# Online Appendix for *Obviously Strategy-Proof Mechanisms*

Shengwu Li*

July 16, 2019

## 1 Corrigendum for Theorem 3

This section issues a correction for Theorem 3, which characterizes obviously strategy-proof mechanisms in binary allocation problems. Theorem 3 asserts that all obviously strategy-proof mechanisms are personal-clock auctions (Definition 15), which are hybrids of ascending auctions ('In-Transfer Falls') and descending-price 'procurement' auctions ('Out-Transfer Falls').

Clause *1.d.iv.* of Definition 15 deals with a corner case. It asserts that *whenever there is more than one 'non-quitting' action available to agent $i$, then there exists an action that, if played, ensures that agent $i$ is in the allocation.* However, the argument offered in the proof establishes a weaker conclusion: whenever there is more than one 'non-quitting' action available to agent $i$, then there exists a *continuation strategy* that, if played, ensures that agent $i$ is in the allocation. If we modify Clause *1.d.iv.* to reflect this, that suffices to correct Theorem 3.

The modification has no apparent economic significance: it does not change the set of choice rules that can be implemented by a personal-clock auction. When an agent encounters one of the information sets covered by *1.d.iv.*, the agent knows that the price will not change in future. Hence, we can alter the mechanism so that, upon hitting the information set in question, the agent immediately reports all information about his type, and the altered mechanism will be obviously strategy-proof. Thus, the additional extensive forms permitted under the amended definition do not allow us to implement additional choice rules.

*shengwu_li@fas.harvard.edu

## 1.1 Original definition and theorem

For ease of reference, we reproduce the original definition and theorem below, using the same notation.

**Definition 15.** *$G$ is a **personal-clock auction** if, for every $i \in N$, at every earliest information set $I_i^*$ such that $|A(I_i^*)| > 1$:*

1. *Either (**In-Transfer Falls**): There exists a fixed transfer $\bar{t}_i \in \mathbb{R}$, a going transfer $\tilde{t}_i : \{I_i \mid I_i^* \preceq I_i\} \to \mathbb{R}$, and a set of 'quitting' actions $A^q$ such that:*
   - *(a) For all $z$ where $I_i^* \prec z$:*
     - *i. Either: $i \notin g_y(z)$ and $g_{t,i}(z) = \bar{t}_i$.*
     - *ii. Or: $i \in g_y(z)$ and*
     $$g_{t,i}(z) = \inf_{I_i \mid I_i^* \preceq I_i \prec z} \tilde{t}_i(I_i) \tag{1}$$
   - *(b) For all $a \in A^q$, for all $z$ such that $a \in \psi_i(z)$: $i \notin g_y(z)$*
   - *(c) $A^q \cap A(I_i^*) \neq \emptyset$.*
   - *(d) For all $I_i', I_i'' \in \{I_i \mid I_i^* \preceq I_i\}$:*
     - *i. If $I_i' \prec I_i''$, then $\tilde{t}_i(I_i') \geq \tilde{t}_i(I_i'')$.*
     - *ii. If $I_i' \prec I_i''$, $\tilde{t}_i(I_i') > \tilde{t}_i(I_i'')$, and there does not exist $I_i'''$ such that $I_i' \prec I_i''' \prec I_i''$, then $A^q \cap A(I_i'') \neq \emptyset$.*
     - *iii. If $I_i' \prec I_i''$ and $\tilde{t}_i(I_i') > \tilde{t}_i(I_i'')$, then $|A(I_i') \setminus A^q| = 1$.*
     - *iv. If $|A(I_i') \setminus A^q| > 1$, then there exists $a \in A(I_i')$ such that: For all $z$ such that $a \in \psi_i(z)$: $i \in g_y(z)$.*
2. *Or (**Out-Transfer Falls**): As above, but we substitute every instance of "$i \in g_y(z)$" with "$i \notin g_y(z)$" and vice versa.*

**Theorem 3.** *If $(G, \mathbf{S}_N)$ OSP-implements $f_y$, then $\mathcal{P}(G, \mathbf{S}_N)$ is a personal-clock auction. If $G$ is a personal-clock auction, then there exist $\mathbf{S}_N$ and $f_y$ such that $(G, \mathbf{S}_N)$ OSP-implements $f_y$.*

## 1.2 Counterexample

The following is a counterexample to Theorem 3 as originally stated. There are two agents, with values in the interval $[0, 2]$ for the object. Agents do not make payments unless they win the object. First, agent i reports whether $\theta_i = 1$ or $\theta_i \in [0, 2] \setminus \{1\}$.

If $i$ reports type $\theta_i = 1$ , then agent $j$ reports either $\theta_j < 1$ or $\theta_j \geq 1$. If $\theta_j < 1$, then $i$ wins the object and pays 1. If $\theta_j \geq 1$, then no agent wins.

If $i$ reports $\theta_i \in [0,2] \setminus \{1\}$, then immediately thereafter he reports either $\theta_i > 1$ or $\theta_i < 1$. If $i$ reports $\theta_i > 1$, then he wins and pays 1. Otherwise, no agent wins.

The above mechanism is obviously strategy-proof. However, clause *1.d.iv.* of Definition 15 is violated when agent $i$ reports whether $\theta_i = 1$ or $\theta_i \in [0,2] \setminus \{1\}$. Both actions are not quitting actions, since they do not rule out that $i$ wins the object. Neither action ensures that $i$ wins the object. However, there exists a continuation strategy for $i$ that ensures he wins: Report that $\theta_i \in [0,2] \setminus \{1\}$, then report $\theta_i > 1$.

## 1.3 Correction

The following step in the proof of Theorem 3 (page 17 of the Online Appendix) is incorrect:

> Suppose that there does not exist $a \in A(I'_i)$ such that, for all $z$ such that $a \in \psi_i(z)$, $i \in g_y(z)$. Then there must exist $(h'' \in I'_i, S''_{-i})$ such that
>
> $$i \notin g_y(h'', S_i^{\theta'_i}, S''_{-i})$$
>
> $$g_{t,i}(h'', S_i^{\theta'_i}, S''_{-i}) = \bar{t}_i$$

The first sentence supposes that no action ensures that $i$ wins the object. The second sentence deduces (incorrectly) that the continuation strategy of type $\theta'_i$ does not ensure that $i$ wins the object.

We now amend clause *1.d.iv.* to read as follows:

> If $|A(I'_i) \setminus A^q| > 1$, then there exists $S_i$ such that: For all $h \in I'_i$, for all $S_{-i}$: $i \in g_y(z)$ where $z = z^G(h, S_i, S_{-i})$.

We can then modify the proof to suppose the second step directly, replacing the part quoted above, as follows:

> Suppose there exists $(h'' \in I'_i, S''_{-i})$ such that $i \notin g_y(h'', S_i^{\theta'_i}, S''_{-i})$. Then $g_{t,i}(h'', S_i^{\theta'_i}, S''_{-i}) = \bar{t}_i$.

We then tweak the other direction of the proof, to show that all personal-clock auctions are obviously strategy-proof under the amended Definition 15. This requires us to modify the construction of the agent's obviously dominant strategy. If the agent faces

In-Transfer Falls, encounters an information set with more than one non-quitting action, and finds it strictly profitable to be in the allocation, the original construction required that the agent *play the action* that ensures he is in the allocation. Now, we instead require the agent *adopts the continuation strategy* that ensures he is in the allocation. The construction for Out-Transfer Falls proceeds symmetrically.

These modifications suffice to correct Theorem 3 and its proof.

### 1.4 Amended definition and theorem

The amended definition and theorem are as follows. The only modification is to Clause *i.d.iv.* of the definition.

**Definition 15A** (correct)**.** *$G$ is a **personal-clock auction** if, for every $i \in N$, at every earliest information set $I_i^*$ such that $|A(I_i^*)| > 1$:*

1. *Either (**In-Transfer Falls**): There exists a fixed transfer $\bar{t}_i \in \mathbb{R}$, a going transfer $\tilde{t}_i : \{I_i \mid I_i^* \preceq I_i\} \to \mathbb{R}$, and a set of 'quitting' actions $A^q$ such that:*
   (a) *For all $z$ where $I_i^* \prec z$:*
      i. *Either: $i \notin g_y(z)$ and $g_{t,i}(z) = \bar{t}_i$.*
      ii. *Or: $i \in g_y(z)$ and*
$$g_{t,i}(z) = \inf_{I_i \mid I_i^* \preceq I_i \prec z} \tilde{t}_i(I_i) \tag{2}$$
   (b) *For all $a \in A^q$, for all $z$ such that $a \in \psi_i(z)$: $i \notin g_y(z)$*
   (c) *$A^q \cap A(I_i^*) \neq \emptyset$.*
   (d) *For all $I_i', I_i'' \in \{I_i \mid I_i^* \preceq I_i\}$:*
      i. *If $I_i' \prec I_i''$, then $\tilde{t}_i(I_i') \geq \tilde{t}_i(I_i'')$.*
      ii. *If $I_i' \prec I_i''$, $\tilde{t}_i(I_i') > \tilde{t}_i(I_i'')$, and there does not exist $I_i'''$ such that $I_i' \prec I_i''' \prec I_i''$, then $A^q \cap A(I_i'') \neq \emptyset$.*
      iii. *If $I_i' \prec I_i''$ and $\tilde{t}_i(I_i') > \tilde{t}_i(I_i'')$, then $|A(I_i') \setminus A^q| = 1$.*
      iv. *If $|A(I_i') \setminus A^q| > 1$, then there exists $S_i$ such that: For all $h \in I_i'$, for all $S_{-i}$: $i \in g_y(z)$ where $z = z^G(h, S_i, S_{-i})$.*
2. *Or (**Out-Transfer Falls**): As above, but we substitute every instance of "$i \in g_y(z)$" with "$i \notin g_y(z)$" and vice versa.*

**Theorem 3A** (correct)**.** *If $(G, \mathbf{S}_N)$ OSP-implements $f_y$, then $\mathcal{P}(G, \mathbf{S}_N)$ is a personal-clock auction. If $G$ is a personal-clock auction, then there exist $\mathbf{S}_N$ and $f_y$ such that $(G, \mathbf{S}_N)$ OSP-implements $f_y$.*

## 2 Proofs omitted from the main text

### 2.1 Proof of Theorem 1

*Proof.* First we prove the "if" direction. Fix agent 1 and preferences $\theta_1$. Suppose that $S_1$ is not obviously dominant in $G = \langle H, \prec, A, \mathcal{A}, P, \delta_c, (\mathcal{I}_i)_{i \in N}, g \rangle$. We need to demonstrate that there exists $\tilde{G}$ that is $i$-indistinguishable from $G$, such that $\lambda_{G,\tilde{G}}(S_1)$ is not weakly dominant in $\tilde{G}$ . We proceed by construction.

Let $(S_1', I_1, h^{\sup}, S_{-1}^{\sup}, d_c^{\sup}, h^{\inf}, S_{-1}^{\inf}, d_c^{\inf})$ be such that $I_1 \in \alpha(S_1, S_1')$, $h^{\inf} \in I_1$, $h^{\sup} \in I_1$, and

$$u_1^G(h^{\sup}, S_1', S_{-1}^{\sup}, d_c^{\sup}, \theta_1) > u_1^G(h^{\inf}, S_1, S_{-1}^{\inf}, d_c^{\inf}, \theta_1) \tag{3}$$

Since $G$ is a game of perfect recall, we can pick $(S_{-1}^{\inf}, d_c^{\inf})$ such that $h^{\inf} \prec z^G(h_\emptyset, S_1, S_{-1}^{\inf}, d_c^{\inf})$, by specifying that $(S_{-1}^{\inf}, d_c^{\inf})$ plays in a way consistent with $h^{\inf}$ at any $h \prec h^{\inf}$. Likewise for $h^{\sup}$ and $(S_{-1}^{\sup}, d_c^{\sup})$. Suppose we have so done.

