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

\[
\max_{\alpha} \left\{ \int_{\tilde{s}}^{s} \left[ u_1(\alpha'(s); s) - c(\alpha'(s)) \right] dF \right\} \quad \text{s.t. (IC)}.
\]

Ignoring (IC), this problem has a unique solution (up to \( \{\tilde{s}, \bar{s}\} \)): \( \alpha' \equiv e \). Since \( e \) is increasing and \( t > 0 \), by standard arguments, there is \( \pi'_e \) such that \( (e, \pi'_e) \) satisfies (IC). Specifically, for every \( s \),

\[
\pi'_e(s) = u_2(e(s); s, t) - \int_{\tilde{s}}^{s} t b(e(y)) dy - k,
\]

where \( k \in \mathbb{R} \). Since \( e \) is differentiable,

\[
\frac{d\pi'_e(s)}{ds} = \frac{\partial u_2(e(s); s, t)}{\partial a} \frac{de(s)}{ds},
\]

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

B.2. Proof of Corollary 4.2

Being increasing, \( a'_{sh} \) is differentiable a.e. on \( [v, \bar{v}] \). If \( \frac{da'_{sh}}{dv} > 0 \) at \( v \), then using condition (E),

\[
\frac{dp'_{sh}/dv}{da'_{sh}/dv} = vb'(a'_{sh}(v)) - 1 \quad \text{and} \quad \frac{dp'_{fb}/dv}{da'_{fb}/dv} = vb'(a'_{fb}(v)) - 1.
\]

The result follows from \( b'' < 0 \) and Theorem 4.1(a).
(Continuity in $x$). Suppose $r^C$. For $x \in (0, 1) \setminus \{x^n\}$, $z$ 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 $\Omega$ is convex and $\Omega \leq Z$, their right and left derivatives satisfy $\Omega^+(x) \leq Z^+(x)$ and $\Omega^-(x) \geq Z^-(x)$. Since $\Omega^-(x) \leq \Omega^+(x)$ and $Z$ is differentiable at $x$, $\Omega^-(x) = \Omega^+(x)$; so $\omega$ is continuous at $x$. Finally, consider $x_m^n$. If $v^n = \overline{w}^\dagger$, then $x_m^n = 1$ and we are done. For $x_m^n \in (0, 1)$, $\omega$ is continuous if $\Omega(x_m^n) < Z(x_m^n)$ when $z$ jumps at $x_m^n$. Recall that $z(x_m^n) - \lim_{r \downarrow x_m^n} w^d(v; r^C)$ and $z(x_m^n+) = z(x_m^n) - \lim_{r \uparrow x_m^n} w^d(v; r^C)$. By expression (A.8), $z$ can only jump down at $x_m^n$, so $z(x_m^n -) < Z(x_m^n)$. Suppose $\Omega(x_m^n) = Z(x_m^n)$. By the previous argument, $\Omega^+(x_m^n) \leq Z^+(x_m^n) = z(x_m^n)$. By convexity, $\omega(x) \leq \Omega^-(x_m^n)$ for $x \leq x_m^n$. So, for $x$ close to $x_m^n$ from the left, we get the following contradiction:

$$\Omega(x) = \Omega(x_m^n) - \int_x^{x_m^n} \omega(y) dy > Z(x_m^n) - \int_x^{x_m^n} z(y) dy = Z(x).$$

(Continuity in $r^C$). Given $x$, $Z(x; r^C)$ is continuous in $r^C$. So $\Omega$ 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 $r^C \geq 0$, 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 $\{r^n\}$ with $r^n \to r^C$. Since $\Omega(x; r^n) \to \Omega(x; r^C)$, Theorem 25.7, p. 248, of Rockafellar (1970) implies $\omega(x; r^n) \to \omega(x; r^C)$.

B.4. Proof of Lemma A.6

Recall that $\overline{w}^d(\overline{w}^\dagger) = \omega(0)$ and $w^d(\overline{w}^\dagger) = z(0)$. If $\omega(0) > z(0)$, since $z$ is continuous on $[0, x_m^n]$ and $\omega$ is increasing, there is $x > 0$ such that $\omega(y) > z(y)$ for $y \leq x$. Since $Z(0) = 0$, 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 $v_0 < (F^\dagger)^{-1}(\hat{x}) > v'$. 

2 SIMONE GALPERTI
Let $t^C = 1$. Since $F$ is uniform, $F'(v) = v - v'$. Using (A.7),

$$w^d(v; r^C) = \begin{cases} 
\frac{(v/t')(1 + r^C(1 - 2t')) + r^C v'}{1 + r^C(t' - 1)^2}, & \text{if } v \in [v', v^C], \\
\frac{(v/t')(1 + r^C(t' - 1)^2)}{1 + r^C(t' - 1)^2}, & \text{if } v \in [v^C, v'].
\end{cases}$$

The function $w^d$ is continuous at $v^C$. It is strictly increasing and greater than $v/t'$ on $[v^C, v']$, as $r^C > 0$ and $t' < 1$; $w^d$ is strictly increasing on $[v', v^C]$ if and only if $t' \leq 1/2$ or $r^C < (2t' - 1)^{-1} = \bar{r}^C$.

