Econometrica Supplementary Material

SUPPLEMENT TO "COMMITMENT, FLEXIBILITY, AND OPTIMAL SCREENING OF TIME INCONSISTENCY"

(*Econometrica*, Vol. 83, No. 4, July 2015, 1425–1465)

BY SIMONE GALPERTI

Appendix B contains all omitted proofs of the main paper. Appendix C contains the calculations for the illustrative example. Appendix D discusses the case of outside options with type-dependent values. Appendix E discusses the case of finitely many states.

APPENDIX B: OMITTED PROOFS

B.1. *Proof of Proposition 3.1 and Corollary 3.1*

IF $\sigma > 0$, (IR) MUST BIND; if $\sigma = 0$, assume w.l.o.g. that (IR) holds with equality. The problem becomes

$$\max_{\boldsymbol{\alpha}^t}\left\{\int_{\underline{s}}^{\overline{s}}\left[u_1\big(\boldsymbol{\alpha}^t(s);s\big)-c\big(\boldsymbol{\alpha}^t(s)\big)\right]dF\right\}\quad \text{s.t.}\quad \text{(IC)}.$$

Ignoring (IC), this problem has a unique solution (up to $\{\underline{s},\overline{s}\}$): $\boldsymbol{\alpha}^t \equiv \mathbf{e}$. Since $\mathbf{e}$ is increasing and $t > 0$, by standard arguments, there is $\boldsymbol{\pi}^t_{\mathbf{e}}$ such that $(\mathbf{e}, \boldsymbol{\pi}^t_{\mathbf{e}})$ satisfies (IC). Specifically, for every s,

$$\boldsymbol{\pi}^t_{\mathbf{e}}(s) = u_2\big(\mathbf{e}(s);s,t\big) - \int_{\underline{s}}^{s} tb\big(\mathbf{e}(y)\big)\,dy - k,$$

where $k \in \mathbb{R}$. Since $\mathbf{e}$ is differentiable,

$$\frac{d\boldsymbol{\pi}^t_{\mathbf{e}}(s)}{ds} = \frac{\partial u_2\big(\mathbf{e}(s);s,t\big)}{\partial a}\frac{d\mathbf{e}(s)}{ds},$$

which equals $c'(\mathbf{e}(s))\frac{d\mathbf{e}(s)}{ds}$ if and only if $t = 1$ by the definition of $\mathbf{e}$ and Assumption 2.1. The expression of $\frac{d\mathbf{q}^t}{ds}$ follows from the definition of u_1 and u_2.

B.2. *Proof of Corollary 4.2*

Being increasing, $\mathbf{a}^I_{sb}$ is differentiable a.e. on $[\underline{v}, \overline{v}]$. If $\frac{d\mathbf{a}^I_{sb}}{dv} > 0$ at v, then using condition (E),

$$\frac{d\mathbf{p}^I_{sb}/dv}{d\mathbf{a}^I_{sb}/dv} = vb'\big(\mathbf{a}^I_{sb}(v)\big) - 1 \quad\text{and}\quad \frac{d\mathbf{p}^I_{fb}/dv}{d\mathbf{a}^I_{fb}/dv} = vb'\big(\mathbf{a}^I_{fb}(v)\big) - 1.$$

The result follows from $b'' < 0$ and Theorem 4.1(a).

 DOI: 10.3982/ECTA11851

B.3. *Proof of Lemma A.2*

(Continuity in x). Suppress r^C. For $x \in (0,1) \setminus \{x^m\}$, z is continuous, so $Z'(x) = z(x)$. If $\Omega(x) < Z(x)$, by definition, $\omega(\cdot)$ is constant in a neighborhood of x. Suppose $\Omega(x) = Z(x)$. Since Ω is convex and $\Omega \le Z$, their right and left derivatives satisfy $\Omega^+(x) \le Z^+(x)$ and $\Omega^-(x) \ge Z^-(x)$. Since $\Omega^-(x) \le \Omega^+(x)$ and Z is differentiable at x, $\Omega^-(x) = \Omega^+(x)$; so ω is continuous at x. Finally, consider x^m. If $v^m = \overline{v}^I$, then $x^m = 1$ and we are done. For $x^m \in (0,1)$, ω is continuous if $\Omega(x^m) < Z(x^m)$ when z jumps at x^m. Recall that $z(x^m-) = \lim_{v\uparrow v^m} w^I(v; r^C)$ and $z(x^m+) = z(x^m) = \lim_{v\downarrow v^m} w^I(v; r^C)$. By expression (A.8), z can only jump down at x^m, so $z(x^m-) > z(x^m)$. Suppose $\Omega(x^m) = Z(x^m)$. By the previous argument, $\Omega^+(x^m) \le Z^+(x^m) = z(x^m)$. By convexity, $\omega(x) \le \Omega^-(x^m)$ for $x \le x^m$. So, for x close to x^m from the left, we get the following contradiction:

$$\Omega(x) = \Omega\big(x^m\big) - \int_x^{x^m} \omega(y)\,dy > Z\big(x^m\big) - \int_x^{x^m} z(y)\,dy = Z(x).$$

(Continuity in r^C). Given x, $Z(x; r^C)$ is continuous in r^C. So Ω is continuous if $x \in \{0,1\}$, since $\Omega(0; r^C) = Z(0; r^C)$ and $\Omega(1; r^C) = Z(1; r^C)$. Consider $x \in (0,1)$. For $r^C \ge 0$, by definition, $\Omega(x; r^C) = \min\{\tau Z(x_1; r^C) + (1-\tau)Z(x_2; r^C)\}$ over all $\tau, x_1, x_2 \in [0,1]$ such that $x = \tau x_1 + (1-\tau)x_2$. By continuity of $Z(x; r^C)$ and the Maximum Theorem, $\Omega(x, \cdot)$ is continuous in r^C for every x. Moreover, $\Omega(\cdot; r^C)$ is differentiable in x with derivative $\omega(\cdot; r^C)$. Fix $x \in (0,1)$ and any sequence $\{r_n^C\}$ with $r_n^C \to r^C$. Since $\Omega(x; r_n^C) \to \Omega(x; r^C)$, Theorem 25.7, p. 248, of Rockafellar (1970) implies $\omega(x; r_n^C) \to \omega(x; r^C)$.

B.4. *Proof of Lemma A.6*

Recall that $\overline{w}^I(\underline{v}^I) = \omega(0)$ and $w^I(\underline{v}^I) = z(0)$. If $\omega(0) > z(0)$, since z is continuous on $[0, x^m]$ and ω is increasing, there is $x > 0$ such that $\omega(y) > z(y)$ for $y \le x$. Since $Z(0) = \Omega(0)$, we get the contradiction

$$Z(x) = Z(0) + \int_0^x z(y)\,dy < \Omega(0) + \int_0^x \omega(y)\,dy = \Omega(x).$$

If $\omega(0) < z(0)$, let $\hat{x} = \sup\{x \mid \forall x' < x,\ \omega(x') < z(x')\}$. By continuity, $\hat{x} > 0$. Then, for $0 < x < \hat{x}$,

$$Z(x) = Z(0) + \int_0^x z(y)\,dy > \Omega(0) + \int_0^x \omega(y)\,dy = \Omega(x).$$

It follows that $v_b \ge (F^I)^{-1}(\hat{x}) > \underline{v}^I$.

B.5. *Proof of Corollary 4.3*

Let $t^C = 1$. Since F is uniform, $F^i(v) = v - \underline{v}^i$. Using (A.7),

$$w^I(v; r^C) = \begin{cases} (v/t^I)(1 + r^C(1 - 2t^I)) + r^C\underline{v}^I, & \text{if } v \in [\underline{v}^I, \underline{v}^C), \\ (v/t^I)(1 + r^C(t^I - 1)^2), & \text{if } v \in [\underline{v}^C, \overline{v}^I]. \end{cases} \tag{B.1}$$

The function w^I is continuous at $\underline{v}^C$. It is strictly increasing and greater than v/t^I on $[\underline{v}^C, \overline{v}^I]$, as $r^C > 0$ and $t^I < 1$; w^I is strictly increasing on $[\underline{v}^I, \underline{v}^C)$ if and only if $t^I \leq 1/2$ or $r^C < (2t^I - 1)^{-1} = \overline{r}^C$.

