THE WALKER-GROVES-LEDYARD MECHANISM: IMPROVING INDIVIDUAL RATIONALITY WITHOUT SACRIFICING SIMPLICITY OR STABILITY[†]

PAUL J. HEALY[*] AND RENKUN YANG[**]

ABSTRACT. It is known that no public goods mechanism can be Pareto efficient in Nash equilibrium, individually rational (IR), simple (using a one-dimensional message space), and dynamically stable. The Walker mechanism satisfies all but stability, while the Groves-Ledyard mechanism satisfies all but IR. Here we show that a hybrid between these two mechanisms maintains all of the properties of the Groves-Ledyard mechanism, but with fewer IR failures in expectation.

Keywords: Public goods; mechanism design; implementation; individual rationality.

JEL Classification: D02, D47, D61, D62, D71

[†]The authors thank the participants of the 2020 Decentralization Conference and 2022 Econometric Society Summer Meetings for helpful comments and feedback.

[*]Ohio State University; healy.52@osu.edu.

[**]Jinan University; renkunyang@jnu.edu.cn.

I. INTRODUCTION

For public goods economies, Groves and Ledyard (1977, henceforth, GL) provide a mechanism that implements a Pareto efficient outcome in Nash equilibrium, thus offering a solution to the classic free-rider problem. Two additional advantages are that it is simple—agents need only to submit a single number as their message—and it can be made stable: if the free parameter is chosen to be large enough then adaptive players will converge quickly to the efficient equilibrium. The main shortcoming of the GL mechanism is that it violates individual rationality (IR): in some economies some agents would prefer the initial endowment over the mechanism's equilibrium outcome.

A natural question is whether the IR failure could be eliminated. Hurwicz (1979a) shows that, under continuity assumptions, IR and Pareto optimality of Nash equilibrium (PO) are obtained if and only if the mechanism implements the Lindahl allocations for any given economy. The GL mechanism doesn't implement Lindahl allocations, which is why it must violate IR. But Walker (1981) provides a remarkably simple mechanism that does achieves this goal, and is therefore both PO and IR. This mechanism was generalized by Tian (1990). Although the Walker-Tian mechanisms are PO and IR, unfortunately they sacrifice stability. In fact, Healy and Mathevet (2012) prove that any simple mechanism—meaning, any mechanism in which agents submit a single number as their message—cannot satisfy PO, IR, and stability. Thus, among PO mechanisms, one must choose either IR failures, stability failures, or more complex message spaces.

Stability has been shown to be crucial in laboratory studies (Chen and Plott, 1996; Chen and Tang, 1998; Healy, 2006, e.g.), so the literature has mostly focused on sacrificing simplicity. Complex mechanisms have been found that are PO, IR, and stable, but they are either balanced only in equilibrium or contain very complicated payment rules (Hurwicz, 1979b; Kim, 1993; Chen, 2002; Healy and Mathevet, 2012; Van Essen, 2013).[1]

In this paper we take a different direction: We ask whether we can retain PO, simplicity, and stability, and somehow minimize the resulting IR failures. We show that a simple hybrid between the GL mechanism and Walker-Tian mechanisms can reduce IR failures "in expectation" while remaining stable, Pareto efficient, and budget balanced both on and off equilibrium path. In particular, we find that when the tax is composed of a linear combination of the GL tax and a simple permutation of the Walker tax, the agent whose equilibrium utility decreases the most relative to consuming the endowment (hence, the agent with the largest potential IR violation) is always better off in

[1]Van Essen et al. (2012) test the two-dimensional Kim (1993) and Chen (2002) mechanisms against the one-dimensional Walker mechanism in the lab. They find that Kim's mechanism performs the best of the three according to almost every measure considered.

expectation in the hybrid mechanism as long as the weight on the Walker-Tian payment is sufficiently small compared to the GL weight.

Our positive result, however, is subject to three limits. First, it is sensitive to which neighbors affect one's tax rate. Certain permutations of neighbors do not always lead to an improvement, though we do identify one that does guarantee improvement. Second, the result holds only in expectation. As shown in Section IV, for any hybrid mechanism there always exist some economies such that the IR violation is more severe for the hybrid mechanism than the GL mechanism. And this event is not "rare" in the sense that there exists an open set of economies wherein the hybrid performs worse. This effect is offset, however, when we aggregate across economies while fixing the identity and preference of the lowest type agent. Third, in the hybrid mechanism (like in the Walker mechanism) agents must treat the other agents asymmetrically, which requires agents to react to individual strategies of others rather than their aggregate contribution. This may be politically undesirable and may hurt performance. These limits highlight the underlying tension between implementing the Lindahl outcome and achieving dynamic stability with simple mechanisms.

II. The Framework

We consider an n-agent public goods economy with one private good and one public good. Let $i \in N = \{1, \ldots, n\}$ index the agents. We assume n is finite and $n \geq 3$. Each agent i is initially endowed with $\omega_i \geq 0$ units of private good and pays a tax $t_i \in \mathbb{R}$ towards the production of the public good (where $t_i < 0$ if they receive a subsidy), which leaves $x_i = \omega_i - t_i$ as their final consumption of the private good.

Initially there is no public good. The government (or planner) receives taxes from the agents and uses this revenue to produce $y \in \mathbb{R}$ units of the public good at a constant marginal cost of $\kappa > 0$. We say that y is is *feasible* if $\kappa y \leq \sum_i t_i$ and *budget balanced* if $\kappa y = \sum_i t_i$.

Each agent has preferences $\succeq_i$ over $\mathbb{R}^2$ representable by a continuously-differentiable utility functional $u_i(x_i, y)$. When there is no confusion, let $u_i(t_i, y) = u_i(\omega_i - t_i, y)$. The marginal rate of substitution of u_i is given by

$$MRS_i(t_i, y) = \frac{\partial u_i(t_i, y)/\partial y}{-\partial u_i(t_i, y)/\partial t_i}.$$

Utility is quasilinear if $u_i(t_i, y) = v_i(y) + \omega_i - t_i$ for some function v_i. In this case, $MRS_i(t_i, y) = v_i'(y)$. Without loss of generality we can set $v_i(0) = 0$.

An economy is given by $E = ((u_i, \omega_i)_{i=1}^n, \kappa)$. We are focused on constructing stable mechanisms, but Kim (1987) shows that stability is impossible for general economies. But stability is possible in quasilinear economies if concavity of preferences is bounded.

Thus, we assume that each u_i is quasilinear and twice differentiable, and that there is some $\eta > 0$ such that $v_i(y)'' \in (-\eta, -1/\eta)$ for all i and y.[2] Let $\mathscr{E}^{QL}$ denote the set of all such economies.

The government uses a mechanism $\Gamma = (M, y, t)$ to determine allocations, where M is a message space with messages of the form $m = (m_1, \ldots, m_n) \in M = \times_i M_i$, the function $y(m) \in \mathbb{R}$ identifies the resulting public good level, and $t(m) = (t_1(m), \ldots, t_n(m)) \in \mathbb{R}^n$ the resulting vector of taxes for each message profile $m \in M$.

We assume Nash equilibrium behavior by the agents, given the mechanism. The best response correspondence for agent i (given Γ and E) is given by

$$\beta_i(m_{-i}) = \{m_i \in M_i : (\forall m'_i \in M_i)\ u_i(t_i(m_i, m_{-i}), y(m_i, m_{-i})) \geq u_i\left(t_i(m'_i, m_{-i}), y(m'_i, m_{-i})\right)\}.$$

The set of Nash equilibrium messages is then given by $NE = \{m^* \in M : (\forall i)\ m^*_i \in \beta_i(m^*_{-i})\}$. Note that, depending on the preferences and the mechanism, it could be that $NE = \emptyset$.

Definition 1. Γ is **budget balanced in equilibrium** if $\sum_i t_i(m^*) = \kappa y(m^*)$ for every $m^* \in NE$, and **budget balanced** if this equality holds for every $m \in M$.

