Supplement to "Dynamic contracting: An irrelevance theorem"

(*Theoretical Economics*, Vol. 12, No. 1, January 2017, 109–139)

Péter Eső
Department of Economics, Oxford University

Balázs Szentes
Department of Economics, London School of Economics

Appendix B

Proof of Proposition 2

Since in a pure adverse selection model $\widetilde{u}_{ta_t} \equiv 0$ for all t, throughout this section, we remove a_t from the arguments of $\widetilde{u}_t$, that is, $\widetilde{u}_t : \Theta_t \times X^t \to \mathbb{R}$. We first inspect the consequences of Assumptions 1 and 2 on the orthogonalized model. Note that since θ_t does not depend on x_{t-1}, the ψ_t inference functions do not depend on the decisions either, so $\psi_t : \mathcal{E}^t \to \Theta_t$. The time-separability of the agent's payoff (part (i) of Assumption 2) is preserved in the orthogonalized model, except that the flow utility at t, $u_t : \mathcal{E}^t \times X^t \to \mathbb{R}$, now depends on the *history of types* up to and including time t:

$$u_t(\varepsilon^t, x^t) = \widetilde{u}_t(\psi_t(\varepsilon^t), x^t). \tag{S1}$$

Part (iii) of Assumption 1 implies that the larger is the type history in the orthogonalized model up to time t, the larger is the corresponding period-t type in the original model. This, coupled with part (ii) of Assumption 2, implies that u_t is weakly increasing in ε^{t-1} and strictly in ε_t. Monotonicity in x^t as well as single crossing (part (iii) of Assumption 2) are also preserved in the orthogonalized model. We state these properties formally in the following lemma.

Lemma S1. *(i) For all $t \in \{0, \dots, T\}$ and $\widehat{\varepsilon}^t, \varepsilon^t \in \mathcal{E}^t$,*

$$\widehat{\varepsilon}^t \le \varepsilon^t \quad \Rightarrow \quad \psi_t(\widehat{\varepsilon}^t) \le \psi_t(\varepsilon^t), \tag{S2}$$

and the inequality is strict whenever $\widehat{\varepsilon}_t < \varepsilon_t$.

(ii) The flow utility, u_t defined by (S1), is weakly increasing in ε^{t-1} and x^{t-1}, and strictly increasing in x_t and ε_t.

(iii) For all $t \in \{0, \dots, T\}$, $u_{t\varepsilon_t}(\varepsilon^t, x^t) \ge u_{t\varepsilon_t}(\varepsilon^t, \widehat{x}^t)$ whenever $x^t \ge \widehat{x}^t$.

Péter Eső: peter.eso@economics.ox.ac.uk
Balázs Szentes: b.szentes@lse.ac.uk

DOI: 10.3982/TE2127

Proof. (i) We have $\varepsilon_t = H_t^{-1}(\theta_t|\theta_{t-1})$, therefore $\psi_0(\varepsilon_0) = H_0^{-1}(\varepsilon_0)$ and ψ_t for $t > 0$ is defined recursively by $\psi_t(\varepsilon^t) = H_t^{-1}(\varepsilon^t|\psi^{t-1}(\varepsilon^{t-1}))$. We prove the statement of this part by induction. For $t = 0$, we have $H_0^{-1}(\varepsilon_0) \geq H_0^{-1}(\widehat{\varepsilon}_0)$ whenever $\varepsilon_0 \geq \widehat{\varepsilon}_0$ and the inequality is strict if $\varepsilon_0 > \widehat{\varepsilon}_0$.

Suppose that (S2) holds for t, that is, $\psi_t(\widehat{\varepsilon}^t) \leq \psi_t(\varepsilon^t)$ whenever $\widehat{\varepsilon}^t \leq \varepsilon^t$ and the inequality is strict whenever $\widehat{\varepsilon}^t < \varepsilon_t$. Note that $\psi_{t+1}(\widehat{\varepsilon}_{t+1}) = H_{t+1}^{-1}(\widehat{\varepsilon}_{t+1}|\psi_t(\widehat{\varepsilon}^t))$ and $\psi_{t+1}(\varepsilon^{t+1}) = H_{t+1}^{-1}(\varepsilon_{t+1}|\psi_t(\varepsilon^t))$. Since $\psi_t(\widehat{\varepsilon}^t) \leq \psi_t(\varepsilon^t)$ by the inductive hypothesis, part (ii) of Assumption 1 implies that $\psi_{t+1}(\widehat{\varepsilon}^{t+1}) \leq \psi_{t+1}(\varepsilon^{t+1})$. In addition, if $\varepsilon_{t+1} > \widehat{\varepsilon}_{t+1}$, then $H_{t+1}^{-1}(\varepsilon_{t+1}|\psi_t(\varepsilon^t)) > H_{t+1}^{-1}(\varepsilon'_{t+1}|\psi_t(\widehat{\varepsilon}^t))$.

(ii) The function u_t is strictly increasing in x_t and weakly increasing in x^{t-1} because of part (ii) of Assumption 2 and (S1). Equalities (S1) and (S2) imply that u_t is strictly increasing in ε_t and weakly increasing in ε^{t-1}.

(iii) Fix a $t \in \{0, \dots, T\}$ and note that by (S1),

$$u_{t\varepsilon_t}(\varepsilon^t, x^t) = \widetilde{u}_{t\theta_t}(\psi_t(\varepsilon^t), x^t)\frac{\partial \psi_t(\varepsilon^t)}{\partial \varepsilon_t}.$$

The result follows from (S2) and part (iii) of Assumption 2. □

Another important consequence of part (i) of Assumption 1 is that for all ε^{t+1} and $\widehat{\varepsilon}_t$, there exists a type $\sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t) \in \mathcal{E}_t$ such that, fixing the principal's past and future decisions as well as the realizations of the agent's types beyond period $t+1$, the agent's utility flow from period $t+1$ on is the *same* with type history ε^{t+1} as it is with $(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t))$. We will show below that σ_{t+1}, interpreted in Eső and Szentes (2007) as the agent's "correction of a lie," defines an optimal strategy for the agent at time $t+1$ after a deviation from truthtelling in an incentive compatible direct mechanism at t. This is formally stated in the following lemma.

Lemma S2. *For all $t \in \{0, \dots, T-1\}$, $\varepsilon^{t+1} \in \mathcal{E}^{t+1}$, and $\widehat{\varepsilon}_t \in \mathcal{E}_t$, there exists a unique $\sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t) \in \mathcal{E}_{t+1}$ such that for all $k = t+1, \dots, T$, all $\widehat{\varepsilon}^k \in \mathcal{E}^k$, and $\widehat{x}^k \in X^k$,*

$$u_k(\varepsilon^{t-1}, \varepsilon_t, \varepsilon_{t+1}, \widehat{\varepsilon}_{t+2}, \dots, \widehat{\varepsilon}_k, \widehat{x}^k) = u_k(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}, \widehat{\varepsilon}_{t+2}, \dots, \widehat{\varepsilon}_k, \widehat{x}^k). \tag{S3}$$

The function σ_{t+1} is increasing in ε_t, strictly increasing in ε_{t+1}, and decreasing in $\widehat{\varepsilon}_t$.