Consider first $v^b$ and $v_b$, when $v^b > v_b$. If $t' \leq 1/2$ or $r^C < \bar{r}^C$, then $w^d$ is strictly increasing and equals $\bar{w}$ (see the proof of Theorem 4.1); so $\bar{v}'$ (see (A.9)) is strictly increasing on $[v^b, v']$, and $v_b = v'$. Otherwise, $v_b \geq v^C > v^b$ and $v_b$ is characterized by (A.16):

$$\int_{v^b}^{v^C} [w^d(y; r^C) - w^d(v^b; r^C)] dy = -\bar{w}'(v^b) r^C \int_{v^b}^{v^C} g^C(y) dy.$$

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

$$r^C \left[2t' (\bar{v}' - v^b) K \right] = (1 + r^C(t' - 1)^2)(\bar{v}' - v^b)^2$$

if $v^b \geq v^C$, and

$$r^C \left[2t' (\bar{v}' - v^b) K \right] = r^C (t')^2 (\bar{v}' - v^C)^2 + (1 + r^C(1 - 2t'))(\bar{v}' - v^b)^2$$

if $v^b < v^C$. So, if $t' > 1/2$, the function $v_b(r^C)$ is constant at $v^b$ for $r^C < \bar{r}^C$, and at $\bar{r}^C$, it jumps from $v^b$ to $v^C$. Monotonicity for $r^C > \bar{r}^C$ follows by applying the Implicit Function Theorem to (B.2):

$$\frac{dv_b}{dr^C} = \frac{1}{2} \left[ \frac{t'}{1 + r^C(t' - 1)^2} \right]^2 \frac{(v^C - v^b)^2}{(v_b - v^b)} > 0.$$
Similarly,

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

for the second inequality, recall that $v_b < v^b < v^b$ if and only if $t^l \leq 1/2$ or $r^C < \bar{r}^C$.

Consider now the behavior of $b'(r^C) = b(a', b)$, which matches that of $a_{sb}'$ for any $r^C$. By Theorem 4.1 and Assumption 2.1, $b'(v; r^C) \in (b(a), b(\bar{a}))$. Also, $b'(v; r^C)$ solves $\max_{y \in [b(a), b(\bar{a})]} \{y b'(v; r^C) + \xi(y)\}$. By strict concavity of $\xi(y)$, it is enough to study how $\overline{w}(r^C)$ relates to $v/t^l$. The function $\overline{w}(r^C)$ crosses $v/t^l$ only once at $v^* \in (\overline{v'}, \overline{v'})$. Also, $\overline{w}(v; r^C) = w'(v; r^C)$ on $[v_b, v^b]$. So, it is enough to show that, as $r^C$ rises, $w'(v^b(r^C); r^C)$ falls and $w'(v_b(r^C); r^C)$ rises.

**Lemma B.1:** *Suppose $v^b$ and $v_b$ are characterized by (A.15) and (A.16). If $w'_e(v_b; r^C) > 0$ and $w'_e(v_b; r^C) > 0$, then $\frac{dv}{dr} w'(v^b(r^C); r^C) < 0$ and $\frac{dv}{dr} w'(v_b(r^C); r^C) > 0$.*

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

Consider $w'(v_b(r^C); r^C)$. If $t^l \leq 1/2$ or $r^C < \bar{r}^C$, then $v_b(r^C) = v^*$ and $w'_e(v^*; r^C) = (1 - t^l)(v^*/t^l) > 0$. If $t^l > 1/2$, then $w'(v^*; r^C) \uparrow w'(v^*/\bar{r}^C) = w'(v^*/\bar{r}^C, \bar{r}^C)$ as $r^C \uparrow \bar{r}^C$. By Lemma B.1, $w'(v_b(r^C); r^C)$ increases in $r^C$, for $r^C > \bar{r}^C$, because $w'_e(v_b(r^C); r^C) = 0$ when $v_b = \overline{v'}$. Similarly, $w'(v^b(r^C); r^C)$ decreases in $r^C$, because $w'_e(v^b(r^C); r^C) = 0$ when $v^b = v_b$.

**B.6. Proof of Corollary 4.4**

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

$$\frac{[b(a) - b(a_{fb}(v^b))] \int_{v'}^v g'(v) dv}{c}$$

$$\geq R^C(a_{sb}') + \int_{v'}^{c} b(a_{fb}(v)) G^C(v) dF^C$$

$$- b(a_{fb}(v^b)) \int_{v'}^v g'(v) dv.$$
Since \(a'_{j|b}\) and \(a'_{sb}\) are infeasible, the right-hand side is positive. \(R^c(a'_{sb})\) has been minimized. The result follows, since \(\int_{\bar{u}} g'(v) \, dv < 0\).

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

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

\[
VS^C(b^{-1}(b), r') = \int_{\bar{v}} \left[ b(v)u^C(v, r') + \xi(b(v)) \right] dF^C
\]

\[
+ r' \int_{\bar{v}} b(v)g'(v) \, dv,
\]

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

\[
\bar{b}^C(v; r') = \arg \max_{y \in [b(a), b(\pi)]} \phi(y, v; r').
\]

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

Suppose \([\bar{v}'^C, \bar{v}']\) is \(\emptyset\). Then \(\bar{b}^C\) is increasing and an argument similar to that in Lemma A.4 establishes that \(\bar{b}^C\) maximizes \(VS^C\). Since such a \(\bar{b}^C\) is pointwise continuous in \(r'\), so is \(VS^C(b^{-1}(\bar{b}^C(r'))), r'\).