Definition 2. Given a set of economies $\mathscr{E}$, Γ is

(2.a) **y-optimal** if, for every $E \in \mathscr{E}$ and every $m^* \in NE$, $\sum_i MRS_i(t_i(m^*), y(m^*)) = \kappa$,

(2.b) **conditionally Pareto optimal** if Γ is y-optimal and, for every $E \in \mathscr{E}$, budget balanced in equilibrium, and

(2.c) **Pareto optimal** (PO) if it is conditionally Pareto optimal and, for every $E \in \mathscr{E}$, $NE \neq \emptyset$.

Definition 2.a is the familiar condition of Samuelson (1954), required to hold at any equilibrium message. But in a quasilinear economy y-optimality does not guarantee Pareto optimality since the transfers may be wasteful or infeasible. Definition 2.b therefore adds budget balance, which gives full Pareto optimality at any Nash equilibrium. This requirement is vacuous, however, if the mechanism has no Nash equilibria, so 2.c further requires that an equilibrium exist in every economy $E \in \mathscr{E}$.

One desirable criterion is that the mechanism never make agents worse off when they play equilibrium strategies. Formally, it should lead to equilibrium allocations that are weakly preferred by every i to their initial endowment point of $(t_i, y) = (0, 0)$.

Definition 3. Γ is **individually rational** (IR) if, for every $m^* \in NE$ and every $i \in N$, $u_i(t_i(m^*), y(m^*)) \geq u_i(0, 0)$.

We define the **minimum utility** of a mechanism to be $\min_{i \in N, m^* \in NE} u_i(t_i(m^*), y(m^*))$; IR thus requires that the minimum utility be non-negative.

[2] See Healy and Mathevet (2012) for a discussion of this assumption.

If $E \in \mathscr{E}^{QL}$ then y-optimality simply requires that $\sum_i v_i'(y(m^*)) = \kappa$. Budget balance requires that $\sum_i t_i(m^*) = \kappa$. And individual rationality reduces to $v_i(y) \geq t_i$.

Another desideratum is that the mechanism be "simple." Following Healy and Mathevet (2012), we say that a mechanism is simple if it only requires agents to submit a single number. The Walker and Groves-Ledyard mechanisms (defined below) are both simple. Examples of non-simple mechanisms include the canonical mechanism of Maskin (1999) and the dynamically stable mechanisms of Chen (2002) and Kim (1996).

Definition 4. Γ is **simple** if $M \subseteq \mathbb{R}^1$.

Finally, experimental research has demonstrated the importance of dynamically stable mechanisms. Chen and Tang (1998) and Healy (2006) show that theoretically-unstable mechanisms indeed perform poorly as subjects' strategies fail to converge to Nash equilibrium. Healy and Mathevet (2012) argue that contractiveness of the best-response functions is an appealing notion of stability for the mechanism design setting, so we apply that notion here.

Definition 5. Given $\mathscr{E}$, a mechanism Γ is **stable** if, for every $E \in \mathscr{E}$ and every $i \in N$, the best response correspondence β_i is a single-valued contraction mapping.

We will focus on cases where the best response function is differentiable, in which case stability has a simple necessary condition due to Conlisk (1973); see Healy and Mathevet (2012) for details.

Lemma 1. If Γ is simple and each β_i is differentiable then Γ is stable in E if for every $i \in N$ and $m \in M$,

$$\sum_{j \neq i} \left| \frac{\partial \beta_i(m)}{\partial m_j} \right| < 1.$$

Unfortunately, requiring stability causes two major difficulties when combined with Pareto optimality and individual rationality. First, Hurwicz (1979a) shows that, under richness and continuity assumptions, if Γ is both PO then Γ implements the Lindahl correspondence. But Kim (1987) (following Jordan, 1986) shows that if the set of allowable preferences is sufficiently rich then *any* mechanism that implements the Lindahl correspondence must be unstable for some economy $E \notin \mathscr{E}^{QL}$. For this reason, the subsequent literature on stability has focused on quasilinear preferences. We shall do the same here. The second problem, however, is that even if we restrict to quasilinear preferences, stability and simplicity are incompatible when also requiring PO and IR (Healy and Mathevet, 2012). In other words, there does not exist a mechanism that satisfies PO, IR, simplicity, and stability.

Lemma 2 (Healy and Mathevet (2012))**.** Under the richness and continuity assumptions of Hurwicz (1979a), Pareto optimality, individual rationality, stability, and simplicity are jointly incompatible. This is true even when restricted to the space of quasilinear economies with $v_i'' < 0$ for all i.

The mechanisms we focus on in this paper all have $M_i = \mathbb{R}^1$ and $y(m) = \sum_i m_i$. Thus, all mechanisms are differentiable and *responsive*, meaning each agent can always change the level of the public good by varying their announcement.[3]

Definition 6. A mechanism Γ is **responsive** if it is simple, y is differentiable, and there exists some $\varepsilon > 0$ such that for any i and m_{-i}, $|\partial y(m)/\partial m_i| > \varepsilon$.

Assume Γ is responsive. Since we are applying best-response logic in our equilibrium analysis, we can view each agent i as choosing a public good level $y_i(m_{-i})$ in response to the message profile of others. Formally, the choice of $y_i(m_{-i})$ represents the choice of the message m_i such that $y(m_i, m_{-i}) = y_i(m_{-i})$. We often drop the dependence on m_{-i}—writing it simply as y_i—when there is no confusion. For example, if $y(m) = \sum_i m_i$ then y_i represents the message $m_i = y_i - \sum_{j \neq i} m_j$. Let $t_i(y_i, m_{-i})$ denote the corresponding tax function. If (m_i^*, m_{-i}^*) is a Nash equilibrium then we can equivalently denote it by (y_i^*, m_{-i}^*), where y_i^* corresponds to the best-response choice of y_i, given m_{-i}^*.

We now specify necessary conditions for Pareto optimality of a mechanism. Following Healy and Mathevet (2012), we can without loss of generality write any mechanism's tax function as

$$t_i(m) = q_i(m_{-i})y(m) + g_i(m), \tag{1}$$

where $q_i(m_{-i})$ is a personalized price per unit charged to agent i (which they cannot affect through their own message) such that $\sum_i q_i(m_{-i}) = \kappa$ for all m, and $g_i(m)$ is an additional penalty or subsidy.[4]

Nash equilibrium implies the individual first-order condition of

$$-\frac{\partial u_i(t_i(m^*), y(m^*))}{\partial t_i}\frac{\partial t_i(m^*)}{\partial m_i} = \frac{\partial u_i(t_i(m^*), y(m^*))}{\partial y}\frac{\partial y(m^*)}{\partial m_i}.$$

Inserting the functional form from (1) and assuming $\partial y(m^*)/\partial m_i \neq 0$ (which is true if the mechanism is responsive) we have

$$q_i(m_{-i}^*) + \frac{\partial g_i(m^*)/\partial m_i}{\partial y(m^*)/\partial m_i} = MRS_i(t_i(y_i^*, m_{-i}^*), y_i^*).$$

[3]Responsiveness is a slight strengthening of agent sovereignty (Moulin and Shenker, 2001; Marchant and Mishra, 2015), which also requires that each agent is able to select any level of the public good regardless of m_{-i}, but may allow $\partial y(m)/\partial m_i = 0$ for some m and i. Agent sovereignty is part of the definition of a "quasi-direct" mechanism in Healy and Jain (2017).

[4]This is without loss of generality because $g_i(m)$ is unrestricted and therefore can contain the term $-q_i(m_{-i})y(m)$ plus any additional terms.

