

SUPPLEMENT TO “COALITIONS IN REPEATED GAMES”

S. NAGEEB ALI

Department of Economics, Pennsylvania State University

CE LIU

Department of Economics, Michigan State University

The Supplementary Appendix is organized as follows:

- [Section 1](#) completes the proof of Theorem 5.
- [Section 2](#) proves results for labor market matching (Propositions 1 to 4).
- [Section 3](#) formalizes the concept of Perfect Coalition Proof Equilibrium and establishes the counterpart of Theorem 1.
- [Section 4](#) describes our application to bargaining games with veto players.

1. COMPLETING THE PROOF OF THEOREM 5

1.1. Preliminary Results

We first prove a few preliminary results used in the argument. The first result shows that when checking profitable blocking plans, we can WLOG focus on those with bounded total continuation values.

**LEMMA 7:** *Let  $\sigma$  be a PCE. Suppose coalition  $C$  has blocking plan  $(\alpha, \beta)$  such that  $\sum_{i \in C} U_i(\bar{h}|\alpha, \beta) > \sum_{i \in C} U_i(\bar{h}|\sigma)$  for some  $\bar{h} \in \bar{\mathcal{H}}$ , then  $C$  has blocking plan  $(\alpha', \beta')$  such that  $\sum_{i \in C} U_i(\bar{h}|\alpha', \beta') > \sum_{i \in C} U_i(\bar{h}|\sigma)$ , and  $\{\sum_{i \in C} U_i(h|\alpha', \beta') : h \in \bar{\mathcal{H}}\}$  is bounded.*

**PROOF:** We break this argument into two parts.

**Part 1:** We show that the set  $\{\sum_{i \in C} U_i(h|\alpha, \beta) : h \in \bar{\mathcal{H}}\}$  is bounded from above. To this end, it suffices to show that the set of stage-game payoffs from the blocking plan,  $\{\sum_{i \in C} u_i(\alpha(h), \beta(h)) : h \in \bar{\mathcal{H}}\}$  is bounded from above.

Consider an arbitrary coalition  $C \in \mathcal{C}$  and an arbitrary history  $h \in \bar{\mathcal{H}}$ . Let  $\tilde{a} = a(h|\sigma)$  denote the default alternative specified by  $\sigma$  and  $\tilde{T} = T(h|\sigma)$  denote the default transfers. By the definition of a blocking plan,  $\alpha(h) \in E_C(\tilde{a})$  and  $\beta(h) = (T'_C, \chi^C(\tilde{T}))$  for some  $T'_C \in \mathcal{T}_C$ . Since the transfers  $T'_C$  may involve nonzero transfers to players outside of  $C$ , we have

$$\begin{aligned} \sum_{i \in C} u_i(\alpha(h), \beta(h)) &= \sum_{i \in C} v_i(\alpha(h)) + \sum_{i \in C, j \notin C} \chi_{ji}^C(\tilde{T}) - \sum_{i \in C, j \notin C} T'_{ij} \\ &\leq \sum_{i \in C} v_i(\alpha(h)) + \sum_{i \in C, j \notin C} \chi_{ji}^C(\tilde{T}) \end{aligned} \quad (6)$$

Now suppose the coalition  $C$  blocks at history  $h$  and chooses alternative  $\alpha(h) \in E_C(\tilde{a})$ ; however, instead of  $\beta(h) = (T'_C, \chi^C(\tilde{T}))$ ,  $C$  chooses transfers  $(T''_C, \chi^C(\tilde{T}))$ , where the transfers  $T''_C$  are such that members of  $C$  make zero payment to players outside of  $C$  while splitting the total payoff within  $C$  evenly. If  $C$  carries out this block, each member  $i \in C$  obtains a

discounted utility of at least

$$(1 - \delta) \frac{1}{|C|} \left[ \sum_{i \in C} v_i(\alpha(h)) + \sum_{i \in C, j \notin C} \chi_{ji}^C(\tilde{T}) \right] + \delta \inf_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma),$$

whereas adhering to  $\sigma$  at  $h$  yields each member at most  $\sup_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma)$ . Since  $\sigma$  is a PCE,  $(\alpha(h), T_C'', \chi^C(\tilde{T}))$  cannot be a profitable block for  $C$ , so it must be true that

$$(1 - \delta) \frac{1}{|C|} \left[ \sum_{i \in C} v_i(\alpha(h)) + \sum_{i \in C, j \notin C} \chi_{ji}^C(\tilde{T}) \right] + \delta \inf_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma) \leq \sup_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma).$$

Combining the inequality above with (6) yields

$$(1 - \delta) \frac{1}{|C|} \left[ \sum_{i \in C} u_i(\alpha(h), \beta(h)) \right] + \delta \inf_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma) \leq \sup_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma).$$

Rearranging terms, we have

$$\begin{aligned} \sum_{i \in C} u_i(\alpha(h), \beta(h)) &\leq \frac{|C|}{1 - \delta} \left[ \sup_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma) - \delta \inf_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma) \right] \\ &\leq \frac{|C|}{1 - \delta} \left| \sup_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma) \right| + \frac{|C|\delta}{1 - \delta} \left| \inf_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma) \right|. \end{aligned}$$

Since  $\{U(h|\sigma) : h \in \overline{\mathcal{H}}\}$  is bounded by Assumption 2, there exists  $L > 0$  such that

$$\left| \sup_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma) \right| \leq L \quad \text{and} \quad \left| \inf_{h \in \overline{\mathcal{H}}, i \in N} U_i(h|\sigma) \right| \leq L.$$

Therefore,

$$\sum_{i \in C} u_i(\alpha(h), \beta(h)) \leq \frac{1 + \delta}{1 - \delta} |C| L.$$

Note that the inequality above holds for all  $h \in \overline{\mathcal{H}}$  while the right hand side does not depend on  $h$ , so our claim follows.

**Part 2:** It is without loss to assume that  $\{\sum_{i \in C} U_i(h|\alpha, \beta) : h \in \overline{\mathcal{H}}\}$  is bounded from below. If not, we can construct another blocking plan  $(\alpha', \beta')$  such that  $\sum_{i \in C} U_i(\hat{h}|\alpha', \beta') > \sum_{i \in C} U_i(\hat{h}|\sigma)$  while ensuring  $\{\sum_{i \in C} U_i(h|\alpha', \beta') : h \in \overline{\mathcal{H}}\}$  is bounded from below: if  $\sum_{i \in C} U_i(\hat{h}|\alpha, \beta)$  falls below  $\min_{a \in A} \sum_{i \in C} v_i(a)$  for some history  $\hat{h} \in \overline{\mathcal{H}}$ , we will ask  $C$  to refuse all outgoing transfers at all histories following  $\hat{h}$ .

Formally, for a history  $\hat{h} \in \overline{\mathcal{H}}$ , let  $F(\hat{h}) := \{\hat{h}h : h \in \overline{\mathcal{H}}\}$  denote the set of histories that can follow from  $\hat{h}$ . Let  $\underline{H}_C := \{h \in \overline{\mathcal{H}} : \sum_{i \in C} U_i(h|\alpha, \beta) < \min_{a \in A} \sum_{i \in C} v_i(a)\}$ . Let  $\mathbf{0}_C$  denote the vector of zero-valued transfers made from players in  $C$ . Set  $\alpha' = \alpha$ , and define

$$\beta'(h) = \begin{cases} (\mathbf{0}_C, \chi^C(T(h|\sigma))) & \text{if } h \in F(\hat{h}) \text{ for some } \hat{h} \in \underline{H}_C, \\ \beta(h) & \text{otherwise.} \end{cases}$$

By construction, the blocking plan  $(\alpha', \beta')$  has continuation values bounded below by  $\min_{a \in A} \sum_{i \in C} v_i(a)$ . In addition, compared to  $(\alpha, \beta)$ , the blocking plan  $(\alpha', \beta')$  gives coalition  $C$  weakly higher total continuation value following any history, so  $\sum_{i \in C} U_i(\bar{h} | \alpha', \beta') > \sum_{i \in C} U_i(\bar{h} | \sigma)$ . ■

Next we argue that there exists a finite set of payoff vectors whose convex hull contains the set  $\mathcal{U}_{IR}$ .

LEMMA 8: *For each alternative  $a \in A$  let  $\mathcal{U}(a) := \{u \in \mathbb{R}^n : \sum_i u_i = \sum_i v_i(a)\}$  denote the set of payoff profiles that can be generated by playing alternative  $a$  and redistributing through transfers. Let  $\bar{a} \in \arg \max_{a \in A} \sum_{i \in N} v_i(a)$  and  $\underline{a} \in \arg \min_{a \in A} \sum_{i \in N} v_i(a)$  be two alternatives that maximize and minimize players' total generated payoffs, respectively. There exist payoff vectors  $\{\tilde{u}^1, \dots, \tilde{u}^M\} \subseteq \mathcal{U}(\bar{a}) \cup \mathcal{U}(\underline{a})$ , such that  $\mathcal{U}_{IR} \subseteq \text{co}(\tilde{u}^1, \dots, \tilde{u}^M)$ .*

PROOF: By definition,

$$\mathcal{U}_{IR} \subseteq \bar{\mathcal{U}}_{IR} := \left\{ u \in \mathbb{R}^n : \sum_{i \in N} v_i(\underline{a}) \leq \sum_{i \in N} u_i \leq \sum_{i \in N} v_i(\bar{a}) \text{ and } u_i \geq \underline{v}_i \forall i \in N \right\}.$$

Since  $\bar{\mathcal{U}}_{IR}$  is a bounded polyhedron, it is also a polytope. Let  $x^1, \dots, x^K$  be its vertices. Any point inside  $\mathcal{U}_{IR}$  can then be expressed as convex combinations of these vertices. Since  $x^k \in \text{co}(\mathcal{U}(\bar{a}) \cup \mathcal{U}(\underline{a}))$  for all  $1 \leq k \leq K$ , for each  $k$ , there exist  $\{\tilde{u}^{k,1}, \dots, \tilde{u}^{k,m_k}\} \subseteq \mathcal{U}(\bar{a}) \cup \mathcal{U}(\underline{a})$  such that  $x^k \in \text{co}(\tilde{u}^{k,1}, \dots, \tilde{u}^{k,m_k})$ . As a result  $\mathcal{U}_{IR} \subseteq \text{co}(\cup_{1 \leq k \leq K} \{\tilde{u}^{k,1}, \dots, \tilde{u}^{k,m_k}\})$ . ■

### 1.2. Proof of Part 2 of Theorem 5.

Here, we establish that for every  $u \in \mathcal{U}_{CR}(\mathcal{S})$ , there  $\underline{\delta} < 1$  such that for all  $\delta \in (\underline{\delta}, 1)$ , there exists a public PCE supporting  $u$ .