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

\[
(b.5) \quad b^C(v^d)H(v^d, r') + \xi(b^C(v^d))F^C(v^d) + \int_{v^d} \phi(b^C(v), v; r') \, dF^C,
\]
where
\[
H(v^d, r') = r' \int_{v_a}^{v^d} g^I(v) \, dv + \int_{v^d}^{v^C} \overline{w}^C(v, r') \, dF^C.
\]

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

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

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

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

**Case 4:** \(b^C\) is constant at \(\overline{y}\) on \([v^a, \overline{v}^C]\). Then \(\overline{y} \in [b(\overline{a}), b(\overline{a})]\) has to maximize
\[
\overline{y} \left[ r' \int_{v_a}^{v^d} g^I(v) \, dv + \int_{v^d}^{v^C} \overline{w}^C(v, r') \, dF^C \right] + \xi(\overline{y}).
\]

The unique solution to this problem identifies a unique constant \(b^*_4\) and value \(\Phi(b^*_4; r')\), which is again continuous in \(r'\).

Now, let \(\hat{b}^C\) be the function that solves \(\max_{j=1,2,3,4} \Phi(b^*_j; r')\). An argument similar to that in Lemma A.5 establishes that
\[
\max_{b \in B} VSC(b^{-1}(b), r') = \Phi(\hat{b}^C; r') + b(a)r' \int_{v_a}^{v^d} g^I(v) \, dv,
\]
which is therefore continuous in \(r'\).

Now, let \(b^*_u = b(a^u)\) and let \(B^*\) be the set of \(b^C \in B\) that equal \(b^*_u\) on \([v^C, \overline{v}^C]\). By construction, \(VSC(b^{-1}(b^*_u), r') = \max_{b \in B^*} VSC(b^{-1}(b), r')\). I claim that there is \(b^C \in B^*\) such that \(VSC(b^{-1}(b^C), r') > VSC(b^{-1}(b^*_u), r')\). 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 \(r' > 0\), \(G^C\) implies \(w^C(v, r') > v/t^C\) for \(v \in [v_m, \overline{v}^C]\). I claim that \(\overline{w}^C(v_m, r') > v_m/t^C\). By the logic in Lemma A.6, \(\overline{w}^C(v_m, r') \leq w^C(v_m, r')\). If \(\overline{w}^C(v_m, r') = w^C(v_m, r')\), the claim follows. If \(\overline{w}^C(v_m, r') < w^C(v_m, r')\), then there is \(v_0 > v_m\) such that \(\overline{w}^C(v, r') = w^C(v, r')\) on \([v_m, v_0]\); so, \(\overline{w}^C(v_m, r') = w^C(v_0, r') \geq v_0/t^C > v_m/t^C\). Since \(\overline{w}^C\) is continuous and increasing, in either case there is
\[ v_1 > v_m \text{ such that } \mathcal{W}^C(v, r') > v/t^C \text{ on } [v_m, v_1]. \]

Construct \( \hat{b}^C \) by letting \( \hat{b}^C(v) = \arg\max_{y \in [b(a), b(m)]} \phi(y, v; r') \) if \( v \in [v_m, \bar{v}] \), and \( b^C_{un}(v) \) if \( v \in [v', v_m) \). Then, \( \hat{b}^C \in \mathcal{B} \), but \( \hat{b}^C(v) > b^C_{un}(v) \) on \( [v_m, v_1] \); so \( \hat{b}^C \notin \mathcal{B}^* \). Finally, \( VSC(b - 1(\hat{b}^C)/r') - VSC(b - 1(b^C_{un})/r') \) equals

\[
\int_{v_m}^{v_1} \left\{ [\hat{b}^C(v)w^C(v, r') + \xi(\hat{b}^C(v))] - [b^C_{un}(v)w^C(v, r') + \xi(b^C_{un}(v))] \right\} dF^C > 0.
\]

**B.8. Proof of Proposition 4.5**

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

\[
\mathcal{P}^N = \left\{ \max_{(a^j, h^j)} \sum_{j=1}^{N} \gamma^j W^j(a^j) + \sigma \sum_{j=1}^{N} \gamma^j [W^j(a^j) - h^j] \right\}
\]

s.t. \( a^j \) increasing, \( h^j \geq h^i + R^i(a^i) \), and \( h^j \geq 0 \), for all \( j, i \).