We now define another game $\tilde{G}$ is 1-indistinguishable from $G$. Intuitively, we construct this as follows:

1. We add a chance move at the start of the game; chance can play $L$ or $R$.
2. Agent 1 does not at any history know whether chance played $L$ or $R$.
3. If chance plays $L$, then the game proceeds as in $G$.
4. If chance plays $R$, then the game proceeds mechanically as though all players in $N \setminus 1$ and chance played according to $S_{-1}^{\sup}, d_c^{\sup}$ in $G$, with one exception:
5. If chance played $R$, we reach the information set corresponding to $I_1$, and agent 1 plays $S_1(I_1)$, then the game henceforth proceeds mechanically as though all players in $N \setminus 1$ and chance played according to $S_{-1}^{\inf}, d_c^{\inf}$ in $G$.

Formally, the construction proceeds as such: $\tilde{A} = A \cup \{L, R\}$, where $A \cap \{L, R\} = \emptyset$. There is a new starting history $\tilde{h}_\emptyset$, with two successors $\sigma(\tilde{h}_\emptyset) = \{\tilde{h}_L, \tilde{h}_R\}$, $\tilde{\mathcal{A}}(\tilde{h}_L) = L$, $\tilde{\mathcal{A}}(\tilde{h}_R) = R$, $\tilde{P}(\tilde{h}_\emptyset) = c$. The subtree $\tilde{H}_L \subset \tilde{H}$ starting from $\tilde{h}_L$ ordered by $\tilde{\prec}$ is the same as the arborescence $(H, \prec)$. $(\tilde{\mathcal{A}}, \tilde{P}, \tilde{\delta}_c, \tilde{g})$ are defined on $\tilde{H}_L$ exactly as $(\mathcal{A}, P, \delta_c, g)$ are on $H$. For $j \neq 1$, $\tilde{\mathcal{I}}_j$ is defined as on $H$.

We now construct the subtree starting from $\tilde{h}_R$. Let $h^*$ be such that $h^* \in \sigma(h^{\sup})$, $h^* \preceq z^G(h^{\sup}, S_1, S_{-1}^{\sup}, d_c^{\sup})$.

$$
\begin{aligned}
H' \equiv \{h \in H \mid \exists S_1'' : h \preceq z^G(h_\emptyset, S_1'', S_{-1}^{\text{sup}}, d_c^{\text{sup}})\}\\
\cap[\{h \in H \mid P(h) = 1\} \cup \{h \in Z\}]\\
\setminus\{h \in H \mid h^* \preceq h\}
\end{aligned}
\tag{4}
$$

In words, these are the histories that can be reached by *some* $S_1''$ when facing $S_{-1}^{\text{sup}}, d_c^{\text{sup}}$, where either agent 1 is called to play or that history is terminal, and such that those histories are not $h^*$ or its successors.

Let $h^{**}$ be such that $h^{**} \in \sigma(h^{\text{inf}})$, $h^{**} \preceq z^G(h^{\text{inf}}, S_1, S_{-1}^{\text{inf}}, d_c^{\text{inf}})$.

$$
\begin{aligned}
H'' \equiv \{h \in H \mid \exists S_1'' : h \preceq z^G(h_\emptyset, S_1'', S_{-1}^{\text{inf}}, d_c^{\text{inf}})\}\\
\cap[\{h \in H \mid P(h) = 1\} \cup \{h \in Z\}]\\
\cap\{h \in H \mid h^{**} \preceq h\}
\end{aligned}
\tag{5}
$$

In words, these are the histories that can be reached by *some* $S_1''$ when facing $S_{-1}^{\text{inf}}, d_c^{\text{inf}}$, where either agent 1 is called to play or that history is terminal, and such that those histories are $h^{**}$ or its successors.

We now paste these together. Let $\tilde{H}_R$ be the rooted subtree ordered by $\tilde{\prec}$, for some bijection $\gamma : \tilde{H}_R \to H' \cup H''$, such that for all $\tilde{h}, \tilde{h}' \in \tilde{H}_R$, $\tilde{h} \tilde{\prec} \tilde{h}'$ if and only if

1. EITHER: $\gamma(\tilde{h}), \gamma(\tilde{h}') \in H'$ and $\gamma(\tilde{h}) \prec \gamma(\tilde{h}')$

2. OR: $\gamma(\tilde{h}), \gamma(\tilde{h}') \in H''$ and $\gamma(\tilde{h}) \prec \gamma(\tilde{h}')$

3. OR: $\gamma(\tilde{h}) \prec h^*$ and $h^{**} \preceq \gamma(\tilde{h}')$

The root of this subtree exists and is unique; it corresponds to $\gamma^{-1}(h)$, where $h$ is the earliest history preceding $z^G(h_\emptyset, S_1, S_{-1}^{\text{sup}}, d_c^{\text{sup}})\}$ where 1 is called to play. Let $\tilde{h}_R$ be the root of $\tilde{H}_R$. This completes the specification of $\tilde{H}$.

For all $\tilde{h} \in \tilde{H}_R$, we define:

1. $\tilde{g}(\tilde{h}) = g(\gamma(\tilde{h}))$ if $\tilde{h}$ is a terminal history.

2. $\tilde{P}(\tilde{h}) = 1$ if $\tilde{h}$ is not a terminal history.

For all $\tilde{h} \in \tilde{H}_R \setminus \tilde{h}_R$, we define $\tilde{A}(\tilde{h}) = A(h)$, for the unique $(\tilde{h}', h)$ such that:

1. $\tilde{h} \in \sigma(\tilde{h}')$

2. $h \in \sigma(\gamma(\tilde{h}'))$

3. $h \preceq \gamma(\tilde{h})$

We now specify the information sets for agent 1. Every $\tilde{h} \in \tilde{H}_L$ corresponds to a unique history in $H$. We use $\gamma_L$ to denote the bijection from $\tilde{H}_L$ to $H$. Let $\hat{\gamma}$ be defined as $\gamma_L$ on $\tilde{H}_L$ and $\gamma$ on $\tilde{H}_R$.

1's information partition $\tilde{\mathcal{I}}_1$ is defined as such: $\forall \tilde{h}, \tilde{h}' \in \tilde{H}$ :

$$\exists \tilde{I}_1' \in \tilde{\mathcal{I}}_1 : \tilde{h}, \tilde{h}' \in \tilde{I}_1' \tag{6}$$

if and only if

$$\exists I_1' \in \mathcal{I}_1 : \hat{\gamma}(\tilde{h}), \hat{\gamma}(\tilde{h}') \in \tilde{I}_1' \tag{7}$$

All that remains is to define $\delta_c$; we need only specify that at $\tilde{h}_\emptyset$, $c$ plays $R$ with certainty.[1]

$\tilde{G} = \langle \tilde{H}, \tilde{\prec}, \tilde{A}, \tilde{\mathcal{A}}, \tilde{P}, \tilde{\delta}_c, (\tilde{\mathcal{I}}_i)_{i \in N}, \tilde{g} \rangle$ is 1-indistinguishable from $G$. Every experience at some history in $\tilde{H}_L$ corresponds to some experience in $G$, and *vice versa*. Moreover, any experience at some history in $\tilde{H}_R$ could also be produced by some history in $\tilde{H}_L$.

Let $\lambda_{G,\tilde{G}}$ be the appropriate bijection from 1's information sets and actions in $G$ onto 1's information sets and actions in $\tilde{G}$. Take arbitrary $\tilde{S}_{-1}$. Observe that since $I_1 \in \alpha(S_1, S_1')$, $\lambda_{G,\tilde{G}}(S_1)$ and $\lambda_{G,\tilde{G}}(S_1')$ result in the same histories following $\tilde{h}_R$, until they reach information set $\lambda_{G,\tilde{G}}(I_1)$. Having reached that point, $\lambda_{G,\tilde{G}}(S_1)$ leads to outcome $g(z^G(h^{\text{inf}}, S_1, S_{-1}^{\text{inf}}, d_c^{\text{inf}}))$ and $\lambda_{G,\tilde{G}}(S_1')$ leads to outcome $g(z^G(h^{\text{sup}}, S_1', S_{-1}^{\text{sup}}, d_c^{\text{sup}}))$. Thus,

$$\begin{aligned}
&\mathbb{E}_{\tilde{\delta}_c}[u_1^{\tilde{G}}(\tilde{h}_\emptyset, \lambda_{G,\tilde{G}}(S_1'), \tilde{S}_{-1}, \tilde{d}_c, \theta_1)] \\
&\quad = u_1^G(h^{\text{sup}}, S_1', S_{-1}^{\text{sup}}, d_c^{\text{sup}}, \theta_1) \\
&\quad > u_1^G(h^{\text{inf}}, S_1, S_{-1}^{\text{inf}}, d_c^{\text{inf}}, \theta_1) \\
&= \mathbb{E}_{\tilde{\delta}_c}[u_1^{\tilde{G}}(\tilde{h}_\emptyset, \lambda_{G,\tilde{G}}(S_1), \tilde{S}_{-1}, \tilde{d}_c, \theta_1)]
\end{aligned} \tag{8}$$

So $\lambda_{G,\tilde{G}}(S_1)$ is not weakly dominant in $\tilde{G}$.

We now prove the "only if" direction. Take arbitrary $\tilde{G}$. Suppose $\lambda_{G,\tilde{G}}(S_1) \equiv \tilde{S}_1$ is not weakly dominant in $\tilde{G}$ (given type $\theta_1$). We want to show that $S_1$ is not obviously dominant in $G$.

[1]If one prefers to avoid $\delta_c$ without full support, an alternative proof puts $\epsilon$ probability on $L$, for $\epsilon$ close to 0, or for games with $|N| > 2$ assigns $\tilde{P}(\tilde{h}_\emptyset) = 2$.

There exist $\tilde{S}'_1$ and $\tilde{S}'_{-1}$ such that:

$$\mathbb{E}_{\tilde{\delta}_c}[u_1^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}'_1, \tilde{S}'_{-1}, \tilde{d}_c, \theta_1)] > \mathbb{E}_{\tilde{\delta}_c}[u_1^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}_1, \tilde{S}'_{-1}, \tilde{d}_c, \theta_1)] \tag{9}$$

This inequality must hold for some realization of the chance function, so there exists $\tilde{d}_c$ such that:

$$u_1^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}'_1, \tilde{S}'_{-1}, \tilde{d}_c, \theta_1) > u_1^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}_1, \tilde{S}'_{-1}, \tilde{d}_c, \theta_1) \tag{10}$$

Fix $(\tilde{S}'_1, \tilde{S}'_{-1}, \tilde{d}_c, \theta_1)$.

$$z^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}'_1, \tilde{S}'_{-1}, \tilde{d}_c) \neq z^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}_1, \tilde{S}'_{-1}, \tilde{d}_c) \tag{11}$$

Define:

$$\tilde{H}^* \equiv \{\tilde{h} \in \tilde{H} \mid \tilde{h} \prec z^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}'_1, \tilde{S}'_{-1}, \tilde{d}_c) \text{ and } \tilde{h} \prec z^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}_1, \tilde{S}'_{-1}, \tilde{d}_c)\} \tag{12}$$

$$\tilde{h}^* \equiv \tilde{h} \in \tilde{H}^* \text{ such that } \forall \tilde{h}' \in \tilde{H}^* : \tilde{h}' \preceq \tilde{h} \tag{13}$$

Since the opponent strategies and chance moves are held constant across both sides of Equation 11, $P(\tilde{h}^*) = 1$ and $\tilde{h}^* \in \tilde{I}_1$, where $\tilde{S}_1(\tilde{I}_1) \neq \tilde{S}'_1(\tilde{I}_1)$. Moreover, $\tilde{I}_1 \in \alpha(\tilde{S}_1, \tilde{S}'_1)$ and $\lambda_{\tilde{G},G}(\tilde{I}_1) \in \alpha(S_1, S'_1)$, where we denote $S'_1 \equiv \lambda_{\tilde{G},G}(\tilde{S}'_1)$.

Since $G$ and $\tilde{G}$ are 1-indistinguishable, consider the experiences $\lambda_{\tilde{G},G}(\psi_1(z^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}_1, \tilde{S}'_{-1}, \tilde{d}_c)))$ and $\lambda_{\tilde{G},G}(\psi_1(z^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}'_1, \tilde{S}'_{-1}, \tilde{d}_c)))$.