Consider first v^{b} and v_{b}, when $v^{\mathrm{b}} > v_{\mathrm{b}}$. If $t^I \leq 1/2$ or $r^C < \overline{r}^C$, then w^I is strictly increasing and equals $\overline{w}^I$ (see the proof of Theorem 4.1); so $\bar{\mathbf{a}}^I$ (see (A.9)) is strictly increasing on $[\underline{v}^I, \overline{v}^I]$, and $v_{\mathrm{b}} = \underline{v}^I$. Otherwise, $v_{\mathrm{b}} \geq \underline{v}^C > \underline{v}^I$ and v_{b} is characterized by (A.16):

$$(v_{\mathrm{b}} - \underline{v}^I)^2 = \frac{r^C(t^I)^2}{1 + r^C(t^I - 1)^2}(\underline{v}^C - \underline{v}^I)^2. \tag{B.2}$$

Since w^I is strictly increasing on $[v_{\mathrm{b}}, \overline{v}^I]$, it equals $\overline{w}^I$. Using (A.15), v^{b} must satisfy

$$\int_{v^{\mathrm{b}}}^{\overline{v}^I} [w^I(y; r^C) - w^I(v^{\mathrm{b}}; r^C)]\, dy = -(\overline{v}^I - \underline{v}^I) r^C \int_{\overline{v}^I}^{\overline{v}^C} g^C(y)\, dy. \tag{B.3}$$

The derivative of the right-hand side of (B.3) with respect to v^{b} is $-w^I_v(v^{\mathrm{b}}; r^C) \times (\overline{v}^I - v^{\mathrm{b}}) < 0$. So, for $r^C > 0$, there is a unique $v^{\mathrm{b}} > v_{\mathrm{b}}$ that satisfies (B.3). Letting $K = \int_{\overline{v}^I}^{\overline{v}^C} g^C(y)\, dy < 0$, (B.3) becomes

$$-r^C[2t^I(\overline{v}^I - \underline{v}^I)K] = (1 + r^C(t^I - 1)^2)(\overline{v}^I - v^{\mathrm{b}})^2 \tag{B.4}$$

if $v^{\mathrm{b}} \geq \underline{v}^C$, and

$$-r^C[2t^I(\overline{v}^I - \underline{v}^I)K] = r^C(t^I)^2(\overline{v}^I - \underline{v}^C)^2 + (1 + r^C(1 - 2t^I))(\overline{v}^I - v^{\mathrm{b}})^2$$

if $v^{\mathrm{b}} < \underline{v}^C$. So, if $t^I > 1/2$, the function $v_{\mathrm{b}}(r^C)$ is constant at $\underline{v}^I$ for $r^C < \overline{r}^C$, and at $\overline{r}^C$, it jumps from $\underline{v}^I$ to $\underline{v}^C$. Monotonicity for $r^C > \overline{r}^C$ follows by applying the Implicit Function Theorem to (B.2):

$$\frac{dv_{\mathrm{b}}}{dr^C} = \frac{1}{2}\left[\frac{t^I}{1 + r^C(t^I - 1)^2}\right]^2 \frac{(\underline{v}^C - \underline{v}^I)^2}{(v_{\mathrm{b}} - \underline{v}^I)} > 0.$$

Similarly,

$$\frac{dv^{\mathrm{b}}}{dr^C} = \begin{cases} -\dfrac{\overline{v}^I - v^{\mathrm{b}}}{2r^C\big[1 + r^C\big(t^I - 1\big)^2\big]} < 0, & \text{if } v^{\mathrm{b}} \geq \underline{v}^C, \\ -\dfrac{\overline{v}^I - v^{\mathrm{b}}}{2r^C\big(1 + r^C\big(1 - 2t^I\big)\big)} < 0, & \text{if } v^{\mathrm{b}} < \underline{v}^C; \end{cases}$$

for the second inequality, recall that $v_{\mathrm{b}} < v^{\mathrm{b}} < \underline{v}^C$ if and only if $t^I \leq 1/2$ or $r^C < \overline{r}^C$.

Consider now the behavior of $\mathbf{b}^I(r^C) = b(\mathbf{a}^I_{sb})$, which matches that of $\mathbf{a}^I_{sb}$ for any r^C. By Theorem 4.1 and Assumption 2.1, $\mathbf{b}^I(v; r^C) \in (b(\underline{a}), b(\overline{a}))$. Also, $\mathbf{b}^I(v; r^C)$ solves $\max_{y \in [b(\underline{a}), b(\overline{a})]}\{y\overline{w}^I(v; r^C) + \xi(y)\}$. By strict concavity of $\xi(y)$, it is enough to study how $\overline{w}^I(r^C)$ relates to v/t^I. The function $\overline{w}^I(\cdot; r^C)$ crosses v/t^I only once at $v^* \in (\underline{v}^I, \overline{v}^I)$. Also, $\overline{w}^I(v; r^C) = w^I(v; r^C)$ on $[v_{\mathrm{b}}, v^{\mathrm{b}}]$. So, it is enough to show that, as r^C rises, $w^I(v^{\mathrm{b}}(r^C); r^C)$ falls and $w^I(v_{\mathrm{b}}(r^C); r^C)$ rises.

LEMMA B.1: *Suppose* v^{b} *and* v_{b} *are characterized by* (A.15) *and* (A.16). *If* $w^I_v(v^{\mathrm{b}}; r^C) > 0$ *and* $w^I_v(v_{\mathrm{b}}; r^C) > 0$, *then* $\frac{d}{dr^C} w^I(v^{\mathrm{b}}(r^C); r^C) < 0$ *and* $\frac{d}{dr^C} w^I(v_{\mathrm{b}}(r^C); r^C) > 0$.

PROOF: It follows by applying the Implicit Function Theorem to (A.15) and (A.16). *Q.E.D.*

Consider $w^I(v_{\mathrm{b}}(r^C); r^C)$. If $t^I \leq 1/2$ or $r^C < \overline{r}^C$, then $v_{\mathrm{b}}(r^C) = \underline{v}^I$ and $w^I_r(\underline{v}^I; r^C) = (1 - t^I)(\underline{v}^I/t^I) > 0$. If $t^I > 1/2$, then $w^I(\underline{v}^I; r^C) \uparrow w^I(\underline{v}^I, \overline{r}^C) = w^I(\underline{v}^C, \overline{r}^C)$ as $r^C \uparrow \overline{r}^C$. By Lemma B.1, $w^I(v_{\mathrm{b}}(r^C); r^C)$ increases in r^C, for $r^C > \overline{r}^C$, because $w^I_v(v_{\mathrm{b}}(r^C); r^C) > 0$ when $v_{\mathrm{b}} > \underline{v}^C$. Similarly, $w^I(v^{\mathrm{b}}(r^C); r^C)$ decreases in r^C, because $w^I_v(v^{\mathrm{b}}(r^C); r^C) > 0$ when $v^{\mathrm{b}} < v_{\mathrm{b}}$.

B.6. *Proof of Corollary 4.4*

Fix $\mathbf{a}^I_{sb}$ and recall that it minimizes $R^C(\mathbf{a}^I)$ among all increasing $\mathbf{a}^I$ equal to $\mathbf{a}^I_{sb}$ on $[\underline{v}^I, \overline{v}^I]$. Using (A.18) and $\mathbf{a}^C_{un}$ from Proposition 4.3, condition (R) becomes

$$\begin{aligned} &\big[b(\underline{a}) - b\big(\mathbf{a}^C_{fb}\big(\underline{v}^C\big)\big)\big] \int_{\underline{v}^I}^{v_{\mathrm{u}}} g^I(v)\, dv \\ &\quad \geq R^C\big(\mathbf{a}^I_{sb}\big) + \int_{\underline{v}^C}^{\overline{v}^C} b\big(\mathbf{a}^C_{fb}(v)\big) G^C(v)\, dF^C \\ &\qquad - b\big(\mathbf{a}^C_{fb}\big(\underline{v}^C\big)\big) \int_{\underline{v}^I}^{\underline{v}^C} g^I(v)\, dv. \end{aligned}$$

Since $\mathbf{a}^C_{fb}$ and $\mathbf{a}^I_{sb}$ are infeasible, the right-hand side is positive. $R^C(\mathbf{a}^I_{sb})$ has been minimized. The result follows, since $\int_{\underline{v}^I}^{v^{\mathrm{u}}} g^I(v)\,dv < 0$.

B.7. *Proof of Lemma A.8*

The proof uses $\mathbf{b} \in \mathcal{B}$ (see the proof of Lemma A.1). Suppose $r^I > 0$. Using $\widetilde{R}^I(\mathbf{b}) = R^I(b^{-1}(\mathbf{b}))$ in (A.18), write $\widetilde{W}^C(\mathbf{b}) - r^I\widetilde{R}^I(\mathbf{b})$ as

$$VS^C\big(b^{-1}(\mathbf{b}), r^I\big) = \int_{\underline{v}^C}^{\overline{v}^C} \big[\mathbf{b}(v)w^C\big(v, r^I\big) + \xi\big(\mathbf{b}(v)\big)\big]\,dF^C + r^I \int_{\underline{v}^I}^{\underline{v}^C} \mathbf{b}(v) g^I(v)\,dv,$$

where $w^C(v, r^I) = v/t^C - r^I G^C(v)$. Note that w^C is continuous in v, except possibly at $\overline{v}^I$ if $\overline{v}^I \geq \underline{v}^C$, where it can jump up. Using the method in the proof of Theorem 4.1, let $\overline{w}^C(v; r^I)$ be the generalized version of w^C. By the argument in Lemma A.2, $\overline{w}^C(v; r^I)$ is continuous in v over $[\underline{v}^C, \overline{v}^C]$—except possibly at $\overline{v}^I$, where we can assume right- or left-continuity w.l.o.g.—and in r^I. Now, on $[\underline{v}^C, \overline{v}^C]$, let $\phi(y, v; r^I) = y\overline{w}^C(v; r^I) + \xi(y)$ and

$$\overline{\mathbf{b}}^C\big(v; r^I\big) = \arg\max_{y \in [b(\underline{a}), b(\overline{a})]} \phi\big(y, v; r^I\big).$$

Since $\overline{w}^C$ is increasing by construction, $\overline{\mathbf{b}}^C$ is increasing on $[\underline{v}^C, \overline{v}^C]$ and continuous in r^I. On $[\underline{v}^I, \underline{v}^C]$, let $\overline{\mathbf{b}}^C$ be the pointwise maximizer of the second integral in VS^C. By Proposition 4.3's proof, $\overline{\mathbf{b}}^C(v; r^I)$ equals $b(\underline{a})$ on $[\underline{v}^I, v^{\mathrm{u}})$ and $b(\overline{a})$ on $[v^{\mathrm{u}}, \underline{v}^C)$.

