation of account freezing guarantees the zero-penalty equilibrium in which all consumers consume in both periods while never using — or incurring a cost from using — consumption tracking. Meanwhile, total welfare is maximized when consumers consume (thus deriving value from the service) in both periods while never incurring the monitoring cost.[26] By guaranteeing the zero-penalty equilibrium, the social planner would therefore choose to implement an account freezing feature (in cases where the social planner is not indifferent) as this can only lead to higher total welfare than would otherwise be attained.

The prediction that an account freezing feature could increase total welfare despite mutual opposition by the firm and consumers can be reconciled in light of consumers' potentially inaccurate expectations. Namely, in these cases an account freezing feature leads to an increase in total welfare through an increase in consumer surplus that consumers fail to anticipate. In other words, our analysis indicates that an account freezing feature can only help consumers even though consumers incorrectly expect to be hurt by such a feature.

[25] Specifically, this can happen (in the original model without account freezing) in cases where trackers and non-trackers are both served yet differ in their initial willingness-to-pay for a subscription. In these cases, consumers in the higher willingness-to-pay segment will expect positive consumer surplus.

[26] This is true because all other determinants of the two inputs in total welfare — i.e. Π and CS — are simply transfers (whether through p or ϕ) that only affect the distribution of total welfare. Formally, this can be seen from the fact that total welfare reduces to $\Pi + CS = \mathrm{E}[d_1 v_1 + d_2 v_2 - d_\tau k]$.