Now, y-optimality requires that $\sum_i MRS_i = \kappa = \sum_i q_i(m_{-i})$, so if Γ is y-optimal then

$$\text{(y-OPT)} \qquad \sum_i \frac{\partial g_i(m^*)/\partial m_i}{\partial y(m^*)/\partial m_i} = 0.$$

For Pareto optimality of Γ we also need budget balance at any equilibrium profile, which further implies that

$$\text{(g-BAL)} \qquad \sum_i g_i(y_i^*, m_{-i}^*) = 0.$$

Conversely, if a mechanism satisfies these two conditions at every equilibrium point, then it is conditionally Pareto optimal.

Lemma 3. Suppose Γ is responsive. Conditions (y-OPT) and (g-BAL) are satisfied at every Nash equilibrium m^* if and only if Γ is conditionally Pareto optimal.

We now give three examples of mechanisms that are conditionally Pareto optimal. All three have $M_i = \mathbb{R}$ and $y(m) = \sum_i m_i$ and therefore are responsive.

Definition 7. The **Proportional Tax mechanism** is given by

(1) $M_i = \mathbb{R}$,
(2) $y(m) = \sum_i m_i$, and
(3) $t_i(m) = \alpha_i \kappa y(m)$.

where $(\alpha_i)_{i=1}^n$ are cost shares that must satisfy $\alpha_i \geq 0$ for all i and $\sum_i \alpha_i = 1$.

The Proportional Tax mechanism uses a fixed personal price ($q_i(m_{-i}) = \alpha_i \kappa$) and no additional penalty ($g_i(m) = 0$). Since $g_i \equiv 0$, Lemma 3 immediately gives conditional Pareto optimality. Unfortunately, Nash equilibrium rarely exists for this mechanism. The following two mechanisms modify the Proportional Tax mechanism to give existence.

Definition 8. The **Walker mechanism** is given by

(1) $M_i = \mathbb{R}$,
(2) $y(m) = \sum_i m_i$, and
(3) $t_i(m) = (\alpha_i \kappa + \lambda(m_{i+2} - m_{i+1}))y(m)$.

where $(\alpha_i)_{i=1}^n$ are cost shares, $\lambda > 0$ is a free parameter, and the indices $i+2$ and $i+1$ are taken to be modulo n.[5]

Compared to the Proportional Tax mechanism, the Walker mechanism modifies the price function $q_i(m_{-i})$, but maintains the property that $\sum_i q_i(m_{-i}) = \kappa$ and $g_i \equiv 0$. Thus it remains conditionally Pareto optimal. Furthermore, Walker (1981) shows that equilibrium exists for any economy, so the mechanism is fully Pareto optimal. Since every

[5]The original mechanism of Walker (1981) considered only $\lambda = 1$.

agent has the ability to set $y(m) = 0$ and $t_i(m) = 0$ by choosing $m_i = -\sum_{j \neq i} m_j$, any Nash equilibrium outcome must be preferred to the endowment. Thus, the Walker mechanism is also individually rational. It is, however, highly unstable (Chen and Tang, 1998; Healy, 2006).

Definition 9. The **Groves-Ledyard mechanism** is given by

(1) $M_i = \mathbb{R}$,
(2) $y(m) = \sum_i m_i$, and
(3) $t_i(m) = \alpha_i \kappa y(m) + \frac{\gamma}{2}\left[\frac{n-1}{n}(m_i - \bar{m}_{-i})^2 - \sigma^2_{-i}\right]$.

where $(\alpha_i)_{i=1}^n$ are cost shares, $\gamma > 0$ is a free parameter, $\bar{m}_{-i} = \frac{1}{n-1}\sum_{j \neq i} m_j$, and $\sigma^2_{-i} = \frac{1}{n-2}\sum_{j \neq i}(m_j - \bar{m}_{-i})^2$.

Whereas the Walker mechanism modified the individual price function $q_i(m_{-i})$, the Groves-Ledyard mechanism instead modifies the added penalty function $g_i(m)$, but does so such that conditions (y-OPT) and (g-BAL) remain intact. Groves and Ledyard (1980) prove existence for a wide range of economies, so the mechanism is Pareto optimal. And Page and Tassier (2010) find that the equilibrium is unique when preferences are quasilinear. It does not implement Lindahl allocations, however, so it violates individual rationality in many economies.

Unlike the Walker mechanism, the Groves-Ledyard mechanism becomes stable when γ is large relative to n (Muench and Walker, 1983; Bergstrom et al., 1983; Chen and Tang, 1998; Page and Tassier, 2004, 2010; Healy, 2006). Experiments verify that stability is an important property for agents to obtain equilibrium outcomes, so our motivation is to maintain the stability of the Groves-Ledyard mechanism while trying to reduce the magnitude of its individual rationality failures. We do so by blending together the Walker and Groves-Ledyard mechanisms.

III. The Walker-Groves-Ledyard (WGL) Mechanism

Here we introduce our new mechanism, which is a blend of the Walker and Groves-Ledyard mechanisms. If we simply combine their tax functions, we would have a new mechanism with $y(m) = \sum_i m_i$ and

$$t_i(m) = (\alpha_i \kappa + \lambda(m_{i+2} - m_{i+1}))y(m) + \frac{\gamma}{2}\left[\frac{n-1}{n}(m_i - \bar{m}_{-i})^2 - \sigma^2_{-i}\right].$$

We will see, however, that which opponents' indices are used in the personalized price term will alter our results. The original Walker mechanism uses $m_{i+2} - m_{i+1}$, but we will show that, surprisingly, using $m_{i+1} - m_{i-1}$ provides superior performance. For our general exposition we will follow Tian (1990) and consider an arbitrary $1 \times n$ vector of weights A_i such that i's personalized price is given by $q_i(m_{-i}) = (\alpha_i \kappa + \lambda A_i m)$. This vector must satisfy three conditions:

(1) $A_{ii} = 0$,
(2) $\sum_j A_{ij} = 0$, and
(3) $A_{ij} \neq 0$ for some $j \neq i$.

The first guarantees that $q_i(m_{-i})$ is not affected by m_i. The second guarantees that $\sum_i q_i(m_{-i}) = \kappa$, which is needed for Pareto optimality. The third ensures the mechanism differs from the Proportional Tax mechanism.

We construct each A_i as a permutation of A_1, which is a $1 \times n$ row vector of weights for agent 1 satisfying the above three conditions. For example, in the original mechanism, $A_1 = (0, -1, 1, 0, \ldots, 0)$. The weights for agent 2 are given by the same vector, but shifted by one coordinate. If we define the cyclic permutation matrix S by

$$S = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

then the weights for agent 2 are $A_2 = A_1 S$. Inductively, the weights for any agent $i > 1$ are given by $A_i = A_{i-1} S = A_1 S^{i-1}$. We can now define the WGL mechanism for general weight vectors.

Definition 10. The **Walker-Groves-Ledyard (WGL) mechanism** is given by

(1) $M_i = \mathbb{R}$,
(2) $y(m) = \sum_i m_i$,
(3) $t_i(m) = (\alpha_i \kappa + \lambda A_i m) y(m) + \frac{\gamma}{2} \left(\frac{n-1}{n} (m_i - \bar{m}^2_{-i})^2 - \sigma^2_{-i} \right)$,

where $m = (m_1, \ldots, m_n)^T$ is an $n \times 1$ column vector, $A_i = A_1 S^{i-1}$ for some admissible weighting vector A_1, $(\alpha_i)_i$ is a vector of non-negative cost shares that sum to one, $\lambda \geq 0$, and $\gamma > 0$.[6]

Since $y(m) = \sum_i m_i$ the WGL mechanism is responsive. And conditions (y-OPT) and (g-BAL) are inherited directly from the original Walker and Groves-Ledyard mechanisms, so the WGL mechanism is conditionally Pareto optimal. The key questions are whether it can be made stable—as in the Groves-Ledyard mechanism—and whether it

[6]The mechanism would more appropriately be called the Walker-Tian-Groves-Ledyard mechanism; in an effort to simplify the name we omit Tian with apologies.

improves the minimum utility of the agents, therefore reducing IR failures. To answer these questions we first characterize the equilibrium messages.