Suppose $[v^{\mathrm{u}}, \underline{v}^C) = \emptyset$. Then $\overline{\mathbf{b}}^C$ is increasing and an argument similar to that in Lemma A.4 establishes that $\overline{\mathbf{b}}^C$ maximizes VS^C. Since such a $\overline{\mathbf{b}}^C$ is pointwise continuous in r^I, so is $VS^C(b^{-1}(\overline{\mathbf{b}}^C(r^I)), r^I)$.

Suppose $[v^{\mathrm{u}}, \underline{v}^C) \neq \emptyset$. Let $v_m = \max\{\overline{v}^I, \underline{v}^C\}$. By an argument similar to that in Lemma A.3, any optimal $\mathbf{b}^C \in \mathcal{B}$ can take only three forms on $[v^{\mathrm{u}}, \overline{v}^C]$: (1) it is constant at $\overline{\mathbf{b}}^C(v^{\mathrm{d}})$ on $[v^{\mathrm{u}}, v^{\mathrm{d}}]$, where $v^{\mathrm{d}} \in (\underline{v}^C, v_m) \cup (v_m, \overline{v}^C)$ and equals $\overline{\mathbf{b}}^C$ otherwise; (2) it is constant at $\overline{y} \in [\overline{\mathbf{b}}^C(v_m-), \overline{\mathbf{b}}^C(v_m+)]$ on $[v^{\mathrm{u}}, v^{\mathrm{d}}]$ with $v^{\mathrm{d}} = v_m$ and equals $\overline{\mathbf{b}}^C$ otherwise; (3) it is constant on $[v^{\mathrm{u}}, \overline{v}^C]$. We can first find an optimal $\mathbf{b}^C$ within each class and then pick an overall maximizer. Note that in both case (1) and (2), $\mathbf{b}^C$ has to maximize

$$\text{(B.5)} \qquad \mathbf{b}^C\big(v^{\mathrm{d}}\big)H\big(v^{\mathrm{d}}, r^I\big) + \xi\big(\mathbf{b}^C\big(v^{\mathrm{d}}\big)\big)F^C\big(v^{\mathrm{d}}\big) + \int_{v^{\mathrm{d}}}^{\overline{v}^C} \phi\big(\overline{\mathbf{b}}^C(v), v; r^I\big)\,dF^C,$$

where

$$H\left(v^{\mathrm{d}}, r^I\right) = r^I \int_{v^{\mathrm{u}}}^{\underline{v}^C} g^I(v)\,dv + \int_{\underline{v}^C}^{v^{\mathrm{d}}} \overline{w}^C\left(v, r^I\right) dF^C.$$

Note that, since $\overline{w}^C(v, r^I)$ is continuous in r^I, so is (B.5).

Case 1: Let $\overline{\mathbf{b}}^C(v_m) = \overline{\mathbf{b}}^C(v_m-)$, so that $\overline{\mathbf{b}}^C$ is continuous on $[\underline{v}^C, v_m]$. Then, (B.5) is continuous in v^{d} for $v^{\mathrm{d}} \in [\underline{v}^C, v_m]$. Hence, there is an optimal v^{d}. By an argument similar to that in Lemma A.4, there is a unique optimal $\mathbf{b}_1^C$ within this case. Let $\Phi(\mathbf{b}_1^C; r^I)$ be the value of (B.5) at $\mathbf{b}_1^C$, which is continuous in r^I.

Case 2: Let $\overline{\mathbf{b}}^C(v_m) = \overline{\mathbf{b}}^C(v_m+)$, so that $\overline{\mathbf{b}}^C$ is continuous on $[v_m, \overline{v}^C]$. Then, (B.5) is continuous in v^{d} for $v^{\mathrm{d}} \in [v_m, \overline{v}^C]$. As before, there is an optimal v^{d} and a unique optimal $\mathbf{b}_2^C$ within this case. Let $\Phi(\mathbf{b}_2^C; r^I)$ be the value of (B.5) at $\mathbf{b}_2^C$, which is continuous in r^I.

Case 3: Let $v^{\mathrm{d}} = v_m$. Then, there is a unique $\mathbf{b}^C(v^{\mathrm{d}}) \in [\overline{\mathbf{b}}^C(v_m-), \overline{\mathbf{b}}^C(v_m+)]$ which maximizes (B.5). This identifies a function $\mathbf{b}_3^C$ and value $\Phi(\mathbf{b}_3^C; r^I)$. Since $\overline{\mathbf{b}}^C(v_m-; r^I)$ and $\overline{\mathbf{b}}^C(v_m+; r^I)$ are continuous in r^I, so is $\Phi(\mathbf{b}_3^C; r^I)$.

Case 4: $\mathbf{b}^C$ is constant at $\overline{y}$ on $[v^{\mathrm{u}}, \overline{v}^C]$. Then $\overline{y} \in [b(\underline{a}), b(\overline{a})]$ has to maximize

$$\overline{y}\left[r^I \int_{v^{\mathrm{u}}}^{\underline{v}^C} g^I(v)\,dv + \int_{\underline{v}^C}^{\overline{v}^C} \overline{w}^C\left(v, r^I\right) dF^C\right] + \xi(\overline{y}).$$

The unique solution to this problem identifies a unique constant $\mathbf{b}_4^C$ and value $\Phi(\mathbf{b}_4^C; r^I)$, which is again continuous in r^I.

Now, let $\hat{\mathbf{b}}^C$ be the function that solves $\max_{j=1,2,3,4} \Phi(\mathbf{b}_j^C; r^I)$. An argument similar to that in Lemma A.5 establishes that

$$\max_{\mathbf{b}\in\mathcal{B}} VS^C\left(b^{-1}(\mathbf{b}), r^I\right) = \Phi\left(\hat{\mathbf{b}}^C; r^I\right) + b(\underline{a}) r^I \int_{\underline{v}^I}^{v^{\mathrm{u}}} g^I(v)\,dv,$$

which is therefore continuous in r^I.

Now, let $\mathbf{b}_{un}^C = b(\mathbf{a}_{un}^C)$ and let $\mathcal{B}^*$ be the set of $\mathbf{b}^C \in \mathcal{B}$ that equal $\mathbf{b}_{un}^C$ on $[\underline{v}^C, \overline{v}^C]$. By construction, $VS^C(b^{-1}(\mathbf{b}_{un}^C), r^I) = \max_{\mathbf{b}\in\mathcal{B}^*} VS^C(b^{-1}(\mathbf{b}), r^I)$. I claim that there is $\hat{\mathbf{b}}^C \in \mathcal{B}\backslash\mathcal{B}^*$ such that $VS^C(b^{-1}(\hat{\mathbf{b}}^C), r^I) > VS^C(b^{-1}(\mathbf{b}_{un}^C), r^I)$. Focus on $[v_m, \overline{v}^C]$ and recall that (w.l.o.g.) $\overline{w}^C$ is continuous on $[v_m, \overline{v}^C]$. Since $r^I > 0$, G^C implies $w^C(v, r^I) > v/t^C$ for $v \in [v_m, \overline{v}^C)$. I claim that $\overline{w}^C(v_m, r^I) > v_m/t^C$. By the logic in Lemma A.6, $\overline{w}^C(v_m, r^I) \leq w^C(v_m, r^I)$. If $\overline{w}^C(v_m, r^I) = w^C(v_m, r^I)$, the claim follows. If $\overline{w}^C(v_m, r^I) < w^C(v_m, r^I)$, then there is $v_0 > v_m$ such that $\overline{w}^C(v, r^I) = w^C(v_0, r^I)$ on $[v_m, v_0]$; so, $\overline{w}^C(v_m, r^I) = w^C(v_0, r^I) \geq v_0/t^C > v_m/t^C$. Since $\overline{w}^C$ is continuous and increasing, in either case there is

$v_1 > v_m$ such that $\overline{w}^C(v, r^I) > v/t^C$ on $[v_m, v_1]$. Construct $\hat{\mathbf{b}}^C$ by letting $\hat{\mathbf{b}}^C(v) = \arg\max_{y \in [b(\underline{a}), b(\overline{a})]} \phi(y, v; r^I)$ if $v \in [v_m, \overline{v}^C]$, and $\mathbf{b}^C_{un}(v)$ if $v \in [\underline{v}^I, v_m)$. Then, $\hat{\mathbf{b}}^C \in \mathcal{B}$, but $\hat{\mathbf{b}}^C(v) > \mathbf{b}^C_{un}(v)$ on $[v_m, v_1]$; so $\hat{\mathbf{b}}^C \notin \mathcal{B}^*$. Finally, $VS^C(b^{-1}(\hat{\mathbf{b}}^C), r^I) - VS^C(b^{-1}(\mathbf{b}^C_{un}), r^I)$ equals

$$\int_{v_m}^{\overline{v}^C} \big\{ \big[\hat{\mathbf{b}}^C(v) w^C\big(v, r^I\big) + \xi\big(\hat{\mathbf{b}}^C(v)\big)\big] - \big[\mathbf{b}^C_{un}(v) w^C\big(v, r^I\big) + \xi\big(\mathbf{b}^C_{un}(v)\big)\big]\big\}\, dF^C > 0.$$

B.8. *Proof of Proposition 4.5*

Recall that, by (E), the j-device is fully defined by $\mathbf{a}^j$ up to k^j. Given $\mathbf{a}^j$, define $h^j = U^j(\mathbf{a}^j, \mathbf{p}^j)$. Then, IC_1^{ji} becomes $h^j \geq h^i + R^j(\mathbf{a}^i)$ and (IR^j) becomes $h^j \geq 0$. Since $\Pi^j(\mathbf{a}^j, \mathbf{p}^j) = W^j(\mathbf{a}^j) - U^j(\mathbf{a}^j, \mathbf{p}^j)$, the provider solves

$$\mathcal{P}^N = \begin{cases} \max\limits_{(\mathbf{a}^j, h^j)_{j=1}^N} (1-\sigma) \sum\limits_{j=1}^N \gamma^j W^j\big(\mathbf{a}^j\big) + \sigma \sum\limits_{j=1}^N \gamma^j \big[W^j\big(\mathbf{a}^j\big) - h^j\big)\big] \\ \text{s.t.} \quad \mathbf{a}^i \text{ increasing}, \quad h^j \geq h^i + R^j\big(\mathbf{a}^i\big), \quad \text{and} \\ h^j \geq 0, \quad \text{for all } j, i. \end{cases}$$

As in the proof of Lemma A.1 and Theorem 4.1, it is convenient to work with the functions $\mathbf{b} \in \mathcal{B}$. Recall that $\widetilde{W}^j(\mathbf{b}^j) = W^j(b^{-1}(\mathbf{b}^j))$ and $\widetilde{R}^j(\mathbf{b}^i) = R^j(b^{-1}(\mathbf{b}^i))$.