In $G$, $\lambda_{\tilde{G},G}(\psi_1(z^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}_1, \tilde{S}'_{-1}, \tilde{d}_c)))$ could lead to outcome $\tilde{g}(z^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}_1, \tilde{S}'_{-1}, \tilde{d}_c))$. We use $(S^{\text{inf}}_{-1}, d^{\text{inf}}_c)$ to denote the corresponding opponent strategies and chance realizations that lead to that outcome. We denote $h^{\text{inf}} \equiv h \in \lambda_{\tilde{G},G}(\tilde{I}_1) : h \prec z^G(h_\emptyset, S_1, S^{\text{inf}}_{-1}, d^{\text{inf}}_c)$ .

In $G$, $\lambda_{\tilde{G},G}(\psi_1(z^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}'_1, \tilde{S}'_{-1}, \tilde{d}_c)))$ could lead to outcome $\tilde{g}(z^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}'_1, \tilde{S}'_{-1}, \tilde{d}_c))$. We use $(S^{\text{sup}}_{-1}, d^{\text{sup}}_c)$ to denote the corresponding opponent strategies and chance realizations that lead to that outcome. We denote $h^{\text{sup}} \equiv h \in \lambda_{\tilde{G},G}(\tilde{I}_1) : h \prec z^G(h_\emptyset, S'_1, S^{\text{sup}}_{-1}, d^{\text{sup}}_c)$ .

$$\begin{aligned} &u_1^G(h^{\text{sup}}, S'_1, S^{\text{sup}}_{-1}, d^{\text{sup}}_c, \theta_1) \\ &= u_1^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}'_1, \tilde{S}'_{-1}, \tilde{d}_c, \theta_1) \\ &> u_1^{\tilde{G}}(\tilde{h}_\emptyset, \tilde{S}_1, \tilde{S}'_{-1}, \tilde{d}_c, \theta_1) \\ &= u_1^G(h^{\text{inf}}, S_1, S^{\text{inf}}_{-1}, d^{\text{inf}}_c, \theta_1) \end{aligned} \tag{14}$$

where $h^{\sup}, h^{\inf} \in \lambda_{\tilde{G},G}(\tilde{I}_1)$ and $\lambda_{\tilde{G},G}(\tilde{I}_1) \in \alpha(S_1, S_1')$. Thus $S_1$ is not obviously dominant in $G$.

$\square$

## 2.2 Proof of Theorem 2

*Proof.* The key is to see that, for every $G \in \mathcal{G}$, there is a corresponding $\tilde{S}_0^{\Delta}$, and *vice versa*. We use $\bar{\mathbb{S}}_0$ to denote the support of $\tilde{S}_0^{\Delta}$. In particular, observe the following isomorphism:

Table 1: Equivalence between extensive game forms and Planner mixed strategies

| | |
|---|---|
| $G$ | $\tilde{S}_0^{\Delta}$ |
| $d_c$ | $\tilde{S}_0 \in \bar{\mathbb{S}}_0$ |
| $\delta_c$ | the probability measure specified by $\tilde{S}_0^{\Delta}$ |
| $g(z)$ for $z \in Z$ | the Planner's choice of outcome when she ends the game |
| $I_i$ | $((m_k, R_k, r_k)_{k=1}^{t-1}, m_t, R_t)$ consistent with some $\tilde{S}_0 \in \bar{\mathbb{S}}_0$ |
| $A(I_i)$ | $R_t$ |
| $\psi_i(z)$ | $o_i^c$ consistent with some $\tilde{S}_0 \in \bar{\mathbb{S}}_0$ and $\tilde{S}_N$ |

Information sets in $G$ are equivalent to sequences of past communication $((m_k, R_k, r_k)_{k=1}^{t-1}, m_t, R_t)$ under $\tilde{S}_0^{\Delta}$. Available actions at some information set $A(I_i)$ are equivalent to acceptable responses $R_t$. Thus, for any strategy in some game $G$, we can construct an equivalent strategy given appropriate $\tilde{S}_0^{\Delta}$, and vice versa.

Furthermore, fixing a chance realization $d_c$ and agent strategies $S_N$ uniquely results in some outcome. Similarly, fixing a realization of the planner's mixed strategy $\tilde{S}_0 \in \bar{\mathbb{S}}_0$ and agent strategies $\tilde{S}_N$ uniquely determines some outcome. Consequently, for any $G \in \mathcal{G}$, there exists $\tilde{S}_0^{\Delta}$ with the same strategies available for each agent and the same resulting (probability measure over) outcomes, and *vice versa*.[2] Table 1 summarizes.

The next step is to see that a bilateral commitment $\hat{\mathbb{S}}_0^i$ is equivalent to the Planner promising to 'run' only games in some equivalence class that is $i$-indistinguishable.

Suppose that there is some $G$ that OSP-implements $f$. Pick some equivalent $\tilde{S}_0^{\Delta}$ with support $\bar{\mathbb{S}}_0$. For each $i \in N$, specify the bilateral commitment $\hat{\mathbb{S}}_0^i \equiv \Phi_i^{-1}(\Phi_i(\bar{\mathbb{S}}_0))$. These bilateral commitments support $f$.

To see this, take any $\tilde{S}_0^{\Delta\prime} \in \Delta\hat{\mathbb{S}}_0^i$, with support $\bar{\mathbb{S}}_0'$. For any $\tilde{S}_0' \in \bar{\mathbb{S}}_0'$, for any $\tilde{S}_N'$, there exists $\tilde{S}_0 \in \bar{\mathbb{S}}_0$ and $\tilde{S}_N$ such that $\phi_i(\tilde{S}_0', \tilde{S}_N') = \phi_i(\tilde{S}_0, \tilde{S}_N)$. By construction, $G$ is such that: There exists $z \in Z$ where $\psi_i(z)$ and $g(z)$ are equivalent to $\phi_i(\tilde{S}_0, \tilde{S}_N)$. Thus, for

[2] Implicitly, this relies on the requirement that both $G$ and $\tilde{S}_0^{\Delta}$ have finite length. If one had finite length but the other could be infinitely long, the resulting outcome would not be well defined and the equivalence would not hold.

$G'$ that is equivalent to $\tilde{S}_0^{\Delta\prime}$, every terminal history in $G'$ results in the same experience for $i$ and the same outcome as some terminal history in $G$. Consequently, $G$ and $G'$ are $i$-indistinguishable. Thus, by Theorem 1, the strategy assigned to agent $i$ with type $\theta_i$ is weakly dominant in $G'$, which implies that it is a best response to $\tilde{S}_0^{\Delta\prime}$ and any $\tilde{S}_{N\setminus i}$ in the bilateral commitment game. Thus, if $f$ is OSP-implementable, then $f$ can be supported by bilateral commitments.

Suppose that $f$ can be supported by bilateral commitments $(\hat{\mathbb{S}}_0^i)_{i\in N}$, with requisite $\tilde{S}_0^{\Delta}$ (with support $\bar{\mathbb{S}}_0$) and $\tilde{\mathbf{S}}_N$. Without loss of generality, let us suppose these are 'minimal' bilateral commitments, *i.e.* $\hat{\mathbb{S}}_0^i = \Phi_i^{-1}(\Phi_i(\bar{\mathbb{S}}_0))$. Pick $G$ that is equivalent to $\tilde{S}_0^{\Delta}$. $G$ OSP-implements $f$.

To see this, consider any $G'$ such that $G$ and $G'$ are $i$-indistinguishable. Let $\tilde{S}_0^{\Delta\prime}$ denote the Planner strategy that corresponds to $G'$. At any terminal history $z'$ in $G'$, the resulting experience $\psi_i(z')$ and outcome $g'(z')$ are equivalent to the experience $\psi_i(z)$ and outcome $g(z)$ for some terminal history $z$ in $G$. These in turn correspond to some observation $o_i \in \Phi_i(\bar{\mathbb{S}}_0)$. Thus $\tilde{S}_0^{\Delta\prime} \in \Delta\hat{\mathbb{S}}_0^i$. Since $f$ is supported by $(\hat{\mathbb{S}}_0^j)_{j\in N}$, $\tilde{\mathbf{S}}_i^{\theta_i}$ is a best response (for type $\theta_i$) to $\tilde{S}_0^{\Delta\prime}$ and any $\tilde{S}_{N\setminus i}$. Thus, the equivalent strategy $\mathbf{S}_i(\theta_i)$ is weakly dominant in $G'$. Since this argument holds for all $i$-indistinguishable $G'$, by Theorem 1, $\mathbf{S}_i(\theta_i)$ is obviously dominant in $G$. Thus, if $f$ can be supported by bilateral commitments, then $f$ is OSP-implementable.

□

### 2.3 Proof of Proposition 2

*Proof.* We prove the contrapositive. Suppose $(\tilde{G}, \tilde{\mathbf{S}}_N)$ does not OSP-implement $f$. Then there exists some $(i, \theta_i, \tilde{S}_i', \tilde{I}_i)$ such that $\tilde{I}_i \in \alpha(\tilde{\mathbf{S}}_i(\theta_i), \tilde{S}_i')$ and

$$u_i^{\tilde{G}}(\tilde{h}, \tilde{\mathbf{S}}_i(\theta_i), \tilde{S}_{-i}, \tilde{d}_c, \theta_i,) < u_i^{\tilde{G}}(\tilde{h}', \tilde{S}_i', \tilde{S}_{-i}', \tilde{d}_c', \theta_i,) \tag{15}$$

for some $(\tilde{h}, \tilde{S}_{-i}, \tilde{d}_c)$ and $(\tilde{h}', \tilde{S}_{-i}', \tilde{d}_c')$.

Notice that $\tilde{h}$ and $\tilde{h}'$ correspond to histories $h$ and $h'$ in $G$. Moreover, we can define $S_i' = \tilde{S}_i'$ at information sets containing histories that are shared by $G$ and $\tilde{G}$, and specify $S_i'$ arbitrarily elsewhere. We do the same for $(\tilde{S}_{-i}, \tilde{d}_c)$ and $(\tilde{S}_{-i}', \tilde{d}_c')$, to construct $(S_{-i}, d_c)$ and $(S_{-i}, d_c)$. But, starting from $h$ and $h'$ respectively, these result in the same outcomes as their partners in $\tilde{G}$. Thus,

$$u_i^{G}(h, \mathbf{S}_i(\theta_i), S_{-i}, d_c, \theta_i,) < u_i^{G}(h', S_i', S_{-i}', d_c', \theta_i,) \tag{16}$$

$h, h' \in I_i$, for $I_i \in \alpha(\mathbf{S}_i(\theta_i), S_i')$. Thus $(G, \mathbf{S}_N)$ does not OSP-implement $f$.

$\square$

## 2.4 Proof of Theorem 3

We split this proof into two parts.

**Proposition 1.** *If* $(G, \mathbf{S}_N)$ *OSP-implements* $f_y$*, then* $\mathcal{P}(G, \mathbf{S}_N)$ *is a personal-clock auction.*

*Proof.* Throughout this proof, we use the following notation: Given some type-strategy $\mathbf{S}_i$, $S_i^{\theta_i}$ denotes the strategy assigned to $\theta_i$ by $\mathbf{S}_i$.