WGL Equilibrium Characterization

We now derive both the equilibrium messages and utility in the WGL mechanism when A_1 puts weight on exactly two adjacent neighbors to $i = 1$. To do so, we first define some variables that simplify these expressions. Fix any WGL mechanism with parameters λ and γ and weighting vector A_1. Define

- scalar $\rho = \lambda/\gamma$,
- matrix $A = \sum_i e_i A_i$, where e_i is the ith standard basis vector for $\mathbb{R}^n$, and
- matrix $W = (I + \rho A)$, where I is the $n \times n$ identity matrix.

We will show that W^{-1} (the inverse of W) plays an important role in our characterization; to that end, let W_i^{-1} be the ith row of W^{-1}.

Next, recall that for any $E \in \mathscr{E}^{QL}$ there is a unique y^o satisfying y-optimality. Define

$$\Delta_i = v_i(y^o) + \omega_i - \alpha_i \kappa y^o \quad \text{and}$$
$$\delta_i = v_i'(y^o) - \alpha_i \kappa$$

to be agent i's total surplus and marginal surplus (respectively) at the Pareto optimal allocation y^o if their tax was simply given by $\alpha_i \kappa y^o$. Note that $\sum_i \delta_i = 0$, so if there is an agent i with $\delta_i > 0$ then there is at least one agent j with $\delta_j < 0$.

With this notation we can now state our characterizations.[7]

Proposition 1. For any WGL mechanism with parameters λ and γ, and weighting vector A_1 such that $W = (I + \rho A)$ is invertible, the equilibrium message for each $i \in N$ is given by

$$m_i^* = \frac{1}{\gamma} W_i^{-1} \delta + \frac{y^o}{n}. \tag{2}$$

The resulting equilibrium utility for each agent i is then given by

$$u_i^{WGL}(m^*) = \Delta_i - \rho(A_i W^{-1} \delta) y^o - \frac{1}{2\gamma} \frac{n-1}{n-2} \left((W_i^{-1}\delta)^2 - \overline{(W_{-i}^{-1}\delta)^2} \right), \tag{3}$$

where $\overline{(W_{-i}^{-1}\delta)^2} = \frac{1}{n-1} \sum_{j \neq i} (W_j^{-1}\delta)^2$.

Proof. Since the WGL mechanism is conditionally Pareto optimal, any equilibrium m^* results in a Pareto optimal public good $y(m^*) = y^o$. Individual optimization for player i

[7] Page and Tassier (2010) provide a somewhat similar characterization of Groves-Ledyard equilibria (so, with $\lambda = 0$), though for the class of quasi-additive preferences studied by Bergstrom and Cornes (1983) and Bergstrom et al. (1983).

implies that at any equilibrium we must have

$$
\begin{aligned}
v_i'(y^o) &= \alpha_i \kappa + \lambda A_i m^* + \gamma \left(\frac{n-1}{n} (m_i^* - \bar{m}_{-i}^*) \right) \\
&= \alpha_i \kappa + \lambda A_i m^* + \gamma \left(m_i^* - \frac{y^o}{n} \right),
\end{aligned} \tag{4}
$$

where the second equality follows since $y^o = \sum_i m_i^*$. Rearranging, we get a linear equation of the form

$$
\gamma m_i^* + \lambda A_i m^* = \delta_i + \frac{\gamma}{n} y^o. \tag{5}
$$

We now write the equilibrium system of equations in matrix notation. Define $\delta = (\delta_1, \dots, \delta_n)^T$ to be the $n \times 1$ column vector of marginal surpluses. Recall that A is the matrix whose ith row is A_i. Let $\mathbb{1}$ be the $n \times 1$ vector of ones. The equilibrium system of equations can thus be written as

$$
\gamma m^* + \lambda A m^* = \delta + \frac{\gamma}{n} y^o \mathbb{1}.
$$

If we let $\rho = \lambda/\gamma$ and define

$$
W = (I + \rho A),
$$

then the linear system becomes

$$
\gamma W m^* = \delta + \gamma \frac{y^o}{n} \mathbb{1}.
$$

An explicit equation for equilibrium messages in the WGL mechanism is therefore given by

$$
\begin{aligned}
m^* &= \frac{1}{\gamma} W^{-1} \delta + \frac{y^o}{n} W^{-1} \mathbb{1} \\
&= \frac{1}{\gamma} W^{-1} \delta + \frac{y^o}{n} \mathbb{1}.
\end{aligned}
$$

The second equality holds because $A\mathbb{1} = \vec{0}$ (each row sums to zero) and $A_{ii} = 0$ for all i, so $W\mathbb{1} = \mathbb{1}$, which gives $W^{-1}\mathbb{1} = \mathbb{1}$. Recall that W_i^{-1} is the ith row of W^{-1}, so the individual messages can be written as

$$
m_i^* = \frac{1}{\gamma} W_i^{-1} \delta + \frac{y^o}{n}.
$$

This is exactly equation (2).

Using the derived expression for m^*, the relevant equilibrium quantities in the tax function can be solved to be[8]

$$m_i^* - \bar{m}_{-i}^* = \frac{1}{\gamma}\frac{n}{n-1}W_i^{-1}\delta,$$
$$A_i m^* = \frac{1}{\gamma}A_i W^{-1}\delta, \text{ and}$$
$$\sigma_{-i}^{2*} = \frac{1}{n-2}\frac{1}{\gamma^2}\left(\sum_{j\neq i}(W_j^{-1}\delta)^2 - \frac{1}{n-1}(W_i^{-1}\delta_i)^2\right).$$

Plugging in these values, we derive the equilibrium utility of agent i to be

$$u_i^{WGL}(m^*) = \Delta_i - \rho(A_i W^{-1}\delta)y^o - \frac{1}{2\gamma}\frac{n-1}{n-2}\left((W_i^{-1}\delta)^2 - \overline{(W_{-i}^{-1}\delta)^2}\right),$$

where $\overline{(W_{-i}^{-1}\delta)^2} = \frac{1}{n-1}\sum_{j\neq i}(W_j^{-1}\delta)^2$. This completes the proof. □

Invertibility of W will not be guaranteed for general weighting vectors, but for three simple cases Searle (1979) proves an inverse exists and gives an explicit formula.

Lemma 4. If A_1 puts weight only on two neighbors adjacent to $i = 1$ (meaning there is some $a \neq 0$ such that either $A_1 = (0,0,0,\dots,0,-a,a)$, $A_1 = (0,-a,0,\dots,0,0,a)$, or $A_1 = (0,a,-a,0,\dots,0)$) and if $\rho \neq 1/(2a)$ then $W = (I + \rho A)$ is invertible.

Proof. For any A_1, both matrix A and matrix W are circulant matrices (each row is a shifted version of the previous row). Searle (1979) shows that if the first row of a circulant matrix contains exactly three adjacent elements such that all other elements are zero (as in the three examples of A_1 given) then W is invertible unless either (1) $\sum_i W_{1i} = 0$ or (2) n is even and the middle of the three numbers equals the sum of the other two. The first never holds since $\sum_i W_{1i} = 1$. The second only holds if $W_1 = (1,0,\dots,0,-1/2,1/2)$ or $W_1 = (1,1/2,-1/2,0,\dots,0)$, which are both ruled out by requiring $\rho \neq 1/(2a)$. □

If A_1 puts weight on more than two other agents then we cannot ensure invertibility of W. The set of invertible matrices is dense, however, so slight perturbations to A_1 will yield an invertible W.[9] We proceed assuming that A_1 is chosen so that W^{-1} exists.