Step 1: There is $b(\underline{a})$ low enough so that unused options suffice to satisfy IC_1^{ji} for $j > i$. If $j > i$, $\overline{v}^j < \overline{v}^i$ and

$$\widetilde{R}^i\big(\mathbf{b}^j\big) = -\int_{\overline{v}^j}^{\overline{v}^i} \mathbf{b}^j(v) g^i(v)\, dv - \int_{\underline{v}^j}^{\overline{v}^j} \mathbf{b}^j(v) G^{ji}(v)\, dF^j,$$

where

$$g^i(v) = \frac{t^i - 1}{t^i} v f^i(v) - \big(1 - F^i(v)\big) \quad \text{and}$$

$$G^{ji}(v) = q^j(v) - \frac{f^i(v)}{f^j(v)} q^i(v);$$

if $i > j$, $\underline{v}^j > \underline{v}^i$ and

$$\widetilde{R}^i\big(\mathbf{b}^j\big) = -\int_{\underline{v}^i}^{\underline{v}^j} \mathbf{b}^j(v) \widehat{g}^i(v)\, dv + \int_{\underline{v}^j}^{\overline{v}^j} \mathbf{b}^j(v) \widehat{G}^{ji}(v)\, dF^j,$$

where

$$\widehat{g}^i(v) = \frac{t^i - 1}{t^i} v f^i(v) + F^i(v),$$

$$\widehat{G}^{ji}(v) = \frac{t^j - 1}{t^j} v - \frac{1 - F^j(v)}{f^j(v)} - \frac{f^i(v)}{f^j(v)}\left[\frac{t^i - 1}{t^i} v - \frac{1 - F^i(v)}{f^i(v)}\right].$$

Take $j > i$. Suppose IC_1^{ji} is violated (and all other constraints hold): $h^j < h^i + \widetilde{R}^j(\mathbf{b}^i)$. Fix $\mathbf{b}^i$ for $v \geq \underline{v}^i$, and let $\mathbf{b}^i(v) = b(\underline{a})$ for $v < \underline{v}^i$. Then,

$$R^j(\mathbf{b}^i) = -b(\underline{a}) \int_{\underline{v}^j}^{\underline{v}^i} \widehat{g}^j(v)\, dv + \int_{\underline{v}^i}^{\overline{v}^i} \mathbf{b}^i(v) \widehat{G}^{ij}(v)\, dF^i.$$

LEMMA B.2: $\int_{\underline{v}^j}^{\underline{v}^i} \widehat{g}^j(v)\, dv < 0$.

PROOF: Integrating by parts,

$$\int_{\underline{v}^j}^{\underline{v}^i} \widehat{g}^j(v)\, dv = -\int_{\underline{v}^j}^{\underline{v}^i} \left(v/t^j\right) f^j(v)\, dv + F^j\left(\underline{v}^i\right)\underline{v}^i$$

$$= \int_{\underline{v}^j}^{\underline{v}^i} \left(\underline{v}^i - \left(v/t^j\right)\right) f^j(v)\, dv.$$

Note that $\underline{v}^i \leq \underline{s} \leq v/t^j$, with strict inequality for $v \in (\underline{v}^j, \underline{v}^i)$. *Q.E.D.*

So there is $b(\underline{a})$ small enough so that the $\widetilde{\mathbf{b}}^i$ just constructed satisfies $h^j \geq h^i + \widetilde{R}^j(\widetilde{\mathbf{b}}^i)$. We need to check the other constraints. For $j' < i$, the values $\mathbf{b}^i$ takes for $v < \underline{v}^i$ are irrelevant; so, $\mathrm{IC}_1^{j'i}$ are unchanged. For $\hat{j} > i$ and $\hat{j} \neq j$, it could be that $R^{\hat{j}}(\widetilde{\mathbf{b}}^i) > R^{\hat{j}}(\mathbf{b}^i)$, and $\widetilde{\mathbf{b}}^i$ may violate $\mathrm{IC}_1^{\hat{j}i}$ while $\mathbf{b}^i$ did not. But since Lemma B.2 holds for every $j > i$ and N is finite, there is $b(\underline{a})$ small enough so that IC_1^{ji} for all $j > i$.

Step 2: As usual, (IR^N) and IC_1^{jN} imply (IR^j) for $j < N$. Let $\mathcal{Y} = (\mathcal{B} \times \mathbb{R})^N$ be the subspace of $(\mathcal{X} \times \mathbb{R})^N$, where $\mathcal{X} = \{\mathbf{b} | \mathbf{b} : [\underline{v}, \overline{v}] \to \mathbb{R}\}$. Now, let $\widetilde{\Pi}(\mathbf{B}, \mathbf{h}) = \sum_{j=1}^N \gamma^j [\widetilde{W}^j(\mathbf{b}^j) - h^j]$ and $\widetilde{W}(\mathbf{B}) = \sum_{j=1}^N \gamma^j \widetilde{W}^j(\mathbf{b}^j)$. $\mathcal{P}^N$ is equivalent to

$$\widetilde{\mathcal{P}}^N = \begin{cases} \max\limits_{\{\mathbf{B},\mathbf{h}\} \in \mathcal{Y}} (1-\sigma)\widetilde{W}(\mathbf{B}) + \sigma \widetilde{\Pi}(\mathbf{B}, \mathbf{h}) \\ \text{s.t.} \quad \Gamma(\mathbf{B}, \mathbf{h}) \leq \mathbf{0}, \end{cases}$$

where $\Gamma : (\mathcal{X} \times \mathbb{R})^N \to \mathbb{R}^r$ $(r = 1 + \frac{N(N-1)}{2})$ is given by $\Gamma^1(\mathbf{B}, \mathbf{h}) = -h^N$ and, for $j = 2, \dots, r$, $\Gamma^j(\mathbf{B}, \mathbf{h}) = \widetilde{R}^i(\mathbf{b}^j) + h^j - h^i$ for $i < j$.

Step 3: Existence of interior points.

LEMMA B.3: *In* $\widetilde{\mathcal{P}}^N$, *there is* $\{\mathbf{B}, \mathbf{h}\} \in \mathcal{Y}$ *such that* $\Gamma(\mathbf{B}, \mathbf{h}) < 0$.

PROOF: $\Gamma(\mathbf{B}, \mathbf{h}) < 0$ if and only if $h^N > 0$ and $h^i > h^j + \widetilde{R}^i(\mathbf{b}^j)$ for $i < j$. For $i = 1, \dots, N$, let $\mathbf{b}^i = \mathbf{b}^i_{fb} = b(\mathbf{a}^i_{fb})$ on $[\underline{v}^i, \overline{v}]$ and possibly extend it on $[\underline{v}, \underline{v}^i)$ to include appropriate unused options. Note that these extensions are irrelevant for $\widetilde{R}^j(\mathbf{b}^i)$ if $j < i$. Recall that $\widetilde{R}^j(\mathbf{b}^i) \geq 0$ for $j < i$, and it can be easily shown that $\widetilde{R}^1(\mathbf{b}^i) \geq \widetilde{R}^j(\mathbf{b}^i)$ for $1 < j < i$. Thus, let $h^N = 1$, and for $i < N$, let $h^i = h^{i+1} + \widetilde{R}^1(\mathbf{b}^{i+1}) + 1$. Now, fix $i < N$ and consider any $j > i$. We have

$$h^i = h^j + \sum_{n=1}^{j-i} \widetilde{R}^1(\mathbf{b}^{i+n}) + (j-i) \geq h^j + \widetilde{R}^i(\mathbf{b}^j) + (j-i) > h^j + \widetilde{R}^i(\mathbf{b}^j).$$

Since $\widetilde{R}^i(\mathbf{b}^j)$ are bounded and N is finite, the vector $\mathbf{h}$ so constructed is well defined. *Q.E.D.*