To economize on notation, we define  $\underline{u}_{\{i\}} := \underline{v}_i$  for all  $i \in N$ , so  $\mathcal{U}_{CR}(\mathcal{S})$  can be written as  $\mathcal{U}_{CR}(\mathcal{S}) = \{u \in \mathcal{U} : \sum_{i \in C} u_i > \underline{u}_C \text{ for every } C \in \mathcal{S} \cup \mathcal{N}\}$ . For every  $C \in \mathcal{S} \cup \mathcal{N}$ , let  $\underline{a}_C \in \arg \min_{a \in A} \max_{a' \in E_C(a)} \sum_{i \in C} v_i(a')$  be an alternative that can be used to minmax  $C$ . Observe that, by Berge's maximum theorem,  $\underline{a}_C$  is well-defined because  $A$  is compact,  $v$  is continuous, and  $E_C(\cdot)$  is continuous and compact-valued. By the definition of  $\underline{a}_C$ ,  $\sum_{i \in C} v_i(a') \leq \underline{u}_C$  for all  $a' \in E_C(\underline{a}_C)$ . Given the reflexivity of effectivity correspondences,  $\underline{a}_C \in E_C(\underline{a}_C)$ , and therefore,  $\sum_{i \in C} v_i(\underline{a}_C) \leq \underline{u}_C$ .<sup>25</sup>

Fix any payoff vector  $u^* \in \mathcal{U}_{CR}(\mathcal{S})$ , and let  $\{u^C : C \in \mathcal{S} \cup \mathcal{N}\}$  be the  $(\mathcal{S} \cup \mathcal{N})$ -specific punishments from Lemma 6. Given these punishments, let  $\kappa \in (0, 1)$  be such that for every  $\tilde{\kappa} \in [\kappa, 1]$ , the following is true for all  $C \in \mathcal{S} \cup \mathcal{N}$  and  $C' \in \mathcal{S} \cup \mathcal{N} \setminus \{C\}$ :

$$(1 - \tilde{\kappa}) \sum_{i \in C} v_i(\underline{a}_C) + \tilde{\kappa} \sum_{i \in C} u_i^C > \underline{u}_C \tag{7}$$

$$(1 - \tilde{\kappa}) \sum_{i \in C'} v_i(\underline{a}_C) + \tilde{\kappa} \sum_{i \in C'} u_i^C > (1 - \tilde{\kappa}) \sum_{i \in C'} v_i(\underline{a}_{C'}) + \tilde{\kappa} \sum_{i \in C'} u_i^{C'}. \tag{8}$$

<sup>25</sup>This is the only step of the argument in which we invoke reflexivity. As this step shows, our proof does not require a global form of reflexivity but only that it holds for minmaxing alternatives. In our application to labor market matching, we establish that this weaker form of reflexivity holds, which allows us to apply the argument here.

By an argument identical to that in Theorem 1, there exists  $\kappa \in (0, 1)$  such that the inequalities above hold for all  $\tilde{\kappa} \in [\kappa, 1]$ ,  $C \in \mathcal{S} \cup \mathcal{N}$  and  $C' \in \mathcal{S} \cup \mathcal{N} \setminus \{C\}$ . Let  $L(\delta) := \left\lceil \frac{\log \kappa}{\log \delta} \right\rceil$ . As before, we use the property that  $\lim_{\delta \rightarrow 1} \delta^{L(\delta)} = \kappa$ .

For each alternative  $a \in A$  let  $\mathcal{U}(a) := \{u \in \mathbb{R}^n : \sum_i u_i = \sum_i v_i(a)\}$  denote the set of payoff profiles that can be generated by playing alternative  $a$  and redistributing through transfers. Let  $\bar{a} \in \arg \max_{a \in A} \sum_{i \in N} v_i(a)$  and  $\underline{a} \in \arg \min_{a \in A} \sum_{i \in N} v_i(a)$  be alternatives that maximize and minimize total payoffs, respectively. Since  $\mathcal{U}_{CR}(\mathcal{S}) \subseteq \mathcal{U}_{IR}$ , by Lemma 8 above, there exist payoff vectors  $\{\tilde{u}^1, \dots, \tilde{u}^M\} \subseteq \mathcal{U}(\bar{a}) \cup \mathcal{U}(\underline{a})$  such that  $\mathcal{U}_{CR}(\mathcal{S}) \subseteq \text{co}(\tilde{u}^1, \dots, \tilde{u}^M)$ , where each  $\tilde{u}^m = u(\tilde{a}^m, \tilde{T}^m)$  for some  $\tilde{a}^m \in \{\bar{a}, \underline{a}\}$  and  $\tilde{T}^m$ . Let  $\tilde{\mathcal{T}} := \{\tilde{T}^m\}_{m=1}^M$  be the set comprising these transfer matrices.

By Lemma 1, for any  $\epsilon > 0$ , there exists  $\underline{\delta} \in (0, 1)$  such that for all  $\delta \in (\underline{\delta}, 1)$ , there exist sequences  $\{(a^{d,\tau}, T^{d,\tau})_{\tau=0}^\infty : d \in \mathcal{S} \cup \mathcal{N} \cup \{*\}\} \subseteq (\{\bar{a}, \underline{a}\} \times \tilde{\mathcal{T}})^\infty$  such that for each  $d$  and  $t$ ,  $(1 - \delta) \sum_{\tau=0}^\infty \delta^\tau u(a^{d,\tau}, T^{d,\tau}) = u^d$  and  $\|u^d - (1 - \delta) \sum_{\tau=t}^\infty \delta^{\tau-t} u(a^{d,\tau}, T^{d,\tau})\| < \epsilon$ . We fix an  $\epsilon$  such that

$$\epsilon < (1 - \kappa) \min \left\{ \min_{d \in \mathcal{S} \cup \mathcal{N}, d' \in \mathcal{S} \cup \mathcal{N} \cup \{*\}, d' \neq d} \left( \sum_{i \in d} u_i^{d'} - \sum_{i \in d} u_i^d \right), \min_{d \in \mathcal{S} \cup \mathcal{N}} \sum_{i \in d} u_i^d - \underline{u}_d \right\},$$

and given that  $\epsilon$ , consider  $\delta$  exceeding the appropriate  $\underline{\delta}$ .

We now describe the public plan that we use to support  $u^*$ . Let  $\mathbf{0}$  denote the transfer matrix where all players make no transfers. For any non-singleton coalition  $C \in \mathcal{C} \setminus \mathcal{N}$ , let  $\mathcal{N}(C) := \{\{i\} : i \in C\}$  denote the set of singleton coalitions formed by members of  $C$ . For any collection of blocking coalitions  $B \in \mathcal{B}$ , define  $\widehat{C}(B) = [B \cap (\mathcal{S} \cup \mathcal{N})] \cup [\cup_{C \in B \setminus (\mathcal{S} \cup \mathcal{N})} \mathcal{N}(C)]$ . Note that  $B \cap (\mathcal{S} \cup \mathcal{N})$  is the set of secret or individual coalitions in  $B$ , while  $\cup_{C \in B \setminus (\mathcal{S} \cup \mathcal{N})} \mathcal{N}(C)$  consists of singleton coalitions converted from members of  $\cup_{C \in B \setminus (\mathcal{S} \cup \mathcal{N})} C$ , so  $\widehat{C}(B)$  is the collection of “players” in  $B$  if coalitions in  $\mathcal{S} \cup \mathcal{N}$  are treated as fictitious players. Consider the plan represented by the automaton  $(W, w(*, 0), f, \gamma)$ , where

- $W := \{w(d, \tau) | d \in \mathcal{S} \cup \mathcal{N} \cup \{*\}, \tau \geq 0\} \cup \{\underline{w}(S, \tau) | S \in \mathcal{S} \cup \mathcal{N}, 0 \leq \tau < L(\delta)\}$  is the set of possible states and  $w(*, 0)$  is the initial state;
- $f : W \rightarrow \overline{\mathcal{O}}$  is the output function, where  $f(w(d, \tau)) = (a^{d,\tau}, \emptyset, T^{d,\tau})$  and  $f(\underline{w}(S, \tau)) = (\underline{a}_S, \emptyset, \mathbf{0})$ ;
- $\gamma : W \times \overline{\mathcal{O}} \rightarrow W$  is the transition function. For each  $S \in \mathcal{S}$  let  $u_S(a, T) = \sum_{i \in S} u_i(a, T)$  denote the total utility accruing to  $S$ . For states of the form  $\{\underline{w}(S, \tau) | 0 \leq \tau < L(\delta) - 1, S \in \mathcal{N} \cup \mathcal{S}\}$ , the transition is

$$\gamma(\underline{w}(S, \tau), (a, B, T)) = \begin{cases} \underline{w}(S^*, 0) & \text{where } S^* \in \arg \min_{S' \in \widehat{C}(B) \setminus \{S\}} u_{S'}(a, T), \\ & \text{if } B \neq \emptyset \text{ but either } \{S \notin \widehat{C}(B)\} \\ & \text{or } \{u_S(a, T) > \underline{u}_S\} \text{ is true.} \\ \underline{w}(S, 0) & \text{if } B \neq \emptyset \text{ and both } \{S \in \widehat{C}(B)\} \\ & \text{and } \{u_S(a, T) \leq \underline{u}_S\} \text{ are true.} \\ \underline{w}(S, \tau + 1) & \text{if } B = \emptyset. \end{cases}$$

For states of the form  $\{\underline{w}(S, L(\delta) - 1) | S \in \mathcal{N} \cup \mathcal{S}\}$ , the transition is

$$\gamma(\underline{w}(S, L(\delta) - 1), (a, B, T)) = \begin{cases} \underline{w}(S^*, 0) & \text{where } S^* \in \arg \min_{S' \in \widehat{C}(B) \setminus \{S\}} u_{S'}(a, T), \\ & \text{if } B \neq \emptyset \text{ but either } \{S \notin \widehat{C}(B)\} \\ & \text{or } \{u_S(a, T) > \underline{u}_S\} \text{ is true.} \\ \underline{w}(S, 0) & \text{if } B \neq \emptyset \text{ and both } \{S \in \widehat{C}(B)\} \\ & \text{and } \{u_S(a, T) \leq \underline{u}_S\} \text{ are true.} \\ w(S, 0) & \text{if } B = \emptyset. \end{cases}$$

For states of the form  $w(d, \tau)$ , the transition is

$$\gamma(w(d, \tau), (a, B, T)) = \begin{cases} \underline{w}(S^*, 0) & \text{if } B \neq \emptyset, \\ w(d, \tau + 1) & \text{if } B = \emptyset, \end{cases}$$

where  $S^* \in \arg \min_{S' \in \widehat{C}(B)} u_{S'}(a, T)$ .