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

**Step 1:** There is \( b(a) \) low enough so that unused options suffice to satisfy \( IC^{ji}_i \) for \( j > i \). If \( j > i \), \( \bar{v}' < \bar{v} \) and

\[
\tilde{R}^j(b^i) = - \int_{\bar{v}'}^{v^j} b^i(v)g^i(v) \, dv - \int_{v^j}^{\bar{v}'} b^i(v)G^ji(v) \, dF^j,
\]

where

\[
g^i(v) = \frac{t^i - 1}{t^i}vf^i(v) - (1 - F^i(v)) \quad \text{and}
\]

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

if \( i > j \), \( \bar{v}' < \bar{v} \) and

\[
\tilde{R}^j(b^i) = - \int_{\bar{v}'}^{v^i} b^i(v)\bar{g}^i(v) \, dv + \int_{\bar{v}'}^{v^i} b^i(v)\bar{G}^ji(v) \, dF^j,
\]
where

\[
\tilde{g}^i(v) = \frac{t^i - 1}{t^i} v f^i(v) + F^i(v),
\]

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

Take \( j > i \). Suppose IC\(_{ji}^j\) is violated (and all other constraints hold): \( h^j < h^i + \tilde{R}^j(b') \). Fix \( b' \) for \( v \geq v^j \), and let \( b'(v) = b(a) \) for \( v < v^j \). Then,

\[
R^j(b') = -b(a) \int_{v^j}^{v^i} \tilde{g}^i(v) \, dv + \int_{v^i}^{v^j} b'(v) \tilde{G}^{ji}(v) \, dF^j.
\]

**Lemma B.2:** \( \int_{v^i}^{v^j} \tilde{g}^i(v) \, dv < 0 \).

**Proof:** Integrating by parts,

\[
\int_{v^i}^{v^j} \tilde{g}^i(v) \, dv = -\int_{v^i}^{v^j} (v/t^i) f^i(v) \, dv + F^j(v^i) v^i
\]

\[
= \int_{v^i}^{v^j} (v^j - (v/t^i)) f^i(v) \, dv.
\]

Note that \( v^i \leq s \leq v/t^i \), with strict inequality for \( v \in (v^i, v^j) \).

**Q.E.D.**

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

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

\[
\widetilde{\mathcal{P}}^N = \left\{ \left. \max_{(B,h) \in \mathcal{Y}} (1 - \sigma) \tilde{W}(B) + \sigma \tilde{\Pi}(B, h) \right| \Gamma(B, h) \leq 0 \right\},
\]

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

**Step 3:** Existence of interior points.
**LEMMA B.3:** In $\tilde{\mathcal{P}}^N$, there is $\{B, h\} \in \mathcal{Y}$ such that $\Gamma(B, h) < 0$.

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

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

Since $\tilde{R}^i(b^i)$ are bounded and $N$ is finite, the vector $h$ 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 $\tilde{\mathcal{P}}^N$. Note that $(X \times \mathbb{R})^N$ is a linear vector space and $\mathcal{Y}$ is a convex subset of it. By Lemma B.3, $\Gamma$ has interior points. Since $\tilde{I}$ and $\tilde{W}$ are concave ($b'' < 0$ and $c'' \geq 0$), the objective is concave and $\Gamma(B, h)$ is convex. For $\lambda \in \mathbb{R}_+$, define $L(B, h; \lambda)$ as

$$(1 - \sigma)\tilde{W}(B) + \sigma\tilde{I}(B, h) + \lambda^N h^N - \sum_{i=1}^N \sum_{j<i} \lambda^{ji} [\tilde{R}^i(b^i) + h^i - h^j]$$

$$= \sum_{i=1}^N \gamma^i \left[ \tilde{W}^i(b^i) - \sum_{j<i} \frac{\lambda^{ji}}{\gamma^i} \tilde{R}^i(b^i) \right] + \sum_{i=1}^N h^i \mu^i(\lambda, \gamma, \sigma),$$

where

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

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

$$b^i \in \arg \max_{b \in B} \tilde{W}^i(b) - \sum_{j<i} \frac{\lambda^{ji}}{\gamma^i} \tilde{R}^i(b)$$

(B.6)
and

\[(B.7) \quad h^i \in \arg \max_{h \in \mathbb{R}} \mu'(\lambda, \gamma, \sigma)h.\]

The second is equivalent to

\[(B.8) \quad -h^N \leq 0 \quad \text{and} \quad \lambda^N h^N = 0,\]

and, for \(j > i,\)

\[(B.9) \quad \tilde{R}(b^i) + h^i - h^i \leq 0 \quad \text{and} \quad \lambda^i [R^i(b^i) + h^i - h^i] = 0.\]

**Lemma B.4:** If \((B, h, \lambda)\) satisfies (B.6)–(B.9), then \(\mu'(\lambda, \gamma, \sigma) = 0\) for all \(i.\)

**Proof:** By (IR\(^N\)) and IC\(^i\), \(h^i \geq 0\) for all \(i;\) so, \(\mu'(\lambda, \gamma, \sigma) \geq 0\) for all \(i.\) Since \((1 - \sigma)\tilde{W}(B) + \sigma\tilde{H}(B, h)\) is bounded below by \(E(u_1(a^{nt}; s)) - c(a^{nt}) > 0,\) then \(\mu'(\lambda, \gamma, \sigma) \leq 0\) for all \(i.\)

**Corollary B.5:** If \(\sigma = 0,\) then \(\lambda = 0.\) If \(\sigma > 0,\) IR\(^N\) binds and, for every \(i < N,\) there is \(j > i\) such that IC\(^j\) binds.

