Supplementary Material

D Sufficiency of MDR **1**

E Statement of the General Problem and of the Lagrangian Objective **3**

F Binding constraint for Games with Linearity in Types **6**

G Consistent Belief System **7**

H The Belief Management Approach and Ex-post Equilibria **9**

H.1 Simple Lotteries (cf. Hörner, Morelli, and Squintani, 2015) 9

H.2 Generalization: Monotone Lotteries 10

D Sufficiency of MDR

Lemma 4. *Suppose (MDR) holds. Then, upward incentive constraints together with downward local incentive constraints are sufficient for (global) downward incentive constraints.*

Proof. Take a $\mathcal{CM}$ that satisfies downward local incentive constraints and (global) upward incentive constraints. We show that the global downward incentive constraints are necessarily satisfied.

Take i and any θ and θ' such that $\theta > \theta'-1$. We establish that $\Pi_i(\theta;\theta) \geq \Pi_i(\theta';\theta)$. We proceed by induction.

Step 1: Basis. Take any θ and θ' such that $\theta - \theta' = 2$. Then $\Pi(\theta;\theta) \geq \Pi(\theta';\theta)$.

We show that

$$z_i(\theta) - z_i(\theta-2) \geq \gamma_i(\theta-2)U_i(\theta-2;\theta,\mathcal{B}(\sigma)) - \gamma_i(\theta)U_i(\theta;\theta,\mathcal{B}(\sigma)). \tag{A.1}$$

By hypothesis all downward local incentive constraints are satisfied, i.e.,

$$z_i(k) - z_i(k-1) \geq \gamma_i(k-1)U_i(k-1;k,\mathcal{B}(\sigma)) - \gamma_i(k)U_i(k;k,\mathcal{B}(\sigma)).$$

Therefore,

$$\begin{aligned} z_i(\theta) - z_i(\theta-2) \geq & \gamma_i(\theta-1)U_i(\theta-1;\theta,\mathcal{B}(\sigma)) - \gamma_i(\theta)U_i(\theta;\theta,\mathcal{B}(\sigma)) \\ & +\gamma_i(\theta-2)U_i(\theta-2;\theta-1,\mathcal{B}(\sigma)) - \gamma_i(\theta-1)U_i(\theta-1;\theta-1,\mathcal{B}(\sigma)). \end{aligned} \tag{A.2}$$

The RHS of equation (A.2) is larger than the RHS of equation (A.1) if

$$\frac{U_i(\theta-1;\theta-1,\mathcal{B}(\sigma)) - U_i(\theta-1;\theta,\mathcal{B}(\sigma))}{U_i(\theta-2;\theta-1,\mathcal{B}(\sigma)) - U_i(\theta-2;\theta,\mathcal{B}(\sigma))} \leq \frac{\gamma_i(\theta-2)}{\gamma_i(\theta-1)}. \tag{A.3}$$

Upward and downward local incentive constraints of $\theta-1, \theta-2$ jointly imply that

$$\frac{\gamma_i(\theta-2)}{\gamma_i(\theta-1)} \geq \frac{U_i(\theta-1;\theta-2,\mathcal{B}(\sigma)) - U_i(\theta-1;\theta-1,\mathcal{B}(\sigma))}{U_i(\theta-2;\theta-2,\mathcal{B}(\sigma)) - U_i(\theta-2;\theta-1,\mathcal{B}(\sigma))}.$$

Thus, equation (A.3) is true if

$$
\begin{aligned}
&\frac{U_i(\theta{-}1;\theta{-}2,\mathcal{B}(\sigma))-U_i(\theta{-}1;\theta{-}1,\mathcal{B}(\sigma))}{U_i(\theta{-}2;\theta{-}2,\mathcal{B}(\sigma))-U_i(\theta{-}2;\theta{-}1,\mathcal{B}(\sigma))}\\
&\geq \frac{U_i(\theta{-}1;\theta{-}1,\mathcal{B}(\sigma))-U_i(\theta{-}1;\theta,\mathcal{B}(\sigma))}{U_i(\theta{-}2;\theta{-}1,\mathcal{B}(\sigma))-U_i(\theta{-}2;\theta,\mathcal{B}(\sigma))}
\end{aligned}
\tag{A.4}
$$

Equation (A.4) is satisfied by the definition of the MDR condition.

Step 2: Induction Hypothesis. Take θ, θ' such that $\theta-\theta'=n$. Then $\Pi(\theta,\theta)\geq\Pi(\theta',\theta)$.

Step 3: Inductive Step. Next, we show that for any θ and θ' such that $\theta-\theta'=n+1$ it holds that $\Pi(\theta,\theta)\geq\Pi(\theta{-}n-1,\theta)$. That is, we establish

$$
z_i(\theta)-z_i(\theta{-}n-1)\geq\gamma_i(\theta{-}n-1)U_i(\theta{-}n-1;\theta,\mathcal{B}(\sigma))-\gamma_i(\theta)U_i(\theta;\theta,\mathcal{B}(\sigma)). \tag{A.5}
$$

By the induction hypothesis and the fact that local incentive constraints hold, we have that

$$
\begin{aligned}
&z_i(\theta)-z_i(\theta{-}n-1)=z_i(\theta)-z_i(\theta{-}1)+z_i(\theta{-}1)-z_i(\theta{-}n-1)\\
&\geq\gamma_i(\theta{-}1)U_i(\theta{-}1;\theta,\mathcal{B}(\sigma))-\gamma_i(\theta)U_i(\theta;\theta,\mathcal{B}(\sigma))\\
&+\gamma_i(\theta{-}n-1)U_i(\theta{-}n-1;\theta{-}1,\mathcal{B}(\sigma))-\gamma_i(\theta{-}1)U_i(\theta{-}1;\theta{-}1,\mathcal{B}(\sigma)).
\end{aligned}
\tag{A.6}
$$

We will show that the RHS of equation (A.6) is larger than the RHS of equation (A.5). This is true if

$$
\frac{\gamma_i(\theta{-}n-1)}{\gamma_i(\theta{-}1)}\geq\frac{U_i(\theta{-}1;\theta{-}1,\mathcal{B}(\sigma))-U_i(\theta{-}1;\theta,\mathcal{B}(\sigma))}{U_i(\theta{-}n-1;\theta{-}1,\mathcal{B}(\sigma))-U_i(\theta{-}n-1;\theta,\mathcal{B}(\sigma))}. \tag{A.7}
$$