Step 4: We can now use Corollary 1, p. 219, and Theorem 2, p. 221, of Luenberger (1969) to characterize solutions of $\widetilde{\mathcal{P}}^N$. Note that $(\mathcal{X} \times \mathbb{R})^N$ is a linear vector space and $\mathcal{Y}$ is a convex subset of it. By Lemma B.3, Γ has interior points. Since $\widetilde{\Pi}$ and $\widetilde{W}$ are concave ($b'' < 0$ and $c'' \geq 0$), the objective is concave and $\Gamma(\mathbf{B}, \mathbf{h})$ is convex. For $\boldsymbol{\lambda} \in \mathbb{R}^r_+$, define $L(\mathbf{B}, \mathbf{h}; \boldsymbol{\lambda})$ as

$$\begin{aligned}
&(1-\sigma)\widetilde{W}(\mathbf{B}) + \sigma\widetilde{\Pi}(\mathbf{B}, \mathbf{h}) + \lambda^N h^N - \sum_{i=1}^{N} \sum_{j<i} \lambda^{ji}\big[\widetilde{R}^j(\mathbf{b}^i) + h^i - h^j\big] \\
&\quad = \sum_{i=1}^{N} \gamma^i \left[\widetilde{W}^i(\mathbf{b}^i) - \sum_{j<i} \frac{\lambda^{ji}}{\gamma^i} \widetilde{R}^j(\mathbf{b}^i)\right] + \sum_{i=1}^{N} h^i \mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma),
\end{aligned}$$

where

$$\mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma) = \begin{cases} \displaystyle\sum_{j>i} \lambda^{ij} - \sum_{j<i} \lambda^{ji} - \sigma\gamma^i, & \text{if } i < N, \\ \displaystyle\lambda^N - \sum_{j<N} \lambda^{jN} - \sigma\gamma_N, & \text{if } i = N. \end{cases}$$

Then, $\{\mathbf{B}, \mathbf{h}\}$ solves $\widetilde{\mathcal{P}}^N$ if and only if there is $\boldsymbol{\lambda} \geq \mathbf{0}$ such that $L(\mathbf{B}, \mathbf{h}; \boldsymbol{\lambda}) \geq L(\mathbf{B}', \mathbf{h}'; \boldsymbol{\lambda})$ and $L(\mathbf{B}, \mathbf{h}; \boldsymbol{\lambda}') \geq L(\mathbf{B}, \mathbf{h}; \boldsymbol{\lambda})$ for all $\{\mathbf{B}', \mathbf{h}'\} \in \mathcal{Y}$, $\boldsymbol{\lambda}' \geq \mathbf{0}$. The first inequality is equivalent to

$$\text{(B.6)} \qquad \mathbf{b}^i \in \arg\max_{\mathbf{b} \in \mathcal{B}} \widetilde{W}^i(\mathbf{b}) - \sum_{j<i} \frac{\lambda^{ji}}{\gamma^i} \widetilde{R}^j(\mathbf{b})$$

and

$$h^i \in \arg\max_{h\in\mathbb{R}} \mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma)h. \tag{B.7}$$

The second is equivalent to

$$-h^N \le 0 \quad \text{and} \quad \lambda^N h^N = 0, \tag{B.8}$$

and, for $j > i$,

$$\widetilde{R}^i(\mathbf{b}^j) + h^j - h^i \le 0 \quad \text{and} \quad \lambda^{ij}\left[R^i(\mathbf{b}^j) + h^j - h^i\right] = 0. \tag{B.9}$$

LEMMA B.4: *If* $(\mathbf{B}, \mathbf{h}, \boldsymbol{\lambda})$ *satisfies* (B.6)–(B.9), *then* $\mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma) = 0$ *for all* i.

PROOF: By (IR^N) and IC^{iN}, $h^i \ge 0$ for all i; so, $\mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma) \ge 0$ for all i. Since $(1-\sigma)\widetilde{W}(\mathbf{B}) + \sigma\widetilde{\Pi}(\mathbf{B}, \mathbf{h})$ is bounded below by $\mathbb{E}(u_1(a^{\text{nf}}; s)) - c(a^{\text{nf}}) > 0$, then $\mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma) \le 0$ for all i. *Q.E.D.*

COROLLARY B.5: *If* $\sigma = 0$, *then* $\boldsymbol{\lambda} = \mathbf{0}$. *If* $\sigma > 0$, IR^N *binds and*, *for every* $i < N$, *there is* $j > i$ *such that* IC^{ij} *binds*.

PROOF: Lemma B.4 implies the second part. For the first part, since $\mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma) = 0$ for all i,

$$\begin{aligned} 0 &= \sum_{i=1}^{N} \mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma) \\ &= \sum_{i=1}^{N-1}\left[\sum_{j>i}\lambda^{ij} - \sum_{j<i}\lambda^{ji}\right] + \lambda^N - \sum_{j<N}\lambda^{jN} - \sigma = \lambda^N - \sigma. \end{aligned}$$

So, if $\sigma = 0 = \lambda^N$, then $\mu^N(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma) = 0$ implies $\sum_{j<N}\lambda^{jN} = 0$. Hence, $\lambda^{jN} = 0$ for $j < N$. Suppose for all $j \ge i+1$, $\lambda^{nj} = 0$ for all $n < j$. Then, $\mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma) = 0$ implies $\sum_{j<i}\lambda^{ji} = \sum_{j>i}\lambda^{ij} = 0$. Hence, $\lambda^{ji} = 0$ for all $j < i$. *Q.E.D.*

So, although by $\mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma) = 0$ any $h^i \in \mathbb{R}$ solves (B.7), the upward binding constraints pin down $\mathbf{h}$, once $\mathbf{B}$ has been chosen.

Thus, $\widetilde{\mathcal{P}}^N$ has a solution if there is $(\mathbf{B}, \boldsymbol{\lambda})$ so that, for every i, $\mathbf{b}^i$ solves (B.6), $\mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma) = 0$, and (B.8) and (B.9) hold. By the arguments in the proof of Theorem 4.1 (see Step 5 below), for $\boldsymbol{\lambda} \ge \mathbf{0}$, a solution $\mathbf{b}^i$ to (B.6) always exists and is unique on $(\underline{v}^i, \overline{v}^i)$ and is pointwise continuous in $\boldsymbol{\lambda}$. Moreover, if $\lambda^{ji} \to +\infty$ for some $j < i$, then $\mathbf{b}^i \to b(a^{\text{nf}})$ on $(\underline{v}^j, \overline{v}^i)$, and $\widetilde{R}^j(\mathbf{b}^i) \to 0$. And since $\mu^i(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma) = 0$, $\lambda^{ij'} \to +\infty$ for some $j' > i$, so that $\widetilde{R}^i(\mathbf{b}^{j'}) \to 0$ and

$h^i \to 0$ (using the binding $\text{IC}_1^{ij'}$). So there is λ^{ji} large enough to make (B.9) hold. Finally, (B.8) always holds with $h^N = 0$.

Step 5: Fix $i > 1$. Using (B.6), the expression of $\widetilde{R}^n(\mathbf{b}^i)$, and $\xi(\cdot) = -b^{-1}(\cdot) - c(b^{-1}(\cdot))$, $\mathbf{b}^i$ must maximize within $\mathcal{B}$

$$VS^i(\mathbf{b}^i; \boldsymbol{\lambda}^i) = \sum_{n=1}^{i-1} \lambda^{ni} \int_{\overline{v}^i}^{\overline{v}^n} \mathbf{b}^i(v) g^n(v)\, dv + \int_{\underline{v}^i}^{\overline{v}^i} \left[\mathbf{b}^i(v) w^i(v, \boldsymbol{\lambda}^i) + \xi(\mathbf{b}^i(v))\right] dF^i,$$

where $\boldsymbol{\lambda}^i \in \mathbb{R}_+^{i-1}$ and

$$w^i(v; \boldsymbol{\lambda}^i) = \frac{v}{t^i} + \sum_{n=1}^{i-1} \lambda^{ni} q^i(v) - \sum_{n=1}^{i-1} \lambda^{ni} \frac{f^n(v)}{f^i(v)} q^n(v).$$

We can apply to $VS^i(\mathbf{b}^i; \boldsymbol{\lambda}^i)$ the method used in the two-type case to characterize $\mathbf{b}^I$ (Theorem 4.1). If $\boldsymbol{\lambda}^i = \mathbf{0}$, $VS^i(\mathbf{b}^i; \mathbf{0}) = \widetilde{W}^i(\mathbf{b}^i)$ and $\mathbf{b}^i = \mathbf{b}^i_{fb} = b(\mathbf{a}^i_{fb})$ on $(\underline{v}^i, \overline{v}^i)$. For $v > \overline{v}^i$, let $\mathbf{b}^i(v) = \mathbf{b}^i(\overline{v}^i)$. For $v < \underline{v}^i$, $\mathbf{b}^i(v)$ may be strictly smaller than $\mathbf{b}^i(\underline{v}^i)$ to satisfy IC_1^{ji} for $j > i$.

Suppose $\lambda^{ni} > 0$ for some $n < i$. Apply the Myerson–Toikka ironing method on $(\underline{v}^i, \overline{v}^i)$, by letting $z^i(x; \boldsymbol{\lambda}^i) = w^i((F^i)^{-1}(x); \boldsymbol{\lambda}^i)$ and $Z^i(x; \boldsymbol{\lambda}^i) = \int_0^x z^i(y; \boldsymbol{\lambda}^i)\, dy$. Let $\Omega^i(x; \boldsymbol{\lambda}^i) = \text{conv}(Z^i(x; \boldsymbol{\lambda}^i))$, and $\omega^i(x; \boldsymbol{\lambda}^i) = \Omega^i_x(x; \boldsymbol{\lambda}^i)$ wherever defined. Extend ω^i by right-continuity, and at 1 by left-continuity. For ω^i to be continuous, it is enough to show that, if z^i is discontinuous at x, then z^i jumps down at x. To see this, note that w^i can be discontinuous only at points like $\underline{v}^j$ for $j < i$ and such that $\underline{v}^j \in (\underline{v}^i, \overline{v}^i)$. At such a point, let $w^i(\underline{v}^j+; \boldsymbol{\lambda}^i) = \lim_{v \downarrow \underline{v}^j} w^i(v; \boldsymbol{\lambda}^i)$ and $w^i(\underline{v}^j-; \boldsymbol{\lambda}^i) = \lim_{v \uparrow \underline{v}^j} w^i(v; \boldsymbol{\lambda}^i)$. For $n < j$, $\underline{v}^n > \underline{v}^j$ and hence $f^n(\underline{v}^j) = 0$. So

$$w^i(\underline{v}^j+; \boldsymbol{\lambda}^i) = \frac{\underline{v}^j}{t^i} + \sum_{n=1}^{i-1} \lambda^{ni} q^i(\underline{v}^j) - \sum_{n=j}^{i-1} \lambda^{ni} \frac{f^n(\underline{v}^j)}{f^i(\underline{v}^j)} q^n(\underline{v}^j),$$

$$w^i(\underline{v}^j-; \boldsymbol{\lambda}^i) = \frac{\underline{v}^j}{t^i} + \sum_{n=1}^{i-1} \lambda^{ni} q^i(\underline{v}^j) - \sum_{n=j+1}^{i-1} \lambda^{ni} \frac{f^n(\underline{v}^j)}{f^i(\underline{v}^j)} q^n(\underline{v}^j).$$