**Proof:** Lemma B.4 implies the second part. For the first part, since \(\mu'(\lambda, \gamma, \sigma) = 0\) for all \(i,\)

\[
0 = \sum_{i=1}^{N} \mu'(\lambda, \gamma, \sigma) = \sum_{i=1}^{N-1} \left[ \sum_{j>i} \lambda^{ij} - \sum_{j<i} \lambda^{ji} \right] + \lambda^N - \sum_{j=N} \lambda^{jn} - \sigma = \lambda^N - \sigma.
\]

So, if \(\sigma = 0 = \lambda^N,\) then \(\mu^N(\lambda, \gamma, \sigma) = 0\) implies \(\sum_{j<N} \lambda^{jn} = 0.\) Hence, \(\lambda^{jn} = 0\) for \(j < N.\) Suppose for all \(j \geq i + 1, \lambda^{ij} = 0\) for all \(n < j.\) Then, \(\mu'(\lambda, \gamma, \sigma) = 0\) implies \(\sum_{j<i} \lambda^{ji} = \sum_{j>i} \lambda^{ij} = 0.\) Hence, \(\lambda^{ij} = 0\) for all \(j < i.\)

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

Thus, \(\tilde{P}^N\) has a solution if there is \((B, \lambda)\) so that, for every \(i, b^i\) solves (B.6), \(\mu'(\lambda, \gamma, \sigma) = 0,\) and (B.8) and (B.9) hold. By the arguments in the proof of Theorem 4.1 (see Step 5 below), for \(\lambda \geq 0,\) a solution \(b^i\) to (B.6) always exists and is unique on \((\nu^i, \nu')\) and is pointwise continuous in \(\lambda.\) Moreover, if \(\lambda^i \to +\infty\) for some \(j < i,\) then \(b^i \to b(a^{nt})\) on \((\nu^i, \nu'),\) and \(\tilde{R}^i(b^i) \to 0.\) And since \(\mu'(\lambda, \gamma, \sigma) = 0, \lambda^{ij} \to +\infty\) for some \(j' > i,\) so that \(\tilde{R}(b^i) \to 0\) and
\( h^i \to 0 \) (using the binding IC_{i-1}^{ij}). So there is \( \lambda^{ii} \) 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 \( \tilde{R}^n(b') \), and \( \xi(\cdot) = -b^{-1}(\cdot) - c(b^{-1}(\cdot)) \), \( b' \) must maximize within \( B \)

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

where \( \lambda^i \in \mathbb{R}_+^{i-1} \) and

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

We can apply to \( VS'(b'; \lambda^i) \) the method used in the two-type case to characterize \( b' \) (Theorem 4.1). If \( \lambda^i = 0 \), \( VS'(b'; 0) = \tilde{\nu}^i(b') \) and \( b' = b_{j\lambda} = b(a_{j\lambda}) \) on \( (v', \tilde{v}') \). For \( v > \tilde{v}' \), let \( b'(v) = b'(\tilde{v}') \). For \( v < \tilde{v}' \), \( b'(v) \) may be strictly smaller than \( b'(\tilde{v}') \) to satisfy IC_i^{ij} for \( j > i \).

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

\[
w'(\psi^+_j; \lambda^i) = \frac{\psi_j}{t^i} + \sum_{n=1}^{i-1} \lambda^{ni} q^n(\psi) - \sum_{n=1}^{i-1} \lambda^{ni} \frac{f^n(\psi)}{f^i(\psi)} q^n(\psi),
\]

\[
w'(\psi^-_j; \lambda^i) = \frac{\psi_j}{t^i} + \sum_{n=1}^{i-1} \lambda^{ni} q^n(\psi) - \sum_{n=j+1}^{i-1} \lambda^{ni} \frac{f^n(\psi)}{f^i(\psi)} q^n(\psi).
\]

Then,

\[
w'(\psi^-_j; \lambda^i) - w'(\psi^+_j; \lambda^i) = \lambda^{ij} \frac{f^i(\psi)}{f^i(\psi)} q'(\psi) \geq 0,
\]
since \( q'(v') = (1 - t')(v'/t') \geq 0 \). Letting \( \overline{w'}(v; \lambda') = \omega'(F'(v); \lambda') \) for \( v \in (v', \overline{v'}) \), construct \( \overline{VS}' \) as in the proof of Theorem 4.1.

Note that \( g''(v) < 0 \) for \( v \in (\overline{v'}, \overline{v''}) \). So, since \( \lambda'^{ni} > 0 \) for some \( n < i \), the first term in \( VS' \) is strictly negative. Letting \( w(v; \lambda i) = \omega(F(v); \lambda i) \) for \( v \in (v_i, v_{ni}) \), construct \( VS_i \) as in the proof of Theorem 4.1. Note that \( g''(v) < 0 \) for \( v \in (v_i, v_{ni}) \). So, since \( \lambda ni > 0 \) for some \( n < i \), the first term in \( VS_i \) is strictly negative. Let \( n = \min\{n : \lambda ni > 0\} \). Then, on \( (v_i, v_{ni}) \), the characterization of Lemma A.3 extends to \( VS_i \). So \( b_i \) must be constant at \( v_i(b) \) on \( (v_i, v_{ni}) \), where \( v_i(b) \leq v_i \) and \( v_i(b) = b_i(v) \) for \( v \in [v', v_{ni}] \). Moreover, \( b_i(v) = \overline{b}(v) \) if \( v_{ni} > v' \); and \( b_i(v) = b_i(v) \) for \( v \in [v', v_{ni}] \). The argument in Lemma A.4 yields that there is a (unique) maximizer of \( VS_i \). The argument in Lemma A.5 implies that the (unique) maximizer of \( VS_i \) is also the (unique) maximizer of \( VS' \).