To verify equation (A.7) we use that upward local incentive constraints and downward local incentive constraints jointly imply that

$$
\frac{\gamma_i(k)}{\gamma_i(k{+}1)}\geq\frac{U_i(k{+}1;k,\mathcal{B}(\sigma))-U_i(k{+}1;k{+}1,\mathcal{B}(\sigma))}{U_i(k;k,\mathcal{B}(\sigma))-U_i(k;k{+}1,\mathcal{B}(\sigma))}. \tag{A.8}
$$

By iteratively applying equation (A.8) the LHS of equation (A.7) can be bounded from below by

$$
\frac{\gamma_i(\theta{-}n-1)}{\gamma_i(\theta{-}1)}\geq\prod_{k=\theta-n}^{\theta-1}\frac{U_i(k,k{-}1,\mathcal{B}(\sigma))-U_i(k;k,\mathcal{B}(\sigma))}{U_i(k{-}1;k{-}1,\mathcal{B}(\sigma))-U_i(k{-}1;k,\mathcal{B}(\sigma))} \tag{A.9}
$$

We thus need to establish that

$$
\prod_{k=\theta-n}^{\theta-1}\frac{U_i(k;k{-}1,\mathcal{B}(\sigma))-U_i(k;k,\mathcal{B}(\sigma))}{U_i(k{-}1;k{-}1,\mathcal{B}(\sigma))-U_i(k{-}1;k,\mathcal{B}(\sigma))}\geq\frac{U_i(\theta{-}1;\theta{-}1,\mathcal{B}(\sigma))-U_i(\theta{-}1;\theta,\mathcal{B}(\sigma))}{U_i(\theta{-}n{-}1;\theta{-}1,\mathcal{B}(\sigma))-U_i(\theta{-}n{-}1;\theta,\mathcal{B}(\sigma))},
$$

or equivalently that

$$\begin{aligned} &\frac{U_i(\theta - n - 1; \theta - 1, \mathcal{B}(\sigma)) - U_i(\theta - n - 1; \theta, \mathcal{B}(\sigma))}{U_i(\theta - 1; \theta - 1, \mathcal{B}(\sigma)) - U_i(\theta - 1; \theta, \mathcal{B}(\sigma))} \\ &\geq \prod_{k=\theta-n}^{\theta-1} \frac{U_i(k-1; k-1, \mathcal{B}(\sigma)) - U_i(k-1; k, \mathcal{B}(\sigma))}{U_i(k; k-1, \mathcal{B}(\sigma)) - U_i(k; k, \mathcal{B}(\sigma))}. \end{aligned} \tag{A.10}$$

The MDR allows us to bound the RHS of equation (A.10) from above by

$$\prod_{k=\theta-n}^{\theta-1} \frac{U_i(k-1; \theta-1, \mathcal{B}(\sigma)) - U_i(k-1; \theta, \mathcal{B}(\sigma))}{U_i(k; \theta-1, \mathcal{B}(\sigma)) - U_i(k; \theta, \mathcal{B}(\sigma))}.$$

Straightforward algebra verifies that this bound is equal to the LHS of equation (A.10).
□

E Statement of the General Problem and of the Lagrangian Objective

For any i, θ, the constraints to the minimization problem are

$$\begin{aligned} \forall \theta' \neq \theta \qquad -(z_i(\theta) - z_{-i}(\theta')) - y_i(\theta; \theta) + y_i(\theta'; \theta) &\leq 0, && (IC) \\ -z_i(\theta) - y_i(\theta; \theta) + V_i(\theta, (p, \rho^V)) &\leq 0, && (PC_i) \\ -1 + \sum_i \sum_\theta p(\theta) z_i(\theta) + Pr(\mathcal{G}) &\leq 0, && (RC) \\ -z_i(\theta) &\leq 0, && (EPI) \\ \gamma(\theta_1, \theta_2) - 1 &\leq 0, && (F) \\ \text{and } \forall Q \subset \Theta & \\ \sum_i \sum_{\theta \in Q_i} z_i(\theta) p(\theta) + \sum_{(\theta_1, \theta_2) \in \bar{Q}} (1 - \gamma(\theta_1, \theta_2)) p(\theta_1) p(\theta_2) - 1 + Pr(\mathcal{G}) &\leq 0. && (EP) \end{aligned}$$

We now derive the Lagrangian representation of the optimization problem. First, we state the complementary slackness conditions and the respective Lagrangian multipliers

$$\begin{aligned} [z_i(\theta) - z_{-i}(\theta') + y_i(\theta; \theta) - y_i(\theta'; \theta)] \nu^i_{\theta, \theta'} &= 0, & \nu^i_{\theta, \theta'} &\geq 0; \\ [z_i(\theta) + y_i(\theta; \theta) - V_i(\theta, (p, \rho^V))] \lambda^i_\theta &= 0, & \lambda^i_\theta &\geq 0; \\ \Big[1 - \sum_i \sum_\theta p(\theta) z_i(\theta) - Pr(\mathcal{G})\Big] \delta &= 0, & \delta &\geq 0; \\ z_i(\theta) \zeta^i_\theta &= 0, & \zeta^i_\theta &\geq 0; \\ [1 - \gamma(\theta_1, \theta_2)] \mu_{\theta_1, \theta_2} &= 0, & \mu_{\theta_1, \theta_2} &\geq 0; \\ \Big[-\sum_i \sum_{\theta \in Q_i} z_i(\theta) p(\theta) - \sum_{\substack{(\theta_1, \theta_2) \\ \in \bar{Q}}} (1 - \gamma(\theta_1, \theta_2)) p(\theta_1) p(\theta_2) + 1 - Pr(\mathcal{G})\Big] \eta_Q &= 0, & \eta_Q &\geq 0. \end{aligned}$$

For any Lagrangian multiplier, say t, we introduce the following notation $\tilde{t} \equiv \frac{t}{\delta}$. Let

Q^2 be the set of all combinations of Q. Let $\tilde{e}^i_\theta := p(\theta) \sum_{Q \in Q^2 | \theta \in Q_i} \tilde{\eta}_Q$ and define

$$\tilde{\Lambda}^i(\theta) := \sum_{k=1}^{\theta} \tilde{\lambda}^i_k, \quad \tilde{E}^i(\theta) := \sum_{k=1}^{\theta} \tilde{e}^i_k, \quad \tilde{Z}^i(\theta) := \sum_{k=1}^{\theta} \tilde{\zeta}^i_\theta \tag{A.11}$$

Next, we characterize the solution in terms of the Lagrangian objective. As stated directly above Corollary 2 we can maximize with respect to $\{Pr(\sigma), \boldsymbol{\rho}(\sigma)\}_{\sigma \in \Sigma}$. Further, consider the symmetrized problem, i.e., the problem in which for each signal σ and associated probabilities $\rho(\theta_1, \theta_2|\sigma)$ there exists signal realization σ' such that $Pr(\sigma) = Pr(\sigma')$ and $\rho(\theta_1, \theta_2|\sigma) = \rho(\theta_2, \theta_1|\sigma')$ for all (θ_1, θ_2).[1]