Take any $(G, \mathbf{S}_N)$ that implements $(f_y, f_t)$. For any history $h$, we define

$$\Theta_h \equiv \{\theta_N \mid h \preceq z^G(h_\emptyset, (S_i^{\theta_i})_{i \in N}\} \tag{17}$$

$$\Theta_{h,i} \equiv \{\theta_i \mid \exists \theta_{-i} : (\theta_i, \theta_{-i}) \in \Theta_h\} \tag{18}$$

For information set $I_i$, we define

$$\Theta_{I_i} \equiv \cup_{h \in I_i} \Theta_h \tag{19}$$

$$\Theta_{I_i,i} \equiv \{\theta_i \mid \exists \theta_{-i} : (\theta_i, \theta_{-i}) \in \Theta_{I_i}\} \tag{20}$$

$$\Theta^1_{I_i,i} \equiv \{\theta_i \mid \exists \theta_{-i} : (\theta_i, \theta_{-i}) \in \Theta_{I_i} \text{ and } i \in f_y(\theta_i, \theta_{-i})\} \tag{21}$$

$$\Theta^0_{I_i,i} \equiv \{\theta_i \mid \exists \theta_{-i} : (\theta_i, \theta_{-i}) \in \Theta_{I_i} \text{ and } i \notin f_y(\theta_i, \theta_{-i})\} \tag{22}$$

Some observations about this construction:

1. Since player $i$'s type-strategy depends only on his own type, $\Theta_{I_i,i} = \Theta_{h,i}$ for all $h \in I_i$.
2. $\Theta_{I_i,i} = \Theta^1_{I_i,i} \cup \Theta^0_{I_i,i}$
3. Since SP requires $1_{i \in f_y(\theta)}$ weakly increasing in $\theta_i$, $\Theta^1_{I_i,i}$ dominates $\Theta^0_{I_i,i}$ in the strong set order.

**Lemma 1.** *Suppose* $(G, \mathbf{S}_N)$ *OSP-implements* $(f_y, f_t)$*, where* $G \equiv \langle H, \prec, A, \mathcal{A}, P, \delta_c, (\mathcal{I}_i)_{i \in N}, g \rangle$*. For all* $i$*, for all* $I_i \in \mathcal{I}_i$*, if:*

*1.* $\theta_i < \theta'_i$

*2.* $\theta_i \in \Theta^1_{I_i,i}$

*3.* $\theta'_i \in \Theta^0_{I_i,i}$

*then* $S_i^{\theta_i}(I_i) = S_i^{\theta'_i}(I_i)$.

*Equivalently, for any* $I_i$*, there exists* $a^*_{I_i}$ *such that for all* $\theta_i \in \Theta^1_{I_i,i} \cap \Theta^0_{I_i,i}$*,* $S_i^{\theta_i}(I_i) = a^*_{I_i}$.

Suppose not. Take $(i, I_i, \theta_i, \theta'_i)$ constituting a counterexample to Lemma 1. Since $\theta_i \in \Theta^1_{I_i,i}$, there exists $h \in I_i$ and $S_{-i}$ such that $i \in g_y(z^G(h, S_i^{\theta_i}, S_{-i}))$. Fix $t_i \equiv g_{t,i}(z^G(h, S_i^{\theta_i}, S_{-i}))$. Since $\theta'_i \in \Theta^0_{I_i,i}$, there exists $h' \in I_i$ and $S'_{-i}$ such that $i \notin g_y(z^G(h', S_i^{\theta'_i}, S'_{-i}))$. Fix $t'_i \equiv g_{t,i}(z^G(h', S_i^{\theta'_i}, S'_{-i}))$. Since $S_i^{\theta_i}(I_i) \neq S_i^{\theta'_i}(I_i)$ and $\theta_i \cup \theta'_i \subseteq \Theta_{I_i,i}$, $I_i \in \alpha(S_i^{\theta_i}, S_i^{\theta'_i})$. Thus, OSP requires that

$$u_i(\theta_i, h, S_i^{\theta_i}, S_{-i}) \geq u_i(\theta_i, h', S_i^{\theta'_i}, S'_{-i}) \tag{23}$$

which implies

$$\theta_i + t_i \geq t'_i \tag{24}$$

and

$$u_i(\theta'_i, h, S_i^{\theta_i}, S_{-i}) \leq u_i(\theta'_i, h', S_i^{\theta'_i}, S'_{-i}) \tag{25}$$

which implies

$$\theta'_i + t_i \leq t'_i \tag{26}$$

But $\theta'_i > \theta_i$, so

$$\theta'_i + t_i > t'_i \tag{27}$$

a contradiction. This proves Lemma 1.

**Lemma 2.** *Suppose* $(G, \mathbf{S}_N)$ *OSP-implements* $(f_y, f_t)$ *and* $\mathcal{P}(G, \mathbf{S}_N) = G$*. Take any* $I_i$ *such that* $\Theta^1_{I_i,i} \cap \Theta^0_{I_i,i} \neq \emptyset$*, and associated* $a^*_{I_i}$*.*

*1. If there exists* $\theta_i \in \Theta^0_{I_i,i}$ *such that* $S_i^{\theta_i}(I_i) \neq a^*_{I_i}$*, then there exists* $t^0_i$ *such that:*

   *(a) For all* $\theta_i \in \Theta^0_{I_i,i}$ *such that* $S_i^{\theta_i}(I_i) \neq a^*_{I_i}$*, for all* $h \in I_i$*, for all* $S_{-i}$*,* $g_{t,i}(z^G(h, S_i^{\theta_i}, S_{-i})) = t^0_i$*.*

(b) *For all* $\theta_i \in \Theta_{I_i,i}$ *such that* $S_i^{\theta_i}(I_i) = a^*_{I_i}$*, for all* $h \in I_i$*, for all* $S_{-i}$*, if* $i \notin g_y(z^G(h, S_i^{\theta_i}, S_{-i}))$*, then* $g_{t,i}(z^G(h, S_i^{\theta_i}, S_{-i})) = t_i^0$*.*

2. *If there exists* $\theta_i \in \Theta^1_{I_i,i}$ *such that* $S_i^{\theta_i}(I_i) \neq a^*_{I_i}$*, then there exists* $t_i^1$ *such that:*

   (a) *For all* $\theta_i \in \Theta^1_{I_i,i}$ *such that* $S_i^{\theta_i}(I_i) \neq a^*_{I_i}$*, for all* $h \in I_i$*, for all* $S_{-i}$*,* $g_{t,i}(z^G(h, S_i^{\theta_i}, S_{-i})) = t_i^1$*.*

   (b) *For all* $\theta_i \in \Theta_{I_i,i}$ *such that* $S_i^{\theta_i}(I_i) = a^*_{I_i}$*, for all* $h \in I_i$*, for all* $S_{-i}$*, if* $i \in g_y(z^G(h, S_i^{\theta_i}, S_{-i}))$*, then* $g_{t,i}(z^G(h, S_i^{\theta_i}, S_{-i})) = t_i^1$*.*

Take any type $\theta'_i \in \Theta^0_{I_i,i}$ such that $S_i^{\theta'_i}(I_i) \neq a^*_{I_i}$. Take any type $\theta''_i \in \Theta^0_{I_i,i}$ such that $S_i^{\theta''_i}(I_i) = a^*_{I_i}$. (By $\Theta^1_{I_i,i} \cap \Theta^0_{I_i,i} \neq \emptyset$ there exists at least one such type.) Notice that $I_i \in \alpha(S_i^{\theta'_i}, S_i^{\theta''_i})$.

By Lemma 1, $\theta'_i \notin \Theta^1_{I_i,i}$, and the game is pruned. Thus,

$$\forall h \in I_i : \forall S_{-i} : i \notin g_y(z^G(h, S_i^{\theta'_i}, S_{-i})) \tag{28}$$

Since $\theta''_i \in \Theta^0_{I_i,i}$,

$$\exists h \in I_i : \exists S_{-i} : i \notin g_y(z^G(h, S_i^{\theta''_i}, S_{-i})) \tag{29}$$

OSP requires that type $\theta'_i$ does not want to (inf-sup) deviate. Thus,

$$\begin{split} \inf_{h \in I_i, S_{-i}} g_{t,i}(z^G(h, S_i^{\theta'_i}, S_{-i})) \geq \\ \sup_{h \in I_i, S_{-i}} \{g_{t,i}(z^G(h, S_i^{\theta''_i}, S_{-i})) : i \notin g_y(z^G(h, S_i^{\theta''_i}, S_{-i}))\} \end{split} \tag{30}$$

OSP also requires that type $\theta''_i$ does not want to (inf-sup) deviate. This implies

$$\begin{split} \inf_{h \in I_i, S_{-i}} \{g_{t,i}(z^G(h, S_i^{\theta''_i}, S_{-i})) : i \notin g_y(z^G(h, S_i^{\theta''_i}, S_{-i}))\} \\ \geq \sup_{h \in I_i, S_{-i}} g_{t,i}(z^G(h, S_i^{\theta'_i}, S_{-i})) \end{split} \tag{31}$$

The RHS of Equation 30 is weakly greater than the LHS of Equation 31. The RHS of Equation 31 is weakly greater than the LHS of Equation 30. Consequently all four terms are equal. Moreover, this argument applies to every $\theta'_i \in \Theta^0_{I_i,i}$ such that $S_i^{\theta'_i}(I_i) \neq a^*_{I_i}$, and every $\theta''_i \in \Theta^0_{I_i,i}$ such that $S_i^{\theta''_i}(I_i) = a^*_{I_i}$. Since the game is pruned, $\theta''_i$ satisfies

(1b) iff $\theta_i'' \in \Theta^0_{I_i,i}$ and $S_i^{\theta_i''}(I_i) = a^*_{I_i}$. This proves part 1 of Lemma 2. Part 2 follows by symmetry; we omit the details since they involve only small changes to the above argument.

**Lemma 3.** *Suppose $(G, \mathbf{S}_N)$ OSP-implements $(f_y, f_t)$ and $\mathcal{P}(G, \mathbf{S}_N) = G$. Take any $I_i$ such that $\Theta^1_{I_i,i} \cap \Theta^0_{I_i,i} \neq \emptyset$, and associated $a^*_{I_i}$. Let $t_i^1$ and $t_i^0$ be defined as before.*

1. *If there exists $\theta_i \in \Theta^0_{I_i,i}$ such that $S_i^{\theta_i}(I_i) \neq a^*_{I_i}$, then for all $(h \in I_i, S_i, S_{-i})$, if $i \in g_y(z^G(h, S_i\, S_{-i}))$, then $g_{t,i}(z^G(h, S_i, S_{-i})) \leq t_i^0 - \sup\{\theta_i \in \Theta^0_{I_i,i} : S_i^{\theta_i}(I_i) \neq a^*_{I_i}\}$.*

2. *If there exists $\theta_i \in \Theta^1_{I_i,i}$ such that $S_i^{\theta_i}(I_i) \neq a^*_{I_i}$, then for all $(h \in I_i, S_i, S_{-i})$, if $i \notin g_y(z^G(h, S_i\, S_{-i}))$, then $g_{t,i}(z^G(h, S_i, S_{-i})) \leq \inf\{\theta_i \in \Theta^1_{I_i,i} : S_i^{\theta_i}(I_i) \neq a^*_{I_i}\} + t_i^1$.*

Suppose that part 1 of Lemma 3 does not hold. Fix $(h \in I_i, S_i, S_{-i})$ such that $i \in g_y(z^G(h, S_i\, S_{-i}))$ and $g_{t,i}(z^G(h, S_i, S_{-i})) > t_i^0 - \sup\{\theta_i \in \Theta^0_{I_i,i} : S_i^{\theta_i}(I_i) \neq a^*_{I_i}\}$. Since $G$ is pruned, we can find some $\theta_i' \in \Theta_{I_i,i}$ such that for every $\tilde{I}_i \in \{I_i' \in \mathcal{I}_i : I_i \in \psi(I_i')\}$, $S_i^{\theta_i'}(\tilde{I}_i) = S_i(\tilde{I}_i)$. Fix that $\theta_i'$.

Fix $\theta_i'' \in \Theta^0_{I_i,i}$ such that $S_i^{\theta_i}(I_i) \neq a^*_{I_i}$ and $\theta_i'' \geq \sup\{\theta_i \in \Theta^0_{I_i,i} : S_i^{\theta_i}(I_i) \neq a^*_{I_i}\} - \epsilon$. Since $G$ is pruned and $\theta_i'' \notin \Theta^1_{I_i,i}$ (by Lemma 1), it must be that $S_i^{\theta_i''}(I_i) \neq S_i^{\theta_i'}(I_i)$.