The Groves-Ledyard and Walker mechanism equilibria are special cases of this characterization. In the Groves-Ledyard mechanism $\lambda = 0$ and W is simply the identity matrix, so $W_i^{-1}\delta = \delta_i$. For the Walker mechanism, the matrix represented by γW in

[8]To derive the expression for $m_i^* - \bar{m}_{-i}^*$ note that $\bar{m}_{-i}^* = (y^0 - m_i^*)/(n-1)$ and plug in m_i^*. For the σ_{-i}^{2*} expression, expand $\sum_{j\neq i}\left(W_j^{-1}\delta + \frac{1}{n-1}W_i^{-1}\delta\right)^2$, and note that $\frac{2}{n-1}W_i^{-1}\delta\sum_{j\neq i}W_j^{-1}\delta = -\frac{2}{n-1}\left(W_i^{-1}\delta\right)^2$ since $\sum_{j\neq i}W_j^{-1}\delta = \mathbb{1}^T W^{-1}\delta - W_i^{-1}\delta$, but $\mathbb{1}^T W^{-1}\delta = \mathbb{1}^T\delta = 0$, giving the result.

[9]If W is near a singular matrix then it may be ill-conditioned, meaning W^{-1} would be poorly-behaved, numerically. This may hinder mechanism performance in applications.

equation (2) reduces to λA, which is also circulant. This matrix is not invertible, but Walker (1981) shows that the equilibrium messages can be obtained by replacing the nth row of the matrix with the equation $\sum_i m_i = y^o$, which restores invertibility.

We will use equation (3) extensively in our analysis of individual rationality. But first we verify that the mechanism can be made stable.

Dynamic Stability

The Groves-Ledyard mechanism is dynamically stable for large γ, but the Walker mechanism is unstable for all λ. We now show that the WGL mechanism is stable for large γ, provided that λ stays small relative to γ. In other words, stability requires that the Groves-Ledyard component dominates the Walker component.

Proposition 2. Recall that $\rho = \lambda/\gamma$ and $v_i''(y) \geq -\eta$, and let $\max(A)$ be the greatest single entry in A. If $\rho < 1/(n \max(A))$ and

$$\gamma > \frac{n\eta}{1 - n\rho\max(A)}$$

then the WGL mechanism is stable.

The proof of Proposition 2 appears in the appendix. The specific examples of WGL mechanisms we will discuss all have $\max(A) = 1$, so the stability condition on ρ simplifies to $\rho^* < 1/n$. Even when $\max(A) \neq 1$, we can always normalize it to 1 and use $\rho\max(A)$ in place of ρ. The other condition implies that γ has to be sufficiently large. As $1 - n\rho\max(A) \in (0,1)$, the only one-size-fits-all condition is $\gamma \to \infty$ if we want to allow for arbitrary n and/or richness in terms of the curvature bound η. This does not mean $\gamma \to \infty$ is always necessary: a finite γ is sufficient if it is allowed to depend on n and if η is known. We focus on the infinite γ case in the main results to obtain sharp predictions regarding IR improvements. We shall come back to the possibility of finite γ in Section V.

IV. Individual Rationality: WGL versus GL

We know that the GL and WGL mechanisms do not implement Lindahl allocations, and therefore must violate IR for at least some economies.[10] But can we rank them in terms of the "severity" of their IR violations? To do so, we look at the welfare of the lowest

[10] For an example economy where WGL violates IR, let $n \geq 3$ and $v_i(y) = a_i y^2/2 + b_i y$ for each i, where $\sum_i b_i = \kappa$. This means $y^o = 0$, so $\Delta_i = 0$. Furthermore, if there is at least one agent with $b_i \neq a_i\kappa$ then equation (3) can be used to show that IR must be violated for any agent with $(W_i^{-1}\delta)^2 > \overline{(W_{-i}^{-1}\delta)^2}$. And it can be verified that at least one such agent must exist.

type, which is the agent i whose value $\delta_i = v_i(y^o) - \alpha_i \kappa$ is the lowest.[11] Henceforth we assume $\alpha_i = 1/n$ for all i, so the lowest type is also the agent with the lowest $v_i'(y^o)$.

We will define a weakening of the usual single-crossing condition such that, under this condition, the lowest-type agent will be the one hurt most in equilibrium. We will then show that the WGL mechanism doesn't always improve the lowest-type's welfare compared to the GL mechanism, but it does improve their welfare both "more often" and "in expectation." We formalize these concepts below.

First, we derive a useful expression for the difference between the WGL and GL equilibrium utilities.

Lemma 5. The WGL equilibrium utility can be rewritten as

$$u_i^{WGL}(m^*) = \Delta_i + \left[W_i^{-1}\delta - \delta_i\right] y^o - \frac{1}{2\gamma}\frac{n-1}{n-2}\left((W_i^{-1}\delta)^2 - \overline{(W_{-i}^{-1}\delta)^2}\right). \tag{6}$$

The difference in equilibrium utility between the WGL and GL mechanisms is therefore given by

(7)

$$u_i^{WGL}(m^*) - u_i^{GL}(m^*) = \left[W_i^{-1}\delta - \delta_i\right] y^o - \frac{1}{2\gamma}\frac{n-1}{n-2}\left(\left[(W_i^{-1}\delta)^2 - \delta_i^2\right] - \frac{1}{n-1}\sum_{j\neq i}\left[(W_j^{-1}\delta)^2 - \delta_j^2\right]\right).$$

Proof. Take equation (3) and note that $\rho A W^{-1} = I - W^{-1}$ because

$$\begin{aligned} WW^{-1} &= I \\ IW^{-1} + \rho A W^{-1} &= I \\ \rho A W^{-1} &= I - W^{-1}. \end{aligned}$$

Thus, we have

$$u_i^{WGL}(m^*) = \Delta_i - (e_i^T - W_i^{-1})\delta y^o - \frac{1}{2\gamma}\frac{n-1}{n-2}\left((W_i^{-1}\delta)^2 - \overline{(W_{-i}^{-1}\delta)^2}\right), \tag{8}$$

where e_i^T is the ith row of the identity matrix. The GL equilibrium utility is found by setting $\lambda = 0$ (so $\rho = 0$) and A to the matrix of zeros, so that W is simply the identity matrix and $W_i^{-1} = e_i^T$. This gives

$$u_i^{GL}(m^*) = \Delta_i - \frac{1}{2\gamma}\frac{n-1}{n-2}\left(\delta_i^2 - \frac{1}{n-1}\sum_{j\neq i}\delta_j^2\right). \tag{9}$$

Subtracting (9) from (8) gives the result.[12] □

[11]The usage of the word "type" is non-standard here since it is not a primitive construct specific to a single agent, but instead depends on y^o. And y^0 is determined by the collection of all agents' preferences. Regardless, for any economy the vector of "types" is well-defined and easily derived from primitives.

[12]We note that m^* varies by mechanism, though our notation for it does not.

Next, we define our weakening of the single-crossing condition, which we call *minimum-surplus sorting*. In the typical single-crossing condition with a quasilinear economy there would be a one-dimensional type space and we would assume that $v_i'(y)$ is strictly increasing in one's type at every y. Thus, the lowest type has the lowest $v_i'(y)$ at every y. Our condition only requires that this "lowest-type" property hold at the Pareto optimal point y^o, and we define one's "type" simply as $v_i(y^o) - \kappa/n$. In other words, all we require is that the agents with the lowest $v_i(y^o)$ also have the lowest $v_i'(y^o)$.

Definition 11. Assume $\alpha_i = 1/n$ for all i. A quasilinear economy $E \in \mathscr{E}^{QL}$ with Pareto optimal public good level y^o satisfies *minimum-surplus sorting* if

$$\arg\min_{i\in N}\ \Delta_i = \arg\min_{i\in N}\ v_i(y^o) \subseteq \arg\min_{i\in N}\ v_i'(y^o) = \arg\min_{i\in N}\ \delta_i.$$

Let $\mathscr{E}^{MS} \subset \mathscr{E}^{QL}$ be the set of quasilinear economies satisfying minimum-surplus sorting.