Lemma 5. *The lottery $\{Pr(\sigma), \boldsymbol{\rho}(\sigma)\}_\sigma$ is an optimal solution to the designers problem if and only if there are Lagrangian multipliers that satisfy complementary slackness given the lottery and the lottery includes every $\boldsymbol{\rho}(\sigma)$ that maximizes*

$$\begin{aligned}\widehat{\mathcal{L}}(\mathcal{B}(\sigma)) :=& \mathcal{T}(\mathcal{B}(\sigma)) + \sum_i \Bigg[\sum_{\theta=1}^{K} \rho_i(\theta|\sigma)\Big(\frac{\mathsf{m}^{\mathsf{i}}_\theta}{p(\theta)}\Big) U_i(\theta;\theta,\mathcal{B}(\sigma)) \\ &+ \sum_{\theta=1}^{K-1} \sum_{k=\theta+1}^{K} \frac{\mathsf{M}^{\mathsf{i}}(\theta) + \nu^i_{k,\theta} - \nu^i(\theta)}{p(\theta)} \rho_i(\theta|\sigma)(U_i(\theta;\theta,\mathcal{B}(\sigma)) - U_i(\theta;k,\mathcal{B}(\sigma))) \\ &- \sum_{\theta=1}^{K} \sum_{k=1}^{\theta-1} \frac{\mathsf{M}^{\mathsf{i}}(\theta) + \nu^i_{k,\theta} - \nu^i(\theta)}{p(\theta)} \rho_i(\theta|\sigma)[U_i(\theta;k,\mathcal{B}(\sigma)) - U_i(\theta;\theta,\mathcal{B}(\sigma))] \Bigg],\end{aligned} \tag{A.12}$$

where $\mathsf{m}^{\mathsf{i}}_\theta := p(\theta) + \tilde{e}^i_\theta - \tilde{\zeta}^i_\theta$, $\mathsf{M}^{\mathsf{i}}(\theta) := \tilde{\Lambda}^i(\theta) - \sum_{k=1}^{k=\theta} p(k) - \tilde{E}^i(\theta) + Z^i(\theta)$, $\nu^i(\theta) := \sum_{k=1}^{\theta} \sum_{v=\theta+1}^{K} [\tilde{\nu}^i_{k,v} - \tilde{\nu}^i_{v,k}]$ *and*

$$\begin{aligned}\mathcal{T}(\mathcal{B}(\sigma)) :=& \sum_{Q \in Q^2} \sum_{(\theta_1,\theta_2) \in \tilde{Q}} [\rho(\theta_1|\sigma)\beta_1(\theta_2|\theta_1,\sigma)]\tilde{\eta}_Q \\ &- \sum_{\theta_1 \times \theta_2} \frac{\rho_1(\theta_1|\sigma)\beta_1(\theta_2|\theta_1,\sigma)}{p(\theta_1)p(\theta_2)} \tilde{\mu}_{\theta_1,\theta_2}.\end{aligned} \tag{A.13}$$

Hence, $\boldsymbol{\rho} = \sum_\sigma Pr(\sigma)\boldsymbol{\rho}(\sigma)$ is a maximizer of the concave hull of the above function. Moreover, the following is true at the optimum:

- *The* (2) *constraint is always binding, i.e., $\delta > 0$.*
- *If the Border constraints are redundant, then $\tilde{e}^i_\theta = \tilde{E}_i(\theta) = 0 = \tilde{Z}^i(\theta) = \tilde{\zeta}^i_\theta$.*
- *If $\tilde{\Lambda}^i(\theta) + \tilde{Z}^i(\theta) - \sum_{v=1}^{v=\theta} p(v) - \tilde{E}^i(\theta) > 0$, then the* downward incentive constraints *are* binding. *If in addition the upward incentive constraints are redundant, then $\tilde{v}^i_{\theta,k} = 0$ for all $k \geq \theta$.*
- *If $\tilde{\Lambda}^i(\theta) + \tilde{Z}^i(\theta) - \sum_{v=1}^{v=\theta} p(v) - \tilde{E}^i(\theta) < 0$, the* upward incentive *constraints are* binding. *If in addition the downward incentive constraints are redundant, then $\tilde{v}^i_{k,\theta} = 0$ for all $k < \theta$.*
- *If local incentive constraints are sufficient, then $\nu^i(\theta) = \tilde{\nu}^i_{\theta,\theta+1} - \tilde{\nu}^i_{\theta+1,\theta}$. In this case, $\tilde{\nu}^i_{k,\theta} - \nu^i(\theta) = -M^i(\theta)$ for any k such that $k > \theta + 1$ or $k < \theta + 1$.*

Proof. The first part of the proof is along the heuristics below Proposition 3. We

[1] We can symmetrize the problem in this way without loss of generality as players are symmetric ex-ante.

manipulate the Lagrangian, $\mathcal{L}$, and derive a tractable dual problem. The second part verifies that the optimum is on the concave hull of the objective. The Lagrangian takes the form

$$\begin{aligned}
\mathcal{L} = {} & Pr(\mathcal{G}) + \delta[-1 + \sum_i \sum_{\theta=1}^{K} p(\theta) z_i(\theta) + Pr(\mathcal{G})] \\
& + \sum_i \sum_{\theta=1}^{K} [-z_i(\theta) - y_i(\theta;\theta) + V_i(\theta,(p,\rho^V))]\lambda^i_\theta \\
& + \sum_i \sum_{\theta=1}^{K} \sum_{k\in\Theta\backslash\theta} [z_i(\theta) - z_i(k) - y_i(k;k) + y_i(\theta;k)]\nu^i_{k,\theta} \\
& + \sum_{Q\in Q^2} \Big[\sum_i \sum_{\theta\in Q_i} z_i(\theta)p(\theta) + \sum_{(\theta_1,\theta_2)\in\bar{Q}} (1-\gamma(\theta_1,\theta_2))p(\theta_1)p(\theta_2) - 1 + Pr(\mathcal{G})\Big]\eta_Q \\
& + \sum_{\theta_1\times\theta_2} [\gamma(\theta_1,\theta_2) - 1]\mu_{\theta_1,\theta_2} - \sum_i \sum_\theta z_i(\theta)\zeta^i_\theta
\end{aligned} \tag{A.14}$$