Then,

$$w^i(\underline{v}^j-; \boldsymbol{\lambda}^i) - w^i(\underline{v}^j+; \boldsymbol{\lambda}^i) = \lambda^{ji} \frac{f^j(\underline{v}^j)}{f^i(\underline{v}^j)} q^j(\underline{v}^j) \geq 0,$$

since $q^j(\underline{v}^j) = (1-t^j)(\underline{v}^j/t^j) \geq 0$. Letting $\overline{w}^i(v;\boldsymbol{\lambda}^i) = \omega^i(F^i(v);\boldsymbol{\lambda}^i)$ for $v \in (\underline{v}^i, \overline{v}^i)$, construct $\overline{VS}^i$ as in the proof of Theorem 4.1.

Note that $g^n(v) < 0$ for $v \in (\overline{v}^i, \overline{v}^n)$. So, since $\lambda^{ni} > 0$ for some $n < i$, the first term in VS^i is strictly negative. Let $\underline{n} = \min\{n : \lambda^{ni} > 0\}$. Then, on $(\underline{v}^i, \overline{v}^{\underline{n}})$, the characterization of Lemma A.3 extends to $\overline{VS}^i$. So $\mathbf{b}^i$ must be constant at y^{ib} on $(v^{ib}, \overline{v}^{\underline{n}})$, where $v^{ib} \leq \overline{v}^i$ and $y^{ib} \leq \overline{\mathbf{b}}^i(\overline{v}^i)$. Moreover, $y^{ib} = \overline{\mathbf{b}}^i(v^{ib})$, if $v^{ib} > \underline{v}^i$; and $\mathbf{b}^i(v) = \overline{\mathbf{b}}^i(v)$ for $v \in [\underline{v}^i, v^{ib}]$. The argument in Lemma A.4 yields that there is a (unique) maximizer of $\overline{VS}^i$. The argument in Lemma A.5 implies that the (unique) maximizer of $\overline{VS}^i$ is also the (unique) maximizer of VS^i.

Step 6: Properties of the solutions to (B.6). Suppose $\lambda^{ni} > 0$ for some $n < i$ and define $\underline{n}$ as before. The analog of the ironing condition for v^{b} applies to v^{ib}:

$$\int_{v^{ib}}^{\overline{v}^i} \left[w^i(y;\boldsymbol{\lambda}^i) - w^i(v^{ib};\boldsymbol{\lambda}^i)\right] dF^i = -\sum_{n=\underline{n}}^{i-1} \lambda^{ni} \int_{\overline{v}^i}^{\overline{v}^n} g^n(v)\, dv.$$

Since the sum is negative, $v^{ib} < \overline{v}^i$. This condition can be written as

$$\begin{aligned}\int_{v^{ib}}^{\overline{v}^i} &\left[w^i(v^{ib};\boldsymbol{\lambda}^i) - (v/t^i)\right] dF^i \\ &= \sum_{n=\underline{n}}^{i-1} \lambda^{ni}\left[\int_{v^{ib}}^{\overline{v}^i} G^{in}(v)\, dF^i + \int_{\overline{v}^i}^{\overline{v}^n} g^n(v)\, dv\right].\end{aligned}$$

To prove that $w^i(v^{ib};\boldsymbol{\lambda}^i) < \overline{v}^i/t^i$, it is enough to observe that the right-hand side is negative by (A.14). So, $\mathbf{b}^i$ exhibits bunching on $[v^{ib}, \overline{v}^{\underline{n}}]$ at value $y^{ib} < \mathbf{b}^i_{fb}(\overline{v}^i)$.

Now consider the bottom of $[\underline{v}^i, \overline{v}^i]$. By the logic in Lemma A.6, $\overline{w}^i(\underline{v}^i;\boldsymbol{\lambda}^i) \leq w^i(\underline{v}^i;\boldsymbol{\lambda}^i)$, with strict inequality if $v^i_{b} > \underline{v}^i$. Moreover, for $v < \underline{v}^{i-1}$, $w^i(v,\boldsymbol{\lambda}^i) = v/t^i + \sum_{n=1}^{i-1} \lambda^{ni} q^i(v)$ and $w^i(\underline{v}^i;\boldsymbol{\lambda}^i) = (\underline{v}^i/t^i)[1 + (1-t^i)\sum_{n=1}^{i-1} \lambda^{ni}] > \underline{v}^i/t^i$. So, if $\overline{w}^i(\underline{v}^i;\boldsymbol{\lambda}^i) = w^i(\underline{v}^i;\boldsymbol{\lambda}^i)$, then $\mathbf{b}^i(\underline{v}^i;\boldsymbol{\lambda}^i) > \mathbf{b}^i_{fb}(\underline{v}^i)$. Otherwise, ironing occurs on $[\underline{v}^i, v^i_{b}] \neq \emptyset$ and

$$\int_{\underline{v}^i}^{v^i_{b}} \left[w^i(y;\boldsymbol{\lambda}^i) - \overline{w}^i(v^i_{b};\boldsymbol{\lambda}^i)\right] dF^i = 0,$$

which corresponds to

$$\int_{\underline{v}^i}^{v^i_{b}} \left[y/t^i - \overline{w}^i(v^i_{b};\boldsymbol{\lambda}^i)\right] dF^i = -\sum_{n=1}^{i-1} \lambda^{ni} \int_{\underline{v}^i}^{v^i_{b}} G^{in}(y)\, dF^i.$$

Now, for $n < i$,

$$\int_{\underline{v}^i}^{v_b^i} G^{in}(y)\,dF^i = \int_{\underline{v}^i}^{v_b^i} q^i(y)\,dF^i - \int_{\underline{v}^i}^{v_b^i} q^n(y)\,dF^n$$
$$= \int_{v_b^i/t^n}^{v_b^i/t^i} \left(s - v_b^i\right) dF > 0.$$

So $\overline{w}^i(v_b^i; \boldsymbol{\lambda}^i) > \underline{v}^i/t^i$, and $\mathbf{b}^i(\underline{v}^i; \boldsymbol{\lambda}^i) > \mathbf{b}^i_{fb}(\underline{v}^i)$.

Finally, note that for $v < v' < \underline{v}^{i-1}$,

$$w^i\left(v'; \boldsymbol{\lambda}^i\right) - w^i\left(v; \boldsymbol{\lambda}^i\right) = \frac{v' - v}{t^i}\left[1 + \sum_{n=1}^{i-1} \lambda^{ni}\left(1 - t^i\right)\right]$$
$$+ \sum_{n=1}^{i-1} \lambda^{ni}\left[\frac{F^i\left(v'\right)}{f^i\left(v'\right)} - \frac{F^i(v)}{f^i(v)}\right].$$

So, $w^i(\cdot; \boldsymbol{\lambda}^i)$ will be decreasing in a neighborhood of $\underline{v}^i$ if, for $s' > s$ in $[\underline{s}, s^\dagger]$,

$$\frac{F\left(s'\right)/f\left(s'\right) - F(s)/f(s)}{s' - s} \geq \frac{1}{t^i}\left[\left(1 - t^i\right) + \left(\sum_{n=1}^{i-1} \lambda^{ni}\right)^{-1}\right].$$

Hence, bunching at the bottom is more likely if t^i is closer to 1 and $\sum_{n=1}^{i-1} \lambda^{ni}$ is large, that is, if the provider assigns large shadow value to *not* increasing the rents of types below i.

APPENDIX C: ILLUSTRATIVE EXAMPLE'S CALCULATIONS

Let $\underline{s} = 10$, $\overline{s} = 15$, and $t = 0.9$. We first characterize the first-best C- and I-device. By Corollary 3.1, $p^C_{\mathbf{e}}$ must be constant; by Proposition 3.1, it must extract the entire surplus that C derives from the C-device, thereby leaving C with expected utility m. With regard to the I-device, again by Corollary 3.1, for $a \in [100, 225]$ we have $p^I_{\mathbf{e}}(a) = p^C_{\mathbf{e}} + q^I(a)$ such that $q^I(\mathbf{e}(s)) = \mathbf{q}^{0.9}(s)$ for every $s \in [\underline{s}, \overline{s}]$. Therefore, using the formula in Corollary 3.1,

$$\frac{dq^I(a)}{da} = \frac{d\mathbf{q}^{0.9}(s)/ds}{d\mathbf{e}(s)/ds} = -0.1.$$

So $q^I(a) = k - 0.1a$, where k is set so that I expects to pay $p^C_{\mathbf{e}}$ (Proposition 3.1).

Consider now the difference between C's and I's expected utility from the efficient I-device (i.e., $R^C(a^I_{fb})$). Recall that $p^I_{\mathbf{e}}(a) = +\infty$ for $a \notin [100, 225]$.