By construction, $I_i \in \alpha(S_i^{\theta_i'}, S_i^{\theta_i''})$.

OSP requires that, for all $h'' \in I_i, S''_{-i}$:

$$u_i(\theta_i'', h'', S_i^{\theta_i''}, S''_{-i}) \geq u_i(\theta_i'', h, S_i^{\theta_i'}, S_{-i}) \tag{32}$$

which entails

$$t_i^0 \geq \theta_i'' + g_{t,i}(z^G(h, S_i, S_{-i})) \tag{33}$$

which entails

$$t_i^0 - \sup\{\theta_i \in \Theta^0_{I_i,i} : S_i^{\theta_i}(I_i) \neq a^*_{I_i}\} + \epsilon \geq g_{t,i}(z^G(h, S_i, S_{-i})) \tag{34}$$

But, by hypothesis,

$$t_i^0 - \sup\{\theta_i \in \Theta^0_{I_i,i} : S_i^{\theta_i}(I_i) \neq a^*_{I_i}\} < g_{t,i}(z^G(h, S_i, S_{-i})) \tag{35}$$

Since this argument holds for all $\epsilon > 0$, we can pick $\epsilon$ small enough to create a contradiction. This proves part 1 of Lemma 3. Part 2 follows by symmetry.

**Lemma 4.** *Suppose $(G, \mathbf{S}_N)$ OSP-implements $(f_y, f_t)$ and $\mathcal{P}(G, \mathbf{S}_N) = G$. Take any $I_i$ such that $|\Theta^1_{I_i,i} \cap \Theta^0_{I_i,i}| > 1$ and associated $a^*_{I_i}$.*

1. *If there exists* $\theta_i \in \Theta^0_{I_i,i}$ *such that* $S_i^{\theta_i}(I_i) \neq a^*_{I_i}$*, then for all* $\theta'_i \in \Theta^1_{I_i,i}$*,* $S_i^{\theta'_i}(I_i) = a^*_{I_i}$*.*

2. *(Equivalently) If there exists* $\theta_i \in \Theta^1_{I_i,i}$ *such that* $S_i^{\theta_i}(I_i) \neq a^*_{I_i}$*, then for all* $\theta'_i \in \Theta^0_{I_i,i}$*,* $S_i^{\theta'_i}(I_i) = a^*_{I_i}$*.*

Suppose Part 1 of Lemma 4 does not hold. Fix $I_i$, and choose $\theta'_i < \theta''_i$ such that $\{\theta'_i\} \cup \{\theta''_i\} \subseteq \Theta^1_{I_i,i} \cap \Theta^0_{I_i,i}$. Fix $\theta'''_i \in \Theta^1_{I_i,i}$ such that $S_i^{\theta'''_i}(I_i) \neq a^*_{I_i}$. By Lemma 1, if $\theta'''_i \in \Theta^0_{I_i,i}$, then $S_i^{\theta'''_i}(I_i) = a^*_{I_i}$, a contradiction. Thus, $\theta'''_i \in \Theta^1_{I_i,i} \setminus \Theta^0_{I_i,i}$, and since $\Theta^1_{I_i,i}$ dominates $\Theta^0_{I_i,i}$ in the strong set order, $\theta''_i < \theta'''_i$.

Since $\theta'_i \in \Theta^1_{I_i,i}$, there exists $h' \in I_i$ and $\theta'_{-i}$ such that $(\theta'_i, \theta'_{-i}) \in \Theta_{I_i}$ and $i \in g_y(z^G(h', S_i^{\theta'_i}, S_{-i}^{\theta'_{-i}}))$. By Lemma 2, there exists $a_i \in A_i(I_i)$ such that $a_i \neq a^*_{I_i}$ and choosing $a_i$ ensures $i \notin y$ and $t_i = t_i^0$. Thus, by $G$ SP

$$\theta'_i + g_{t,i}(z^G(h', S_i^{\theta'_i}, S_{-i}^{\theta'_{-i}})) \geq t_i^0 \tag{36}$$

By $\theta''_i \in \Theta^0_{I_i,i}$, there exists $h'' \in I_i$ and $\theta''_{-i}$ such that $i \notin g_y(z^G(h'', S_i^{\theta''_i}, S_{-i}^{\theta''_{-i}}))$. By Lemma 2

$$g_{t,i}(z^G(h'', S_i^{\theta''_i}, S_{-i}^{\theta''_{-i}})) = t_i^0 \tag{37}$$

By $G$ SP, $i \in g_y(z^G(h', S_i^{\theta'''_i}, S_{-i}^{\theta'_{-i}}))$ and $g_{t,i}(z^G(h', S_i^{\theta'''_i}, S_{-i}^{\theta'_{-i}})) = g_{t,i}(z^G(h', S_i^{\theta'_i}, S_{-i}^{\theta'_{-i}}))$.

Notice that $I_i \in \alpha(S_i^{\theta''_i}, S_i^{\theta'''_i})$. Thus, OSP requires that $\theta''_i$ does not want to (inf-sup) deviate to $\theta'''_i$'s strategy, which entails:

$$g_{t,i}(z^G(h'', S_i^{\theta''_i}, S_{-i}^{\theta''_{-i}})) \geq \theta''_i + g_{t,i}(z^G(h', S_i^{\theta'''_i}, S_{-i}^{\theta'_{-i}})) \tag{38}$$

$$\begin{aligned} t_i^0 &\geq \theta''_i + g_{t,i}(z^G(h', S_i^{\theta'''_i}, S_{-i}^{\theta'_{-i}})) \\ &> \theta'_i + g_{t,i}(z^G(h', S_i^{\theta'_i}, S_{-i}^{\theta'_{-i}})) \end{aligned} \tag{39}$$

which contradicts Equation 36.

Part 2 is the contrapositive of Part 1. This proves Lemma 4.

**Lemma 5.** *Suppose* $(G, \mathbf{S}_N)$ *OSP-implements* $(f_y, f_t)$ *and* $\mathcal{P}(G, \mathbf{S}_N) = G$*.*

*For all* $I_i$*, if* $|\Theta^1_{I_i,i} \cap \Theta^0_{I_i,i}| \leq 1$ *and* $|A(I_i)| \geq 2$*, then there exists* $t_i^1$ *and* $t_i^0$ *such that:*

1. *For all* $\theta_i \in \Theta_{I_i,i}$*,* $h \in I_i$*,* $S_{-i}$*:*

   (a) *If* $i \notin g_y(z^G(h, S_i^{\theta_i}, S_{-i}))$ *then* $g_{t,i}(z^G(h, S_i^{\theta_i}, S_{-i})) = t_i^0$

(b) *If* $i \in g_y(z^G(h, S_i^{\theta_i}, S_{-i}))$ *then* $g_{t,i}(z^G(h, S_i^{\theta_i}, S_{-i})) = t_i^1$

2. *If* $|\Theta^1_{I_i,i}| > 0$ *and* $|\Theta^0_{I_i,i}| > 0$*, then* $t_i^1 = -\inf\{\theta_i \in \Theta^1_{I_i,i}\} + t_i^0$

By $G$ pruned, $\Theta_{I_i,i} \neq \emptyset$. By the Green-Laffont-Holmström Theorem,

$$f_{t,i}(\theta_i, \theta_{-i}) = -1_{i \in f_y(\theta)} \inf\{\theta_i' \mid i \in f_y(\theta_i', \theta_{-i})\} + r_i(\theta_{-i}) \tag{40}$$

Consider the case where $\Theta^1_{I_i,i} = \emptyset$. Pick some $\theta_i' \in \Theta^0_{I_i,i}$ and some $h' \in I_i, S_{-i}'$. Fix $t_i^0 \equiv g_{t,i}(z^G(h', S_i^{\theta_i'}, S_{-i}'))$. Suppose there exists some $(\theta_i'', \theta_{-i}'') \in \Theta_{I_i}$ such that $f_{t,i}(\theta_i'', \theta_{-i}'') = t_i^{0''} \neq t_i^0$. Pick $h'' \in I_i$ such that $h'' \preceq z^G(\emptyset, S_i^{\theta_i''}, S_{-i}^{\theta_{-i}''})$. By Equation 40, for all $\theta_i \in \Theta_{I_i,i}$, $f_{t,i}(\theta_i, \theta_{-i}'') = t_i^{0''}$. By $G$ pruned and $|A(I_i)| \geq 2$, we can pick $\theta_i''' \in \Theta^0_{I_i,i}$ such that $S_i^{\theta_i'''}(I_i) \neq S_i^{\theta_i'}(I_i)$. Notice that $I_i \in \alpha(S_i^{\theta_i'''}, S_i^{\theta_i'})$. If $t_i^{0''} > t_i^0$, then

$$u_i(\theta_i', h', S_i^{\theta_i'}, S_{-i}') = t_i^0 < t_i^{0''} = u_i(\theta_i', h'', S_i^{\theta_i'''}, S_{-i}^{\theta_{-i}''}) \tag{41}$$

so $S_i^{\theta_i'}$ is not obviously dominant for $(i, \theta_i')$. If $t_i^{0''} < t_i^0$, then

$$u_i(\theta_i''', h', S_i^{\theta_i'}, S_{-i}') = t_i^0 > t_i^{0''} = u_i(\theta_i''', h'', S_i^{\theta_i'''}, S_{-i}^{\theta_{-i}''}) \tag{42}$$

so $S_i^{\theta_i'''}$ is not obviously dominant for $(i, \theta_i''')$. By contradiction, this proves Lemma 5 for this case. A symmetric argument proves Lemma 5 for the case where $\Theta^0_{I_i,i} = \emptyset$.

Note that, if Lemma 5 holds at some information set $I_i$, it holds at all information sets $I_i'$ that follow $I_i$. Thus, we need only consider some earliest information set $I_i^*$ at which $|\Theta^1_{I_i^*,i} \cap \Theta^0_{I_i^*,i}| \leq 1$ and $|A(I_i^*)| \geq 2$.

Now we consider the case where $\Theta^1_{I_i^*,i} \neq \emptyset$ and $\Theta^0_{I_i^*,i} \neq \emptyset$.

At every prior information set $I_i$ prior to $I_i^*$, $|\Theta^1_{I_i,i} \cap \Theta^0_{I_i,i}| > 1$. Since $\Theta^1_{I_i^*,i} \neq \emptyset$ and $\Theta^0_{I_i^*,i} \neq \emptyset$, by Lemma 4, $I_i^*$ is reached by some interval of types all taking the same action. Thus $\sup\{\theta_i \in \Theta^0_{I_i^*,i}\} = \inf\{\theta_i \in \Theta^1_{I_i^*,i}\}$.

Fix $\hat{\theta}_i \in \Theta^0_{I_i^*,i}$ such that $\hat{\theta}_i \geq \sup\{\theta_i \in \Theta^0_{I_i^*,i}\} - \epsilon$. Choose corresponding $\hat{h} \in I_i^*$ and $\hat{\theta}_{-i} \in \Theta_{I_i^*,-i}$ such that $i \notin g_y(z^G(\hat{h}, S_i^{\hat{\theta}_i}, S_{-i}^{\hat{\theta}_{-i}}))$. Define $t_i^0 \equiv g_{t,i}(z^G(\hat{h}, S_i^{\hat{\theta}_i}, S_{-i}^{\hat{\theta}_{-i}}))$.

Fix $\check{\theta}_i \in \Theta^1_{I_i^*,i}$ such that $\check{\theta}_i \leq \inf\{\theta_i \in \Theta^1_{I_i^*,i}\} + \epsilon$. Choose corresponding $\check{h} \in I_i^*$ and $\check{\theta}_{-i} \in \Theta_{I_i^*,-i}$ such that $i \in g_y(z^G(\check{h}, S_i^{\check{\theta}_i}, S_{-i}^{\check{\theta}_{-i}}))$. Define $t_i^1 \equiv g_{t,i}(z^G(\check{h}, S_i^{\check{\theta}_i}, S_{-i}^{\check{\theta}_{-i}}))$.