The plan represented by this automaton yields payoff profile  $u^*$ . The plan is also public since the transition relies only on  $B$  and  $\{u_S(a, T) : S \in \widehat{C}(B)\}$ , both of which are public information. By construction,  $\|u^d - V(w(d, \tau))\| < \epsilon$  and  $V(\underline{w}(S, \tau)) = (1 - \delta^{L(\delta) - \tau})v(\underline{a}_S) + \delta^{L(\delta) - \tau}V(w(S, 0))$  for all  $\tau$  in  $\{0, \dots, L(\delta) - 1\}$  and  $S \in \mathcal{S} \cup \mathcal{N}$ . As the arguments from here on are standard, we verify in the Supplementary Appendix that no coalition can profitably block in any state of this automaton.

Recall that  $\widetilde{\mathcal{T}} = \{\widetilde{T}^m\}_{m=1}^M$  and  $\{(a^{d, \tau}, T^{d, \tau}) : \tau \geq 0, d \in \mathcal{S} \cup \mathcal{N} \cup \{*\}\} \subseteq \{\bar{a}, \underline{a}\} \times \widetilde{\mathcal{T}}$ , so all default transfers in the plan are selected from a finite set. By Assumption 1, if a coalition  $C$  blocks a default transfers matrix  $T \in \widetilde{\mathcal{T}}$ , there exists  $\tilde{b} > 0$  such that

$$\sum_{i \notin C, j \in C} \chi_{ij}^C(T) \leq \tilde{b} \text{ for all } T \in \widetilde{\mathcal{T}} \text{ and } C \in \mathcal{C}. \quad (9)$$

In addition, since  $A$  is compact and  $v(\cdot)$  is continuous, there exists  $\widehat{b}$  such that

$$\max_{C \in \mathcal{C}} \max_{a \in A} \sum_{i \in C} v_i(a) \leq \widehat{b}. \quad (10)$$

We verify the incentives in the automaton states below.

**States of the form  $w(d, \tau)$ :** Suppose coalition  $C$  blocks and the outcome  $(\widehat{a}, \{C\}, \widehat{T})$  is realized. Recall that  $\widehat{C}(B) = [B \cap (\mathcal{S} \cup \mathcal{N})] \cup [\cup_{C \in B \setminus (\mathcal{S} \cup \mathcal{N})} \mathcal{N}(C)]$ , so if  $C \in \mathcal{S} \cup \mathcal{N}$ , then  $\widehat{C}(\{C\}) = \{C\}$  is a singleton set containing  $C$  as a unitary player; however, if  $C \notin \mathcal{S} \cup \mathcal{N}$ , then  $\widehat{C}(\{C\}) = \mathcal{N}(C)$ , which is a set of singleton coalitions made up of players in  $C$ . The plan punishes  $S^* \in \arg \min_{S' \in \widehat{C}(\{C\})} u_{S'}(a, T)$ , so  $S^*$  is either  $C \in \mathcal{S}$  as a unitary player or some singleton coalition  $\{i\}$  where  $i \in C$ . In either case, the (total) stage-game payoff for  $S^*$  satisfies

$$u_{S^*}(\widehat{a}, \widehat{T}) \leq \frac{1}{|\widehat{C}(\{C\})|} \sum_{S' \in \widehat{C}(\{C\})} u_{S'}(\widehat{a}, \widehat{T}) \leq \max_{a \in A} \sum_{j \in C} v_j(a) + \sum_{j \in C} \sum_{k \notin C} \chi_{kj}^C(T^{d, \tau}) \leq \widehat{b} + \tilde{b},$$

where the first inequality follows since the minimum among a set of numbers is less than their average; the second inequality follows since  $C$ 's total payoff comes from the generated payoffs

plus the net transfers paid by players outside of  $C$ ; lastly, the third inequality follows from (9) and (10) and the fact that all  $T^{d,\tau}$  are drawn from  $\{\tilde{T}^m\}_{m=1}^M$ .

Thus, we can find a uniform bound  $\bar{b}_1 := \hat{b} + \bar{b}$  such that the total stage-game payoff of  $S^*$  satisfies  $u_{S^*}(\hat{a}, \hat{T}) \leq \bar{b}$  for all  $C$ ,  $\delta$  and  $(d, \tau)$ . Following the same steps as those in the analogous part of Theorem 1, we can show that  $S^*$  obtains lower total payoff after coalition  $C$  blocks. Since  $S^* \subseteq C$ , there exists player  $i \in C$  who is not better off, so this is not a profitable block for  $C$ .

**States of the form  $\underline{w}(S, \tau)$  where  $S \in \mathcal{N} \cup \mathcal{S}$ :** Suppose coalition  $C$  blocks and the outcome is  $(\hat{a}, \{C\}, \hat{T})$ . Just like above, depending on whether  $C \in \mathcal{S} \cup \mathcal{N}$ ,  $\hat{C}(\{C\})$  is either  $\{C\}$  containing  $C$  as a unitary player or the set  $\mathcal{N}(C)$  of singleton coalitions made up of  $C$ 's members. There are 2 cases to consider.

*Case I:  $S \in \hat{C}(\{C\})$ , and  $u_S(\hat{a}, \hat{T}) \leq \underline{u}_S$ .* In this case, the plan punishes the current scapegoat  $S$ , where  $S$  is either  $C$  or a singleton set containing a member of  $C$ . Using (7) for sufficiently high  $\delta$  and following steps identical to the analogous argument in Theorem 1, we can show that

$$(1 - \delta^{L(\delta)-\tau})v_S(\underline{a}_S) + \delta^{L(\delta)-\tau}u_S^S \geq (1 - \delta)\underline{u}_S + \delta(1 - \delta^{L(\delta)})v_S(\underline{a}_S) + \delta^{L(\delta)+1}u_S^S,$$

where  $v_S(\cdot) = \sum_{i \in S} v_i(\cdot)$  and  $u_S^S = \sum_{i \in S} u_i^S$  denote  $S$ 's total payoff. If  $S$  contains only a member of  $C$ , then the inequality above shows that the blocking is not profitable for  $C$ ; if  $S$  is  $C$  itself, it implies that blocking does not improve  $C$ 's total value, so there exist  $i \in C$  who is not better off, so again the blocking is not profitable for  $C$ .

*Case II: either  $S \notin \hat{C}(\{C\})$  or  $u_S(\hat{a}, \hat{T}) > \underline{u}_S$ .* In this case the plan punishes

$$S^* \in \arg \min_{S' \in \hat{C}(\{C\}) \setminus \{S\}} u_{S'}(\hat{a}, \hat{T})$$

as scapegoat. First observe that if  $C$  blocks in state  $\underline{w}(S, \tau)$  and the stage-game payoff satisfies  $u_S(\hat{a}, \hat{T}) > \underline{u}_S$ , then no matter if  $S \in \mathcal{N}$  or  $S \in \mathcal{S}$ , it must be that  $C \neq S$  and therefore  $\hat{C}(\{C\}) \neq \{S\}$ ; otherwise the definition of  $\underline{u}_S$  would ensure  $u_S(\hat{a}, \hat{T}) \leq \underline{u}_S$ . As a result, under either of the conditions defining the current case (i.e.  $S \notin \hat{C}(\{C\})$  or  $u_S(\hat{a}, \hat{T}) > \underline{u}_S$ ),  $\hat{C}(\{C\}) \neq \{S\}$  must be true, so the scapegoat  $S^* \in \arg \min_{S' \in \hat{C}(\{C\}) \setminus \{S\}} u_{S'}(\hat{a}, \hat{T})$  is well defined (i.e. the arg min is not taken over an empty set).

We show that the (total) stage-game payoff of  $S^*$  is bounded. If  $S \notin \hat{C}(\{C\})$ , then

$$S^* \in \arg \min_{S' \in \hat{C}(\{C\}) \setminus \{S\}} u_{S'}(\hat{a}, \hat{T}) = \arg \min_{S' \in \hat{C}(\{C\})} u_{S'}(\hat{a}, \hat{T}), \text{ and}$$

$$u_{S^*}(\hat{a}, \hat{T}) \leq \frac{1}{|\hat{C}(\{C\})|} \sum_{S' \in \hat{C}(\{C\})} u_{S'}(\hat{a}, \hat{T}) \leq \frac{1}{|\hat{C}(\{C\})|} \max_{a \in A} \sum_{j \in C} v_j(a), \quad (11)$$

where the last inequality above follows from Assumption 1 and the fact that the default transfers are 0 in states  $\underline{w}(S, \tau)$ . Alternatively, if  $S \in \hat{C}(\{C\})$  and  $u_S(\hat{a}, \hat{T}) > \underline{u}_S$ , then

$$u_{S^*}(\hat{a}, \hat{T}) \leq \frac{1}{|\hat{C}(\{C\})| - 1} \sum_{S' \in \hat{C}(\{C\}) \setminus \{S\}} u_{S'}(\hat{a}, \hat{T})$$

$$\begin{aligned}
&= \frac{1}{|\widehat{C}(\{C\})| - 1} \left[ \sum_{S' \in \widehat{C}(\{C\})} u_{S'}(\widehat{a}, \widehat{T}) - u_S(\widehat{a}, \widehat{T}) \right] \\
&\leq \frac{1}{|\widehat{C}(\{C\})| - 1} \left[ \sum_{S' \in \widehat{C}(\{C\})} u_{S'}(\widehat{a}, \widehat{T}) - \underline{u}_S \right],
\end{aligned}$$

where the last inequality follows because we are considering the case  $u_S(\widehat{a}, \widehat{T}) > \underline{u}_S$ . Since the plan specifies zero default transfers, Assumption 1 ensures

$$u_{S^*}(\widehat{a}, \widehat{T}) \leq \frac{1}{|\widehat{C}(\{C\})| - 1} \left[ \max_{a \in A} \sum_{i \in C} v_i(a) - \underline{u}_S \right]. \quad (12)$$

Comparing the RHS of (11) and (12) to the bounds obtained in (9) and (10), it follows that we can find  $\bar{b}_2$  such that  $u_{S^*}(\widehat{a}, \widehat{T}) < \bar{b}_2$ .