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

\[
\int_{v_{ni}}^{v'} \left[ w_i(y; \lambda i) - w_i(v_{ni}; \lambda i) \right] dF_i = - \sum_{n=0}^{i-1} \lambda ni \int_{v_{ni}}^{v'} g''(v) \ dv.
\]

Since the sum is negative, \( v_{ni} < \overline{v'} \). This condition can be written as

\[
\int_{v_{ni}}^{v'} \left[ w(v_{ni}; \lambda i) - (v/t') \right] dF_i = - \sum_{n=0}^{i-1} \lambda ni \left[ \int_{v_{ni}}^{v'} G''(v) dF_i + \int_{v_{ni}}^{v'} g''(v) \ dv \right].
\]

To prove that \( w_i(v_{ni}; \lambda i) < \overline{v'} / t' \), it is enough to observe that the right-hand side is negative by (A.14). So, \( b_i \) exhibits bunching on \( [v_{ni}, \overline{v''}] \) at value \( y_{ni} < b_i(v) \).

Now consider the bottom of \( [v', \overline{v''}] \). By the logic in Lemma A.6, \( \overline{w}(v'; \lambda') \leq w_i(v'; \lambda i) \), with strict inequality if \( v'_{ni} > v' \). Moreover, for \( v < v'_{ni} \), \( w_i(v, \lambda') = v/t' + \sum_{n=1}^{i-1} \lambda ni q_i(v) \) and \( w_i(v'; \lambda i) = (v'/t')(1 + (1 - t') \sum_{n=1}^{i-1} \lambda ni) > v'/t' \). So, if \( \overline{w}(v'; \lambda') = w_i(v'; \lambda i) \), then \( b_i(v'; \lambda') > b_{ni}(v') \). Otherwise, ironing occurs on \( [v', v'_{ni}] \neq \emptyset \) and

\[
\int_{v'}^{v_{ni}} \left[ w_i(y; \lambda i) - \overline{w}(v_{ni}; \lambda i) \right] dF_i = 0,
\]

which corresponds to

\[
\int_{v'}^{v_{ni}} \left[ y/t' - \overline{w}(v_{ni}; \lambda i) \right] dF_i = - \sum_{n=1}^{i-1} \lambda ni \int_{v'}^{v_{ni}} G''(y) \ dv.
\]
Now, for $n < i$,
\[
\int_{v_i^n}^{v_i^b} G^{n}(y) \, dF^n = \int_{v_i^n}^{v_i^b} q^{n}(y) \, dF^n - \int_{v_i^n}^{v_i^b} q^{n}(y) \, dF^n
\]
\[
= \int_{v_i^n/t_i}^{v_i^n/t_n} (s - v_i^b) \, dF > 0.
\]

So $\tilde{w'}(v_i^n; \lambda^i) > v_i^n/t_i$, and $b^i(v_i^n; \lambda^i) > b_i(v_i^n)$.

Finally, note that for $v < v' < v_i^{t_i-1}$,
\[
w^i(v'; \lambda^i) - w^i(v; \lambda^i) = \frac{v' - v}{t_i} \left[ 1 + \sum_{n=1}^{i-1} \lambda^n (1 - t_i^n) \right]
\]
\[
+ \sum_{n=1}^{i-1} \lambda^n \left[ \frac{F^i(v') - F^i(v)}{f^i(v')} - \frac{f^i(v)}{f^i(v')} \right].
\]

So, $w^i(\cdot; \lambda^i)$ will be decreasing in a neighborhood of $v_i^n$ if, for $s' > s$ in $[s, s_i]$, ...

Hence, bunching at the bottom is more likely if $t_i^n$ is closer to 1 and $\sum_{n=1}^{i-1} \lambda^n$ 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 $s = 10$, $\tilde{s} = 15$, and $t = 0.9$. We first characterize the first-best $C$- and $I$-device. By Corollary 3.1, $p_{C}^e$ 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 $p_{I}^e(a) = p_{C}^e + q^I(a)$ such that $q^I(e(s)) = q^{0.9}(s)$ for every $s \in [s, \tilde{s}]$. Therefore, using the formula in Corollary 3.1,
\[
dq^I(a)/da = dq^{0.9}(s)/ds = -0.1.
\]

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

Consider now the difference between $C$’s and $I$’s expected utility from the efficient $I$-device (i.e., $R^C(a_{fb}^I)$). Recall that $p_{I}^e(a) = +\infty$ for $a \notin [100, 225]$. ...
Under this $I$-device, at time 2 type $C$ chooses $\alpha_C(s) = \frac{s^2}{t^2}$ for $s < \frac{3}{t}$ and $\alpha_C(s) = \bar{s}$ otherwise. Thus

$$R_C(a_{f_b}^C) = m - p_C^e - k + \int_t^\bar{s} \left[ 2s\sqrt{\alpha_C(s) - t\alpha_C(s)} \right] \frac{ds}{\bar{s} - s}$$

$$- \left\{ m - p_C^e - k + \int_t^\bar{s} \left[ 2s\sqrt{\varepsilon(s) - t\varepsilon(s)} \right] \frac{ds}{\bar{s} - s} \right\}$$

$$= \frac{1 - t}{3t(\bar{s} - s)} \left[ \bar{s}^3(3 - t)t - (1 + t)s^3 \right].$$

Substituting the values of $s$, $\bar{s}$, and $t$, we get $R_C(a_{f_b}^C) \approx 33.18$.