Using Theorem 1 we optimize over $\{z_i(\cdot), \gamma^\sigma(1,1), \boldsymbol{\rho}(\sigma)\}$, with $\gamma^\sigma(1,1){:=}Pr(\mathcal{G},\sigma|\theta_1{=}1,\theta_2{=}1)$.
Step 1: Eliminating $z_i(\cdot)$ using first order conditions. Define $\nu^i_{K+1,K} := 0 =: \nu^i_{1,0} = \nu^i_{0,1}$ for ease of notation. The FOC w.r.t. $z_i(k)$ are

$$p(k)\delta - \lambda^i_k - \sum_{k\in\Theta\backslash k} [\nu^i_{k,v} - \nu^i_{v,k}] + p(k) \sum_{Q\in Q^2|k\in Q_i} \eta_Q - \zeta^i_k = 0. \tag{A.15}$$

Summing over all K conditions in (A.15) and recalling definition (A.11) yields

$$1 = \tilde{\Lambda}^i(K) - \tilde{E}^i(K) + \tilde{Z}^i(k), \tag{A.16}$$

(A.15) holds for all θ if and only if

$$\nu^i(\theta) := \sum_{k=1}^{\theta} \sum_{v=\theta+1}^{K} [\tilde{\nu}^i_{k,v} - \tilde{\nu}^i_{v,k}] = \sum_{v=1}^{v=\theta} p(v) - \tilde{\Lambda}^i(\theta) + \tilde{E}^i(\theta) - Z^i(\theta). \tag{A.17}$$

Moreover, all terms involving $z_i(\cdot)$ cancel out from (A.14) via (A.15).
Step 2: Reformulation Lagrangian. Given the above necessary conditions, we manipulate the Lagrangian objective to derive a more tractable maximization problem. Define $\eta := \sum_{Q\in Q^2} \eta_Q$ and $\tilde{\eta} := \sum_{Q\in Q^2} \tilde{\eta}_Q$. Next, using Bayes' rule together with the homogeneity established in the proof of Theorem 1 (step 1), applying algebra and using the first-order-conditions it is straightforward to show that (A.14) admits the following representation

$$\mathcal{L} = Pr(\mathcal{G})(1+\delta+\eta) - \delta C - \delta \sum_\sigma \Pr(\mathcal{G},\sigma)\widehat{\mathcal{L}}(\mathcal{B}(\sigma)), \tag{A.18}$$

where C is a constant that is independent of the choice variables,

$$C{:=}1 + \tilde{\eta} - \sum_{Q\in Q^2} \sum_{(\theta_1,\theta_2)\in\tilde{Q}} p(\theta_1)p(\theta_2)\tilde{\eta}_Q - \sum_i \sum_\theta \tilde{\lambda}_\theta V_i(\theta,(p,\rho^V)) + \sum_{\theta_1\times\theta_2} \tilde{\mu}_{\theta_1,\theta_2} < 0.$$

Define $\gamma^\sigma(\theta_1,\theta_2) := Pr(\mathcal{G},\sigma|\theta_1,\theta_2)$. From the proof of Theorem 1 (step 1) with $\alpha^\sigma = \gamma^\sigma(1,1)$ it follows that $\gamma^\sigma(\theta_1,\theta_2) = f(\mathcal{B}(\sigma),\theta_1,\theta_2)\gamma^\sigma(1,1)$, where $f(\mathcal{B}(\sigma),\theta_1,\theta_2)$ is a positive real number. Thus, $Pr(\mathcal{G},\sigma) = \gamma^\sigma(1,1)R(\mathcal{B}(\sigma))$ with $R(\mathcal{B}(\sigma)) := \sum_{\theta_1\times\theta_2} p(\theta_1)p(\theta_2)f(\mathcal{B}(\sigma),\theta_1,\theta_2)$. Plugging into (A.18) yields

$$\mathcal{L} = \sum_\sigma \gamma^\sigma(1,1)R(\mathcal{B}(\sigma))(1+\delta+\tilde{\mu}) - \delta C - \delta\sum_\sigma \gamma^\sigma(1,1)R(\mathcal{B}(\sigma))\widehat{\mathcal{L}}(\mathcal{B}(\sigma)). \tag{A.19}$$

The FOC of (A.19) w.r.t. $\gamma^\sigma(1,1)$ is

$$R(\mathcal{B}(\sigma))\Big((1+\delta+\eta) - \delta\widehat{\mathcal{L}}(\mathcal{B}(\sigma))\Big) = 0, \tag{A.20}$$

for each signal. By Assumption 1 $R(\mathcal{B}(\sigma)) > 0$ and thus, $\widehat{\mathcal{L}}(\mathcal{B}(\sigma)) - 1 > 0$ if $\gamma^\sigma(1,1) > 0$. Therefore, $\delta = (1+\eta)(\widehat{\mathcal{L}}(\mathcal{B}(\sigma)) - 1)^{-1}$. As δ is independent of σ, $\widehat{\mathcal{L}}(\mathcal{B}(\sigma))$ takes the same value for each signal realization. Substituting into (A.19) and simplifying yields

$$\mathcal{L} = \frac{-C(1+\eta)}{\widehat{\mathcal{L}}(\mathcal{B}(\sigma)) - 1} \tag{A.21}$$

which is minimized if and only if $\widehat{\mathcal{L}}(\mathcal{B}(\sigma))$ is maximized.[2]

Thus, for the optimal multipliers one constructs the concave hull of $\widehat{\mathcal{L}}$ by taking spreads over those $\mathcal{B}(\sigma)$ that are a global maximum of $\widehat{\mathcal{L}}$. If there are multipliers and a unique maximizer, $\mathcal{B}(\sigma)$, that satisfies the complementary slackness conditions, signals do not improve. If there are multiple global optima and there is a spread that satisfies the complementary slackness conditions, signals improve. □

F Binding constraint for Games with Linearity in Types

In this part we state a more general version of Lemma 1 if games satisfy linearity in types, but Assumption 3 and MDR may not hold. The proof is analogous to that of Lemma 1. We then characterize the set of binding constraints C_R under this relaxed condition.

Lemma 6. *The following holds for the optimal mechanism*

i. all incentive compatibility constraints not concerning adjacent types are redundant,
ii. if both local incentive compatibility constraints are redundant for type θ_i, then her participation constraint is satisfied with equality or $z_i(\theta_i) = 0$,
iii. the participation constraints for at least one type of every player is binding.

Although the set of binding constraints may depend on the exact location of the optimum, Lemma 6 provides enough structure to determine settlement values as a function of escalation values. This reduces the dimensionality of the choice set.