Finally, to show that the blocking by  $C$  is not profitable, note that the (total) payoff of  $S^*$  is not improved by blocking if

$$(1 - \delta^{L(\delta) - \tau})v_{S^*}(\underline{a}_S) + \delta^{L(\delta) - \tau}u_{S^*}^S \geq (1 - \delta)\bar{b}_2 + \delta(1 - \delta^{L(\delta)})v_{S^*}(\underline{a}_{S^*}) + \delta^{L(\delta) + 1}u_{S^*}^{S^*}.$$

This inequality follows for sufficiently high  $\delta$  from the same steps as that of the analogous part of Theorem 1. Based on the same arguments as in previous cases, the blocking is not profitable for  $C$ . ■

## 2. PROOFS FOR LABOR MARKET MATCHING

### 2.1. Proof of Proposition 1

To establish this conclusion, we apply Theorem 4, or more specifically, the proof of Theorem 5 in which  $\mathcal{S} = \emptyset$ . Note that for each player  $i \in \mathcal{F} \cup \mathcal{W}$ , the individual minmax is  $\underline{v}_i = 0$ . In this setting, to minmax a player, it suffices to use any assignment in which that player is unmatched. For specificity, we consider the assignment in which all players are unmatched, and denote it by  $\phi^\circ$ . All steps of the argument of Theorem 5 go through without adaptation but one, namely that which invokes the reflexivity of the effectivity correspondence. Observe that, as discussed in Footnote 25, the weaker local reflexivity holds for the minmaxing alternative: for every player  $i$ ,  $E_{\{i\}}(\phi^\circ) = \{\phi^\circ\}$ . Therefore, this step also goes through. ■

### 2.2. Proof of Proposition 2

As we argued in the proof of Proposition 1, the minmax for an individual player  $i$  is  $\underline{v}_i = 0$ , attained using the assignment  $\phi^\circ$ . Like the proof of Theorem 5, we use  $\underline{u}_{\{i\}} := \underline{v}_i = 0$  to denote the minmax for singleton coalitions consisting of individual firms and workers, so  $\underline{u}_C = 0$  for all  $C \in \mathcal{N}$ . For an essential coalition  $C = \{f\} \cup W \in \mathcal{E}$ ,  $\underline{u}_C = \sum_{w \in W} v_w(\{f\}) + v_f(W)$ , which is the total value generated when  $C$  matches.

*Preliminary Results.* We establish several results that allow us to characterize  $\mathcal{K}$ .

LEMMA 9: A matching  $(\phi, T)$  is a core allocation, as defined in Definition 8, if and only if it admits no profitable blocking by any singleton firm, singleton worker, or essential coalition.

PROOF: The “only if” direction is immediate. We prove the “if” direction by contradiction. Suppose that the matching  $(\phi, T)$  admits no profitable blocking by any coalition  $C \in \mathcal{E} \cup \mathcal{N}$  but can be profitably blocked by some larger coalition  $C \notin \mathcal{E} \cup \mathcal{N}$ . It follows then that there exists an assignment  $\phi' \in E_C(\phi)$  and transfers  $T'_C = [T'_{ij}]_{i \in C, j \in N}$  such that every member of  $C$  is better off from the matching  $(\phi', T'_C, \chi^C(T))$ , i.e.  $u_i(\phi', T'_C, \chi^C(T)) > u_i(\phi, T)$  for all  $i \in C$ . By the definition of the effectivity function  $E_C$ ,  $\phi'$  induces a partition  $\pi'_C$  of  $C$  such that  $\pi'_C \subseteq \mathcal{E} \cup \mathcal{N}$ . Let  $u_{C'}(m) := \sum_{i \in C'} u_i(m)$  denote the total utility a coalition  $C'$  obtains from a matching  $m$ , we have

$$\sum_{C' \in \pi'_C} u_{C'}(\phi', T'_C, \chi^C(T)) > \sum_{C' \in \pi'_C} u_{C'}(\phi, T),$$

so there exists some  $\widehat{C} \in \mathcal{E} \cup \mathcal{N}$  such that  $u_{\widehat{C}}(\phi', T'_C, \chi^C(T)) > u_{\widehat{C}}(\phi, T)$ . Since in the matching  $(\phi', T'_C, \chi^C(T))$  transfers are only made between matched firm and workers,  $u_{\widehat{C}}(\phi', T'_C, \chi^C(T)) = \sum_{i \in \widehat{C}} v_i(\phi'(i))$ , which is the total value generated by coalition  $\widehat{C}$ . This value can be secured by  $\widehat{C}$  when it blocks alone, so  $\widehat{C} \in \mathcal{E} \cup \mathcal{N}$  can profitably block the matching  $(\phi, T)$ , which is a contradiction. ■

LEMMA 10: *Let  $\widehat{x} = \max_{\phi \in A} \sum_{i \in N} v_i(\phi)$ . If  $u \in \mathbb{R}^n$  satisfies  $\sum_{i \in N} u_i \leq \widehat{x}$  and  $\sum_{i \in C} u_i \geq \underline{u}_C$  for all  $C \in \mathcal{E} \cup \mathcal{N}$ , then  $\sum_{i \in N} u_i = \widehat{x}$ .*

PROOF: Take any  $\tilde{u} \in \mathbb{R}^n$  satisfying  $\sum_{i \in N} \tilde{u}_i \leq \widehat{x}$  and  $\sum_{i \in C} \tilde{u}_i \geq \underline{u}_C$  for all  $C \in \mathcal{E} \cup \mathcal{N}$ . Towards a contradiction, suppose that  $\tilde{u}$  is not utilitarian efficient; that is, suppose  $\sum_{i \in N} \tilde{u}_i < \widehat{x}$ . Then there exists an assignment  $\phi' \in A$  such that  $\sum_{i \in N} v_i(\phi') > \sum_{i \in N} \tilde{u}_i$ . Let  $\pi'$  denote the partition of players induced by the matching  $\phi'$ . Note that  $\pi'$  consists of either essential coalitions or singletons so  $\pi' \subseteq \mathcal{E} \cup \mathcal{N}$ . It follows that there exists  $C' \in \pi' \subseteq \mathcal{E} \cup \mathcal{N}$  such that  $\underline{u}_{C'} = \sum_{i \in C'} v_i(\phi') > \sum_{i \in C'} \tilde{u}_i$ , which is a contradiction to the assumption that  $\sum_{i \in C} \tilde{u}_i \geq \underline{u}_C$  for all  $C \in \mathcal{E} \cup \mathcal{N}$ . So  $\tilde{u}$  must be utilitarian efficient. ■

LEMMA 11: *Let  $\widehat{x} = \max_{\phi \in A} \sum_{i \in N} v_i(\phi)$ . The set  $\mathcal{K}$  is characterized by*

$$\mathcal{K} = \{u \in \mathbb{R}^n : \sum_{i \in N} u_i = \widehat{x}, \sum_{i \in C} u_i \geq \underline{u}_C \text{ for all } C \in \mathcal{E} \cup \mathcal{N}\}. \quad (13)$$

PROOF: Take any  $\tilde{u} \in \mathcal{K}$ . Suppose, for the sake of contradiction, that there exists some  $C \in \mathcal{E} \cup \mathcal{N}$  such that  $\sum_{i \in C} \tilde{u}_i < \underline{u}_C$ , then  $\tilde{u}$  would be blocked by  $C$ , which contradicts the assumption that  $\tilde{u} \in \mathcal{K}$ . So  $\sum_{i \in C} \tilde{u}_i \geq \underline{u}_C$  must hold for all  $C \in \mathcal{E} \cup \mathcal{N}$ . Lemma 10 then implies that  $\tilde{u}$  is utilitarian efficient, so  $\tilde{u}$  satisfies the conditions in (13).

For the converse, take any  $\tilde{u}$  that satisfies the conditions in (13). We will show that  $\tilde{u} \in \mathcal{K}$ , i.e., there exists a core allocation  $(\phi, T)$  such that  $\tilde{u} = u(\phi, T)$ .

By Lemma 9, the set of core allocations is the set of matchings that cannot be profitably blocked by singletons or essential coalitions, which is nonempty under our assumptions on firm preferences (Kelso and Crawford 1982). Since  $\mathcal{K}$  is nonempty, there exists a core allocation  $(\tilde{\phi}, \tilde{T})$ , which by the arguments above must satisfy  $\sum_{i \in N} v_i(\tilde{\phi}) = \widehat{x}$ . Since  $\sum_{i \in N} \tilde{u}_i = \widehat{x}$ , there exists  $\tilde{T}' \in \mathcal{T}$  such that  $\tilde{u} = u(\tilde{\phi}, \tilde{T}')$ . Note however that  $\tilde{T}'$  may involve nonzero transfers between players who are not in an employment relationship, so  $(\tilde{\phi}, \tilde{T}')$  may not constitute a matching. Nevertheless, let  $\tilde{\pi}$  denote the partition of players induced by  $\tilde{\phi}$ . For every  $C \in \tilde{\pi}$ , it

must hold that

$$\sum_{i \in C, j \notin C} \tilde{T}'_{ij} - \sum_{i \in C, j \notin C} \tilde{T}'_{ji} = 0,$$

for otherwise we would have  $\sum_{i \in C'} \tilde{u}_i < \sum_{i \in C'} v_i(\tilde{\phi})$  for some  $C' \in \tilde{\pi}$ , contradicting the fact that  $\tilde{u}$  satisfies (13). Therefore, we can construct  $\tilde{T}'' \in \mathcal{T}$  such that

$$\tilde{u} = u(\tilde{\phi}, \tilde{T}''), \quad \text{and} \quad \tilde{T}''_{ij} \neq 0 \text{ only if } i = \tilde{\phi}(j) \text{ or } i \in \tilde{\phi}(j),$$

so  $(\tilde{\phi}, \tilde{T}'')$  is a matching that induces payoff profile  $\tilde{u}$ . Since  $\sum_{i \in C} \tilde{u}_i \geq \underline{u}_C$  for all  $C \in \mathcal{E} \cup \mathcal{N}$ ,  $(\tilde{\phi}, \tilde{T}'')$  cannot be blocked by any coalition  $C \in \mathcal{E} \cup \mathcal{N}$ , so by Lemma 9  $(\tilde{\phi}, \tilde{T}'')$  is a core allocation, and therefore  $\tilde{u} \in \mathcal{K}$ . ■

LEMMA 12: *Let*

$$\mathcal{U}^{\mathcal{M}} := \text{co} \left( \left\{ u \in \mathbb{R}^n : \exists (\phi, T) \in \mathcal{M} \text{ such that } u = u(\phi, T) \right\} \right)$$

denote the convex hull of all feasible matching payoffs. Then

$$\left\{ u \in \mathcal{U}^{\mathcal{M}} : \sum_{i \in C} u_i \geq \underline{u}_C \text{ for all } C \in \mathcal{E} \cup \mathcal{N} \right\} = \mathcal{K}.$$

PROOF: The fact that  $\mathcal{K} \subseteq \left\{ u \in \mathcal{U}^{\mathcal{M}} : \sum_{i \in C} u_i \geq \underline{u}_C \text{ for all } C \in \mathcal{E} \cup \mathcal{N} \right\}$  follows from the definition of  $\mathcal{K}$ . To show  $\left\{ u \in \mathcal{U}^{\mathcal{M}} : \sum_{i \in C} u_i \geq \underline{u}_C \text{ for all } C \in \mathcal{E} \cup \mathcal{N} \right\} \subseteq \mathcal{K}$ , take any  $\tilde{u} \in \mathcal{U}^{\mathcal{M}}$ , since  $\tilde{u}$  is a convex combination of feasible payoff vectors, it must be that

$$\sum_{i \in \mathcal{N}} \tilde{u}_i \leq \hat{x} := \max_{\phi \in A} \sum_{i \in \mathcal{N}} v_i(\phi).$$

Lemma 10 then implies that  $\sum_{i \in \mathcal{N}} \tilde{u}_i = \hat{x}$ , so  $\tilde{u} \in \mathcal{K}$  by Lemma 11. ■

*Proof of Proposition 2.* A model in which all essential coalitions can offer private wage terms corresponds to restricting attention to public PCEs in the model with secret transfers (Section 4.3 of Ali and Liu 2026) that do not condition on any past wage terms.