Proof. Fix a $t \in \{0, \dots, T-1\}$, $\varepsilon^{t+1} \in \mathcal{E}^{t+1}$, and $\widehat{\varepsilon}_t \in \mathcal{E}_t$. Let

$$\sigma_{t+1} = H_{t+1}(\psi_{t+1}(\varepsilon^{t+1})|\psi_t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)). \tag{S4}$$

By the full support assumption in part (i) of Assumption 1, it follows that

$$\psi_{t+1}(\varepsilon^{t+1}) = \psi_{t+1}(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}),$$

that is, the computed time-$(t+1)$ type of the original model is the same after ε^{t+1} and $(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1})$. Therefore, the inferred type in the original model is also the same after any future observations, that is,

$$\psi_k(\varepsilon^{t-1}, \varepsilon_t, \varepsilon_{t+1}, \widehat{\varepsilon}_{t+2}, \dots, \widehat{\varepsilon}_k) = \psi_k(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}, \widehat{\varepsilon}_{t+2}, \dots, \widehat{\varepsilon}_k)$$

for all $k = t+1, \dots, T$, all $\widehat{\varepsilon}^k \in \mathcal{E}^k$. This equality and (S1) imply (S3). Also note that $\sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t)$, defined by (S4), is increasing in ε_t, strictly increasing in ε_{t+1} by part (i) of Lemma S1, and decreasing in $\widehat{\varepsilon}_t$ by part (i) of Lemma S1 and part (iii) of Assumption 1.

It remains to show that there does not exist any other σ_{t+1} that satisfies (S3). This follows from part (ii) of Lemma S1, which states that u_{t+1} is strictly increasing in ε_{t+1}, which implies that (S3) with $k = t+1$ cannot hold for two different σ_{t+1}'s. □

The statement of the previous lemma might appear somewhat complicated at first glance, but its meaning and its intuitive proof are quite straightforward. Part (i) of Assumption 1 requires the support of θ_t to be independent of θ_{t-1}. Therefore, if the type of the agent is $\psi_t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)$ at time t, there is a chance that the period-$(t+1)$ type will be $\psi^{t+1}(\varepsilon^{t+1})$. The type $\sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t)$ denotes the orthogonalized information of the agent at $t+1$ that induces the transition from $\psi_t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)$ to $\psi_{t+1}(\varepsilon^{t+1})$, that is,

$$\psi_{t+1}(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t)) = \psi_{t+1}(\varepsilon^{t+1}).$$

This means that the inferred type in the original model is the same after the histories $(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t))$ and ε^{t+1}. Part (i) of Assumption 1 and part (ii) of Assumption 2 imply that, given the decisions, the flow utilities in the future only depend on current type, which, in turn, implies (S3).

The decision rule in the orthogonalized model, $\{x_t : \mathcal{E}^t \to X_t\}_{t=0}^T$, which corresponds to $\{\widetilde{x}_t\}_{t=0}^T$, is defined by $x_t(\varepsilon^t) = \widetilde{x}_t(\psi^t(\varepsilon^t))$ for all t and ε^t. Note that, by (S2), if $\{\widetilde{x}_t\}_{t=0}^T$ is increasing in type ($\widetilde{x}_t$ is increasing in θ^t for all t), then the corresponding decision rule $\{x_t\}_{t=0}^T$ in the orthogonalized model is also increasing in type.[1]

In fact, the monotonicity of $\{\widetilde{x}_t\}_{t=0}^T$ implies a stronger monotonicity condition on $\{x_t\}_{t=0}^T$. Consider the two type histories ε^k and $(\varepsilon_1, \dots, \varepsilon_{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t), \varepsilon_{t+2}, \dots, \varepsilon_k)$. Note that the inferred types in the original model are exactly the same along these histories except at time t. At time t, the inferred type is smaller after ε^t if and only if $\varepsilon_t \le \widehat{\varepsilon}_t$. Since $\widetilde{x}_k$ is increasing in θ_t, the decision is smaller after ε^k if and only if $\varepsilon_t \le \widehat{\varepsilon}_t$. This is formally stated as follows.

Remark S1. If $\{\widetilde{x}_t\}_{t=0}^T$ is increasing, then for all $k = 1, \dots, T$, $t < k$, $\varepsilon^k \in \mathcal{E}^k$,

$$x^k(\varepsilon^k) \le x^k(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t), \varepsilon_{t+2}, \dots, \varepsilon_k) \quad \Leftrightarrow \quad \widehat{\varepsilon}_t \ge \varepsilon_t. \tag{S5}$$

Proof. Recall from the proof of Lemma S2 that for all $k = t+1, \dots, T$,

$$\psi_k(\varepsilon^{t+1}, \varepsilon_{t+2}, \dots, \varepsilon_k) = \psi_k(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t), \varepsilon_{t+2}, \dots, \varepsilon_k).$$

By (S2), $\psi_t(\varepsilon^t) \le \psi_t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)$ if and only if $\widehat{\varepsilon}_t \ge \varepsilon_t$. Then (S5) follows from the monotonicity of $\{\widetilde{x}_t\}_0^T$ and the definition of $\{x\}_0^T$. □

[1]To see this, note that if $v^t \ge \widehat{v}^t$, then $x_t(\widehat{v}^t) = \widetilde{x}_t(\psi^t(\widehat{v}^t)) \le \widetilde{x}_t(\psi^t(v^t)) = x_t(v^t)$, where the inequality follows from the monotonicity of $\{\widetilde{x}_t\}_0^T$ and (S2).

To simplify the exposition, we introduce the following notation for $t = 0, \dots, T$, $k \geq t$:

$$\zeta_t^k(\varepsilon^k, y) = (\varepsilon^{t-1}, y, \varepsilon_{t+1}, \dots, \varepsilon_k)$$
$$\rho_t^k(\varepsilon^k, y, \widehat{\varepsilon}_t) = (\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t-1}, y, \varepsilon_{t+1}, \widehat{\varepsilon}_t), \varepsilon_{t+2}, \dots, \varepsilon_k).$$

The vectors $\zeta_t^k(\varepsilon^k, y)$ and $\rho_t^k(\varepsilon^k, y, \widehat{\varepsilon}_t)$ are type histories up to period k, true or reported, which are different from ε^k only at t or at t and $t+1$. For $k = t$ these are appropriately truncated, e.g., $\rho_t^t(\varepsilon^t, y, \widehat{\varepsilon}_t) = (\varepsilon^{t-1}, \widehat{\varepsilon}_t)$.