Suppose there exists some $(\theta_i', \theta_{-i}') \in \Theta_{I_i^*}$ such that $i \notin f_y(\theta_i', \theta_{-i}')$ and $f_{t,i}(\theta_i', \theta_{-i}') = t_i^{0'} \neq t_i^0$. Since $\sup\{\theta_i \in \Theta^0_{I_i^*,i}\} = \inf\{\theta_i \in \Theta^1_{I_i^*,i}\}$, it follows that for all $\theta_{-i} \in \Theta_{I_i^*,-i}$, $\inf\{\theta_i : i \in f_y(\theta_i, \theta_{-i})\} = \inf\{\theta_i \in \Theta^1_{I_i^*,i}\}$. Thus, by Equation 40, for all $\theta_i \in \Theta_{I_i^*,i}$: $f_{t,i}(\theta_i, \theta_{-i}') = -1_{i \in f_y(\theta_i, \theta_{-i}')} \inf\{\theta_i \in \Theta^1_{I_i^*,i}\} + t_i^{0'}$. Fix $h' \in I_i^*$ such that $h' \preceq z^G(\emptyset, S_i^{\theta_i'}, S_{-i}^{\theta_{-i}'})$.

By $A(I_i) \geq 2$, we can pick some $\theta_i'' \in \Theta_{I_i^*,i}$ such that $S_i^{\theta_i''}(I_i) \neq S_i^{\hat{\theta}_i}(I_i)$. Notice that $I_i^* \in \alpha(S_i^{\theta_i''}, S_i^{\hat{\theta}_i})$. Either $\theta_i'' \in \Theta^0_{I_i^*}$ or $\theta_i'' \in \Theta^1_{I_i^*} \setminus \Theta^0_{I_i^*}$. Suppose $\theta_i'' \in \Theta^0_{I_i^*}$. Suppose $t_i^{0\prime} > t_i^0$. By OSP,

$$u_i(\hat{\theta}_i, \hat{h}, S_i^{\hat{\theta}_i}, S_{-i}^{\hat{\theta}_{-i}}) \geq u_i(\hat{\theta}_i, h', S_i^{\theta_i''}, S_{-i}^{\theta_{-i}'}) \tag{43}$$

which entails

$$\begin{aligned} t_i^0 &\geq t_i^{0\prime} + 1_{i \in f_y(\theta_i'', \theta_{-i}')}(\hat{\theta}_i - \inf\{\theta_i \in \Theta^1_{I_i^*,i}\}) \\ &\geq t_i^{0\prime} + 1_{i \in f_y(\theta_i'', \theta_{-i}')}(-\epsilon) \\ &\geq t_i^{0\prime} - \epsilon \end{aligned} \tag{44}$$

and we can pick $\epsilon$ small enough to constitute a contradiction. Suppose $t_i^{0\prime} < t_i^0$. By OSP

$$u_i(\theta_i'', \hat{h}, S_i^{\hat{\theta}_i}, S_{-i}^{\hat{\theta}_{-i}}) \leq u_i(\theta_i'', h', S_i^{\theta_i''}, S_{-i}^{\theta_{-i}'}) \tag{45}$$

which entails

$$\begin{aligned} t_i^0 &\leq t_i^{0\prime} + 1_{i \in f_y(\theta_i'', \theta_{-i}')}(\theta_i'' - \inf\{\theta_i \in \Theta^1_{I_i^*,i}\}) \\ &= t_i^{0\prime} + 1_{i \in f_y(\theta_i'', \theta_{-i}')}(\theta_i'' - \sup\{\theta_i \in \Theta^0_{I_i^*,i}\}) \leq t_i^{0\prime} \end{aligned} \tag{46}$$

which is a contradiction.

The case that remains is $\theta_i'' \in \Theta^1_{I_i^*} \setminus \Theta^0_{I_i^*}$. Then $i \in f_y(\theta_i'', \theta_{-i}')$ and $f_{t,i}(\theta_i'', \theta_{-i}') = -\inf\{\theta_i \in \Theta^1_{I_i^*,i}\} + t_i^{0\prime}$. Suppose $t_i^{0\prime} > t_i^0$. OSP requires:

$$u_i(\hat{\theta}_i, \hat{h}, S_i^{\hat{\theta}_i}, S_{-i}^{\hat{\theta}_{-i}}) \geq u_i(\hat{\theta}_i, h', S_i^{\theta_i''}, S_{-i}^{\theta_{-i}'}) \tag{47}$$

which entails

$$\begin{aligned} t_i^0 &\geq \hat{\theta}_i - \inf\{\theta_i \in \Theta^1_{I_i^*,i}\} + t_i^{0\prime} \\ &\geq t_i^{0\prime} - \epsilon \end{aligned} \tag{48}$$

and we can pick $\epsilon$ small enough to constitute a contradiction.

Suppose $t_i^{0\prime} < t_i^0$. Since $S_i^{\theta_i''}(I_i) \neq S_i^{\hat{\theta}_i}(I_i)$, either $S_i^{\check{\theta}_i}(I_i) \neq S_i^{\hat{\theta}_i}(I_i)$ or $S_i^{\check{\theta}_i}(I_i) \neq S_i^{\theta_i''}(I_i)$. Moreover, $f_{t,i}(\check{\theta}_i, \theta_i') = -1_{i \in f_y(\check{\theta}_i, \theta_{-i}')} \inf\{\theta_i \in \Theta^1_{I_i^*,i}\} + t_i^{0\prime}$.

Suppose $S_i^{\check{\theta}_i}(I_i) \neq S_i^{\hat{\theta}_i}(I_i)$. OSP requires:

$$u_i(\check{\theta}_i, h', S_i^{\check{\theta}_i}, S_{-i}^{\theta'_{-i}}) \geq u_i(\check{\theta}_i, \hat{h}, S_i^{\hat{\theta}_i}, S_{-i}^{\hat{\theta}_{-i}}) \tag{49}$$

which entails

$$1_{i \in f_y(\check{\theta}_i, \theta'_{-i})}(\check{\theta}_i - \inf\{\theta_i \in \Theta^1_{I_i^*, i}\}) + t_i^{0\prime} \geq t_i^0 \tag{50}$$

which entails

$$1_{i \in f_y(\check{\theta}_i, \theta'_{-i})}\epsilon + t_i^{0\prime} \geq t_i^0 \tag{51}$$

and we can pick $\epsilon$ small enough to yield a contradiction. Suppose $S_i^{\check{\theta}_i}(I_i) \neq S_i^{\theta''_i}(I_i)$. By Equation 40, $f_{t,i}(\check{\theta}_i, \theta'_{-i}) = -1_{i \in f_y(\check{\theta}_i, \theta'_{-i})} \inf\{\theta_i \in \Theta^1_{I_i^*, i}\} + t_i^{0\prime}$, and $f_{t,i}(\theta''_i, \hat{\theta}_{-i}) = -\inf\{\theta_i \in \Theta^1_{I_i^*, i}\} + t_i^0$. OSP requires

$$u_i(\check{\theta}_i, h', S_i^{\check{\theta}_i}, S_{-i}^{\theta'_{-i}}) \geq u_i(\check{\theta}_i, \hat{h}, S_i^{\theta''_i}, S_{-i}^{\hat{\theta}_{-i}}) \tag{52}$$

which entails

$$1_{i \in f_y(\check{\theta}_i, \theta'_{-i})}(\check{\theta}_i - \inf\{\theta_i \in \Theta^1_{I_i^*, i}\}) + t_i^{0\prime} \geq (\check{\theta}_i - \inf\{\theta_i \in \Theta^1_{I_i^*, i}\}) + t_i^0 \tag{53}$$

which entails

$$1_{i \in f_y(\check{\theta}_i, \theta'_{-i})}\epsilon + t_i^{0\prime} \geq \epsilon + t_i^0 \tag{54}$$

which entails

$$t_i^{0\prime} \geq t_i^0 \tag{55}$$

a contradiction. By the above argument, for all $I_i$ satisfying the assumptions of Lemma 5, there is a unique transfer $t_i^0$ for all terminal histories $z$ passing through $I_i$ such that $i \notin g_y(z)$. Equation 40 thus implies that there is a unique transfer $t_i^1$ for all terminal histories $z$ passing through $I_i$ such that $i \in g_y(z)$. Moreover, $t_i^1 = -\inf\{\theta_i \in \Theta^1_{I_i, i}\} + t_i^0$. This proves Lemma 5.

Now to bring this all together. We leave showing parts (1.d.iv) and (2.d.iv) of Definition 15 to the last. Take any $(\hat{G}, \hat{\mathbf{S}}_N)$ that OSP-implements $(f_y, f_t)$. Define $G \equiv \mathcal{P}(\hat{G}, \hat{\mathbf{S}}_N)$ and $\mathbf{S}_N$ as $\hat{\mathbf{S}}_N$ restricted to $G$. By Proposition 2, $(G, \mathbf{S}_N)$ OSP-implements $(f_y, f_t)$.

We now characterize $(G, \mathbf{S}_N)$. For any player $i$, consider any information set $I_i^*$ such that $|A(I_i^*)| \geq 2$, and for all prior information sets $I_i' \prec I_i^*$, $|A(I_i')| = 1$. By Lemma 1, there is a unique action $a^*_{I_i^*}$ taken by all types in $\Theta^1_{I_i^*,i} \cap \Theta^0_{I_i^*,i}$.

Either $|\Theta^1_{I_i^*,i} \cap \Theta^0_{I_i^*,i}| > 1$ or $|\Theta^1_{I_i^*,i} \cap \Theta^0_{I_i^*,i}| \leq 1$.

If $|\Theta^1_{I_i^*,i} \cap \Theta^0_{I_i^*,i}| > 1$, then by Lemma 4, $\tilde{G}$ pruned and $|A(I_i)| \geq 2$,

1. EITHER: There exists $\theta_i \in \Theta^0_{I_i^*,i}$ such that $S_i^{\theta_i}(I_i^*) \neq a^*_{I_i^*}$, and for all $\theta_i' \in \Theta^1_{I_i^*,i}$, $S_i^{\theta_i'}(I_i^*) = a^*_{I_i^*}$.

2. OR: There exists $\theta_i \in \Theta^1_{I_i^*,i}$ such that $S_i^{\theta_i}(I_i^*) \neq a^*_{I_i^*}$, and for all $\theta_i' \in \Theta^0_{I_i^*,i}$, $S_i^{\theta_i'}(I_i^*) = a^*_{I_i^*}$.

In the first case, then by Lemma 2, there is some $\bar{t}_i$ such that, for all $(S_i, S_{-i})$, for all $h \in I_i$, if $i \notin g_y(z^G(h, S_i, S_{-i})$, then $g_{t,i}(z^G(h, S_i, S_{-i}) = \bar{t}_i$. Moreover, we can define a 'going transfer' at all information sets $I_i'$ such that $I_i^* \preceq I_i'$:

$$\tilde{t}_i(I_i') \equiv \min_{I_i'' : I_i^* \preceq I_i'' \preceq I_i'} [\bar{t}_i - \sup\{\theta_i \in \Theta^0_{I_i'',i} : S_i^{\theta_i}(I_i'') \neq a^*_{I_i''}\}] \tag{56}$$

Notice that this function falls monotonically as we move along the game tree; for any $I_i', I_i''$ such that $I_i' \preceq I_i''$, $\tilde{t}_i(I_i') \geq \tilde{t}_i(I_i'')$. Moreover, by construction, at any $I_i', I_i''$ such that $I_i' \prec I_i''$, $\tilde{t}_i(I_i') > \tilde{t}_i(I_i'')$, and there does not exist $I_i'''$ such that $I_i' \prec I_i''' \prec I_i''$, then there exists $a \in A(I_i'')$ that yields $i \notin y$, and by Lemma 2 this yields transfer $\bar{t}_i$. We define $A^q$ to include all such quitting actions; i.e. $A^q$ is the set of all actions such that:

1. $a \in I_i$ for some $I_i \in \mathcal{I}_i$

2. For all $z$ such that $a \in \psi_i(z)$: $i \notin g_y(z)$ and $g_{t,i}(z) = \bar{t}_i$

Lemma 3 and SP together imply that, at any terminal history $z$, if $i \in g_y(z)$, then

$$g_{t,i}(z) = \inf_{I_i : I_i^* \preceq I_i \prec z} \tilde{t}_i(I_i) \tag{57}$$

This holds because, if $g_{t,i}(z) < \inf_{I_i : I_i^* \preceq I_i \prec z} \tilde{t}_i(I_i)$, then type $\theta_i$ such that $\bar{t}_i - \inf_{I_i : I_i^* \preceq I_i \prec z} \tilde{t}_i(I_i)) < \theta_i < \bar{t}_i - g_{t,i}(z)$ could profitably deviate to play $a \in A^0$ at information set $I_i^*$.