We first prove that every payoff vector in  $\mathcal{K}$  can be supported by a public PCE that satisfies this measurability condition. For any  $\tilde{u} \in \mathcal{K}$ , there exists a core allocation  $(\phi, T)$  such that  $\tilde{u} = u(\phi, T)$ . Consider the plan  $\tilde{\sigma}$  in which  $\tilde{\sigma}(h) = (\phi, T)$  for every history  $h \in \tilde{\mathcal{H}}$ . This plan is public and does not condition on past wages, and produces discounted payoff profile  $\tilde{u}$ . Given that  $(\phi, T)$  is a core allocation,  $\tilde{\sigma}$  is also a PCE.

We now prove that for every  $\delta \geq 0$ , every public PCE in the model with private wages implements a discounted payoff profile in  $\mathcal{K}$ . Given that every such public PCE is also a public PCE with secret transfers, it suffices to show that this conclusion holds for public PCE with secret transfers. We apply Theorem 5 to establish this conclusion. All steps of the proof of Theorem 5 carry over to the matching environment without adaptation except the step that invokes reflexivity of the effectivity correspondences (see Footnote 25). However, the assignment  $\phi^\circ$  in which all players are unmatched satisfies  $\max_{\phi' \in E_C(\phi^\circ)} \sum_{i \in C} v_i(\phi') = \min_{\phi \in A} \max_{\phi' \in E_C(\phi)} \sum_{i \in C} v_i(\phi') = \underline{u}_C$  for every coalition  $C$ . Hence, the PCE can specify  $\phi^\circ$  to minmax any coalition  $C \in \mathcal{S} \cup \mathcal{N}$ . Moreover,  $\phi^\circ \in E_C(\phi^\circ)$  for every coalition  $C$ . Thus, the argument for this step also goes through. Since  $\mathcal{E} \subseteq \mathcal{S}$ , Theorem 5 implies that

for every  $\delta \geq 0$ , every discounted payoff profile  $\tilde{u}$  produced by a public PCE must satisfy  $\sum_{i \in C} \tilde{u}_i \geq \underline{u}_C$  for all  $C \in \mathcal{E} \cup \mathcal{N}$ , so  $\tilde{u} \in \{u \in \mathcal{U}^M : \sum_{i \in C} u_i \geq \underline{u}_C \text{ for all } C \in \mathcal{E} \cup \mathcal{N}\}$ . By Lemma 12,  $\tilde{u}$  must be an element of  $\mathcal{K}$ . ■

### 2.3. Proof of Proposition 3

*A Preliminary Result.* The following lemma characterizes the set of static stable matchings in the stage game.

**LEMMA 13:** *All static stable matchings fill slots  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$  while leaving other slots vacant; more over, all workers receive the same payoff  $r$  where  $\max\{0, \eta(L+1)\} \leq r \leq \max\{0, \eta(L)\}$ .*

**PROOF:** We break down the proof into two parts.

**Part 1:** All static stable matchings fill slots  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$  while leaving other slots vacant.

Let  $m$  be any static stable matching. We first show that  $m$  must be utilitarian efficient. Suppose, for the sake of contradiction, that  $m$  is not utilitarian efficient. Then there exists a reassignment of players that increases players' total payoff, which implies the existence of  $f \in \mathcal{F}$  and  $W \subseteq \mathcal{W}$  such that  $v_f(W) + \sum_{w \in W} v_w(f) > u_f(m) + \sum_{w \in W} u_w(m)$ . But this implies that  $m$  is profitably blocked by  $(f, W)$ , contradicting the stability of  $m$ .

Next, we show that since  $m$  is utilitarian efficient, it fills all slots in  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$ . Suppose, for the sake of contradiction, that there exists a slot  $(\tilde{f}, \tilde{l}) \in \{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$  that is not filled. Let  $\tilde{l}^* = \min\{l : (\tilde{f}, l) \text{ is unfilled}\}$  denote the first unfilled position at firm  $\tilde{f}$ . Since firms have diminishing marginal products, we have  $\rho(\tilde{f}, \tilde{l}^*) \geq \rho(\tilde{f}, \tilde{l})$ , so  $(\tilde{f}, \tilde{l}^*)$  is an open slot in  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$  that is immediately accessible by workers. Since there are  $L$  workers in total, if not all slots in  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$  are filled, there exists  $w' \in W$  who is either unemployed or filling a slot outside of  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$ . In the first scenario, matching  $w'$  to the unfilled slot  $(\tilde{f}, \tilde{l}^*)$  would strictly increase the total surplus. In the second scenario, let  $(\hat{f}, \hat{l})$  be the slot filled by  $w'$ , and let  $\hat{l}^* = \max\{l : (\hat{f}, l) \text{ is filled}\}$  denote the last occupied slot at firm  $\hat{f}$ , and  $\hat{w}^*$  denote the worker filling  $(\hat{f}, \hat{l}^*)$ . It follows from decreasing marginal product that  $(\hat{f}, \hat{l}^*)$  is also outside of  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$ , so matching  $\hat{w}^*$  to the unfilled slot  $(\tilde{f}, \tilde{l}^*)$  instead would strictly increase the total surplus, again contradicting the utilitarian efficiency of  $m$ . Thus, all stable matchings must fill the slots in  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$ .

To show that all slots outside of  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$  are vacant, there are two cases to consider. If  $\eta(L) > 0$ , we know from the arguments above that the  $L$  slots in  $\{(f, l) : \rho(f, l) \geq \eta(L)\}$  are filled, so all other slots must be vacant. If  $\eta(L) < 0$ , then the set  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}$  becomes  $\{(f, l) : \rho(f, l) \geq 0\}$ , and let us suppose, for the sake of a contradiction, that some slot  $(\tilde{f}, \tilde{l})$  with  $\rho(\tilde{f}, \tilde{l}) < 0$  is filled. Let  $\tilde{l}^* = \max\{l : (\tilde{f}, l) \text{ is filled}\}$  denote the last filled slot at firm  $\tilde{f}$ , and let  $\tilde{w}$  denote the worker matched to this position. Due to decreasing marginal returns, we have  $\rho(\tilde{f}, \tilde{l}^*) < 0$  as well, so simply unmatching  $\tilde{w}$  from  $(\tilde{f}, \tilde{l}^*)$  will increase the total surplus. This contradicts the efficiency of  $m$ , which implies that  $m$  would not be stable. It follows that all stable matchings must leave slots outside  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$  vacant.

**Part 2:** All workers receive the same payoff  $r$ , where  $\max\{0, \eta(L+1)\} \leq r \leq \max\{0, \eta(L)\}$ .

First we establish that all workers receive the same payoff. Take any static stable matching  $m$ . From Part 1, all positions in  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}$  are filled. We prove that all workers have the same payoff under two separate cases.

First, suppose  $\eta(L) < 0$ . It follows that  $|\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\}| < L$ , so in the stable matching  $m$  there exists a worker  $\tilde{w}$  who is unmatched. This means  $\tilde{w}$  receives 0 payoff in the stable matching  $m$ . It then follows that any other employed worker must also receive 0 payoff, since otherwise there is a profitable block where their employer replaces them with  $\tilde{w}$ .

Second, suppose  $\eta(L) > 0$ . Since by assumption  $\rho(f, L) < \max\{0, \eta(L)\}$ , there exists  $f_1$  and  $f_2$  such that both  $f_1$  and  $f_2$  employ workers in  $m$ . Since workers are identical, each worker working for  $f_1$  must receive the same payoff as any worker at  $f_2$  in  $m$ . This implies that workers at  $f_1$  and  $f_2$  all have the same payoff. The same argument applies to workers employed by any other firm, so all workers receive the same payoff.

Let  $r$  denote the payoff that workers receive, we next show that  $\max\{0, \eta(L+1)\} \leq r \leq \max\{0, \eta(L)\}$ . It is obvious that  $r \geq 0$  by workers' individual rationality. To complete the arguments, it suffices to demonstrate the validity of three statements: A.  $r \geq \eta(L+1)$  if  $\eta(L+1) > 0$ ; B.  $r \leq \eta(L)$  if  $\eta(L) > 0$ ; and C.  $r = 0$  if  $\eta(L) \leq 0$ .

Statement A: if  $\eta(L+1) > 0$ , then decreasing marginal return implies  $\eta(L) > 0$ , so from Part 1 we know all  $L$  workers are assigned to  $\{(f, l) : \rho(f, l) \geq \max\{0, \eta(L)\}\} = \{(f, l) : \rho(f, l) \geq \eta(L)\}$ . Let  $(\tilde{f}, \tilde{l})$  denote the slot with value  $\rho(\tilde{f}, \tilde{l}) = \eta(L+1)$ . By decreasing marginal return, any slot in  $\{(\tilde{f}, l) : l < \tilde{l}\}$  at  $\tilde{f}$  is in  $\{(f, l) : \rho(f, l) \geq \eta(L)\}$  and already filled. It follows that  $r \geq \eta(L+1)$ , since otherwise  $\tilde{f}$  can profitably block  $m$  by poaching a worker from other firms, which generates additional surplus  $\eta(L+1)$ , while offering her a payoff  $r'$  satisfying  $\eta(L+1) > r' > r$ .

Statement B: if  $\eta(L) > 0$ , again from Part 1 we know that all  $L$  workers are assigned to  $\{(f, l) : \rho(f, l) \geq \eta(L)\}$ . Let  $(\tilde{f}, \tilde{l})$  be the slot such that  $\rho(\tilde{f}, \tilde{l}) = \eta(L)$ . By decreasing marginal return we know that  $(\tilde{f}, \tilde{l})$  must be the last filled slot at firm  $\tilde{f}$ . It follows that workers' payoff is no more than  $\eta(L)$  since otherwise  $\tilde{f}$  can profitably block by firing the worker matched to the slot  $(\tilde{f}, \tilde{l})$ .

Statement C: if  $\eta(L) < 0$ , then there are at most  $(L-1)$  slots with a positive surplus, which by Part 1 implies that in any stable matching there exists a worker  $\tilde{w}$  who is unmatched. In this case, workers' payoff must be 0 since otherwise the matching is profitably blocked by a firm replacing one of its employees with worker  $\tilde{w}$ .