Recall that $\Delta_i = v_i(y^o) - \alpha_i \kappa y^o$ and $\delta_i = v_i'(y^o) - \alpha_i\kappa$, so the equalities follow from the assumption that $\alpha_i = 1/n$. The minimum-surplus sorting condition will be important for establishing our main welfare results, since we will focus on improving the welfare of the consumer with the lowest type (δ_i), and for the relevant parameter ranges of the WGL mechanism this will also be the consumer with the lowest equilibrium utility.

Lemma 6. Fix any $E \in \mathscr{E}^{MS}$ and a WGL mechanism with $\alpha_i = 1/n$ for all i and such that A_1 satisfies the "two neighbors" condition of Lemma 4. There exists a $\bar{\rho} \in (0, 1/n]$ and $\bar{\gamma} \ge 0$ such that if $\rho \in (0, \bar{\rho}]$ and $\gamma \ge \bar{\gamma}$ then the consumers with the lowest type are also the consumers with the lowest equilibrium utility ($\arg\min_i \delta_i \subseteq \arg\min_i u_i^{WGL}(m^*)$).

Proof. Recall the expression for $u_i^{WGL}(m^*)$ from equation (6). Under the conditions of Lemma 4, W is invertible for all $\rho \in (0, \bar{\rho}]$ since $\bar{\rho} \le 1/n < 1/2$. Thus W^{-1} is continuous in ρ, and so $u_i^{WGL}(m^*)$ is continuous in both ρ and γ when $\rho \le \bar{\rho}$ and $\gamma \ge \bar{\gamma}$. The result then follows since

$$\lim_{\rho\to 0}\lim_{\gamma\to\infty} u_i^{WGL}(m^*) = \Delta_i,$$

and so for small ρ and large γ we have $\arg\min_i u_i^{WGL}(m^*) = \arg\min_i \Delta_i = \arg\min_i \delta_i$. □

Realized Utility of the Lowest Type

In this subsection we focus on large γ—which is needed for stability—and establish that the WGL mechanism does not always improve the lowest type's utility compared to the GL mechanism. But, for every case where it reduces their utility there is another case where it increases their utility. Thus, a stable WGL mechanism helps "weakly more often" than it hurts.

First, we establish that any stable WGL mechanism can reduce welfare for the lowest type, and that such examples are not knife-edge.

Proposition 3. Fix $n = 3$ and any stable WGL mechanism with $\rho \max(A) < 1/3$ and $\gamma \to \infty$. There exists an open subset of $\mathscr{E}^{MS}$ for which the agent with the lowest type is worse off in the WGL mechanism than in the GL mechanism.

Corollary 1. By Lemma 6, if ρ is small and γ is large then for each economy in the open subset the minimum utility is lower in the WGL mechanism than in the GL mechanism.

Proof. Given the minimum surplus sorting condition, it is sufficient to show that the equilibrium utility under the WGL is lower for the agent with the lowest δ_i, since that agent also has the lowest Δ_i. Assume without loss that this is agent 1, so that $\delta_1 \leq 0$. If $\delta_1 = 0$ then $\delta = \vec{0}$ and both mechanisms give an equilibrium utility of Δ_i to each i, so consider $\delta_1 < 0$. For $n = 3$ we have that $A_1 = (0, -a, a)$ for some $a \neq 0$, so $W_1 = (1, -\rho a, \rho a)$, and the first stability condition reduces to $\rho|a| < 1/3$. Thus, an explicit solution for W^{-1} is easily obtained. By (6), $u_1^{WGL}(m^*) - u_1^{GL}(m^*) = [W_1^{-1}\delta - \delta_1]y^o$ when $\gamma \to \infty$. For $n = 3$,

$$\begin{aligned}
W_1^{-1}\delta - \delta_1 &= \frac{1}{1+3(\rho a)^2}\left[(1+(\rho a)^2)\delta_1 + (\rho a + (\rho a)^2)\delta_2 + (-\rho a + (\rho a)^2)\delta_3\right] - \delta_1 \\
&= \frac{\rho a}{1+3(\rho a)^2}\left[\delta_2 - \delta_3 + 3\rho a(\delta_2 + \delta_3)\right].
\end{aligned}$$

where the second equality follows from two applications of the fact that $\sum_i \delta_i = 0$. The WGL mechanism performs worse than the GL mechanism if this expression is negative. Since $(\delta_2 + \delta_3) = -\delta_1 > 0$, the condition for this expression to be negative is

$$\begin{aligned}
&\rho a < \frac{1}{3}\frac{\delta_3 - \delta_2}{\delta_3 + \delta_2} \quad \text{if } a > 0, \text{ and} \\
&\rho a > \frac{1}{3}\frac{\delta_3 - \delta_2}{\delta_3 + \delta_2} \quad \text{if } a < 0.
\end{aligned} \tag{10}$$

In the first case ($a > 0$) let $\delta_2 < 0$ and $\delta_3 > -\delta_2 > 0$ and the condition will be satisfied since $\frac{\delta_3 - \delta_2}{\delta_3 + \delta_2} > 1$ and we have already assumed $\rho a < 1/3$. There is an open set of (δ_2, δ_3) satisfying these restrictions, proving the proposition for $a > 0$.

In the case of $a < 0$ let $\delta_3 < 0$ and $\delta_2 > -\delta_3 > 0$ and the condition will be satisfied since $\frac{\delta_3 - \delta_2}{\delta_3 + \delta_2} < -1$ and we have already assumed $\rho a > -1/3$. There is an open set of (δ_2, δ_3) satisfying these restrictions, proving the proposition for $a < 0$. □

Remark. In the proof of Proposition 3, suppose $a > 0$. Since $\rho > 0$, the right-hand side of condition (10) must be positive. But consider a new economy in which agents 2 and 3 (thus, δ_2 and δ_3) are switched. Now the right-hand side of (10) becomes negative and the inequality is reversed. Thus, for this new economy the WGL mechanism gives higher utility to agent 1 than the GL mechanism. Similarly, if $a < 0$ then the economy in which

δ_2 and δ_3 are switched also reverses the inequality, meaning agent 1 prefers the WGL mechanism over the GL mechanism in this case as well.

The above remark shows that any 3-agent quasilinear economy for which the WGL is relatively worse for the lowest type can be permuted (switching the identities of agents 2 and 3) to give an economy where the WGL is relatively better for the lowest type. Thus, the WGL improves the lowest type's utility "more often" than it hurts. The following proposition confirms that this remains true for any $n \geq 3$.

Proposition 4. Let $\gamma \to \infty$. Fix a *WGL* mechanism with $A_1 = (0, -1, 0, \cdots, 0, 1)$. For any $E \in \mathscr{E}^{QL}$ such that the WGL mechanism gives lower utility (than the GL mechanism) to the agent with the lowest type, there is another economy $\tilde{E} \in \mathscr{E}^{QL}$ such that the WGL mechanism gives higher utility to the lowest type.

Corollary 2. For large γ and small ρ, for every economy where the WGL mechanism has the lower minimum utility (compared to GL) there is another economy where it has the higher minimum utility.

A proof appears in the appendix. Again, the main idea is to construct $\tilde{E}$ from E by switching identities of the two neighbors of the lowest type agent.