As we explained, the monotonicity of $\{\widetilde{x}_t\}_{t=0}^T$ implies the monotonicity of both$\{x_t\}_{t=0}^T$ and (S5). Therefore, to prove Proposition 2, it is sufficient to show that any increasing decision rule in the orthogonalized model that satisfies (S5) can be implemented. In what follows, fix a direct mechanism with an increasing decision rule $\{x_t\}_{t=0}^T$ that satisfies (S5). Let $\Pi_t(\varepsilon_t|\varepsilon^{t-1})$ denote a truthful agent's expected payoff at t conditional on ε^t; that is,

$$\Pi_t(\varepsilon_t|\varepsilon^{t-1}) = E\left[\sum_{k=0}^{T} u_k(\varepsilon^k, x^k(\varepsilon^k)) - p(\varepsilon^T)\middle|\varepsilon^t\right].$$

Define the payment function, p, such that for all $t = 0, \dots, T$ and $\varepsilon^t \in \mathcal{E}^t$,

$$\Pi_t(\varepsilon_t|\varepsilon^{t-1}) = \Pi_t(0|\varepsilon^{t-1}) + E\left[\int_0^{\varepsilon_t} \sum_{k=t}^{T} u_{k\varepsilon_t}(\zeta_t^k(\varepsilon^k, y), x^k(\zeta_t^k(\varepsilon^k, y)))\, dy\middle|\varepsilon^t\right]. \tag{S6}$$

It is not hard to show that the integral on the right-hand side of (S6) exists and is finite because of part (ii) of Assumption 1, part (i) of Assumption 2, and the monotonicity of x^k. It should be clear that it is possible to define p such that (S6) holds.

In this mechanism, let $\pi_t(\varepsilon_t, \widehat{\varepsilon}_t|\varepsilon^{t-1})$ denote the expected payoff of the agent at time t whose type history is ε^t and who has reported $(\varepsilon^{t-1}, \widehat{\varepsilon}_t)$. This is the maximum payoff she can achieve from using any reporting strategy from $t+1$ conditional on the type history ε^t and on the reports $(\varepsilon^{t-1}, \widehat{\varepsilon}_t)$. If the mechanism is incentive compatible (IC) then, clearly, $\Pi_t(\varepsilon_t|\varepsilon^{t-1}) = \pi_t(\varepsilon_t, \varepsilon_t|\varepsilon^{t-1})$.

We call a mechanism *IC after time t* if, conditional on telling the truth before and at time $t-1$, it is an equilibrium strategy for the agent to tell the truth afterward, that is, from period t on. By Lemma S2, the continuation utilities of the agent with type ε^{t+1} are the same as those of the agent with type $(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t))$ conditional on the reports and the realization of types after $t+1$. Therefore, if a mechanism is IC after $t+1$, the agent whose type history is ε^{t+1} and who reported $(\varepsilon^{t-1}, \widehat{\varepsilon}_t)$ up to time t maximizes her continuation payoff by reporting $\sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t)$ at time $t+1$ and reporting truthfully afterward. If this were not the case, then the agent with $(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t))$ would have a profitable deviation after truthful reports up to and including t, contradicting the assumption that the mechanism is IC after $t+1$. Therefore, in a mechanism that is IC after $t+1$, we have

$$\begin{aligned}\pi_t(\varepsilon_t, \widehat{\varepsilon}_t|\varepsilon^{t-1}) = {}& u_t(\varepsilon^t, x^t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)) - u_t(\varepsilon^{t-1}, \widehat{\varepsilon}_t, x^t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)) \\ & + \int \Pi_{t+1}(\sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t)|\varepsilon^{t-1}, \widehat{\varepsilon}_t)\, d\varepsilon_{t+1}.\end{aligned} \tag{S7}$$

We use (S7) in the following lemma to characterize the continuation payof of the agent who deviates at t in a mechanism, that is, IC after t.

Lemma S3. *Suppose that the mechanism is IC after time $t+1$ and (S6) is satisfied. Then, for all ε^t and $\widehat{\varepsilon}_t$,*

$$\begin{aligned}
&\pi_t(\varepsilon_t, \widehat{\varepsilon}_t|\varepsilon^{t-1}) - \pi_t(\widehat{\varepsilon}_t, \widehat{\varepsilon}_t|\varepsilon^{t-1}) \\
&\quad = \sum_{k=t}^{T} E\left[\int_{\widehat{\varepsilon}_t}^{\varepsilon_t} u_{k\varepsilon_t}(\zeta_t^k(\varepsilon^k, y), x^k(\rho_t^k(\varepsilon^k, y, \widehat{\varepsilon}_t)))\,dy\middle|\varepsilon^t\right].
\end{aligned} \tag{S8}$$

This lemma is a direct generalization of Lemma 5 of Eső and Szentes (2007).

Proof of Lemma S3. Let $\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, y)$ denote $(\varepsilon^{t-1}, \widehat{\varepsilon}_t, y, \varepsilon_{t+2}, \ldots, \varepsilon_k)$ for $k = t+1, \ldots, T$. Suppose first that $\varepsilon_t > \widehat{\varepsilon}_t$. Then $\sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t) > \varepsilon_{t+1}$, and

$$\begin{aligned}
\pi_t(\varepsilon_t, \widehat{\varepsilon}_t|\varepsilon^{t-1}) &= u_t(\varepsilon^t, x^t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)) - u_t(\varepsilon^{t-1}, \widehat{\varepsilon}_t, x^t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)) \\
&\quad + \int \Pi_{t+1}(\sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t)|\varepsilon^{t-1}, \widehat{\varepsilon}_t)\,d\varepsilon_{t+1} \\
&= u_t(\varepsilon^t, x^t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)) - u_t(\varepsilon^{t-1}, \widehat{\varepsilon}_t, x^t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)) + \Pi_t(\widehat{\varepsilon}_t|\varepsilon^{t-1}) \\
&\quad + \sum_{k=t+1}^{T} \int \cdots \int \int_{\varepsilon_{t+1}}^{\sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t)} u_{\varepsilon_{t+1}}^k(\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, y), \\
&\quad x^k(\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, y)))\,dy\,d\varepsilon_{t+1} \cdots d\varepsilon_k \\
&= u_t(\varepsilon^t, x^t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)) - u_t(\varepsilon^{t-1}, \widehat{\varepsilon}_t, x^t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)) + \pi_t(\widehat{\varepsilon}_t, \widehat{\varepsilon}_t|\varepsilon^{t-1}) \\
&\quad + \sum_{k=t+1}^{T} \int \cdots \int \int_{\varepsilon_{t+1}}^{\sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t)} u_{k\varepsilon_{t+1}}(\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, y), \\
&\quad x^k(\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, y)))\,dy\,d\varepsilon_{t+1} \cdots d\varepsilon_k,
\end{aligned}$$

where the first equality is just (S7), the second one follows from (S6), and the third one follows from $\Pi_t(\widehat{\varepsilon}_t|\varepsilon^{t-1}) = \pi_t(\widehat{\varepsilon}_t, \widehat{\varepsilon}_t|\varepsilon^{t-1})$. So to prove (S8), we only need to show that

$$u_t(\varepsilon^t, x^t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)) - u_t(\varepsilon^{t-1}, \widehat{\varepsilon}_t, x^t(\varepsilon^{t-1}, \widehat{\varepsilon}_t)) = \int_{\widehat{\varepsilon}_t}^{\varepsilon_t} u_{t\varepsilon_t}(\varepsilon_{-t}^t, y, x^t(\rho_t^t(\varepsilon^t, \widehat{\varepsilon}_t)))\,dy \tag{S9}$$

and

$$\begin{aligned}
&\sum_{k=t+1}^{T} \int \cdots \int \int_{\varepsilon_{t+1}}^{\sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t)} u_{k\varepsilon_{t+1}}(\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, y), x^k(\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, y)))\,dy\,d\varepsilon_{t+1} \cdots d\varepsilon_k \\
&\quad = \sum_{k=t+1}^{T} \int \cdots \int \int_{\widehat{\varepsilon}_t}^{\varepsilon_t} u_{t\varepsilon_t}(\zeta_t^k(\varepsilon^k, y), x^k(\rho_t^k(\varepsilon^k, y, \widehat{\varepsilon}_t)))\,dy\,d\varepsilon_{t+1} \cdots d\varepsilon_k.
\end{aligned} \tag{S10}$$