Combining statements A, B, and C lets us conclude that  $\max\{0, \eta(L+1)\} \leq r \leq \max\{0, \eta(L)\}$ . ■

*Proof of Proposition 3.* The first half of Proposition 3 follows from Proposition 1, while the second half of Proposition 3 follows from combining Proposition 3 and Lemma 13. ■

#### 2.4. Proof of Proposition 4

Since by assumption both markets  $M_1$  and  $M_2$  satisfy  $\eta_i(L+1) > 0$ , Lemma 13 implies that in each market  $M_i$ , where  $i = 1$  or  $2$ , all static stable matchings fill the slots in  $\{(f, l) : \rho_i(f, l) \geq \max\{0, \eta_i(L)\}\} = \{(f, l) : \rho_i(f, l) \geq \eta_i(L)\}$ . Moreover, the workers' payoff  $r$  in market  $M_i$  satisfies  $\eta_i(L+1) \leq r \leq \eta_i(L)$ . Recall that the total surplus is  $\Pi_i := \sum_{\ell=1}^L \eta_i(\ell)$ , while the set of potential workers' total surplus is  $\Pi_i^W = [L\eta_i(L+1), L\eta_i(L)]$ , and the set of potential firms' total surplus is

$$\Pi_i^F = \Pi_i - \Pi_i^W = \left[ \sum_{\ell=1}^L \eta_i(\ell) - L\eta_i(L), \sum_{\ell=1}^L \eta_i(\ell) - L\eta_i(L+1) \right].$$

To simplify notation let us denote  $\underline{b}_i^{\mathcal{W}} := L\eta_i(L+1)$  and  $\bar{b}_i^{\mathcal{W}} := L\eta_i(L)$ , so  $\Pi_i^{\mathcal{W}} = [\underline{b}_i^{\mathcal{W}}, \bar{b}_i^{\mathcal{W}}]$ . Similarly, let  $\underline{b}_i^{\mathcal{F}} := \sum_{\ell=1}^L \eta_i(\ell) - L\eta_i(L)$  and  $\bar{b}_i^{\mathcal{F}} = \sum_{\ell=1}^L \eta_i(\ell) - L\eta_i(L+1)$ , so  $\Pi_i^{\mathcal{F}} = [\underline{b}_i^{\mathcal{F}}, \bar{b}_i^{\mathcal{F}}]$ .

Let  $s := \eta_1(1) = \eta_2(1)$ . For each  $2 \leq \ell \leq L$ , define  $\Delta_\ell^i := \eta_i(\ell-1) - \eta_i(\ell)$ , so  $\eta_i(\ell) = s - \sum_{k=2}^{\ell} \Delta_k^i$  for all  $\ell \geq 2$ . It follows that

$$\sum_{\ell=1}^L \eta_i(\ell) = sL - \sum_{\ell=2}^L (L+1-\ell)\Delta_\ell^i,$$

$$L\eta_i(L) = sL - L \sum_{\ell=2}^L \Delta_\ell^i, \quad \text{and} \quad L\eta_i(L+1) = sL - L \sum_{\ell=2}^{L+1} \Delta_\ell^i.$$

This allows us to express the bounds for firms' and workers' aggregate surplus in terms of  $s$  and  $\Delta_\ell^i$ 's, yielding

$$\underline{b}_i^{\mathcal{W}} = sL - L \sum_{\ell=2}^{L+1} \Delta_\ell^i, \quad \text{and} \quad \bar{b}_i^{\mathcal{W}} = sL - L \sum_{\ell=2}^L \Delta_\ell^i, \quad (14)$$

$$\underline{b}_i^{\mathcal{F}} = \sum_{\ell=2}^L (\ell-1)\Delta_\ell^i, \quad \text{and} \quad \bar{b}_i^{\mathcal{F}} = \sum_{\ell=2}^{L+1} (\ell-1)\Delta_\ell^i. \quad (15)$$

Market  $M_2$  exhibiting more steeply decreasing returns than  $M_1$  is equivalent to  $\Delta_\ell^2 \geq \Delta_\ell^1$  for all  $2 \leq \ell \leq L+1$ , which implies  $\eta_2(\ell) \leq \eta_1(\ell)$  for all  $1 \leq \ell \leq L+1$ , so  $\Pi_2 \leq \Pi_1$ .

In (14), all the  $\Delta_\ell^i$ 's enter the bounds for worker surplus with negative coefficients, so  $\underline{b}_2^{\mathcal{W}} \leq \underline{b}_1^{\mathcal{W}}$  and  $\bar{b}_2^{\mathcal{W}} \leq \bar{b}_1^{\mathcal{W}}$ , where the inequalities are strict if  $M_2$  has strictly more steeply decreasing returns than  $M_1$ . By contrast, in (15) the  $\Delta_\ell^i$  terms enter the bounds with positive coefficients, so  $\underline{b}_2^{\mathcal{F}} \geq \underline{b}_1^{\mathcal{F}}$  and  $\bar{b}_2^{\mathcal{F}} \geq \bar{b}_1^{\mathcal{F}}$ , where, again, the inequalities are strict if  $M_2$  has strictly more steeply decreasing returns than  $M_1$ . Together, the directions of change for these bounds imply  $\Pi_2^{\mathcal{W}} \preceq_{SSO} \Pi_1^{\mathcal{W}}$  and  $\Pi_2^{\mathcal{F}} \succeq_{SSO} \Pi_1^{\mathcal{F}}$ , with strict set orders if  $M_2$  has strictly more steeply decreasing returns than  $M_1$ . ■

### 3. PERFECT COALITION-PROOF EQUILIBRIUM

In this subsection, we develop a solution concept that combines ingredients of Perfect Coalitional Equilibrium and coalition-proofness in the sense of [Bernheim, Peleg, and Whinston \(1987\)](#). We focus on strategic form games (Example 1 in [Ali and Liu 2026](#)) and show that Theorem 1 continues to hold without change.<sup>26</sup>

Consider any strategic form game  $G$  with player set  $N$ , action set  $A := \times_{i=1}^n A_i$ , where each player  $i$  has payoff function  $v_i : A \rightarrow \mathbb{R}$ . For an action profile  $a \in A$  and proper coalition  $C \in \mathcal{C} \setminus \{N\}$ , let  $G|_{a_{-C}}$  denote the induced game with player set  $C$ , action set  $A_C = \times_{i \in C} A_i$ , and payoff functions  $\{v_i^C\}_{i \in C}$  where each  $v_i^C : A_C \rightarrow \mathbb{R}$  is defined by  $v_i^C(a'_C) = v_i(a'_C, a_{-C})$  for all  $a'_C \in A_C$ . We first restate [Bernheim, Peleg, and Whinston's](#) definition of coalition-proof Nash equilibrium.

<sup>26</sup>We consider strategic form games for two reasons. First, [Bernheim, Peleg, and Whinston \(1987\)](#) formulate their solution concept for this class of games. Second, our definition below exploits the product structure of the alternative set  $A$  in a strategic form game. A similar construction would also apply to characteristic function and matching games (Example 3 in [Ali and Liu 2026](#)), wherein a player's payoff depends only on the identity of her coalition partners.

DEFINITION 10: (a) In a single player game  $G$ , action  $a_i^*$  is a **coalition-proof Nash Equilibrium** (CPNE) if and only if  $a_i^*$  maximizes  $v_i(a_i)$ .

(b) Suppose the number of players is  $n > 1$  and assume that coalition-proof Nash equilibrium has been defined for games with fewer than  $n$  players. Then,

- (i) For any game  $G$  with  $n$  players,  $a^* \in A$  is **self-enforcing** if, for all proper sub-coalitions  $C \in \mathcal{C} \setminus \{N\}$ ,  $a_C^*$  is a coalition-proof Nash equilibrium in the game  $G|_{a_{-C}^*}$ .
- (ii) For any game  $G$  with  $n$  players,  $a^* \in A$  is a coalition-proof Nash equilibrium if it is self-enforcing and if there does not exist another self-enforcing action profile  $a \in A$  such that  $v_i(a) > v_i(a^*)$  for all  $i = 1, \dots, n$ .

We use this definition to formulate a coalition-proof analog of PCE. In the repeated game, fix a plan  $\sigma$  and history  $h$ . Let  $G(\sigma, h)$  denote the continuation game  $G(\sigma, h)$ , defined as the strategic form game in each player  $i$  chooses an action from  $A_i$  and has the payoff function  $V_i(a) := (1 - \delta)v_i(a) + \delta U_i(h, a|\sigma)$ .<sup>27</sup>

DEFINITION 11: A plan  $\sigma$  is a **perfect coalition-proof equilibrium** (PCPE) if, at every history  $h$ , the default action profile prescribed at history  $h$ ,  $\sigma(h)$ , is a CPNE in the continuation game  $G(\sigma, h)$ .

Translating this definition to profitable blocking, a block by coalition  $C$  invalidates a plan  $\sigma$  from being a PCPE if it is profitable (as in Definition 1) and is itself immune to any further “credible” blocks by a sub-coalition of  $C$ . Hence, every PCE is a PCPE but the converse need not hold. The following result shows, however, that the set of PCPE payoffs coincides with the set of PCE payoffs as  $\delta \rightarrow 1$ .

THEOREM 1\*: *For every  $\delta \geq 0$ , every PCPE gives each player  $i$  a payoff of at least  $\underline{v}_i^\circ$ . Moreover, for every  $v \in \mathcal{V}_{CR}$ , there is a  $\underline{\delta} < 1$  such that for every  $\delta \in (\underline{\delta}, 1)$ , there exists a PCPE with discounted payoff equal to  $v$ .*

Theorem 1\* establishes that for every discount factor  $\delta$ , no player  $i$  can be driven below her coalitional minmax  $\underline{v}_i^\circ$  (defined in Section 3.1 of Ali and Liu 2026). Recall that  $C(i)$  is the (potentially singleton) set of players whose utilities are equivalent to player  $i$ 's, and  $\underline{v}_i^\circ$  corresponds to the minmax  $\min_{a \in A} \max_{a' \in E_{C(i)}(a)} v_i(a')$  that would come if coalition  $C(i)$  behaved as a unitary agent. This lower bound for PCPE payoffs coincides with that which Theorem 1 obtained for PCE payoffs. The second sentence of Theorem 1\* asserts that every feasible payoff profile that gives each player strictly more than her coalitional minmax is supportable as a PCPE.

PROOF OF THEOREM 1\*: We observe that the second sentence of Theorem 1\* follows from Theorem 1 and the fact that every PCE is a PCPE. Therefore, it suffices to establish the first sentence of Theorem 1\*.