To compute the difference between $I$’s and $C$’s expected utilities from the efficient $C$-device (i.e., $R^C(a_{f_b}^C)$), recall that $p_C^e(a) = +\infty$ for $a \notin [100, 225]$. Given this, at time 2 type $I$ chooses $\alpha_I(s) = t^2s^2$ for $s > \frac{3}{t}$ and $\alpha_I(s) = \bar{s}$ otherwise. Thus

$$R^I(a_{f_b}^C) = m - p_C^e + \int_t^\bar{s} \left[ 2s\sqrt{\alpha_I(s) - \alpha_I(s)} \right] \frac{ds}{\bar{s} - s}$$

$$- \left\{ m - p_C^e + \int_t^\bar{s} \left[ 2s\sqrt{\varepsilon(s) - \varepsilon(s)} \right] \frac{ds}{\bar{s} - s} \right\}$$

$$= \frac{(1 - t)^2}{3(\bar{s} - s)} \left[ \bar{s}^3t^2 - \bar{s}^3 \right].$$

Substituting $s$, $\bar{s}$, and $t$, we get $R^I(a_{f_b}^C) \approx -1.43$.

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 $[a_b, a^b]$, we have that $a_b = [w'(v_b; r_C)^2$ and $a^b = [w'(v^b; r_C)^2$, where $w'(v; r)$ 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 $a_{\text{ef}} = (\frac{\gamma + \mu}{2})^2 = 156.25$.

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

$$g^I(v) = \frac{t - 1}{t} v f^I(v) + F_I(v) = \frac{1}{t(\bar{s} - s)} \left[ (2t - 1)s - \frac{s^3}{t} \right],$$

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

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

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 \([s, s/t]\), and hence it equals

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

Using the parameters’ values, this difference is \(-46.91 \). Since it exceeds \( R^C(a')_b \approx 33.18 \), \( I \) 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 \((a_0, p_0)\) using the formalism of Section 4.1. For simplicity, consider the two-type model. By Proposition 4.1, \( U^C(a_0, p_0) \geq U^J(a_0, p_0) \) with equality if and only if \( a_0 \) is constant over \((v, \overline{v})\). So \( C \) and \( I \) value the outside option differently, unless they always end up making the same choice.

When \( U^C(a_0, p_0) > U^J(a_0, p_0) \), the analysis in Section 4 can be adjusted without changing its thrust. The constraints (IR\(_C\)) and (IC\(_C\)) set two lower bounds on \( C \)'s payoff from the \( C \)-device: one endogenous (i.e., \( U^C(a', p') = U^J(a', p') + R^C(a', p') \)) and one exogenous (i.e., \( U^C(a_0, p_0) = U^J(a_0, p_0) + R^C(a_0) \)). The question is which binds first. In Section 4, (IC\(_C\)) always binds first, for (IR\(_C\)) and (IC\(_C\)) imply (IR\(_C\)). Now this is no longer true. Intuitively, if (IC\(_C\)) 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, \( a_0 = a'_I \) over \([v', \overline{v'}]\). In this case, she must grant \( C \) at least the rent \( R^C(a_0) \), which already exceeds \( R^C(a'_I) \). Finally, if (IC\(_I\)) 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 \( R^C(a_0) \) 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, \((a_0, p_0)\) 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 $e$, 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 $\alpha^*_2 = e(s_2) > e(s_1) = \alpha^*_1$, with payments $\pi_1 = \pi^t(s_1)$ and $\pi_2 = \pi^t(s_2)$ that satisfy

$$(E.1) \quad u_2(\alpha^*_2; s_2, t) - u_2(\alpha^*_1; s_2, t) \geq \pi_2 - \pi_1 \geq u_2(\alpha^*_2; s_1, t) - u_2(\alpha^*_1; s_1, t),$$

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 $e$: For any $t$, (E.1) does not pin down $\pi_1$ and $\pi_2$ uniquely. Suppose $t^I$ is close to $t^C$. Intuitively, to sustain $e$ 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 $e$ with each type. Since $t^I < t^C$, $I$ is tempted to pick $\alpha^*_1$ also in $s_2$, and the more so, the lower is $t^I$. So, for $I$ not to choose $\alpha^*_1$ in $s_2$, $\alpha^*_1$ must be sufficiently more expensive than $\alpha^*_2$, and this gap must rise as $t^I$ falls. At some point, this gap must exceed $C$'s willingness to pay for switching from $\alpha^*_2$ to $\alpha^*_1$ 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 $\bar{t} = \max T$.