Equation (S9) directly follows from the fundamental theorem of calculus. We now turn our attention to (S10). By Lemma S2, σ_{t+1} is continuous and monotone. The image of $\sigma_{t+1}(\varepsilon^{t+1}, y)$ on $y \in [\widehat{\varepsilon}_t, \varepsilon_t]$ is $[\varepsilon_{t+1}, \sigma_{t+1}(\varepsilon^{t+1}, \widehat{\varepsilon}_t)]$. Hence, after changing the variables of integration, for all $k = t+1, \dots, T$,

$$\begin{aligned}\int_{\varepsilon_{t+1}}^{\sigma_{t+1}(\varepsilon^{t+1},\widehat{\varepsilon}_t)} & u_{k\varepsilon_{t+1}}(\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, y), x^k(\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, y)))\, dy \\ &= \int_{\widehat{\varepsilon}_t}^{\varepsilon_t} u_{k\varepsilon_{t+1}}(\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t-1}, y, \varepsilon_{t+1}, \widehat{\varepsilon}_t)), \\ &\qquad x^k(\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t-1}, y, \varepsilon_{t+1}, \widehat{\varepsilon}_t)))) \frac{\partial \sigma_{t+1}(\varepsilon^{t-1}, y, \varepsilon_{t+1}, \widehat{\varepsilon}_t)}{\partial y}\, dy.\end{aligned} \tag{S11}$$

Recall that by (S3) the expression

$$u_k(\varepsilon^{t-1}, y, \varepsilon_{t+1}, \dots, \varepsilon_k, x^k) \equiv u_k(\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t-1}, y, \varepsilon_{t+1}, \widehat{\varepsilon}_t)), x^k)$$

is an identity in y, so by the implicit function theorem,

$$\begin{aligned}& u_{k\varepsilon_t}(\varepsilon^{t-1}, y, \varepsilon_{t+1}, \dots, \varepsilon_k, x^k) \\ &\quad = u_{k\varepsilon_{t+1}}(\varepsilon^{t-1}, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t-1}, y, \varepsilon_{t+1}, \widehat{\varepsilon}_t), \dots, \varepsilon_k, x^k) \frac{\sigma_{t+1}(\varepsilon^{t-1}, y, \varepsilon_{t+1}, \widehat{\varepsilon}_t)}{\partial y}.\end{aligned} \tag{S12}$$

Plugging (S12) into (S11) and noting that $\gamma_t^k(\varepsilon^k, \widehat{\varepsilon}_t, \sigma_{t+1}(\varepsilon^{t-1}, y, \varepsilon_{t+1}, \widehat{\varepsilon}_t)) = \rho_t^k(\varepsilon^k, y, \widehat{\varepsilon}_t)$ yields (S10).

An identical argument can be used to deal with the case where $\widehat{\varepsilon}_t > \varepsilon_t$. □

We are now ready to prove Proposition 2.

Proof of Proposition 2. To prove that the transfers defined by (S6) implement $\{x_t\}_{t=0}^T$, it is enough to prove that the mechanism is IC after all $t = 0, \dots, T-1$. We prove this by induction. For $t = T-1$ this follows from the standard result in static mechanism design with the observation that x_T is monotone and (S6) is satisfied for T. Suppose now that the mechanism is IC after $t+1$. We show that the mechanism is IC after t, that is, the agent has no incentive to lie at t if she has told the truth before t.

Consider an agent with type history ε^t and report history ε^{t-1} who is contemplating to report $\widehat{\varepsilon}_t < \varepsilon_t$. We have to show that $\pi_t(\varepsilon_t, \widehat{\varepsilon}_t | \varepsilon^{t-1}) - \pi_t(\varepsilon_t, \varepsilon_t | \varepsilon^{t-1}) \le 0$, which can be written as

$$\pi_t(\varepsilon_t, \widehat{\varepsilon}_t | \varepsilon^{t-1}) - \pi_t(\widehat{\varepsilon}_t, \widehat{\varepsilon}_t | \varepsilon^{t-1}) + \pi_t(\widehat{\varepsilon}_t, \widehat{\varepsilon}_t | \varepsilon^{t-1}) - \pi_t(\varepsilon_t, \varepsilon_t | \varepsilon^{t-1}) \le 0.$$

By (S6) and (S8), the previous inequality can be expressed as

$$\begin{aligned}& \sum_{k=t}^T E\left[\int_{\widehat{\varepsilon}_t}^{\varepsilon_t} u_{k\varepsilon_t}(\zeta_t^k(\varepsilon^k, y), x^k(\zeta_t^k(\varepsilon^k, y)))\, dy \,\middle|\, \varepsilon^t\right] \\ &\qquad \ge \sum_{k=t}^T E\left[\int_{\widehat{\varepsilon}_t}^{\varepsilon_t} u_{k\varepsilon_t}(\zeta_t^k(\varepsilon^k, y), x^k(\rho_t^k(\varepsilon^k, y, \widehat{\varepsilon}_t)))\, dy \,\middle|\, \varepsilon^t\right].\end{aligned} \tag{S13}$$

To prove this inequality it is enough to show that the integrand on the left-hand side is larger than the integrand on the right-hand side. By part (iii) of Lemma S1, to show this, we only need that $x^k(\rho_t^k(\varepsilon^k, y, \widehat{\varepsilon}_t)) \leq x^k(\zeta_t^k(\varepsilon^k, y))$ on $y \in [\widehat{\varepsilon}_t, \varepsilon_t]$, which follows from Remark S1. An identical argument can be used to rule out deviation to $\widehat{\varepsilon}_t > \varepsilon_t$. □

From the proof of Proposition 2 it is clear that in the environment satisfying Assumptions 1 and 2 (i.e., with Markov types and a well behaved agent payoff function), a decision rule $\{\tilde{x}_t\}_{t=0}^T$ is implemented by transfers satisfying (S6) if and only if condition (S13) holds in the orthogonalized model.[2] But (S6) is also a necessary condition of implementation (differentiate it in ε_t and compare that with the condition in Proposition 1); therefore, condition (S13) is indeed the necessary and sufficient condition of implementability of a decision rule in the regular, Markov environment. This is formally stated in the following remark.

Remark S2. Suppose that Assumptions 1 and 2 hold. Then a decision rule, $\{\widetilde{x}_t\}_0^T$, is implementable if and only if (S13) holds in the model with orthogonalized information.