Corollary 3. *Consider the escalation values $\{y_i(\theta_i;\theta_i), y_i(\theta_i{+}1;\theta_i), y_i(\theta_i{-}1;\theta_i)\}$ for all $\theta_i \in \Theta$ of the optimal mechanism. Then there is a partition of $P(\Theta) = \{\Theta^{z_i=0}, \Theta^{PC}, \Theta^{IC^+}, \Theta^{IC^-}\}$ such that*

$$z_i(\theta_i) = z_i(\tilde{\theta}_i) + y_i(\tilde{\theta}_i;\theta_i) - y_i(\theta_i;\theta_i), \quad \forall\theta_i \in \{\Theta^{IC^+}, \Theta^{IC^-}\} \tag{Z}$$

[2] The Lagrangian multipliers are necessarily such that C is negative at the optimum. Otherwise (A.14) and (A.21) imply that $Pr(\mathcal{G})$ is negative, a contradiction to Assumption 1 and /or the fact that a feasible solution to the minimization problem always exists: take a degenerate signal distribution and set $\gamma(\theta_1,\theta_2) = 1$ for all type profiles.

with

$$\tilde{\theta}_i = \begin{cases} \theta_i + 1 & \textit{if } \theta_i \in \Theta^{IC^+} \\ \theta_i - 1 & \textit{if } \theta_i \in \Theta^{IC^-}. \end{cases}$$

Moreover $z_i(\theta_i) = 0$ *for* $\theta_i \in \Theta^{z_i=0}$*, and* $z_i(\theta_i) = V_i(\theta_i, (p, \rho^V)) - y_i(\theta_i; \theta_i)$ *for* $\theta_i \in \Theta^{PC}$*.*

G Consistent Belief System

We first confine ourselves to interior belief systems, i.e., belief systems that do not have zero entries.

We first derive a representation that links the belief system, $\mathcal{B}(\sigma)$ to the escalation rule, γ^σ. We interpret every type profile as a node in a network. Node (θ_1, θ_2) has the value $\gamma^\sigma(\theta_1, \theta_2)$. The (values of the) nodes are linked using a *transition* function $q_i(\theta_1, \theta_2)$.

Observation 5. Consider nodes (θ_i, θ'_{-i}), (θ_i, θ_{-i}) and the *transition* function $q_i(\theta'_{-i}, \theta_{-i}|\theta_i) = \frac{p(\theta_{-i})}{p(\theta'_{-i})} \frac{\beta_i(\theta'_{-i}|\theta_i,\sigma)}{\beta_i(\theta_{-i}|\theta_i,\sigma)}$. Then,

$$\gamma^\sigma(\theta'_1, \theta_2) = q_2(\theta'_1, \theta_1|\theta_2)\gamma^\sigma(\theta_1, \theta_2),$$

$$\gamma^\sigma(\theta_1, \theta'_2) = q_1(\theta'_2, \theta_2|\theta_1)\gamma^\sigma(\theta_1, \theta_2).$$

The result follows using Bayes' rule as

$$q_i(\theta'_{-i}, \theta_{-i}|\theta_i) = \frac{p(\theta_{-i})}{p(\theta'_{-i})} \frac{\gamma^\sigma(\theta_i, \theta'_{-i})}{\gamma^\sigma(\theta_i, \theta_{-i})} \frac{p(\theta'_{-i})}{p(\theta_{-i})}.$$

Fix two nodes in the network, say (θ_1, θ_2) and (k_1, k_2). There are several paths that connect the two nodes. Starting from (θ_1, θ_2) we can go to (k_1, θ_2) and then to (k_1, k_2). Equivalently we can approach (k_1, k_2) through (θ_1, k_2). Bayes' rule implies that both paths have the same length, or, the values of the nodes are the same. Using Observation 5 we have that

$$\gamma^\sigma(k_1, k_2) = q_1(k_2, \theta_2|k_1)q_2(k_1, \theta_1|\theta_2)\gamma^\sigma(\theta_1, \theta_2),$$
$$\gamma^\sigma(k_1, k_2) = q_2(k_1, \theta_1|k_2)q_1(k_2, \theta_2|\theta_1)\gamma^\sigma(\theta_1, \theta_2).$$

Definition 11 (Bayes' consistency)**.** A belief-system is Bayes' consistent if for every (θ_1, θ_2), (k_1, k_2)

$$q_1(k_2, \theta_2|k_1)q_2(k_1, \theta_1|\theta_2) = q_2(k_1, \theta_1|k_2)q_1(k_2, \theta_2|\theta_1). \tag{A.22}$$

Definition 12 ($(1,1)$-Consistent)**.** A belief-system is $(1,1)$-consistent if for every (θ_1, θ_2),

$$q_1(\theta_2, 1|\theta_1)q_2(\theta_1, 1|1) = q_2(\theta_1, 1|\theta_2)q_1(\theta_2, 1|1). \tag{A.23}$$

Lemma 7. *An interior belief system is Bayes' consistent if and only if it is* $(1,1)$*-consistent.*

Proof. Bayes' consistency trivially implies (1,1)-consistency. For the reverse direction take any $\gamma^\sigma(k_1, k_2)$ and $\gamma^\sigma(\theta_1, \theta_2)$. We want to show that (1,1)-consistency implies

$$q_1(k_2, \theta_2|k_1)q_2(k_1, \theta_1|\theta_2) = q_2(k_1, \theta_1|k_2)q_1(k_2, \theta_2|\theta_1).$$

By (1,1)-consistency we know that

$$q_1(\theta_2, 1|k_1)q_2(k_1, 1|1) = q_2(k_1, 1|\theta_2)q_1(\theta_2, 1|1), \tag{A.24}$$
$$q_1(k_2, 1|\theta_1)q_2(\theta_1, 1|1) = q_2(\theta_1, 1|k_2)q_1(k_2, 1|1), \tag{A.25}$$
$$q_1(\theta_2, 1|\theta_1)q_2(\theta_1, 1|1) = q_2(\theta_1, 1|\theta_2)q_1(\theta_2, 1|1), \tag{A.26}$$
$$q_1(k_2, 1|k_1)q_2(k_1, 1|1) = q_2(k_1, 1|k_2)q_1(k_2, 1|1). \tag{A.27}$$

Plugging into $q_i(\theta_{-i}, 1|\theta_i)q_{-i}(\theta_i, 1|1) = q_{-i}(\theta_i, 1|\theta_{-i})q_i(\theta_{-i}, 1|1)$, and rearranging yields

$$\frac{\beta_i(\theta_{-i}|\theta_i, \sigma)}{\beta_{-i}(\theta_i|\theta_{-i}, \sigma)} = \frac{\beta_i(1|\theta_i, \sigma)}{\beta_{-i}(1|\theta_{-i}, \sigma)} \frac{\beta_i(\theta_{-i}|1, \sigma)}{\beta_i(1|1, \sigma)} \frac{\beta_{-i}(1|1, \sigma)}{\beta_{-i}(\theta_i|1, \sigma)}. \tag{A.28}$$