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

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 $\alpha_i = \alpha(s_i)$ and $\pi_i = \pi(s_i)$. 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 $\alpha^* = e$ for all $i$, then $\alpha^*_N > \alpha^*_{N-1} > \cdots > \alpha^*_1$ (Assumption 2.1). To sustain $e$ with $t$, $\Delta_i$ must satisfy

$$(CIC_{i, i-1}) \quad u_2(\alpha^*_i; s_i, t) - u_2(\alpha^*_{i-1}; s_i, t) \geq \Delta_i \geq u_2(\alpha^*_i; s_{i-1}, t) - u_2(\alpha^*_{i-1}; s_{i-1}, t),$$

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 $s_k/s_{k-1} = \min_{s_i/s_{i-1}}, \min_{s_i/s_{i-1}}$, and suppose $\tilde{s}_{k-1} > s_k t$. Then,

$$u_2(\alpha^*_k; s_{k-1}, \tilde{t}) - u_2(\alpha^*_{k-1}; s_{k-1}, \tilde{t}) > u_2(\alpha^*_k; s_k, t) - u_2(\alpha^*_{k-1}; s_k, t),$$
and no $\Delta_k$ satisfies (CIC$_{k,k-1}$) for both $t$ and $\bar{t}$. If instead $t_{s_i} \geq \bar{t}_{s_{i-1}}$ for $i = 2, \ldots, N$, then for every $t$ and $i$,
\[
    u_2(\alpha^*_i; s_i, t) - u_2(\alpha^*_{i-1}; s_i, t) \geq u_2(\alpha^*_i; s_{i-1}, \bar{t}) - u_2(\alpha^*_{i-1}; s_{i-1}, \bar{t})
\geq u_2(\alpha^*_i; s_{i-1}, \bar{t}) - u_2(\alpha^*_{i-1}; s_{i-1}, \bar{t}).
\]
Set $\Delta^*_i = u_2(\alpha^*_i; s_{i-1}, \bar{t}) - u_2(\alpha^*_{i-1}; s_{i-1}, \bar{t})$. Then $\{\Delta^*_i\}_{i=2}^N$ satisfies all (CIC$_{i-1,i}$) for every $t$. The payment rule $\pi^*_i = \pi^*_i + \sum_{j=i}^{N} \Delta^*_j$—with $\pi^*_i$ small to satisfy (IR)—sustains $\mathbf{e}$ with each $t$.

So, if the heterogeneity across types (measured by $\bar{t}/t$) is small, the provider may sustain $\mathbf{e}$ 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 $\mathbf{e}$. Even if $\bar{t}/t$ is large, the provider may be able to design different devices such that each sustains $\mathbf{e}$ 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 $(\pi^*_1, \pi^*_2)$ that satisfy (E.1) and (IR) with equality:

\[
(1 - f)\pi^*_2 + f \pi^*_1 = (1 - f)u_1(\alpha^*_2; s_2) + fu_1(\alpha^*_1; s_1),
\]
where $f = F(s_1)$. Finally, choose $(\pi^*_1, \pi^*_2)$ so that $h$’s self-1 strictly prefers $\alpha^*_2$ in $s_2$—i.e., $u_1(\alpha^*_2; s_2) - \pi^*_2 > u_1(\alpha^*_1; s_2) - \pi^*_1$—and $(\pi^*_1, \pi^*_2)$ so that $l$’s self-1 strictly prefers $\alpha^*_1$ in $s_1$—that is, $u_1(\alpha^*_1; s_1) - \pi^*_1 > u_1(\alpha^*_2; s_1) - \pi^*_2$. Then, the $l$-device (respectively, $h$-device) sustains $\mathbf{e}$ and gives zero expected payoffs to the agent if and only if $l$ (or $h$) chooses it. Moreover, $l$ strictly prefers the $l$-device and $h$ the $h$-device. To see this, note that if self-$l$ of either type had to choose at time 2, under either device he would strictly prefer to implement $\mathbf{e}$. 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 $\mathbf{e}$. 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 $t^k = \min T^k$.

**PROPOSITION E.2**: Suppose $S$ is finite and $s_N > s_{N-1} > \cdots > s_1$. If $\max(\bar{t}^1/t^1, \bar{t}^2/t^2) > \min s_{i+1}/s_i$, 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 $\mathbf{e}$ with $t$, and (iii) all $t$ get the same expected payoff.

**PROOF**: Suppose $\max(\bar{t}^1/t^1, \bar{t}^2/t^2) = \bar{t}^1/t^1$—the other case is similar—and that there exist devices that satisfy (i)–(iii). Let $U$ be each $t$’s expected pay-off and $\mathbf{p}$ be the payment rule in the $t^1$-device. Given $\mathbf{p}$, let $a^*_i(t)$ be an optimal choice of $t \in T^1$ in $s_i$. For $t^1$, $a^*_i(t^1) = \alpha^*_i$ for every $i$. Let $\bar{S} = \{i : s_{i+1}/s_i <
Then, (a) for every \( i \), \( t^1_i s_i > t^1_i s_i \) and hence \( g_i(t^1_i) \geq \alpha_i^* \); (b) for \( i \in S \), \( t^1_i s_i > t^1_i s_{i+1} \), and so \( g_i(t^1_i) \geq \alpha_{i+1}^* > \alpha_i^* \). Since \( t \leq 1 \), (a) and (b) imply

\[
\mathbf{p}(a_i(t^1_i)) - \mathbf{p}(\alpha_i^*) \leq u_2(g_i(t^1_i); s_i, t^1_i) - u_2(\alpha_i^*; s_i, t^1_i) \\
\leq u_1(a_i(t^1_i); s_i) - u_1(\alpha_i^*; s_i),
\]

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

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

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


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

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