In the second case, then by Lemma 2, there is some $\bar{t}_i$ such that, for all $(S_i, S_{-i})$, for all $h \in I_i^*$, if $i \in g_y(z^G(h, S_i, S_{-i})$, then $g_{t,i}(z^G(h, S_i, S_{-i}) = \bar{t}_i$. Moreover, we can define a 'going transfer' at all information sets $I_i'$ such that $I_i^* \preceq I_i'$):

$$\tilde{t}_i(I_i') \equiv \min_{I_i'' : I_i^* \preceq I_i'' \preceq I_i'} [\bar{t}_i + \inf\{\theta_i \in \Theta^0_{I_i'',i} : S_i^{\theta_i}(I_i'') \neq a^*_{I_i''}\}] \tag{58}$$

We define $A^1$ symmetrically for this second case.

Part (1.d.iii) and (and its analog in Clause 2) of Definition 15 follow from Lemma 4. The above constructions suffice to prove Theorem 3 for cases where $|\Theta^1_{I_i^*,i} \cap \Theta^0_{I_i^*,i}| > 1$. Cases where $|\Theta^1_{I_i^*,i} \cap \Theta^0_{I_i^*,i}| \leq 1$ are dealt with by Lemma 5.

Now for the last piece: We prove that parts (1.d.iv) and (and its analog in Clause 2) of Definition 15 hold. The proof of part (1.d.iv) is as follows: Suppose we are facing Clause 1 of Definition 15, and for some $I_i$, $|A(I'_i) \setminus A^0| > 1$. By part (1.d.iii), we know that the going transfer $\tilde{t}_i$ can fall no further. Since $G$ is pruned and $|A(I'_i) \setminus A^0| > 1$, there exist two distinct types of $i$, $\theta_i, \theta'_i \in \Theta_{I'_i,i}$, who do not quit at $I'_i$, and take different actions. Since neither quits at $I'_i$ and the going transfer falls no further, there exist $\theta_{-i}, \theta'_{-i} \in \Theta_{I'_i,-i}$ such that $i \in f_y(\theta_i, \theta_{-i})$ and $i \in f_y(\theta'_i, \theta'_{-i})$. So there exist $(h \in I'_i, S_{-i})$ and $(h' \in I'_i, S'_{-i})$ such that

$$i \in g_y(h, S_i^{\theta_i}, S_{-i}) \tag{59}$$

$$i \in g_y(h', S_i^{\theta'_i}, S'_{-i}) \tag{60}$$

$$g_{t,i}(h, S_i^{\theta_i}, S_{-i}) = g_{t,i}(h', S_i^{\theta'_i}, S'_{-i}) = \tilde{t}_i(I'_i) \tag{61}$$

WLOG suppose $\theta_i < \theta'_i$. Suppose that there does not exist $a \in A(I'_i)$ such that, for all $z$ such that $a \in \psi_i(z)$, $i \in g_y(z)$. Then there must exist $(h'' \in I'_i, S''_{-i})$ such that

$$i \notin g_y(h'', S_i^{\theta'_i}, S''_{-i}) \tag{62}$$

$$g_{t,i}(h'', S_i^{\theta'_i}, S''_{-i}) = \bar{t}_i \tag{63}$$

Note that $I'_i \in \alpha(S_i^{\theta_i}, S_i^{\theta'_i})$. But then $S_i^{\theta'_i}$ is not obviously dominant, a contradiction, since

$$\begin{aligned} u_i^{\tilde{G}}(h'', S_i^{\theta'_i}, S''_{-i}, \theta'_i) &= \bar{t}_i \leq \theta_i + \tilde{t}_i(I'_i) \\ &< \theta'_i + \tilde{t}_i(I'_i) = u_i^{\tilde{G}}(h, S_i^{\theta_i}, S_{-i}, \theta'_i) \end{aligned} \tag{64}$$

(The first inequality holds because of type $\theta_i$'s incentive constraint.) This shows that part (1.d.iv) of Definition 15 holds. Its analog in Clause 2 is proved symmetrically.

□

**Proposition 2.** *If $G$ is a personal-clock auction, then there exist $\mathbf{S}_N$ and $f_y$ such that $(G, \mathbf{S}_N)$ OSP-implements $f_y$.*

*Proof.* Take any personal-clock auction $G$. For any $i$ and any $\theta_i$, the following strategy $S_i$ is obviously dominant:

1. If $i$ encounters an information set consistent with Clause 1 of Definition 15, then, from that point forward:

   (a) If $\theta_i + \tilde{t}_i(I_i) > \bar{t}_i$ and there exists $a \in A(I_i) \setminus A^q$, play $a \in A(I_i) \setminus A^q$.

      i. If $|A(I_i) \setminus A^q| > 1$, then play $a \in A(I_i')$ such that: For all $z$ such that $a \in \psi_i(z)$: $i \in g_y(z)$.

   (b) Else play some $a \in A^q$.[3]

2. If $i$ encounters an information set consistent with Clause 2 of Definition 15, then, from that point forward:

   (a) If $\theta_i + \bar{t}_i < \tilde{t}_i(I_i)$ and there exists $a \in A(I_i) \setminus A^q$, play $a \in A(I_i) \setminus A^q$.

      i. If $|A(I_i) \setminus A^q| > 1$, then play $a \in A(I_i')$ such that: For all $z$ such that $a \in \psi_i(z)$: $i \notin g_y(z)$.

   (b) Else play some $a \in A^q$.

The above strategy is well-defined for any agent in any personal-clock auction, by inspection of Definition 15.

Consider any deviating strategy $S_i'$. At any earliest point of departure, the agent will have encountered an information set consistent with either Clause 1 or Clause 2 of Definition 15. Suppose that the agent has encountered an information set covered by Clause 1.

Take some earliest point of departure $I_i \in \alpha(S_i, S_i')$. Notice that, by (1.a) of Definition 15, *no matter what strategy $i$ plays*, conditional on reaching $I_i$, either agent $i$ is not in the allocation and receives $\bar{t}_i$, or agent $i$ is in the allocation and receives a transfer $\hat{t}_i \leq \tilde{t}_i(I_i)$.

Suppose $\theta_i + \tilde{t}_i(I_i) > \bar{t}_i$. Note that under $S_i$, conditional on reaching $I_i$, the agent either is not in the allocation and receives $\bar{t}_i$, or is in the allocation and receives a transfer strictly above $\bar{t}_i - \theta_i$. If $S_i'(I_i) \in A_q$ (*i.e.* if agent $i$ quits), then the best outcome under $S_i'$ is no better than the worst outcome under $S_i$. If $S_i'(I_i) \notin A_q$, then, since $S_i'(I_i) \neq S_i(I_i)$, $|A(I_i) \setminus A^q| > 1$. Then, by (1.d.iii) of Definition 15, $\tilde{t}_i$ will fall no further. So $S_i(I_i)$ guarantees that $i$ will be in the allocation and receive transfer $\tilde{t}_i(I_i)$. But, by (1.a) of

[3] If $|A^q \cap A(I_i)| > 1$, the agent chooses deterministically but arbitrarily.

Definition 15, the best possible outcome under $S_i'$ conditional on reaching $I_i$ is no better, so the obvious dominance inequality holds.

Suppose $\theta_i + \tilde{t}_i(I_i) \leq \bar{t}_i$. Then, under $S_i$, conditional on reaching $I_i$, agent $i$ is not in the allocation and has transfer $\bar{t}_i$.[4] However, under $S_i'$, either the outcome is the same, or agent $i$ is in the allocation for some transfer $\hat{t}_i \leq \tilde{t}_i(I_i) \leq \bar{t}_i - \theta_i$. Thus, the best possible outcome under $S_i'$ is no better than the worst possible outcome under $S_i$, and the obvious dominance inequality holds.

The argument proceeds symmetrically for Clause 2.

Notice that the above strategies result in *some* allocation and *some* payments, as a function of the type profile. We define these to be $(f_y, f_t)$, such that $G$ OSP-implements $(f_y, f_t)$.

□

# 3 Alternative Empirical Specifications

Here we report alternative empirical specifications for the experiment.

A natural measure of errors would be to take the sum, for $k = 1, 2, 3, 4$, of the absolute difference between the *kth* highest bid and the *kth* highest value. However we do not observe the highest bid under AC, and we often do not observe the highest bid under AC+X. We could instead take the sum for $k = 2, 3, 4$ of the absolute difference between the *kth* highest bid and the *kth* highest value, averaged as before in five-round blocks. Table 2 reports the results.

Another measure of errors would be to take the sum of the absolute difference between each bidder's bid and *that bidder's* value, dropping all highest bidders for symmetry. Table 3 reports the results.

Table 4 reports the results of Table 4 in the main text, except that the $p$-values are calculated using the Wilcoxon rank-sum test.

29.0% of rank-order lists are incorrect under SP-RSD. 2.6% of choices are incorrect under OSP-RSD. However, this is not a fair comparison; rank-order lists mechanically allow us to spot more errors than single choices. To compare like with like, we compute the proportion of incorrect choices we would have observed, if subjects played OSP-RSD as though they were implementing the submitted rank-order lists for SP-RSD. This is a cautious measure; it counts errors under SP-RSD only if they alter the outcome. Table 5 reports the results.

[4] By (1.d.ii) of Definition 15, either $i$ will have quit in the past, or will have an opportunity to quit now, which he exercises.

Table 2: mean(sum(abs($k$th bid - $k$th value))), for $k = 2, 3, 4$

| Format | Rounds | SP | OSP | $p$-value |
|---|---|---|---|---|
| Auction | 1-5 | 32.63 | 9.89 | $< .001$ |
| | | (4.64) | (1.89) | |
| | 6-10 | 16.28 | 5.53 | .001 |
| | | (2.73) | (0.91) | |
| +X Auction | 1-5 | 17.04 | 6.18 | .011 |
| | | (3.70) | (1.06) | |
| | 6-10 | 14.21 | 4.74 | .022 |
| | | (3.70) | (0.75) | |

For each auction, we sum the absolute differences between the $k$th bid and the $k$th value, for $k = 2, 3, 4$. We then take the mean of this over each 5-round block. We then compute standard errors counting each group's 5-round mean as a single observation. (18 observations per cell.) $p$-values are computed using a two-sample $t$-test, allowing for unequal variances.

Table 3: mean(sum(abs($i$'s bid - $i$'s value))), dropping highest bidders

| Format | Rounds | SP | OSP | $p$-value |
|---|---|---|---|---|
| Auction | 1-5 | 35.13 | 10.18 | $< .001$ |
| | | (5.20) | (1.88) | |
| | 6-10 | 15.46 | 4.89 | .002 |
| | | (2.85) | (0.72) | |
| +X Auction | 1-5 | 17.88 | 5.58 | .009 |
| | | (4.10) | (1.01) | |
| | 6-10 | 14.20 | 4.64 | .022 |
| | | (3.72) | (0.83) | |

For each auction, we sum the absolute differences between each bidder's bid and their value, dropping the highest bidder. We then take the mean of this over each 5-round block. We then compute standard errors counting each group's 5-round mean as a single observation. (18 observations per cell.) $p$-values are computed using a two-sample $t$-test, allowing for unequal variances.