Arrange all equations (A.24) to (A.27) according to (A.28). Observe on these transformations that the RHS of (A.24) times that of (A.25) is the same as the RHS of (A.26) times that of (A.27). Using the respective LHS of the equations yields

$$\begin{aligned} & \frac{\beta_1(k_2|k_1, \sigma)}{\beta_1(\theta_2|k_1, \sigma)} \frac{\beta_2(k_1|\theta_2, \sigma)}{\beta_2(\theta_1|\theta_2, \sigma)} = \frac{\beta_2(k_1|k_2, \sigma)}{\beta_2(\theta_1|k_2, \sigma)} \frac{\beta_1(k_2|\theta_1, \sigma)}{\beta_1(\theta_2|\theta_1, \sigma)} \\ \Leftrightarrow \quad & q_1(k_2, \theta_2|k_1)q_2(k_1, \theta_1|\theta_2) = q_2(k_1, \theta_1|k_2)q_1(k_2, \theta_2|\theta_1). \end{aligned}$$

□

Using Lemma 7 consistency is reduced to (1,1)-consistency. Thus, any consistent belief system is implemented by some escalation rule only if every node (θ_1, θ_2) has a value weakly below 1. The value of a node is given by the length of the path connecting the node with the initial node, i.e.,

$$\gamma^\sigma(\theta_1, \theta_2) = q_1(\theta_2, 1|\theta_1)q_2(\theta_1, 1|1)\gamma^\sigma(1, 1). \tag{A.29}$$

Note that the above exposition implies the following:

Observation 6. An interior belief-system can be attained via some escalation rule if and only if it is Bayes' consistent.

Finally, observe that Bayes' consistency is equivalent to the definition of consistency given in the main text.

Finally, we show that it is without loss of generality to focus on interior belief systems.

Lemma 8. *A belief system can be implemented if and only if it can be approximated by a convergent sequence of implementable interior-belief-systems.*

Proof. Take a sequence of consistent $\mathcal{B}_n(\sigma) \to \mathcal{B}(\sigma)$. $\mathcal{B}_n(\sigma)$ is consistent, thus Observation 6 implies some function $f : \mathcal{B}(\sigma) \to [0, 1]^{K\times K}$, such that $f(\mathcal{B}_n(\sigma)) = \gamma_n^\sigma$ with γ_n^σ implementing $\mathcal{B}_n(\sigma)$. Since f is continuous, $\lim_{n\to\infty} f(\mathcal{B}_n(\sigma)) = f(\lim_{n\to\infty} \mathcal{B}_n(\sigma)) = \gamma^\sigma$. Equation (A.23) can be rewritten as

$$g_L(\mathcal{B}(\sigma)) = g_R(\mathcal{B}(\sigma)), \tag{A.30}$$

where both g_L and g_R are continuous functions from belief systems to $\mathbb{R}$. Because $g_L(\mathcal{B}_n(\sigma)) - g_H(\mathcal{B}_n(\sigma)) = 0$, we can conclude that $g_L(\mathcal{B}(\sigma)) - g_R(\mathcal{B}(\sigma)) = \lim_{n\to\infty}[g_L(\mathcal{B}_n(\sigma)) - g_R(\mathcal{B}_n(\sigma))] = 0$ and $\mathcal{B}(\sigma)$ satisfies equation (A.23).

Conversely, take any $\mathcal{B}(\sigma)$ being implemented by some γ^σ. We show that we can find a sequence of interior belief-systems that are consistent and converge to $\mathcal{B}(\sigma)$: Let

$\hat{\gamma}^{\sigma}$ be the escalation rule so that $\mathcal{B}(\sigma)$ is implemented by $\hat{\gamma}^{\sigma}$. Choose a sequence of escalation rules in the interior that converges to $\hat{\gamma}^{\sigma}$. By Bayes' rule every element of the sequence, γ_n^{σ}, corresponds to some belief system $\mathcal{B}_n(\sigma)$. Moreover, Observation 6 implies that there exists a continuous function, say $f^{-1} : [0,1]^{K\times K} \to [0,1]^{K\times K}$, such that $f^{-1}(\gamma_n^{\sigma}) = \mathcal{B}_n(\sigma)$ and $\mathcal{B}_n(\sigma)$ satisfies equation (A.23). Because f^{-1} is continuous, $\lim_{n\to\infty} B_n(\sigma) = \lim_{n\to\infty} f^{-1}(\gamma_n^{\sigma}) = f^{-1}(\hat{\gamma}^{\sigma}) = \mathcal{B}(\sigma)$. □

Lemma 9. *Let $\mathcal{O}$ be a continuous function defined on the domain of $\gamma \in [0,1]\cap C$ where C consists of those γ's that satisfy a given set of weak inequality constraints, each of which is continuous in γ. Then,* $\arg\max_{\gamma\in[0,1]|\gamma\in C} \mathcal{O}(\gamma) = \arg\sup_{\gamma\in(0,1)|\gamma\in C} \mathcal{O}(\gamma)$.

Proof. Without loss of generality suppose the argument that maximizes $\mathcal{O}$, γ^*, gives rise to a non-interior belief system, $\mathcal{B}^*$. Then, Lemma 8 implies that we can approximate $\mathcal{B}^*$ by a convergent sequence of consistent interior belief-systems. Because $\mathcal{O}$ is continuous in γ (and through Observation 1 continuous in $\mathcal{B}$), it follows that $\lim_{n\to\infty} \mathcal{O}(\mathcal{B}_n) = \mathcal{O}(\mathcal{B}^*)$. Moreover, because the constraints are inequality constraints and continuous in γ (and $\mathcal{B}$), there is n' such that every element $\mathcal{B}_n$ with $n > n'$ satisfies the constraints. Therefore, $\max_{\gamma\in[0,1]|\gamma\in C} \mathcal{O}(\gamma) = \sup_{\gamma\in(0,1)|\gamma\in C} \mathcal{O}(\gamma)$ and $B^* = \lim_{n\to\infty} \mathcal{B}_n$. Using Lemma 8 we note that for every $\mathcal{B}_n$ there is γ_n so that $\lim_{n\to\infty} \gamma_n = \gamma^*$. □

H The Belief Management Approach and Ex-post Equilibria

H.1 Simple Lotteries (cf. Hörner, Morelli, and Squintani, 2015)