Under this I-device, at time 2 type C chooses $\boldsymbol{\alpha}^C(s) = \frac{s^2}{t^2}$ for $s < \frac{\overline{s}}{t}$ and $\boldsymbol{\alpha}^C(s) = \overline{s}$ otherwise. Thus

$$R^C\big(a^I_{fb}\big) = m - p^C_{\mathbf{e}} - k + \int_{\underline{s}}^{\overline{s}} \big[2s\sqrt{\boldsymbol{\alpha}^C(s)} - t\boldsymbol{\alpha}^C(s)\big]\frac{ds}{\overline{s}-\underline{s}}$$

$$- \left\{m - p^C_{\mathbf{e}} - k + \int_{\underline{s}}^{\overline{s}} \big[2s\sqrt{\mathbf{e}(s)} - t\mathbf{e}(s)\big]\frac{ds}{\overline{s}-\underline{s}}\right\}$$

$$= \frac{1-t}{3t(\overline{s}-\underline{s})}\big[\overline{s}^3(3-t)t - (1+t)\underline{s}^3\big].$$

Substituting the values of $\underline{s}$, $\overline{s}$, and t, we get $R^C(a^I_{fb}) \approx 33.18$.

To compute the difference between I's and C's expected utilities from the efficient C-device (i.e., $R^I(a^C_{fb})$), recall that $p^C_{\mathbf{e}}(a) = +\infty$ for $a \notin [100, 225]$. Given this, at time 2 type I chooses $\boldsymbol{\alpha}^I(s) = t^2 s^2$ for $s > \frac{\underline{s}}{t}$ and $\boldsymbol{\alpha}^I(s) = \underline{s}$ otherwise. Thus

$$R^I\big(a^C_{fb}\big) = m - p^C_{\mathbf{e}} + \int_{\underline{s}}^{\overline{s}} \big[2s\sqrt{\boldsymbol{\alpha}^I(s)} - \boldsymbol{\alpha}^I(s)\big]\frac{ds}{\overline{s}-\underline{s}}$$

$$- \left\{m - p^C_{\mathbf{e}} + \int_{\underline{s}}^{\overline{s}} \big[2s\sqrt{\mathbf{e}(s)} - \mathbf{e}(s)\big]\frac{ds}{\overline{s}-\underline{s}}\right\}$$

$$= \frac{(1-t)^2}{3(\overline{s}-\underline{s})}\big[\underline{s}^3 t^{-2} - \overline{s}^3\big].$$

Substituting $\underline{s}$, $\overline{s}$, and t, we get $R^I(a^C_{fb}) \approx -1.43$.

The properties of the screening I-device follow from the argument in the proof of Corollary 4.3 above. The thresholds s_{b} and s^{b} can be computed using formulas (B.2) and (B.4) for v_{b} and v^{b}. Regarding the range $[a_{\mathrm{b}}, a^{\mathrm{b}}]$, we have that $a_{\mathrm{b}} = [w^I(v_{\mathrm{b}}; r^C)]^2$ and $a^{\mathrm{b}} = [w^I(v^{\mathrm{b}}; r^C)]^2$, where $w^I(v; r^C)$ is given in (B.1). These formulas depend on $r^C = \frac{\gamma}{1-\gamma} + \frac{\mu}{1-\gamma}$, but in this example $\mu = 0$ because unused options are always enough to deter I from taking the C-device (see below). Varying $\gamma \in (0, 1)$ delivers the values in Figure 1 of the main text. By Proposition 4.2, when the provider completely removes flexibility from the I-device, she induces I to choose the ex ante efficient action $a^{\mathrm{nf}} = (\frac{\overline{s}+\underline{s}}{2})^2 = 156.25$.

The most deterring unused option for the C-device depends on v_{u} in Proposition 4.3. As shown in its proof, $v_{\mathrm{u}} = \sup\{v \in [\underline{v}^I, \underline{v}^C] \mid g^I(v) < 0\}$ where

$$g^I(v) = \frac{t-1}{t} v f^I(v) + F^I(v) = \frac{1}{t(\overline{s}-\underline{s})}\left[(2t-1)s - \frac{\underline{s}}{t}\right],$$

which is strictly increasing since $t > 1/2$. Since $\underline{v}^C = \underline{s}$ and $g^I(\underline{s}) = \frac{2(t-1)\underline{s}}{t^2(\overline{s}-\underline{s})} < 0$, we have $v_{\mathrm{u}} = \underline{s}$. That is, the most deterring C-device induces I to choose the unused option with $\underline{a} = 0$ whenever $s < \frac{\underline{s}}{t}$. The associated payment must render I indifferent at time 2 between saving $\boldsymbol{\alpha}^I(\underline{s}/t) = \underline{s}^2$ and zero in state $\frac{\underline{s}}{t}$:

$$m - p^C(0) = m - p^C(\underline{s}^2) - \underline{s}^2 + 2t\left(\frac{\underline{s}}{t}\right)\sqrt{\underline{s}^2}.$$

Substituting and rearranging, we get $p^C(0) = p^C(100) - 100$.

We can now compute the difference in I's expected utility between the C-device with and without the unused option. This depends only on I's different choices for states in $[\underline{s}, \underline{s}/t)$, and hence it equals

$$\int_{\underline{s}}^{\underline{s}/t} [-p^C(0)]\frac{ds}{\overline{s}-\underline{s}} - \int_{\underline{s}}^{\underline{s}/t} [-p^C(\underline{s}^2) - \underline{s}^2 + 2s\sqrt{\underline{s}^2}]\frac{ds}{\overline{s}-\underline{s}} = \frac{\underline{s}^3(1-t^2)}{t^2(\overline{s}-\underline{s})}.$$

Using the parameters' values, this difference is -46.91. Since it exceeds $R^C(a^I_{fb}) \approx 33.18$, I would never choose the C-device that contains unused option $(0, p^C(0))$.

APPENDIX D: OUTSIDE OPTION WITH TYPE-DEPENDENT VALUES

After rejecting all the provider's devices at time 1, the agent will make certain state-contingent choices at time 2, which can be described with $(\mathbf{a}_0, \mathbf{p}_0)$ using the formalism of Section 4.1. For simplicity, consider the two-type model. By Proposition 4.1, $U^C(\mathbf{a}_0, \mathbf{p}_0) \geq U^I(\mathbf{a}_0, \mathbf{p}_0)$ with equality if and only if $\mathbf{a}_0$ is constant over $(\underline{v}, \overline{v})$. So C and I value the outside option differently, unless they always end up making the same choice.

When $U^C(\mathbf{a}_0, \mathbf{p}_0) > U^I(\mathbf{a}_0, \mathbf{p}_0)$, the analysis in Section 4 can be adjusted without changing its thrust. The constraints (IR^C) and (IC_1^C) set two lower bounds on C's payoff from the C-device: one endogenous (i.e., $U^C(\mathbf{a}^I, \mathbf{p}^I) = U^I(\mathbf{a}^I, \mathbf{p}^I) + R^C(\mathbf{a}^I)$) and one exogenous (i.e., $U^C(\mathbf{a}_0, \mathbf{p}_0) = U^I(\mathbf{a}_0, \mathbf{p}_0) + R^C(\mathbf{a}_0)$). The question is which binds first. In Section 4, (IC_1^C) always binds first, for (IR^I) and (IC_1^C) imply (IR^C). Now this is no longer true. Intuitively, if (IC_1^C) binds first, then we are in a situation similar to Section 4; so the provider will distort the I-device as shown in Section 4.2.[1] If (IR^C) binds first, then obviously the provider has no reason to distort the I-device. For example, she will never distort the I-device, if the outside option sustains the efficient outcome with I—that is, $\mathbf{a}_0 = \mathbf{a}^I_{fb}$ over $[\underline{v}^I, \overline{v}^I]$. In this case, she must grant C at least the rent $R^C(\mathbf{a}_0)$, which already exceeds $R^C(\mathbf{a}^I_{fb})$. Finally, if (IC_1^I) binds, then the provider will design the C-device as shown in Section 4.3.[2]

[1]This case is more likely when the outside option involves little flexibility, so that $R^C(\mathbf{a}_0)$ is small.

[2]We can extend this argument to settings in which, at time 1, the agent has access to other devices if he rejects the provider's ones. In these settings, $(\mathbf{a}_0, \mathbf{p}_0)$ can be type-dependent.

APPENDIX E: FINITELY MANY STATES AND IRRELEVANCE OF ASYMMETRIC INFORMATION

This section shows that if the set of states S is finite, then the provider may be able to always sustain the efficient outcome $\mathbf{e}$, even if she cannot observe the agent's degree of inconsistency. To see the intuition, consider a two-state case with $s_2 > s_1$. If the provider can observe t, she sustains $\alpha_2^* = \mathbf{e}(s_2) > \mathbf{e}(s_1) = \alpha_1^*$, with payments $\pi_1 = \boldsymbol{\pi}^t(s_1)$ and $\pi_2 = \boldsymbol{\pi}^t(s_2)$ that satisfy

$$u_2(\alpha_2^*; s_2, t) - u_2(\alpha_1^*; s_2, t) \geq \pi_2 - \pi_1 \geq u_2(\alpha_2^*; s_1, t) - u_2(\alpha_1^*; s_1, t), \tag{E.1}$$

which follows from (IC). Since $u_2(a; s, t)$ has strictly increasing differences in (a, s), having a discrete S creates some slack in (IC) at $\mathbf{e}$: For any t, (E.1) does not pin down π_1 and π_2 uniquely. Suppose t^I is close to t^C. Intuitively, to sustain $\mathbf{e}$ with each type, the provider should be able to use incentive schemes that are sufficiently alike; also, since discrete states leave some leeway in the payments, she may be able to find *one* scheme that works for both types. If instead t^I is far from t^C, the provider must use different schemes to sustain $\mathbf{e}$ with each type. Since $t^I < t^C$, I is tempted to pick α_1^* also in s_2, and the more so, the lower is t^I. So, for I not to choose α_1^* in s_2, α_1^* must be sufficiently more expensive than α_2^*, and this gap must rise as t^I falls. At some point, this gap must exceed C's willingness to pay for switching from α_2^* to α_1^* in s_1.