Table 4: Proportion of serial dictatorships not ending in dominant strategy outcome, $p$-values calculated using Wilcoxon rank-sum test

| | SP | OSP | $p$-value |
|---|---|---|---|
| Rounds 1-5 | 43.3% | 7.8% | .0001 |
| Rounds 6-10 | 28.9% | 6.7% | .0010 |

This is the same as Table 4 in the main text, except that the $p$-values are calculated using the Wilcoxon rank-sum test.

Table 5: Proportion of incorrect choices under serial dictatorship: SP (imputed) vs OSP (actual)

| | SP | OSP | $p$-value |
|---|---|---|---|
| Rounds 1-5 | 17.8% | 2.6% | $< .001$ |
| | (3.5%) | (1.1%) | |
| Rounds 6-10 | 10.7% | 2.6% | .002 |
| | (2.0%) | (1.3%) | |
| $p$-value | .078 | 1.000 | |

For each group in SP-RSD, for each period, we simulate the three choices that we would have observed under OSP-RSD. For each group, for each 5-round block, we record the proportion of choices that are incorrect. We then compute standard errors counting each group-block pair as a single observation. (18 observations per cell.) When comparing SP to OSP, we compute $p$-values using a two-sample $t$-test. When comparing early to late rounds of the same game, we compute $p$-values using a paired $t$-test. In the sample for OSP-RSD, there are 7 incorrect choices in the first five rounds and 7 incorrect choices in the last five rounds.

In Table 6, we compute the mean difference between the second-highest bid and the second-highest value, by auction format and by five-round blocks. This summarizes the average *direction* of deviations (positive or negative) from equilibrium play. I had no prior hypothesis about these outcome variables, but report this analysis for completeness.

Table 7 breaks down the rate that subjects submit incorrect rank-order lists in one-shot SP-RSD by demographic category.

Table 6: mean(2nd bid - 2nd value)

| Format | Rounds | SP | OSP |
|---|---|---|---|
| Auction | 1-5 | -1.59 | 0.58 |
| | | (2.07) | (0.77) |
| | 6-10 | 4.22 | 0.78 |
| | | (1.28) | (0.40) |
| +X Auction | 1-5 | 1.62 | 0.15 |
| | | (0.70) | (0.47) |
| | 6-10 | 2.97 | 0.02 |
| | | (0.97) | (0.31) |

For each group, we take the mean difference between the second-highest bid and the second-highest value over each 5-round block. We then compute standard errors counting each group's 5-round mean as a single observation. (18 observations per cell, standard errors in parentheses.)

Table 7: Incorrect lists in one-shot SP-RSD, by demographic categories

| | Yes | No | $p$-value |
|---|---|---|---|
| Economics major | 46% | 31% | .131 |
| STEM major | 32% | 34% | .876 |
| Took a course in game theory | 34% | 33% | .876 |
| Male | 37% | 28% | .214 |

This table displays the proportion of subjects who submitted incorrect rank-order lists in one-shot SP-RSD, by self-reported demographic categories. There are 192 subjects in this treatment. $p$-values are computed with Fisher's exact test.

## 4 Quantal Response Equilibrium

Quantal response equilibrium (QRE) is defined for normal form games (McKelvey and Palfrey, 1995). Agent quantal response equilibrium (AQRE) adapts QRE to extensive form games (McKelvey and Palfrey, 1998); agents play a perturbed best response at each information set, given correct beliefs about the (perturbed) play of other agents and their future selves.

Standard specifications of AQRE predict extreme under-bidding in ascending auctions with fine bid increments. I here provide a simple proof that, for any private-value ascending auction, there exist bid increments fine enough that logit-AQRE predicts that all bids will be close to 0 with probability close to 1. The proof also works for a class of related models, that allow agents to have arbitrary beliefs about the strategies of other agents and their future selves.

**Definition 1.** *An* ***equilibrium with $\tau$-logit errors*** *specifies, for every information set, some (arbitrary) beliefs about opponent strategies and the agent's own continuation strategy. Given those beliefs, let $\{v_1, \ldots, v_J\}$ denote the implied continuation values for actions $\{a_1, \ldots, a_J\}$ at information set $I_i$. The probability that agent $i$ chooses $a_j$ at $I_i$ is equal to $\frac{e^{\tau v_j}}{\sum_{l=1}^{J} e^{\tau v_l}}$ for some $\tau \geq 0$, where this randomization is independent across information sets and agents.*

This contains logit-AQRE (McKelvey and Palfrey, 1998) as a special case. It also includes naïve logit models, such that agents believe that their opponents and their future selves play dominant strategies.

We now consider the following environment: There is a single object and some finite set of bidders, with types (private values) in the interval $[0, 1]$ and quasi-linear utility. Types are drawn according to some joint cdf $F : [0, 1]^{|N|} \to [0, 1]$.

**Definition 2.** *In a* ***$k$-increment ascending auction****, the price rises step by step* *through the set* $\{\frac{0}{k}, \frac{1}{k}, \frac{2}{k}, \ldots\}$. *Agents observe only the clock price, and whether the auction has ended. At each price, active agents (in some arbitrary order) choose In or Out. Agents who choose Out become inactive; they make no payments and do not receive the object. At any point when only one agent is active, the auction ends, and the active agent wins the object at the current clock price.*

**Proposition 3.** *For any* $\epsilon > 0$ *and any* $\tau$, *there exists* $\overline{k}$ *such that, for all* $k \geq \overline{k}$, *in any equilibrium with* $\tau$*-logit errors of the* $k$*-increment ascending auction, the probability that the closing price is above* $\epsilon$ *is less than* $\epsilon$.

*Proof.* Let $M_k^\epsilon := \max\{j \in \mathbb{Z} \text{ such that } \frac{j}{k} \leq \epsilon\}$; this is the index of the highest bid increment below $\epsilon$. Take any agent $i$, any type $\theta_i$, any $k$-increment auction, and any equilibrium with $\tau$-logit errors. Let $v^{\text{In}}_{i,j,k,\theta_i}$ denote the continuation value (given $i$'s arbitrary beliefs) of playing In at clock price $\frac{j}{k}$.

$$\begin{aligned}
&P([i \text{ does not quit before } \epsilon] \cap [\text{ some } j \in N \setminus i \text{ does not quit before } \epsilon]) \\
&\qquad \leq P(i \text{ does not quit before } \epsilon | \text{ some } j \in N \setminus i \text{ does not quit before } \epsilon) \\
&\qquad\qquad = \prod_{j=0}^{M_k^\epsilon} \frac{e^{\tau v^{\text{In}}_{i,j,k,\theta_i}}}{e^{\tau v^{\text{In}}_{i,j,k,\theta_i}} + e^0} \leq \prod_{j=0}^{M_k^\epsilon} \frac{e^\tau}{e^\tau + e^0}
\end{aligned} \tag{65}$$

The last inequality holds because, for *any* beliefs, $i$'s payoff from the $k$-increment ascending auction is bounded above by 1. Moreover,

$$\begin{aligned}
&P(\text{closing price above } \epsilon) \\
&\quad = P(\bigcup_{i \in N} [[i \text{ does not quit before } \epsilon] \cap [\text{ some } j \in N \setminus i \text{ does not quit before } \epsilon]]) \\
&\quad \leq \sum_{i \in N} P([i \text{ does not quit before } \epsilon] \cap [\text{ some } j \in N \setminus i \text{ does not quit before } \epsilon]) \\
&\qquad \leq |N| \prod_{j=0}^{M_k^\epsilon} \frac{e^\tau}{e^\tau + e^0}
\end{aligned} \tag{66}$$

$\lim_{k \to \infty} M_k^\epsilon = \infty$, so $\lim_{k \to \infty} \prod_{j=0}^{M_k^\epsilon} \frac{e^\tau}{e^\tau + e^0} = 0$. Consequently, we can pick $\overline{k}$ large enough that, for all $k \geq \overline{k}$, the right-hand side of the above equation is less than $\epsilon$, which completes the proof.

□

Table 8: AQRE computed for simplified auctions

| | mean(abs(2nd bid - 2nd value)) | |
|---|---|---|
| $\tau$ | 2nd-Price | Ascending |
| 0 | 4.57 | 11.61 |
| $2^{-4}$ | 4.34 | 11.52 |
| $2^{-2}$ | 3.79 | 11.06 |
| 1 | 2.69 | 9.32 |
| $2^{2}$ | 1.63 | 5.91 |
| $2^{4}$ | 0.91 | 2.67 |

As Proposition 3 indicates, in any ascending auction with fine enough bid increments, models such as logit-AQRE or naïve logit predict extreme underbidding.

In the auctions in my experiment, the range of types is $120 and bid increments are 25 cents. It is not feasible to compute AQRE for these auctions. Players have 481 possible types and there are 601 possible bids. Consequently, the pure strategy space in 2P and 2P+X is of the order of $601^{481} \approx 10^{1336}$, and larger yet in AC and AC+X, since there are multiple payoff-equivalent strategies of each kind (specifying different actions at information sets that are never reached).

However, I compute logit-AQRE for the following simpler games: There are 4 agents, with types drawn independently and uniformly at random from 0.25 to 20 (in 0.25 increments). Agents either play a second-price auction with 80 possible bids (equal to the types), or an ascending clock auction with 80 clock prices.

I compute symmetric logit-AQRE for $\tau \in \{0, 2^{-4}, 2^{-2}, 1, 2^{2}, 2^{4}\}$, and sample $10^6$ plays for each game-parameter pair. Table 8 displays the average *absolute* difference between the second-highest bid and the second-highest value, by $\tau$ and auction format. For every computed parameter value, AQRE predicts larger errors in the ascending auction than in the second-price auction, which is the opposite of the effect in the experimental data.

For the simplified auctions that I computed, Table 9 displays the average difference between the second-highest bid and the second-highest value predicted by AQRE. Comparing Table 8 and Table 9 indicates that the large errors in ascending auctions under AQRE are driven almost entirely by under-bidding.

As Figure 2 and Figure 3 (in the main text) illustrate, there is no evidence of systematic under-bidding in AC and AC+X. Table 6 shows this rigorously.

For further robustness, I compute a 'naïve' alternative to AQRE: Agents make logistic errors at each information set, while believing that their opponents (and their future selves) play their dominant strategies[5]. Table 10 displays the average *absolute* difference

[5]In ascending auctions, there are multiple payoff-equivalent dominant strategies, which prescribe

Table 9: Directional Errors under AQRE

| | mean(2nd bid - 2nd value) | |
|---|---|---|
| $\tau$ | 2nd-Price | Ascending |
| 0 | 0.01 | -11.61 |
| $2^{-4}$ | -0.33 | -11.52 |
| $2^{-2}$ | -0.83 | -11.06 |
| 1 | -0.75 | -9.32 |
| $2^{2}$ | -0.32 | -5.91 |
| $2^{4}$ | -0.06 | -2.65 |

Table 10: Naïve logistic errors computed for simplified auctions

| | mean(abs(2nd bid - 2nd value)) | |
|---|---|---|
| $\tau$ | 2nd-Price | Ascending |
| 0 | 4.57 | 11.61 |
| $2^{-4}$ | 4.35 | 11.59 |
| $2^{-2}$ | 3.89 | 11.52 |
| 1 | 2.86 | 10.71 |
| $2^{2}$ | 1.68 | 6.02 |
| $2^{4}$ | 0.91 | 2.43 |

between the second-highest bid and the second-highest value, by $\tau$ and auction format. This model also predicts larger mistakes in ascending auctions than in second-price auctions.

The preceding results suggest that models in which agents make payoff-sensitive mistakes *independently* at every information set are substantially at odds with the data. They predict that ascending auctions with fine increments will induce extreme underbidding, and that errors should be larger in ascending auctions than in second-price auctions. One could invent new theories that permit non-independent errors across information sets or different accuracy parameters for different games, but this introduces many new degrees of freedom, and may have little testable content.

# References

McKelvey, R. D. and Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38.

different actions at prices above the agents' values. For the purpose of these computations, I assume that agents believe that they will quit at all such information sets.

McKelvey, R. D. and Palfrey, T. R. (1998). Quantal response equilibria for extensive form games. *Experimental economics*, 1(1):9–41.