We establish this claim by proving its contrapositive: let  $\sigma$  be a plan, and suppose there exists a player  $i$  that satisfies  $U_i(\emptyset|\sigma) < \underline{v}_i^\circ$ , we will prove that  $\sigma$  cannot be a PCPE. Part 1 of the proof of Theorem 1 establishes that coalition  $C(i)$  has a profitable blocking plan: choosing its myopic best response to the default alternative at every history would result in a payoff

<sup>27</sup>Because we operate in a strategic form game, we simplify our notation by recording only the past action profiles and not the identity of blocking coalitions.

no less than  $v_i^\circ$ . Lemma 2 shows that if the coalition  $C(i)$  has a profitable blocking plan, then there exists some history  $h$  at which  $C(i)$  can profitably block. Let  $a^* := \sigma(h)$  denote the action profile prescribed at that history by the plan  $\sigma$ . The existence of a profitable block by coalition  $C(i)$  implies that there exists  $a'_{C(i)} \in A_{C(i)}$  such that for all  $j \in C(i)$ ,

$$\begin{aligned} & (1 - \delta)v_j(a'_{C(i)}, a^*_{-C(i)}) + \delta U_j(h, a'_{C(i)}, a^*_{-C(i)} | \sigma) \\ & > (1 - \delta)v_j(a^*_{C(i)}, a^*_{-C(i)}) + \delta U_j(h, a^*_{C(i)}, a^*_{-C(i)} | \sigma). \end{aligned} \tag{16}$$

Observe that the RHS corresponds to  $V_j(a^*)$  in the game  $G(\sigma, h)$ . We use the following lemma—which we prove afterwards—that this inequality cannot hold in a PCPE.

**LEMMA 14:** *Suppose  $G$  is a strategic form game in which all players share equivalent utilities.<sup>28</sup> An action profile  $a^*$  is a coalition-proof Nash equilibrium of  $G$  if and only if  $a^* \in \arg \max_{a \in A} v_1(a)$ .*

**Lemma 14** asserts that in a strategic form game in which all players share equivalent utilities, every CPNE maximizes players' payoffs. **Definition 11** implies that were  $\sigma$  a PCPE,  $a^* = \sigma(h)$  must be a CPNE in  $G(\sigma, h)$ , so then by **Definition 10(b)(i)**,  $a^*_{C(i)}$  must be a CPNE in the game  $G(\sigma, h) | a^*_{-C(i)}$ . Applying **Lemma 14** to the game  $G(\sigma, h) | a^*_{-C(i)}$  yields that  $a^*_{C(i)}$  maximizes  $V_j(a_{C(i)}, a^*_{-C(i)})$  for all  $j \in C(i)$ . This property contradicts (16), implying that  $\sigma$  is not a PCPE. ■

**PROOF OF LEMMA 14:** We first prove that every  $a^* \in \arg \max_{a \in A} v_1(a)$  is a coalition-proof Nash equilibrium by induction over the number of players  $k$ .

Suppose  $k = 2$ . Given any  $a^* \in \arg \max_{a \in A} v_1(a)$ , since players have equivalent utilities, each player  $i$  maximizes  $v_i$  by playing  $a_i^*$ . Hence  $a^*$  is self-enforcing in the sense of **Definition 10(b)(i)**. Moreover, since  $a^*$  is Pareto efficient among all action profiles, it cannot be strictly Pareto dominated by another self-enforcing profile. Therefore,  $a^*$  is a coalition-proof Nash equilibrium.

Suppose that  $G$  has  $|N| = n$  players. Suppose also that we have established that for any game  $G(k)$  with  $k < n$  players, action space  $A^k$ , and payoff functions  $\{v_i^k\}_{i=1}^k$ , every element of  $\arg \max_{a^k \in A^k} v_1^k(a^k)$  is a coalition-proof Nash equilibrium. Let  $a^* \in \arg \max_{a \in A} v_1(a)$ . Let  $C \in \mathcal{C} \setminus \{N\}$  be any proper coalition and consider the induced  $|C|$ -player game  $G | a^*_{-C}$ . By definition,  $a^*_{C} \in \arg \max_{a_C \in A_C} v_1(a_C, a^*_{-C})$  so the inductive hypothesis implies that  $a^*_{C}$  is a coalition-proof Nash equilibrium in the game  $G | a^*_{-C}$ . Hence  $a^*$  is self-enforcing in  $G$  in the sense of **Definition 10(b)(i)**. Furthermore,  $a^*$  cannot be strictly Pareto dominated by another self-enforcing profile, so  $a^*$  is a coalition-proof Nash equilibrium in  $G$ .

Having proved that every element of  $\arg \max_{a \in A} v_1(a)$  is a coalition-proof Nash equilibrium, it then follows from **Definition 10(b)(ii)** that no action profile outside of  $\arg \max_{a \in A} v_1(a)$  can be a coalition-proof Nash equilibrium. ■

#### 4. AN APPLICATION TO DISTRIBUTIVE POLITICS

This section studies a repeated distribution problem, in which the players repeatedly choose how to divide a dollar. Such division problems feature prominently in the political economy

<sup>28</sup>In other words, for every pair of players  $i, j$ , there exist some  $\alpha > 0$  and  $\beta \in \mathbb{R}$  such that  $v_i(a) = \alpha v_j(a) + \beta$  for all  $a \in A$ .

literature (e.g. Baron and Ferejohn 1989) and relate to the *simple games* (Von Neumann and Morgenstern 1945) studied in cooperative game theory.

In this game,  $A := \{a \in \mathbb{R}_+^N : \sum_{i \in N} a_i = 1\}$ , where player  $i$ 's payoff from alternative  $a$  is  $a_i$ . Divisions are chosen by a "winning" coalition:  $\mathcal{W}$  is a set of coalitions such that for every coalition  $C$  in  $\mathcal{W}$ ,  $E_C(a) = A$  for every division  $a$ , and for every coalition  $C$  not in  $\mathcal{W}$ ,  $E_C(a) = \{a\}$ . As standard,  $\mathcal{W}$  is monotone and proper.<sup>29</sup> A simple-majority rule protocol corresponds to  $\mathcal{W} := \{C \in \mathcal{C} : |C| \geq \frac{n+1}{2}\}$ . This formulation also allows for *veto power*: if a player belongs to every winning coalition ( $\cap_{C \in \mathcal{W}} C$ ), then effectively no block can happen without her approval. We denote the set of veto players by  $D := \cap_{C \in \mathcal{W}} C$ .

Bernheim and Slavov (2009) study simple majority rule, emphasizing how Dynamic Condorcet Winners exist although the stage game lacks a Condorcet Winner. We focus instead on settings with at least one veto and one non-veto player, and in which veto players are not dictators ( $D \notin \mathcal{W}$ ). Absent history dependence, these settings lead to highly unequal splits: the veto players steal the entire dollar, emerging as *de facto* dictators of the game. Formally, the set of core alternatives of the stage game is  $\mathcal{K} := \{a \in \mathbb{R}_+^N : \sum_{i \in D} a_i = 1\}$ . The logic is that any division that gives a positive share to a non-veto player would be profitably blocked by a winning coalition who would extract that share and divide it among themselves.

Against this backdrop, we evaluate how history dependence can counter this tendency towards unequal splits. Consider a three-player example in which player 1 alone has veto power; however, she needs the support of at least one other player to block. Player 1 captures the entire dollar in the stage-game core. Nevertheless, relatively simple schemes in the repeated game can promote equal splits. Consider a core reversion plan that prescribes  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  every period if that has been the division up to now and switches to the stage-game core otherwise. On the equilibrium path, even if player 1 offers the entire dollar to either player 2 or 3, neither finds it profitable to block with her if  $(1 - \delta)(1) + \delta(0) \leq \frac{1}{3}$ . Going further, core-reversion can support any division in the triangle formed by the vertices  $\{(2\delta - 1, 1 - \delta, 1 - \delta), (0, \delta, 1 - \delta), (0, 1 - \delta, \delta)\}$ , which converges to the unit simplex as  $\delta \rightarrow 1$ .

One could go beyond core-reversion to characterize all PCE payoffs. Because the game is convex and exhibits default-independent power, Theorem 2 implies that all PCE payoffs can be supported by stationary PCE. Using this result, we find that if players are sufficiently patient, then every payoff in which each non-veto player obtains up to  $\delta$  can be supported in a PCE. We depict these outcomes in Figure 4(A).

These schemes collapse if the veto player can make and receive secret side-payments. Suppose players 1 and 2 can transfer utility under the table. Theorem 5 implies that player 3 then obtains 0 in all PCE payoffs, as illustrated in Figure 4(B). Even worse, player 1 takes the entire dollar in every period if she can make secret side-payments with each player.

These intuitions generalize to  $n$ -player games in which there are at least one veto and one non-veto player, and veto players are not dictators. We call a coalition  $C$  a *minimal winning coalition* if  $C$  is a winning coalition and every proper subset is not.

PROPOSITION 5: *The following hold:*

- (a) *Absent secret transfers, there exists  $\underline{\delta} \geq 0$  such that if  $\delta \geq \underline{\delta}$ , the set of supportable payoffs are those that give at least  $(1 - \delta)$  to each winning coalition.*
- (b) *A winning coalition  $C$  obtains the entire dollar in every period in every PCE, regardless of  $\delta$ , if it can make secret transfers.*

<sup>29</sup>In other words, if  $C$  is in  $\mathcal{W}$ , then  $\mathcal{W}$  contains every superset of  $C$  but not its complement.

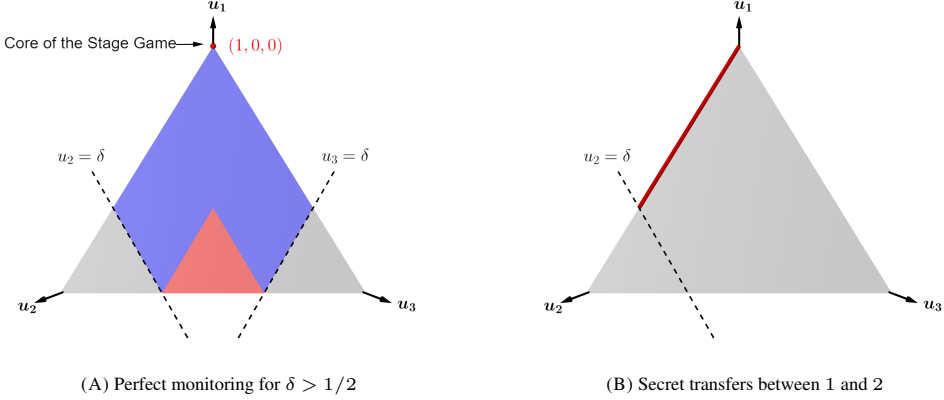


FIGURE 4.—(A) depicts the set of supportable outcome. The red region depicts payoffs supported by core-reversion, and the blue region illustrates those from other PCE. (B) shows the set of supportable payoffs once coalition  $\{1, 2\}$  can make secret transfers; player 3 then obtains 0.