Setting. Consider the general setup from Section 4. However, the escalation game is different. Participation requires a fixed cost $0 \le c < 1/2$ and a player's strategy choice is constant in the information structure.[3] To emphasize this effect, we suppress action choices in the payoff function. It is

$$u(\theta_i, \theta_{-i}) = \begin{cases} 1/2 - c, & \text{if } \theta_i = \theta_{-i} \\ 1/2 - c + \xi, & \text{if } \theta_i > \theta_{-i} \\ 1/2 - c - \xi, & \text{if } \theta_i < \theta_{-i}, \end{cases}$$

where $\xi > c$ is the marginal payoff of a better case. Let $b_i(m_i) := \beta_i(1|m_i)$ be the probability mass on type 1, and $\mathbb{1}_1 = 1$ if θ_i=1 and 0 otherwise. Expected payoffs are

$$U_i(m_i; \theta_i, \mathcal{B}) = b_i(m_i)u(\theta_i, 1) + (1 - b_i(m_i))u(\theta_i, K) = 1/2 - c - b_i(m_i)\xi + \mathbb{1}_1\xi.$$

Optimal conflict management maximizes $\sum_i (\mathbb{E}[U_i|\mathcal{B}] + \mathbb{E}[\Psi_i|\mathcal{B}])$ subject to constraints C_F. K-types receive no virtual rent by definition, thus $\Psi(1, \mathcal{B}) = w(1)D_i(1; K, \mathcal{B})$.

Moreover, the ability premium, $D_i(1; K, \mathcal{B}) = U_i(1; 1, \mathcal{B}) - U_i(1; K, \mathcal{B}) = \xi$, is constant. Hence, $\mathbb{E}[\Psi_i|\mathcal{B}] = \rho_i(1)w(1)\xi$ (linearly) increases in the post-escalation likelihood, $\rho_i(1)$. By design, the joint expected payoffs, $\sum_i \mathbb{E}[U_i|\mathcal{B}] = 1 - 2c$, are independent of $\mathcal{B}$. Thus, the belief system that maximizes $\sum_i (\mathbb{E}[U_i|\mathcal{B}] + \mathbb{E}[\Psi_i|\mathcal{B}])$ is $(b_i(1) = 1, b_i(K) = 0)$ implying $\rho_i(1) = 1$. We denote this belief system by $\widetilde{\mathcal{B}}$.

[3] In fact, we use the term *lottery* to emphasize that we could model the situation without the need of any formulation of a *game*. However, the application suggests existence of a *continuation game* with an ex-post equilibrium. The two are observationally equivalent.

The belief system $\widetilde{\mathcal{B}}$ *sorts* type profiles into the set of *easy-to-settle matches*, $\{(K,K), (1,K), (K,1)\}$, and that of *difficult-to-settle matches*, $\{(1,1)\}$. Under $\widetilde{\mathcal{B}}$ no type-1 player wants to mimic a type-K player. However, it is possible that the reduced-form mechanism implementing $\widetilde{\mathcal{B}}$ is not feasible. This happens if the worst escalation rule leading to $\widetilde{\mathcal{B}}$ does not satisfy the designer's resource constraint. By occasional escalation of $(1,K)$ and $(K,1)$ the designer increases the escalation utility of type-1 players and, in turn, decreases their promised shares until the solution is resource feasible; (K,K)-matches remain settled for sure. The entire program is linear, and signals never improve.

Proposition 5. *Suppose the escalation game is the simple lottery described above. Optimal conflict management induces an on-path belief system in the escalation game with beliefs that are*

- *type-dependent, $b_i(1) \neq b_i(K)$, and*
- *degenerate for K-types, $b_i(K) = 1$.*

Furthermore, it is without loss of generality to assume that beliefs are symmetric, *$b_1(\theta_i) = b_2(\theta_i)$, and to ignore public signals.*

H.2 Generalization: Monotone Lotteries

In this section we provide a general solution algorithm for *monotone lottery games* with an arbitrary number of types. The solution nests that of the previous paragraph, Fey and Ramsay (2011), Hörner, Morelli, and Squintani (2015), and the monotone cases in Bester and Wärneryd (2006).[4] Extension to the non-monotone case is straightforward.

Definition 13 (Lottery Game)**.** $\mathcal{G}$ is a lottery game, if $u(\theta_i, \theta_{-i}, \mathcal{B})$ is constant in $\mathcal{B}$.

Consistent with the main text we suppress the argument $\mathcal{B}$ in u.

Definition 14 (Monotone Lottery Game)**.** A lottery game is monotone if $u(\theta_i, \theta_{-i}) - u(\theta_i + 1, \theta_{-i})$ is weakly decreasing in θ_i and θ_{-i} and the prior p induces a weakly decreasing inverse hazard rate $w(\theta_i)$.

The lottery feature entails that expected continuation payoffs are linear in any $\beta_i(\cdot|m_i)$. Therefore, the problem is reminiscent of a standard mechanism design problem with interdependent values. The solution is thus given by a linear program.

We define the ex-ante probability that the type profile (θ, k) will face each other in litigation as $\rho(\theta, k) := \rho_1(\theta)\beta_1(k|\theta)$ with $\sum_{(\theta,k)} \rho(\theta, k) = 1$. Naturally, $\rho(1,1) = 1$ implies $\beta(1|1) = 1$ and 0 for all other beliefs. Substituting for $\beta_i(\theta_i|\theta_{-i})$ and rearrange yields the function $\widetilde{VV}(\theta_1, \theta_2)$, which is independent of any $\rho(\cdot,\cdot)$. The objective of $(\mathrm{P}^{\mathcal{B}}_{max})$ becomes

$$\Xi(\rho(\cdot,\cdot)) := 2u(K,K) + \sum_{\Theta^2 \backslash (K,K)} \rho(\theta_1, \theta_2)\widetilde{VV}(\theta_1, \theta_2). \tag{A.31}$$

We define the condition

$$(2V(1,(p,\underline{\rho}^V)) - 1)\rho(1,1) \leq (p(1))^2 \left(\Xi(\rho(\cdot,\cdot)) - 1\right). \tag{A.32}$$

Definition 15 (Top-Down Algorithm)**.** Denote the set of type pairs (k, θ) such that $\rho(k,\theta) > 0$ by Θ^2_+. Start with $\Theta^2_+ = \emptyset$.

1. Set $\rho(1,1) = 1$ and check whether condition (A.32) is satisfied. If it is satisfied, then terminate. Otherwise continue at 2.

[4] A matlab program implementing the algorithm is available from the authors.