Proposition E.1 formalizes this intuition. Consider a finite set T of types, which may include both $t > 1$ and $t < 1$; let $\bar{t} = \max T$ and $\underline{t} = \min T$.

PROPOSITION E.1: *Suppose S is finite and $s_N > s_{N-1} > \cdots > s_1$. There is a single commitment device that sustains* $\mathbf{e}$ *with each $t \in T$ if and only if $\bar{t}/\underline{t} \leq \min_i s_{i+1}/s_i$.*

PROOF: With N states, (IC) becomes

$$u_2(\alpha_i; s_i, t) - \pi_i \geq u_2(\alpha_j; s_i, t) - \pi_j$$

for all i, j, where $\alpha_i = \boldsymbol{\alpha}(s_i)$ and $\pi_i = \boldsymbol{\pi}(s_i)$. By standard arguments, it is enough to focus on adjacent constraints. For $i = 2, \ldots, N$, let $\Delta_i = \pi_i - \pi_{i-1}$. If $\boldsymbol{\alpha}^* = \mathbf{e}$ for all i, then $\alpha_N^* > \alpha_{N-1}^* > \cdots > \alpha_1^*$ (Assumption 2.1). To sustain $\mathbf{e}$ with t, Δ_i must satisfy

$$u_2(\alpha_i^*; s_i, t) - u_2(\alpha_{i-1}^*; s_i, t) \geq \Delta_i \geq u_2(\alpha_i^*; s_{i-1}, t) - u_2(\alpha_{i-1}^*; s_{i-1}, t), \tag{CIC$_{i,i-1}$}$$

for $i = 2, \ldots, N$. For any s and t, $u_2(a'; s, t) - u_2(a; s, t) = ts(b(a') - b(a)) - a' + a$. Let $s_k/s_{k-1} = \min_i s_i/s_{i-1}$, and suppose $\bar{t}s_{k-1} > s_k\underline{t}$. Then,

$$u_2(\alpha_k^*; s_{k-1}, \bar{t}) - u_2(\alpha_{k-1}^*; s_{k-1}, \bar{t}) > u_2(\alpha_k^*; s_k, \underline{t}) - u_2(\alpha_{k-1}^*; s_k, \underline{t}),$$

and no Δ_k satisfies ($\text{CIC}_{k,k-1}$) for both $\underline{t}$ and $\bar{t}$. If instead $\underline{t}s_i \geq \bar{t}s_{i-1}$ for $i = 2, \dots, N$, then for every t and i,

$$u_2(\alpha_i^*; s_i, t) - u_2(\alpha_{i-1}^*; s_i, t) \geq u_2(\alpha_i^*; s_{i-1}, \bar{t}) - u_2(\alpha_{i-1}^*; s_{i-1}, \bar{t})$$
$$\geq u_2(\alpha_i^*; s_{i-1}, t) - u_2(\alpha_{i-1}^*; s_{i-1}, t).$$

Set $\Delta_i^* = u_2(a_i^*; s_{i-1}, \bar{t}) - u_2(a_{i-1}^*; s_{i-1}, \bar{t})$. Then $\{\Delta_i^*\}_{i=2}^N$ satisfies all ($\text{CIC}_{i,i-1}$) for every t. The payment rule $\pi_i^* = \pi_1^* + \sum_{j=2}^i \Delta_j^*$—with π_1^* small to satisfy (IR)—sustains **e** with each t. *Q.E.D.*

So, if the heterogeneity across types (measured by $\bar{t}/\underline{t}$) is small, the provider can sustain **e** without worrying about time-1 incentive constraints.

The condition in Proposition E.1, however, is not necessary for the unobservability of t to be irrelevant when sustaining **e**. Even if $\bar{t}/\underline{t}$ is large, the provider may be able to design different devices such that each sustains **e** with one t, and each t chooses the device for himself ('t-device'). To see why, consider an example with two types, $t^h > t^l$, and two states, $s_2 > s_1$. Suppose $t^h > 1 > t^l$, $t^h s_1 > t^l s_2$, but $s_2 > s_1 t^h$ and $s_2 t^l > s_1$. Consider all (π_1, π_2) that satisfy (E.1) and (IR) with equality:

$$(1-f)\pi_2 + f\pi_1 = (1-f)u_1(\alpha_2^*; s_2) + f u_1(\alpha_1^*; s_1),$$

where $f = F(s_1)$. Finally, choose (π_1^h, π_2^h) so that h's self-1 strictly prefers α_2^* in s_2—i.e., $u_1(\alpha_2^*; s_2) - \pi_2^h > u_1(\alpha_1^*; s_2) - \pi_1^h$—and (π_1^l, π_2^l) so that l's self-1 strictly prefers α_1^* in s_1—that is, $u_1(\alpha_1^*; s_1) - \pi_1^l > u_1(\alpha_2^*; s_1) - \pi_2^l$. Then, the l-device (respectively, h-device) sustains **e** and gives zero expected payoffs to the agent if and only if l (h) chooses it. Moreover, l strictly prefers the l-device and h the h-device. To see this, note that if *self-1* of either type had to choose at time 2, under either device he would strictly prefer to implement **e**. So, by choosing the 'wrong' device, either type can only lower his payoff below zero.

Proposition E.2 gives a necessary condition for the unobservability of t to be irrelevant when sustaining **e**. Let $T^1 = T \cap [0, 1]$ and $T^2 = T \cap [1, +\infty)$. For $k = 1, 2$, let $\bar{t}^k = \max T^k$ and $\underline{t}^k = \min T^k$.

PROPOSITION E.2: *Suppose S is finite and $s_N > s_{N-1} > \cdots > s_1$. If* $\max\{\bar{t}^1/\underline{t}^1, \bar{t}^2/\underline{t}^2\} > \min_i s_{i+1}/s_i$, *then there is no set of devices, each designed for a $t \in T$, such that* (i) *t chooses the t-device*, (ii) *the t-device sustains* **e** *with t, and* (iii) *all t get the same expected payoff.*

PROOF: Suppose $\max\{\bar{t}^1/\underline{t}^1, \bar{t}^2/\underline{t}^2\} = \bar{t}^1/\underline{t}^1$—the other case is similar—and that there exist devices that satisfy (i)–(iii). Let U be each t's expected payoff and $\underline{\mathbf{p}}$ be the payment rule in the $\underline{t}^1$-device. Given $\underline{\mathbf{p}}$, let $\underline{a}_i(t)$ be an optimal choice of $t \in T^1$ in s_i. For $\underline{t}^1$, $\underline{a}_i(\underline{t}^1) = \alpha_i^*$ for every i. Let $\overline{S} = \{i : s_{i+1}/s_i <$

$\bar{t}^1/\underline{t}^1\} \neq \emptyset$. Then, (a) for every i, $\bar{t}^1 s_i > \underline{t}^1 s_i$ and hence $\underline{a}_i(\bar{t}^1) \geq \alpha_i^*$; (b) for $i \in \overline{S}$, $\bar{t}^1 s_i > \underline{t}^1 s_{i+1}$, and so $\underline{a}_i(\bar{t}^1) \geq \alpha_{i+1}^* > \alpha_i^*$. Since $t \leq 1$, (a) and (b) imply

$$\underline{\mathbf{p}}\big(\underline{a}_i(\bar{t}^1)\big) - \underline{\mathbf{p}}\big(\alpha_i^*\big) \leq u_2\big(\underline{a}_i(\bar{t}^1); s_i, \bar{t}^1\big) - u_2\big(\alpha_i^*; s_i, \bar{t}^1\big)$$
$$\leq u_1\big(\underline{a}_i(\bar{t}^1); s_i\big) - u_1\big(\alpha_i^*; s_i\big),$$

where the first inequality is strict for $i \in \overline{S}$. The expected payoff of $\bar{t}^1$ from $\underline{\mathbf{p}}$ is then

$$\sum_{i=1}^{N} \big[u_1\big(\underline{a}_i(\bar{t}^1); s_i\big) - \underline{\mathbf{p}}\big(\underline{a}_i(\bar{t}^1)\big)\big] f_i > \sum_{i=1}^{N} \big[u_1\big(\alpha_i^*; s_i\big) - \underline{\mathbf{p}}\big(\alpha_i^*\big)\big] f_i = U,$$

where $f_i = F(s_i) - F(s_{i-1})$ for $i = 2, \ldots, N$ and $f_1 = F(s_1)$. *Q.E.D.*

So, if $T^1 \setminus \{1\} = \emptyset$ or $T^2 \setminus \{1\} = \emptyset$, the condition in Proposition E.1 is also necessary for the provider to be able to sustain $\mathbf{e}$, even if she cannot observe t.

REFERENCES

LUENBERGER, D. G. (1969): *Optimization by Vector Space Methods*. New York: Wiley. [9]
ROCKAFELLAR, R. T. (1970): *Convex Analysis*. Princeton: Princeton University Press. [2]

Dept. of Economics, UCSD, La Jolla, CA 92093, U.S.A.; sgalperti@ucsd.edu.

Manuscript received August, 2013; final revision received March, 2015.