Expected Utility of the Lowest Type

Finally, our main result demonstrates that that the WGL improves welfare of the lowest type relative to the GL "in expectation." Since economies are randomly drawn, each agent's type δ_i is randomly drawn from some joint distribution F over economies that respect the minimum-surplus sorting property. The expected gain in surplus for the lowest type is then taken with respect to this distribution. Given that we condition on the lowest type, this expectation is a conditional expectation, conditional on the realization of the lowest type. Formally, an economy $E \in \mathscr{E}^{MS}$ is drawn and only the lowest value of δ_i is identified. Then we take the expectation of the others' types, conditional only on the fact that the minimal type is δ_i.[13]

[13]Equivalently, we can model the designer as having a hierarchical distribution over economies and "types." Specifically, an economy $E = ((u_i, \omega_i)_i, \kappa)$ is first drawn from a distribution over possible n-agent economies, and then the identities (or, indices i) are randomly permuted. One way to view this is that the vector $(u_i, \omega_i)_i$ represents the n "types" in the chosen economy, and these types are then randomly assigned to the agents. We assume the permutation from types to agents is drawn uniformly from the set of all permutations. Although this hierarchical model is different from the Bayesian model in the main text, they serve the same purpose because all we need is the interim belief $E[\delta_j | \delta_i, \delta_i = \min_k \delta_k] = -\frac{\delta_i}{n-1}$, which holds in both cases.

Our only assumption is that the sequence of random variables $\{\delta_i\}_{i=1}^n$ is exchangeable.[14] We make no other assumptions on F.

We can now state our main result.

Theorem 1. There exists a class of dynamically stable WGL mechanisms with either $A_1 = (0, -1, 0, \ldots, 0, 1)$ or $A_1 = (0, 1, 0, \ldots, 0, -1)$ such that for any $E \in \mathscr{E}^{MS}$ and $n \geq 3$, the consumer with the lowest type gets a higher expected equilibrium payoff in the WGL mechanism than in the GL mechanism.

Corollary 3. Fix any $E \in \mathscr{E}^{MS}$. For large γ, small ρ, and either $A_1 = (0, -1, 0, \ldots, 0, 1)$ or $A_1 = (0, 1, 0, \ldots, 0, -1)$, the WGL mechanism gives a higher expected minimum utility than the GL mechanism.

Proof of Theorem 1. First, we derive an expression for the expected surplus of the lowest type in the WGL mechanism.

Lemma 7. Suppose we know agent i's δ_i is minimal among $(\delta_1, \ldots, \delta_n)$, but we do not know the values of δ_j for any $j \neq i$. As $\gamma \to \infty$ agent i's expected equilibrium utility difference between the WGL and GL mechanisms is given by

$$E[u_i^{WGL}(m^*) - u_i^{GL}(m^*)|\delta_i, \delta_i = \min_k \delta_k] = \frac{n}{n-1}(W_{ii}^{-1} - 1)\delta_i y^o.$$

Proof of Lemma 7. Conditional on the fact that δ_i is minimal (and no other information about $j \neq i$) the interim belief about each $j \neq i$ given by

$$\begin{aligned} E[\delta_j|\delta_i, \delta_i = \min_k \delta_k] &= \frac{1}{n-1}\sum_{j \neq i} E\left[\delta_j|\delta_i, \delta_i = \min_k \delta_k\right] \\ &= \frac{1}{n-1} E\left[\sum_{j \neq i} \delta_j|\delta_i, \delta_i = \min_k \delta_k\right] \\ &= \frac{1}{n-1}(-\delta_i) \geq 0. \end{aligned}$$

The first equality comes from the exchangeability assumption of the joint type distribution, conditional only on δ_i and $\delta_i = \min_j \delta_j$. The third holds because it's true at every realization.

Using this, the expectation of the vector δ can be written as

$$E\left[\delta|\delta_i, \delta_i = \min_k \delta_k\right] = \frac{\delta_i}{n-1}(n e_i - \mathbb{1}),$$

[14]Sequence $\{\delta_i\}_{i=1}^n$ is exchangeable if for any permutation of the indices $1, 2, \cdots, n$ the joint distribution of the permuted sequence is the same as F. Note that the realization of real-valued random variables $\delta_1, \cdots, \delta_n$ is determined by the realization of function-valued random variables $v_1, \cdots, v_n$. In particular, the latter determines y^o through $\sum_i v_i'(y^o) = \kappa$ and $\delta_i = v_i'(y^o) - \alpha_i \kappa$. The exchangeability of δ distribution holds as long as we assume the exchangeability of v distribution.

where e_i is the ith standard basis vector.

From equation (7) we have that an agent i who is assigned the minimal δ_i will have an expected surplus gain of

$$E[u_i^{WGL}(m^*) - u_i^{GL}(m^*)|\delta_i, \delta_i = \min_k \delta_k] =$$

$$\left[W_i^{-1}(ne_i - \mathbb{1})\frac{\delta_i}{n-1} - \delta_i\right] y^o - \frac{1}{2\gamma}\frac{n-1}{n-2}E\left[\left[(W_i^{-1}\delta)^2 - \delta_i^2\right] - \frac{1}{n-1}\sum_{j\neq i}\left[(W_j^{-1}\delta)^2 - \delta_j^2\right]|\cdots\right]$$

$$= \frac{n}{n-1}(W_{ii}^{-1} - 1)\delta_i y^o - \frac{1}{2\gamma}\frac{n-1}{n-2}E\left[\left[(W_i^{-1}\delta)^2 - \delta_i^2\right] - \frac{1}{n-1}\sum_{j\neq i}\left[(W_j^{-1}\delta)^2 - \delta_j^2\right]|\cdots\right],$$

where all expectations are conditional on δ_i and $\delta_i = \min_k \delta_k$. The second equality comes because $W_i^{-1}\mathbb{1} = 1$. When $\gamma \to \infty$ the last term goes to zero, giving the result. □

Given Lemma 7, we simply need to verify that $(W_{ii}^{-1} - 1)\delta_i y^o$ is positive for the lowest type. But the minimal δ_i must be negative (or zero), so this reduces to $(W_{ii}^{-1} - 1)y^o < 0$. Since we assume all admissible economies have $y^o > 0$, then we require only that $W_{ii}^{-1} < 1$. Without loss we let agent 1 be the agent with the lowest type.

We prove the theorem by showing that $W_{ii}^{-1} \le 1$ when $A_1 = (0, -1, 0, \cdots, 0, 1)$ or $A_1 = (0, 1, 0, \cdots, 0, -1)$. To do so, we first prove a useful lemma about the monotonicity of W_{ii}^{-1}.

Lemma 8. $A_1 = (0, -1, 0, \cdots, 0, 1)$ and $A_1 = (0, 1, 0, \cdots, 0, -1)$ result in the same W_{ii}^{-1}. For any $n \ge 3$, and for every i, this W_{ii}^{-1} decreases in ρ.

The proof is in the appendix. Given Lemma 8, proving that $W_{ii}^{-1} \le 1$ simply requires that we prove that it converges to 1 as ρ approaches 0.

Note that $W = I + \rho A$ is a circulant matrix with 3 nonzero elements $(1, -\rho, \rho)$ and we need $\rho < \frac{1}{n}$ for stability. We directly obtain the analytical formula of W_{ii}^{-1} from Searle (1979):

$$W_{ii}^{-1} = W_{11}^{-1} = \frac{1}{\sqrt{1+4\rho^2}}\left(\frac{1}{1-z_1^n} - \frac{1}{1-z_2^n}\right)$$

where

$$z_1 = \frac{-1+\sqrt{1+4\rho^2}}{2\rho}, \quad z_2 = \frac{-1-\sqrt{1+4\rho^2}}{2\rho}.$$

We can analytically verify that $W_{ii}^{-1} \to 1$ as $\rho \to 0$, because $z_1 \to 0$ from above and $z_2 \to -\infty$. Note that the fraction $1/\sqrt{1+4\rho^2}$ converges to 1 from below and the term in parentheses also converges to 1. Thus, their product converges to 1, proving Theorem 1. □

Remark. This framework is related to, but distinct from those in the literature of Bayesian public good mechanisms. First, the aim here is to capture the idea of "improving IR in expectation." Although we are evaluating the IR performance by the interim utility, our mechanisms are still within a Nash implementation framework in terms of incentive compatibility, budget balance, and stability. In contrast, the key ingredient of the Bayesian public good mechanisms is Bayesian incentive compatibility (BIC), wherein agents do not know the realized preference of others when making choices. Second, we evaluate the expected utility for agent i conditional on not only δ_i, but also on the fact that $\delta_i = \min_k \delta_k$. The standard notion of interim IR, on the other hand, conditions only on the realization of δ_i.