Implementability in the Benchmark Case. Suppose that the principal can observe $\varepsilon_1, \dots, \varepsilon_T$. Then, using arguments in standard static mechanism design, a decision rule $\{x_t\}_0^T$ can be implemented if and only if, for all $\widehat{\varepsilon}_0, \varepsilon_0 \in \mathcal{E}_0$, $\widehat{\varepsilon}_0 \leq \varepsilon_0$,

$$E\left[\sum_{k=0}^{T} \int_{\widehat{\varepsilon}_0}^{\varepsilon_0} u_{k\varepsilon_0}(y, \varepsilon_{-0}^k, x^k(y, \varepsilon_{-0}^k))\, dy \middle| \varepsilon_0\right] \geq E\left[\sum_{k=0}^{T} \int_{\widehat{\varepsilon}_0}^{\varepsilon_0} u_{k\varepsilon_0}(y, \varepsilon_{-0}^k, x^k(\widehat{\varepsilon}_0, \varepsilon_{-0}^k))\, dy \middle| \varepsilon_0\right].$$

This inequality is obviously a weaker condition than (S13), so the principal can implement more allocations in the benchmark case.

Proof of Proposition 5

To be able to refer to the additional restrictions required by the proposition, we state the strict single-crossing properties in the following assumption.

Assumption 6. *(i) For all* $t \in \{0, \dots, T\}$, $\theta_t \in \Theta_t$, *and* $a_t \in A_t$, $\widetilde{u}_{t\theta_t}(\theta_t, a_t, x^t) > \widetilde{u}_{t\theta_t}(\theta_t, a_t, \widehat{x}^t)$ *whenever* $x^t > \widehat{x}^t$.

(ii) For all $t \in \{0, \dots, T\}$, $\theta_t \in \Theta_t$, $a_t \in A_t$, *and* $x^t \in X^t$, $\widetilde{u}_{t\theta_t x_\tau}(\theta_t, a_t, x^t) f_{ta_t}(\theta_t, a_t) > \widetilde{u}_{ta_t x_\tau}(\theta_t, a_t, x^t) f_{t\theta_t}(\theta_t, a_t)$.

Recall that in the proof of Proposition 4 we decomposed the gain from any deviation strategy into the sum of two parts. The first part is the difference between the payoff from deviating and truth-telling in the hypothetical model where y is contractible. The second part is the difference between the payoff from the misreporting strategy but

[2] Note that condition (S13) is a joint restriction on $\{x_t\}_0^T$ and the marginal utility of the agent's type, and it is implied by the monotonicity of the decision rule in the environment of Assumptions 1 and 2.

matching the actions with the misreports and the payoff from misreporting and altering the actions optimally. Then we appealed to Proposition 2 to conclude that the first part is negative and proved that the second part is small. The key to the proof of this proposition is to show that if the deviation strategy leads to a decision rule that is far away from $(\widetilde{\mathbf{x}}^T, \widetilde{\mathbf{y}}^T)$, then the first part of the decomposed payoff is not only negative, but also large relative to the possible gain corresponding to the second part. To do so, we follow the standard argument in static mechanism design to estimate the deviation payoffs. This estimation is based on the single-crossing property and we have derived the key formula, (S13), in the orthogonalized model.

In what follows, we use the notation introduced in the proof of Proposition 4. Recall that the payoff difference corresponding to the first part of the decomposition is

$$E_{\theta^T}\left[\sum_{t=0}^{T} w_t(\theta_t, \widetilde{\mathbf{y}}_t(\theta^t), \widetilde{\mathbf{x}}^t(\theta^t)) - \overline{\mathbf{p}}(\theta^T)\Big|\theta_0\right] - E_{\theta^T}\left[\sum_{t=0}^{T} w_t(\theta_t, \widetilde{\mathbf{y}}_t(\rho_t(\theta^t)), \widetilde{\mathbf{x}}^t(\rho_t(\theta^t)) - \overline{\mathbf{p}}(\rho^T(\theta^T))\Big|\theta_0\right]. \tag{S14}$$

Since (S13) is derived in the orthogonalized model, we rewrite the previous inequality in terms of the orthogonalized information structure. To this end, let $(\underline{\widetilde{\mathbf{x}}}_t, \underline{\widetilde{\mathbf{a}}}_t)_{t=0}^T$ denote the allocation corresponding to $(\widetilde{\mathbf{x}}_t, \widetilde{\mathbf{a}}_t)_{t=0}^T$, that is, $(\underline{\widetilde{\mathbf{x}}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{a}}}_t(\varepsilon^t)) \equiv (\widetilde{\mathbf{x}}_t(\psi^t(\varepsilon^t)), \widetilde{\mathbf{a}}_t(\psi^t(\varepsilon^t)))$ for all t, ε^t. Similarly, define $\underline{\widetilde{\mathbf{y}}}_t(\varepsilon^t)$ and $\underline{\overline{\mathbf{p}}}_t(\varepsilon^t)$ to be $\widetilde{\mathbf{y}}_t(\psi^t(\varepsilon^t))$ and $\overline{\mathbf{p}}(\psi^t(\varepsilon^t))$, respectively, for all t and ε^t. Let $\underline{\rho}_t(\varepsilon^t)$ denote the deviation strategy, that is, $\underline{\rho}_t(\varepsilon^t) = \rho_t(\psi^t(\varepsilon^t))$. Finally, let $\omega_t(\varepsilon^t, y, x) \equiv w_t(\psi^t(\varepsilon^t), y, x)$ for all $t = 0, \dots, T$. It is not hard to prove that by Assumption 6 it follows that there exists an $\widetilde{m} \in \mathbb{R}_+$, such that for all $t \in \{0, \dots, T\}$ and $\varepsilon_t \in \Theta_t$,

$$\omega_{t\varepsilon_t}(\varepsilon^t, y_t, x^t) - \omega_{t\varepsilon_t}(\varepsilon^t, \widehat{y}_t, \widehat{x}^t) \geq \widetilde{m}\|(y_t, x^t) - (\widehat{y}_t, \widehat{x}^t)\| \tag{S15}$$

whenever $(y_t, x^t) \geq (\widehat{y}_t, \widehat{x}^t)$ and $(y_t, x^t), (\widehat{y}_t, \widehat{x}^t) \in \{(\underline{\widetilde{\mathbf{y}}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{x}}}^t(\varepsilon^t)) : \varepsilon^t \in \mathcal{E}^t\}$.