2. Identify the set $\Theta_N^2 = \{(\theta_1, \theta_2) | (\theta_1, \theta_2) = \arg\max_{\Theta^2 \setminus \Theta_+^2} \widetilde{VV}(\theta_1, \theta_2)\}$.
   (a) Set $\rho(1,1)$ to the solution of

$$\sum_{\substack{(\theta_1,\theta_2)\in \\ \Theta_+^2 \cup \Theta_N^2}} \frac{p(\theta_1)p(\theta_2)}{(p(1))^2} \rho(1,1) = 1. \tag{A.33}$$

   (b) Replace $\rho(\theta_1, \theta_2) = \frac{p(\theta_1)p(\theta_2)}{(p(1))^2} \rho(1,1)\ \forall (\theta_1, \theta_2) \in \Theta_+^2 \cup \Theta_N^2$.
   (c) Check whether condition (A.32) is satisfied. If it is satisfied, decrease all ρ for the set Θ_N^2 at the expense $\rho(1,1)$ keeping the relation of 2(b) for $\Theta_+^2 \setminus \Theta_N^2$ until the condition holds with equality. Then, terminate. If (A.32) is violated, repeat step 2.

Proposition 6. *Suppose the escalation game is a monotone lottery. Then optimal conflict management is the solution to the top-down algorithm.*

Proof. Define

$$A_i(\theta_i, \theta_{-i}) := u(\theta_i, \theta_{-i}) - u(\theta_i{+}1, \theta_{-i}) \text{ and } W(\theta, k) = u(\theta_1, \theta_2) + u(\theta_2, \theta_1),$$

$$\textit{and } \widetilde{VV}(\theta, k) := \omega(\theta) A_1(\theta, k) + \omega(\theta) A_2(k, \theta) + W(\theta, k) - W(K, K).$$

The objective becomes

$$\sum_i \sum_{\theta_i \in \Theta} \rho(\theta_i) \left(\sum_{\theta_{-i} \in \Theta} \beta_i(\theta_{-i}|\theta_i) u(\theta_i, \theta_{-i}) + \sum_{\theta_{-i} \in \Theta} \beta_i(\theta_{-i}|\theta_i) \omega(\theta_i) A_i(\theta_i, \theta_{-i}) \right). \tag{A.34}$$

Substituting in for $\rho(\theta_1, \theta_2)$ we get

$$\max_{\rho(\cdot,\cdot)} \Big(2u(K,K) + \sum_{(\theta,k) \in \Theta \setminus (K,K)} \rho(\theta, k) \widetilde{VV}(\theta, k) \Big), \tag{A.35}$$

were we treat $A_i(K, \theta_{-i})$ as an arbitrary positive real number to avoid case distinctions. Simple algebra shows that $\widetilde{VV}$ is non-decreasing for monotone lotteries. Substituting $\rho(\theta_1, \theta_2) = (\gamma(\theta_1, \theta_2) p(\theta_1) p(\theta_2)) / Pr(\mathcal{G})$ and rearranging yields that conditions (A.32) and (A.33) are necessary and sufficient for the feasibility constraint (2).

Hence, the top-down algorithm point-wise maximizes (A.35) subject to (2) by construction. What remains is to show that any ignored constraint is satisfied. We show this using monotonicity, i.e., $p(\theta_i)\rho(\theta_i{-}1, \theta_{-i}) \geq p(\theta_i{-}1)\rho(\theta_i, \theta_{-i})$, for all θ_i, θ_{-i}.

Monotonicity trivially holds if $\gamma(1,1) \neq 1$, as it implies an optimum at $\rho(1,1) = 1$. Now, assume $\gamma(1,1) = 1$. By Bayes' rule

$$\gamma(\theta_1, \theta_2) = \frac{\rho(\theta_1, \theta_2)}{p(\theta_1)p(\theta_2)} Pr(\mathcal{G}).$$

When $\gamma(\theta_1, \theta_2) > 0$, then $\gamma(\theta_1{-}1, \theta_2) = 1$ and

$$Pr(\mathcal{G})\rho(\theta_1{-}1, \theta_1) = p(\theta_1{-}1)p(\theta_2),$$
$$Pr(\mathcal{G})\rho(\theta_1, \theta_1) \leq p(\theta_1)p(\theta_2).$$

Thus, monotonicity holds since

$$p(\theta_1)\rho(\theta_1{-}1,\theta_1) \geq p(\theta_1{-}1)\rho(\theta_1,\theta_1).$$

Monotonicity implies that all but the local downward incentive constraints are redundant as

$$\sum_{\theta_{-i}=1}^{K} \left(p(\theta_i')\rho(\theta_i,\theta_{-i}) - p(\theta_i)\rho(\theta_i',\theta_{-i})\right)\left[u(\theta_i,\theta_{-i}) - u(\theta_i',\theta_{-i})\right] \geq 0, \tag{A.36}$$

for all θ_i and $\theta_i' > \theta_i$. Moreover, $z_i(\theta_i) - z_i(\theta_i{-}1) = y_i(\theta_i;\theta_i{-}1) - y_i(\theta_i;\theta_i) \geq 0$ which implies that $z_i(\theta_i) \geq 0$. Finally, we claim that only the lowest types' participation constraint binds at the optimum. It is straightforward to verify that this implies that downward local incentive constraints hold with equality. We verify the claim by induction. We first show that $\Pi_i(2;2) \geq V(2,(p,\rho^V))$. By local incentive compatibility, and $\Pi_i(1;1) \geq V(1,(p,\rho^V))$,[5] we know that $\Pi_i(2;2) = V(2,(p,\rho^V)) - y_i(1;1) + y_i(1;2)$. We want to show that $V(1,(p,\rho^V)) - V(2,(p,\rho^V)) \geq y_i(1;1) - y_i(1;2)$. The game is a simple lottery, so the condition the reduces to

$$\sum_{\theta_{-i}}[p(\theta_{-i}) - \gamma_i(1)\beta_i(\theta_{-i}|1)][u(1,\theta_{-i}) - u(2,\theta_{-i})] \geq 0.$$

Moreover,

$$[p(\theta_{-i}) - \gamma_i(1)\beta_i(\theta_{-i}|1)] = \frac{1}{p(1)}[Pr(1,\theta_{-i}) - Pr(1,\theta_{-i},\mathcal{G})] \geq 0.$$

Assume that $\Pi_i(\theta_i;\theta_i) \geq V(\theta_i,(p,\rho^V))$ for all $\theta_i < K$. Using a similar argument verifies that $V(K-1,(p,\rho^V)) - V(K,(p,\rho^V)) \geq y_i(K-1,K-1) - y_i(K-1,K)$.

Hence, ignoring all participation constraints, but those of type 1 is without loss. As $z_i(\theta_i)$ is weakly increasing, downward local incentive constraints are binding. □

[5] In the optimal mechanism it holds that $\Pi_i(1;1) = V(1,(p,\rho^V))$.