(c) *The veto players obtain the entire dollar in every period in every PCE, regardless of  $\delta$ , if every minimal winning coalition can make secret transfers.*

**Proposition 5(a)** highlights how egalitarian schemes can be supported by history dependence. We use Theorem 2 to obtain this fixed discount factor characterization; it turns out that  $\delta = 0$  if there are at least two veto players so the characterization then is complete. **Proposition 5(b)** and (c) elucidate how secret side-payments destabilize egalitarian schemes: the veto players regain *de facto* dictatorial power if every minimal winning coalition can make secret transfers.

#### 4.1. Proof of Proposition 5

Below, we prove **Proposition 5**. We will use an alternative  $a \in A$  to also represent its generated payoff profile  $v(a)$ .

*Preliminary Results.* We establish two preliminary results. **Lemma 15** establish the existence of “punishment PCEs”  $\{\sigma^i\}_{i=1}^n$  that guarantee  $U_i(\emptyset|\sigma^i) = 0$  for each player  $i$ . **Lemma 16** proves that any PCE can be enforced by punishments where every member of a deviating coalition simultaneously obtains 0.

**LEMMA 15:** *Under perfect monitoring, for every player  $i \in N$ , there is a PCE  $\sigma^i$  such that  $U_i(\emptyset|\sigma^i) = 0$  when  $\delta > \frac{n-2}{n-1}$ .*

**PROOF:** We consider two case,  $|D| = 1$  and  $|D| \geq 2$ . The case where  $|D| = 1$  requires the discount factor to be sufficiently high. The case where there are two or more veto players ( $|D| \geq 2$ ) applies for every discount factor.

**Case 1:**  $|D| = 1$ . Suppose without loss of generality that  $D$  consists of player 1. Let  $\hat{a} := (1, 0, \dots, 0)$  denote the unique core alternative, and  $\bar{a} := (0, \frac{1}{n-1}, \dots, \frac{1}{n-1})$  denote the alternative that equally divides the total payoff among all non-veto players.

For  $i \neq 1$ , let  $\sigma^i$  be the plan that specifies the core alternative  $\hat{a}$  as default after every history, so each  $\sigma^i$  is a PCE that satisfies  $U_i(\emptyset|\sigma^i) = 0$

For  $i = 1$ , let  $\sigma^1$  be the plan that specifies  $\bar{a}$  on path, and  $\hat{a}$  at any history where an alternative distinct from  $\bar{a}$  has been chosen in the past. Note that  $U_1(\emptyset|\sigma^1) = 0$ . We will verify that  $\sigma^1$  is a PCE. No coalition can profitably block once continuation play reverts back to the core alternative. On the path of play, consider a winning coalition  $C \in \mathcal{W}$  blocking and choosing alternative  $a'$ . Since the game is non-dictatorial, if  $C$  is a winning coalition, player 1 cannot be its only member. Let  $j \neq 1$  be a player in  $C$ . Since  $a'_j \leq 1$ , we have

$$(1 - \delta)a'_j + \delta 0 \leq 1 - \delta \leq \frac{1}{n - 1}$$

so player  $j$  prefers following the plan  $\sigma^1$  over blocking and reverting to the core alternative. As a result, no coalition  $C$  can profitably block the plan  $\sigma^1$  at any history, so  $\sigma^1$  is a PCE.

**Case 2:**  $|D| \geq 2$ . Without loss of generality, suppose  $\{1, 2\} \subseteq D$ . Let  $a^1 := (1, 0, \dots, 0)$  and  $a^2 := (0, 1, 0, \dots, 0)$  be two alternatives that allocate all payoff to player 1 and 2, respectively. It follows that both  $a^1$  and  $a^2$  are core alternatives. Let  $\sigma^1$  be the plan that specifies  $a^2$  at all histories; for all  $i \neq 1$ , let  $\sigma^i$  be the plan that specifies  $a^1$  at all histories. Each  $\sigma^i$  is a PCE, and  $U_i(\emptyset|\sigma^i) = 0$  for every  $i \in N$ . ■

**LEMMA 16:** *Suppose  $\mathcal{U}$  is the set of PCE-supportable payoff profiles. For each player  $i \in N$ , let  $\underline{u}_i := \min_{u \in \mathcal{U}} u_i$  be player  $i$ 's smallest possible payoff from PCEs. There is a stationary PCE with payoff profile  $a$  if and only if for every coalition  $C$  and alternative  $a' \in E_C(a)$ , there is a player  $i \in C$  such that*

$$(1 - \delta)a'_i + \delta \underline{u}_i \leq a_i \tag{17}$$

**PROOF:** To see the “only if” direction, suppose (17) fails for some coalition  $C$  and  $a' \in E_C(a)$ . In other words, suppose there exists a coalition  $C$  and alternative  $a'$  such that  $(1 - \delta)a'_i + \delta \underline{u}_i > a_i$  for all  $i \in C$ . Towards a contradiction, suppose also that there exists a stationary PCE  $\sigma$  that supports payoff  $a$ . Since  $\sigma$  is a PCE, it follows that  $U_i(h|\sigma) \geq \underline{u}_i$  for every  $i \in C$  and all  $h \in \mathcal{H}$ . As a result, for every  $i \in C$ ,

$$(1 - \delta)a'_i + \delta U_i(a', \{C\}|\sigma) \geq (1 - \delta)a'_i + \delta \underline{u}_i > a_i.$$

Moreover, since  $\sigma$  is stationary, it always plays  $a$  on path. The inequality above then implies that  $(a', C)$  is a profitable block for coalition  $C$  on path, contradicting  $\sigma$  being a stationary PCE.

For the “if” direction, (17) implies that for every coalition  $C$  and alternative  $a' \in E_C(a)$ , there exists a player  $i[a', C]$  and a PCE  $\sigma[a', C]$  such that

$$(1 - \delta)a'_{i[a', C]} + \delta U_{i[a', C]}(\emptyset|\sigma[a', C]) \leq a_{i[a', C]}. \tag{18}$$

Since the stage game exhibits default-independent power, by Theorem 2, we can without loss assume that each  $\sigma[a', C]$  is a stationary PCE.

Consider a plan  $\sigma^*$  that specifies  $a$  on path, but switches to  $\sigma[a', C]$  if coalition  $C$  blocks to implement  $a'$ . Inequality (18) implies that on path, no coalition can find profitably block. In addition, the fact that each  $\sigma[a', C]$  is a PCE ensures that after any off-path history, no coalition can profitably block. Finally,  $\sigma^*$  is also stationary since it is stationary on path, and each  $\sigma[a', C]$  is also stationary. Therefore,  $\sigma^*$  is a stationary PCE that supports payoff  $a$ . ■

*Proof of Proposition 5.* We prove each statement separately.

**Statement (a).** Set  $\underline{\delta} = \frac{n-2}{n-1}$ . By Lemma 15, there exist PCEs  $\{\sigma^i : i \in N\}$  satisfying  $U_i(\emptyset|\sigma^i) = 0$  for all  $i \in N$ . It is straightforward to see that no players shared aligned payoffs in the stage game; in addition, no single player can form a winning coalition since the game is non-dictatorial. It follows that each player  $i$ 's individual minmax is  $\underline{v}_i = 0$ . Moreover, this minmax payoff is achieved by the PCE  $\sigma^i$ .

By Lemma 16, in order for a payoff profile  $u$  to be supported by a stationary PCE, it is necessary and sufficient that for every winning coalition  $C \in \mathcal{W}$ , there exist no alternative  $a' \in E_C(u)$  such that

$$(1 - \delta)a'_i + \delta \cdot 0 = (1 - \delta)a'_i > u_i \text{ for all } i \in C. \quad (19)$$

Note that the condition above is equivalent to  $\sum_{i \in C} u_i \geq 1 - \delta$  for every  $C \in \mathcal{W}$ , since if  $\sum_{i \in C} u_i < (1 - \delta) \cdot 1$  for some coalition  $C \in \mathcal{W}$ , there would be a certain  $a' \in E_C(u)$  representing a division of total payoff 1 among players in  $C$ , such that (19) holds for every  $i \in C$ . It follows that a payoff profile  $u$  is supportable by a stationary PCE if and only if  $\sum_{i \in C} u_i \geq 1 - \delta$  for every  $C \in \mathcal{W}$ . Finally, Theorem 2 implies that this same set is also the set of PCE-supportable payoff profiles.

**Statement (b).** Because a winning coalition can obtain the entire dollar by blocking, its minmax value is 1. This statement then follows immediately from Theorem 5.

**Statement (c).** Let  $\widehat{\mathcal{W}}$  denote the set of minimal winning coalitions. By definition,  $\widehat{\mathcal{W}} \subseteq \mathcal{W}$  so  $\cap_{C \in \mathcal{W}} C \subseteq \cap_{C \in \widehat{\mathcal{W}}} C$ . Furthermore,  $\cap_{C \in \widehat{\mathcal{W}}} C \subseteq \cap_{C \in \mathcal{W}} C$ , since otherwise there exists  $i \in \cap_{C \in \widehat{\mathcal{W}}} C$  and  $\tilde{C} \in \mathcal{W}$  such that  $i \notin \tilde{C}$ , but this would lead to a contradiction since  $\tilde{C}$  must contain a minimal winning coalition  $\hat{C}$ , and  $i \in \hat{C}$ . So  $\cap_{C \in \widehat{\mathcal{W}}} C = \cap_{C \in \mathcal{W}} C = D$ . By Theorem 5, every  $C \in \widehat{\mathcal{W}}$  obtains total payoff 1. This implies the total payoff for players in  $D$  is 1. ■

#### REFERENCES

- Ali, S. Nageeb and Ce Liu (2026), "Coalitions in repeated games." *Econometrica*. [9, 12, 13]
- Baron, David P. and John A. Ferejohn (1989), "Bargaining in legislatures." *American Political Science Review*, 83 (4), 1181–1206. [15]
- Bernheim, B. Douglas, Bezalel Peleg, and Michael D. Whinston (1987), "Coalition-proof nash equilibria i. concepts." *Journal of Economic Theory*, 42 (1), 1–12. [12]
- Bernheim, B. Douglas and Sita N. Slavov (2009), "A solution concept for majority rule in dynamic settings." *Review of Economic Studies*, 76 (1), 33–62. [15]
- Kelso, Alexander S and Vincent P Crawford (1982), "Job matching, coalition formation, and gross substitutes." *Econometrica*, 1483–1504. [8]
- Von Neumann, John and Oskar Morgenstern (1945), *Theory of Games and Economic Behavior*, Princeton University Press Princeton, NJ. [15]