Using these notations, we can rewrite (S14) as

$$E_{\varepsilon^T}\left[\sum_{t=0}^{T} \omega_t(\varepsilon_t, \underline{\widetilde{\mathbf{y}}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{x}}}^t(\varepsilon^t)) - \underline{\overline{\mathbf{p}}}(\varepsilon^T)\Big|\varepsilon_0\right] - E_{\varepsilon^T}\left[\sum_{t=0}^{T} \omega_t(\varepsilon_t, \underline{\widetilde{\mathbf{y}}}_t(\rho_t(\varepsilon^t)), \underline{\widetilde{\mathbf{x}}}^t(\rho_t(\varepsilon^t)) - \underline{\overline{\mathbf{p}}}(\rho^T(\varepsilon^T))\Big|\varepsilon_0\right].$$

Observe that it is without the loss of generality to assume that there exists a $K > 0$ such that

$$\left\|(\underline{\widetilde{\mathbf{y}}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{x}}}^t(\varepsilon^t)) - (\underline{\widetilde{\mathbf{y}}}_t(\varepsilon'^t), \underline{\widetilde{\mathbf{x}}}^t(\varepsilon'^t))\right\| \leq K\|\varepsilon^t - \varepsilon'^t\| \tag{S16}$$

for all $t \in \{0, \dots, T\}$ and $\varepsilon^t \in \mathcal{E}^t$ because any increasing allocation rule can be approximated arbitrarily well in the L_2 norm with a decision rule that satisfies the previous inequality. Also note that the set $\{(\underline{\widetilde{\mathbf{y}}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{x}}}^t(\varepsilon^t)) : \varepsilon^t \in \mathcal{E}^t\}$ is bounded. For notational convenience, we assume that

$$\|(\underline{\widetilde{\mathbf{y}}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{x}}}^t(\varepsilon^t))\| \le \tfrac{1}{2} \tag{S17}$$

for all $t \in \{0, \dots, T\}$ and $\varepsilon^t \in \mathcal{E}^t$.

Next, we show that for all $\widetilde{\delta} > 0$ and $\tau = 0, \dots, T$, we can construct payment rules so that

$$E_{\varepsilon^t} \big\| (\underline{\widetilde{\mathbf{y}}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{x}}}_t(\varepsilon^t)) - (\underline{\widetilde{\mathbf{y}}}_t(\rho(\varepsilon^t))), \underline{\widetilde{\mathbf{x}}}_t(\rho(\varepsilon^t)) \big\|^2 < \frac{\widetilde{\delta}}{T+1} \tag{S18}$$

for all $t \le \tau$. We prove this statement by induction. Consider $\tau = 0$. As we mentioned before, we use (S13) to estimate the loss from a deviation. In particular, we estimate this loss by the difference between the first term of the summation on the left-hand side and the first term of the summation on the right-hand side of (S13). In other words, we approximate the loss due to a time t deviation by the instantaneous loss and ignore future losses. Hence, by (S13),

$$\begin{aligned} &E_{\varepsilon^T}\Bigg[\sum_{t=0}^{T} \omega_t(\varepsilon_t, \underline{\widetilde{\mathbf{y}}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{x}}}^t(\varepsilon^t)) - \underline{\overline{\mathbf{p}}}(\varepsilon^T) \Big| \varepsilon_0\Bigg] \\ &\qquad - E_{\varepsilon^T}\Bigg[\sum_{t=0}^{T} \omega_t(\varepsilon_t, \underline{\widetilde{\mathbf{y}}}_t(\underline{\rho}_t(\varepsilon^t)), \underline{\widetilde{\mathbf{x}}}^t(\underline{\rho}_t(\varepsilon^t)) - \underline{\overline{\mathbf{p}}}(\underline{\rho}^T(\varepsilon^T)) \Big| \varepsilon_0\Bigg] \\ &\ge E_{\varepsilon_0}\Bigg[\int_{\varepsilon_0}^{\rho(\varepsilon_0)} \omega_{0\varepsilon_0}(z, \underline{\widetilde{\mathbf{y}}}_0(z), \underline{\widetilde{\mathbf{x}}}^0(z)) - \omega_{0\varepsilon_0}(z, \underline{\widetilde{\mathbf{y}}}_0(\underline{\rho}_t(\varepsilon_0)), \widetilde{\mathbf{x}}^0(\underline{\rho}_t(\varepsilon_0)))\, dz\Bigg]. \end{aligned}$$

By (S15), the right-hand side of the previous inequality is weakly larger than

$$\widetilde{m} E_{\varepsilon_0}\Bigg[\int_{\varepsilon_0}^{\rho(\varepsilon_0)} \big\| (\underline{\widetilde{\mathbf{y}}}_0(\varepsilon_0), \underline{\widetilde{\mathbf{x}}}^0(\varepsilon_0)) - (\underline{\widetilde{\mathbf{y}}}_0(\underline{\rho}_t(\varepsilon_0)), \underline{\widetilde{\mathbf{x}}}^0(\underline{\rho}_t(\varepsilon_0))) \big\|\, dz\Bigg].$$

Furthermore, by (S16), the previous expression is larger than

$$\widetilde{m} E_{\varepsilon_0}\Bigg[\frac{\|(\underline{\widetilde{\mathbf{y}}}_0(\varepsilon_0), \underline{\widetilde{\mathbf{x}}}^0(\varepsilon_0)) - (\underline{\widetilde{\mathbf{y}}}_0(\underline{\rho}_t(\varepsilon_0)), \underline{\widetilde{\mathbf{x}}}^0(\underline{\rho}_t(\varepsilon_0)))\|^2}{2K}\Bigg].$$

Recall that in the proof of Proposition 4, we proved that for each $\delta > 0$ it is possible to construct payments so that the second part of the decomposed gain from deviation is less than δ. Therefore, to guarantee that the deviation $(\rho_t)_t$ is profitable it must be that

$$\widetilde{m} E_{\varepsilon_0}\Bigg[\frac{\|(\underline{\widetilde{\mathbf{y}}}_0(\varepsilon_0), \underline{\widetilde{\mathbf{x}}}^0(\varepsilon_0)) - (\underline{\widetilde{\mathbf{y}}}_0(\underline{\rho}_t(\varepsilon_0)), \underline{\widetilde{\mathbf{x}}}^0(\underline{\rho}_t(\varepsilon_0)))\|^2}{2K}\Bigg] < \delta,$$

that is,

$$E_{\varepsilon_0}\big[\big\| (\underline{\widetilde{\mathbf{y}}}_0(\varepsilon_0), \underline{\widetilde{\mathbf{x}}}^0(\varepsilon_0)) - (\underline{\widetilde{\mathbf{y}}}_0(\underline{\rho}_t(\varepsilon_0)), \underline{\widetilde{\mathbf{x}}}^0(\underline{\rho}_t(\varepsilon_0))) \big\|^2\big] < \frac{2K\delta}{\widetilde{m}}. \tag{S19}$$

So, choosing $\delta = \widetilde{m}\widetilde{\delta}/(2K(T+1))$ yields the claim in (S18) for $\tau = 0$.

Suppose that the claim in (S18) is true for $\tau \geq 0$. We show that this claim is also true for $\tau + 1$. The difficulty with the inductive step is that (S13) can only be used to estimate the time-$(\tau+1)$ loss due to a deviation if there were no deviations in previous periods. Let us explain how we overcome this problem. By the inductive hypothesis, there are payments so that the optimal deviation strategy induces a decision rule that is arbitrarily close to the decision rule generated by truth-telling in time periods $0, 1, \ldots, \tau$. Therefore, we can approximate the optimal deviation by a misreporting strategy that specifies truth-telling until period τ and then coincides with the optimal deviation. According to this approximating deviation strategy, the first deviation occurs in period $\tau + 1$ and hence, we can use (S13) to estimate the loss due to this deviation once again. To this end, let $\underline{\rho}_t^\tau(\varepsilon^t)$ be defined as

$$\underline{\rho}_t^\tau(\varepsilon^t) = \begin{cases} \varepsilon^t & \text{if } t \leq \tau \\ \underline{\rho}_t(\varepsilon^t) & \text{if } t > \tau, \end{cases}$$

that is, $\{\underline{\rho}_t^\tau\}_t$ is a deviation strategy that prescribes truth-telling until period t and after period t, it coincides with $\{\underline{\rho}_t\}_t$. Using this notation,

$$\begin{aligned}
&E_{\varepsilon^T}\left[\sum_{t=0}^T \omega_t(\varepsilon_t, \underline{\widetilde{\mathbf{y}}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{x}}}^t(\varepsilon^t)) - \underline{\overline{\mathbf{p}}}(\varepsilon^T)\Big|\varepsilon_0\right] \\
&\qquad - E_{\varepsilon^T}\left[\sum_{t=0}^T \omega_t(\varepsilon_t, \underline{\widetilde{\mathbf{y}}}_t(\underline{\rho}_t(\varepsilon^t)), \underline{\widetilde{\mathbf{x}}}^t(\underline{\rho}_t(\varepsilon^t)) - \underline{\overline{\mathbf{p}}}(\underline{\rho}^T(\varepsilon^T))\Big|\varepsilon_0\right] \\
&= E_{\varepsilon^T}\left[\sum_{t=0}^T \omega_t(\varepsilon_t, \underline{\widetilde{\mathbf{y}}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{x}}}^t(\varepsilon^t)) - \underline{\overline{\mathbf{p}}}(\varepsilon^T)\Big|\varepsilon_0\right] \\
&\qquad - E_{\varepsilon^T}\left[\sum_{t=0}^T \omega_t(\varepsilon_t, \underline{\widetilde{\mathbf{y}}}_t(\underline{\rho}_t^\tau(\varepsilon^t)), \underline{\widetilde{\mathbf{x}}}^t(\underline{\rho}_t^\tau(\varepsilon^t)) - \underline{\overline{\mathbf{p}}}(\underline{\rho}^{\tau T}(\varepsilon^T))\Big|\varepsilon_0\right] \\
&\qquad + E_{\varepsilon^T}\left[\sum_{t=0}^T \omega_t(\varepsilon_t, \underline{\widetilde{\mathbf{y}}}_t(\underline{\rho}_t^\tau(\varepsilon^t)), \underline{\widetilde{\mathbf{x}}}^t(\underline{\rho}_t^\tau(\varepsilon^t)) - \underline{\overline{\mathbf{p}}}(\underline{\rho}^{\tau T}(\varepsilon^T))\Big|\varepsilon_0\right] \\
&\qquad - E_{\varepsilon^T}\left[\sum_{t=0}^T \omega_t(\varepsilon_t, \underline{\widetilde{\mathbf{y}}}_t(\underline{\rho}_t(\varepsilon^t)), \underline{\widetilde{\mathbf{x}}}^t(\underline{\rho}_t(\varepsilon^t)) - \underline{\overline{\mathbf{p}}}(\underline{\rho}^T(\varepsilon^T))\Big|\varepsilon_0\right].
\end{aligned}$$

By the inductive hypothesis and the Lipschitz continuity of payoffs, for all $\delta, \widetilde{\delta} > 0$, there are payment rules so that

$$\left|E_{\varepsilon^T}\left[\sum_{t=0}^T \omega_t(\varepsilon_t, \underline{\mathbf{y}}_t(\underline{\rho}_t^\tau(\varepsilon^t)), \underline{\widetilde{\mathbf{x}}}^t(\underline{\rho}_t^\tau(\varepsilon^t)) - \underline{\overline{\mathbf{p}}}(\underline{\rho}^{\tau T}(\varepsilon^T))\Big|\varepsilon_0\right]\right. \tag{S20}$$

$$- E_{\varepsilon^T}\left[\sum_{t=0}^{T}\omega_t(\varepsilon_t,\underline{\mathbf{y}}_t(\underline{\rho}_t(\varepsilon^t)),\widetilde{\underline{\mathbf{x}}}^t(\underline{\rho}_t(\varepsilon^t))-\overline{\underline{\mathbf{p}}}(\underline{\rho}^T(\varepsilon^T))\middle|\varepsilon_0\right]\Bigg| \le \delta,$$

$$E_{\varepsilon^{\tau+1}}\big[\big\|(\widetilde{\underline{\mathbf{y}}}_{\tau+1}(\underline{\rho}_{\tau+1}(\varepsilon^{\tau+1})),\widetilde{\underline{\mathbf{x}}}^{\tau+1}(\underline{\rho}_{\tau+1}(\varepsilon^{\tau+1}))) - (\widetilde{\underline{\mathbf{y}}}_{\tau+1}(\underline{\rho}^{\tau}_t(\varepsilon^{\tau+1})),\widetilde{\underline{\mathbf{x}}}^{\tau+1}(\underline{\rho}^{\tau}_{\tau+1}(\varepsilon^{\tau+1})))\big\|^{\sigma}\big] \le \frac{\widetilde{m}\delta}{2K} \tag{S21}$$

for $\sigma = 1, 2$ and (S18) is satisfied for all $t \le \tau$. Therefore, by (S20),

$$\begin{aligned}
&E_{\varepsilon^T}\left[\sum_{t=0}^{T}\omega_t(\varepsilon_t,\widetilde{\underline{\mathbf{y}}}_t(\varepsilon^t),\widetilde{\underline{\mathbf{x}}}^t(\varepsilon^t))-\overline{\underline{\mathbf{p}}}(\varepsilon^T)\middle|\varepsilon_0\right]\\
&\quad - E_{\varepsilon^T}\left[\sum_{t=0}^{T}\omega_t(\varepsilon_t,\widetilde{\underline{\mathbf{y}}}_t(\underline{\rho}_t(\varepsilon^t)),\widetilde{\underline{\mathbf{x}}}^t(\underline{\rho}_t(\varepsilon^t))-\overline{\underline{\mathbf{p}}}(\underline{\rho}^T(\varepsilon^T))\middle|\varepsilon_0\right]\\
&\ge E_{\varepsilon^T}\left[\sum_{t=0}^{T}\omega_t(\varepsilon_t,\widetilde{\underline{\mathbf{y}}}_t(\varepsilon^t),\widetilde{\underline{\mathbf{x}}}^t(\varepsilon^t))-\overline{\underline{\mathbf{p}}}(\varepsilon^T)\middle|\varepsilon_0\right]\\
&\quad - E_{\varepsilon^T}\left[\sum_{t=0}^{T}\omega_t(\varepsilon_t,\widetilde{\underline{\mathbf{y}}}_t(\underline{\rho}^{\tau}_t(\varepsilon^t)),\widetilde{\underline{\mathbf{x}}}^t(\underline{\rho}^{\tau}_t(\varepsilon^t))-\overline{\underline{\mathbf{p}}}(\underline{\rho}^{\tau T}(\varepsilon^T))\middle|\varepsilon_0\right] - \delta.
\end{aligned}$$

Note that according to the deviation strategy $\{\underline{\rho}^{\tau}_t\}_t$, the first deviation occurs in period $\tau+1$. Therefore, we appeal to (S13) once again, and using the same arguments leading to (S19), we conclude that

$$\begin{aligned}
&E_{\varepsilon^T}\left[\sum_{t=0}^{T}\omega_t(\varepsilon_t,\widetilde{\underline{\mathbf{y}}}_t(\varepsilon^t),\widetilde{\underline{\mathbf{x}}}^t(\varepsilon^t))-\overline{\underline{\mathbf{p}}}(\varepsilon^T)\middle|\varepsilon_0\right]\\
&\qquad - E_{\varepsilon^T}\left[\sum_{t=0}^{T}\omega_t(\varepsilon_t,\widetilde{\underline{\mathbf{y}}}_t(\underline{\rho}^{\tau}_t(\varepsilon^t)),\widetilde{\underline{\mathbf{x}}}^t(\underline{\rho}^{\tau}_t(\varepsilon^t))-\overline{\underline{\mathbf{p}}}(\underline{\rho}^{\tau T}(\varepsilon^T))\middle|\varepsilon_0\right]\\
&\ge \widetilde{m}E_{\varepsilon^{\tau+1}}\left[\frac{\|(\widetilde{\underline{\mathbf{y}}}_{\tau+1}(\varepsilon^{\tau+1}),\widetilde{\underline{\mathbf{x}}}^{\tau+1}(\varepsilon^{\tau+1}))-(\widetilde{\underline{\mathbf{y}}}_{\tau+1}(\underline{\rho}^{\tau}_{\tau+1}(\varepsilon^{\tau+1})),\widetilde{\underline{\mathbf{x}}}^{\tau+1}(\underline{\rho}^{\tau}_{\tau+1}(\varepsilon^{\tau+1})))\|^2}{2K}\right]\\
&\ge \widetilde{m}E_{\varepsilon^{\tau+1}}\left[\frac{\|(\widetilde{\underline{\mathbf{y}}}_{\tau+1}(\varepsilon^{\tau+1}),\widetilde{\underline{\mathbf{x}}}^{\tau+1}(\varepsilon^{\tau+1}))-(\widetilde{\underline{\mathbf{y}}}_{\tau+1}(\underline{\rho}_{\tau+1}(\varepsilon^{\tau+1})),\widetilde{\underline{\mathbf{x}}}^{\tau+1}(\underline{\rho}_{\tau+1}(\varepsilon^{\tau+1})))\|^2}{2K}\right]\\
&\qquad - 3\delta,
\end{aligned}$$

where the last inequality follows from (S21) and (S17). Hence, to guarantee that the deviation strategy $(\rho_t)_t$ is profitable, we need that

$$E_{\varepsilon^{\tau+1}}\big[\big\|(\widetilde{\underline{\mathbf{y}}}_{\tau+1}(\varepsilon^{\tau+1}),\widetilde{\underline{\mathbf{x}}}^{\tau+1}(\varepsilon^{\tau+1}))-(\widetilde{\underline{\mathbf{y}}}_{\tau+1}(\underline{\rho}_{\tau+1}(\varepsilon^{\tau+1})),\widetilde{\underline{\mathbf{x}}}^{\tau+1}(\underline{\rho}_{\tau+1}(\varepsilon^{\tau+1})))\big\|^2\big] \le \frac{2K(5\delta)}{\widetilde{m}}.$$

So, choosing $\delta = \widetilde{m}\widetilde{\delta}/(10K)$ yields the claim in (S18) for $\tau+1$.

By the proof of Proposition 4, the payment rule can be defined so that

$$E_{\varepsilon^T}\sum_{t=0}^{T}\left\| f_t(\rho_t(\psi^t(\varepsilon^t)), \alpha_t(\psi^t(\varepsilon^t))) - \underline{\mathbf{y}}_t(\underline{\rho}_t(\varepsilon^t))\right\|^{\sigma} \le \widetilde{\delta} \tag{S22}$$

for $\sigma = 1, 2$. Define $(\overline{\mathbf{x}}_t(\psi_t(\varepsilon^t)), \overline{\mathbf{a}}_t(\psi_t(\varepsilon^t))) \equiv (\widetilde{\mathbf{x}}_t(\underline{\rho}_t(\varepsilon^t)), \widetilde{\mathbf{a}}_t(\underline{\rho}_t(\varepsilon^t)))$ for all $t = 0, \dots, T$. Then

$$\begin{aligned}
&\left\|(\widetilde{\mathbf{y}}_t(\theta^t), \widetilde{\mathbf{x}}_t(\theta^t)) - (\overline{\mathbf{y}}_t(\theta^t), \overline{\mathbf{x}}_t(\theta^t))\right\| \\
&\quad \le \left\|(\underline{\mathbf{y}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{x}}}_t(\varepsilon^t)) - (\underline{\mathbf{y}}_t(\underline{\rho}_t(\varepsilon^t)), \underline{\widetilde{\mathbf{x}}}_t(\underline{\rho}_t(\varepsilon^t)))\right\| \\
&\qquad\qquad + \left\| f_t(\rho_t(\psi^t(\varepsilon^t)), \alpha_t(\psi^t(\varepsilon^t))) - \underline{\mathbf{y}}_t(\underline{\rho}_t(\varepsilon^t))\right\|.
\end{aligned}$$

Notice that summing up the inequalities in (S18) for $t = 0, \dots, T$ yields

$$E_{\varepsilon^T}\sum_{t=0}^{T}\left\|(\underline{\mathbf{y}}_t(\varepsilon^t), \underline{\widetilde{\mathbf{x}}}_t(\varepsilon^t)) - (\underline{\mathbf{y}}_t(\underline{\rho}_t(\varepsilon^t)), \underline{\widetilde{\mathbf{x}}}_t(\underline{\rho}_t(\varepsilon^t)))\right\|^2 < \widetilde{\delta}.$$

By (S17), the previous inequality, and (S22), we conclude that

$$E_{\theta^T}\sum_{t=0}^{T}\left\|(\widetilde{\mathbf{y}}_t(\theta^t), \widetilde{\mathbf{x}}_t(\theta^t)) - (\overline{\mathbf{y}}_t(\theta^t), \overline{\mathbf{x}}_t(\theta^t))\right\|^2 \le 4\widetilde{\delta}.$$

Therefore, the allocation $(\overline{\mathbf{x}}_t, \overline{\mathbf{a}}_t)_t$ satisfies desired inequality in the statement of Proposition 5. Finally, note that $(\overline{\mathbf{x}}_t, \overline{\mathbf{a}}_t)_t$ is implementable because it results from the agent's optimal deviation strategy in the mechanism $(\widetilde{\mathbf{x}}^T, \widetilde{\mathbf{a}}^T, \widetilde{\mathbf{p}})$. □

Reference

Eső, Péter and Balázs Szentes (2007), "Optimal information disclosure in auctions and the handicap auction." *Review of Economic Studies*, 74, 705–731. [2, 5]

Co-editor Johannes Hörner handled this manuscript.

Manuscript received 17 March, 2015; final version accepted 18 October, 2015; available online 23 December, 2015.