Remark. Surprisingly, not all WGL mechanisms reduce IR failures in expectation. For example, if we construct the WGL mechanism based on the original Walker mechanism with $A_1 = (0, -1, 1, 0, \cdots, 0)$, then the hybrid mechanism can have worse performance in expectation for the lowest type, relative to the GL mechanism.

V. A Discussion of Finite γ

We obtain our main results under the restriction $\gamma \to \infty$ in order to guarantee dynamic stability in arbitrarily rich environments. For a certain economy or a restricted class of economies that deviates from the richness assumption, however, infinitely large γ is not necessary. Allowing for finite γ gives us further flexibility to improve IR by relaxing the stability constraint. Although a full analysis of IR comparisons is not tractable due to the complexity of the penalty terms, we demonstrate through examples how IR failures can be reduced while keeping γ finite.

Finite γ in the Groves-Ledyard Mechanism

It is easy to verify that for a fixed economy the equilibrium utility of any one agent in the GL mechanism is monotone in γ (either increasing or decreasing). The minimum utility is the lower envelope of these monotone functions, which will be quasiconcave. Thus, the set of γ for which the lowest utility is positive will either be empty or an interval.

Agent	a_i	b_i			
		economy (a)	economy (b)	economy (c)	economy (d)
1	1	40	118	30	30
2	4	50	20	40	60
3	8	78	30	98	78

TABLE I. Four 3-agent economies with quadratic preferences.

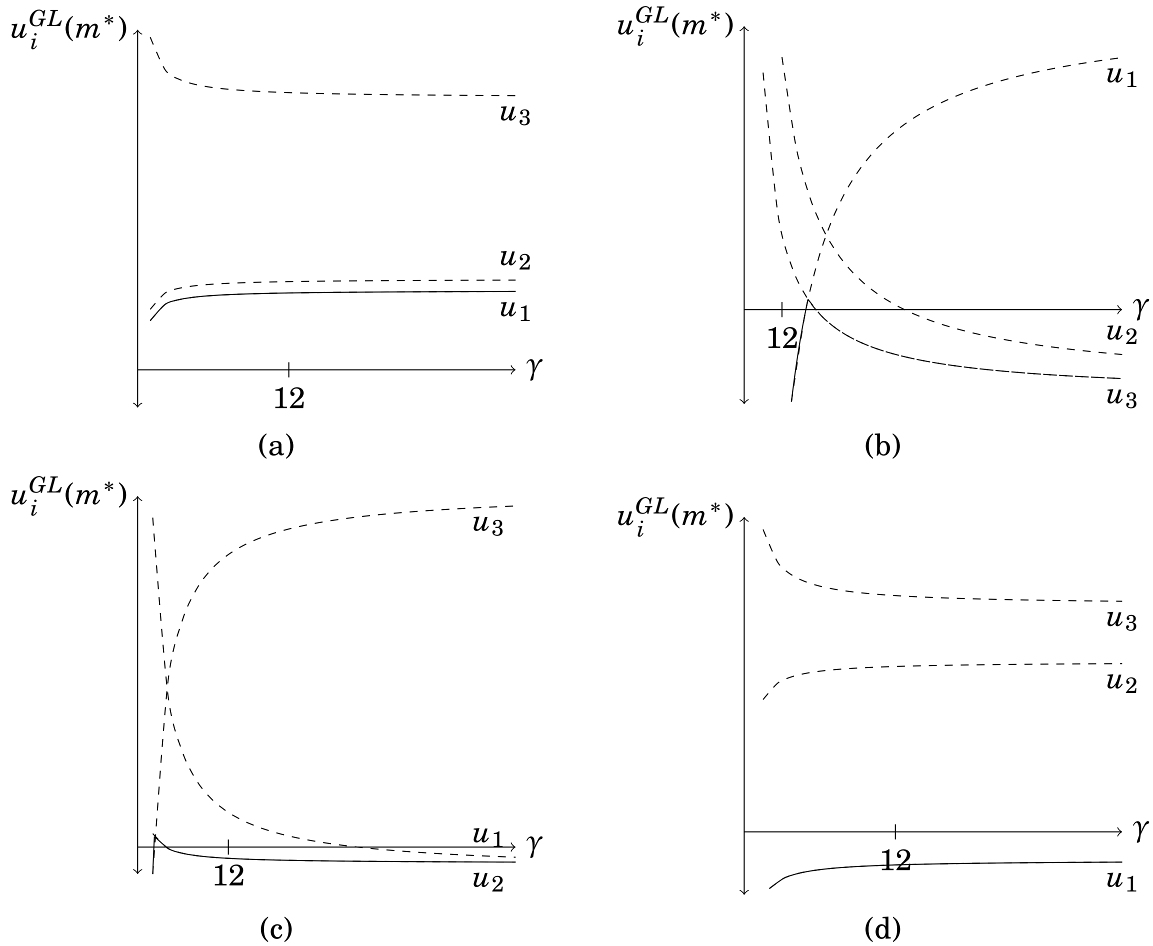

FIGURE I. Equilibrium utilities for each player in the Groves-Ledyard mechanism for the four economies shown in Table I. The solid curve shows the minimum utility, which must be positive for IR to be satisfied.

Figure I depicts the equilibrium utilities of agents as functions of γ in several examples of 3-agent economies, the preference parameters of which are summarized in Table I. In all four economies agents have quadratic preferences of the form $v_i(y) = b_i y - a_i y^2$, equal cost shares ($\alpha_i = \frac{1}{3}$), and a marginal cost of $\kappa = 90$. Furthermore, the economies all have the same quadratic coefficient vector a and differ only in the linear coefficient b's. In all four economies, the PO allocation is $y^o = 3$ and stability requires that $\gamma > 12$. Economy (a) and economy (d) belong to $\mathscr{E}^{MS}$ while the other two do not. In economy (a) a larger γ always improves equilibrium utility of the lowest type and can fully eliminate IR failures when $\gamma \in [\frac{8}{21}, \infty)$, which is compatible with stability. In economy (d), the IR failure cannot be eliminated for any γ, as u_1 remains negative for all γ. Economy (b) and economy (c) suggest that for some economies different values of γ might reverse the ranking of equilibrium utilities. In both cases there exists an interval of γ that guarantees IR for all agents—$\left[\frac{4994}{255}, \frac{409}{18}\right]$ in the former and $\left[\frac{71}{33}, \frac{11}{3}\right]$ in the latter—but IR is compatible with stability only in economy (b).

Finite γ in WGL mechanisms

Now we consider the WGL mechanism with finite γ. To illustrate how a finite γ can improve welfare we focus first on economy 4 mentioned above. Brute-force calculations give the following equilibrium utilities:

$$u_1(m^*) = -9 + 18\frac{\rho + 3\rho^2}{1+3\rho^2} - \frac{18}{\gamma}\frac{1-6\rho-3\rho^2}{(1+3\rho^2)^2},$$
$$u_2(m^*) = 54 + 18\frac{\rho - 3\rho^2}{1+3\rho^2} - \frac{18}{\gamma}\frac{1+6\rho-3\rho^2}{(1+3\rho^2)^2}, \text{ and}$$
$$u_3(m^*) = 72 - 36\frac{\rho}{1+3\rho^2} - \frac{18}{\gamma}\frac{6\rho^2-2}{(1+3\rho^2)^2}.$$

We can verify that IR is violated in the original GL mechanism since $\rho = 0$ leads to $u_1(m^*) \leq -9$ for all γ. But notice that if $\rho > 0$ then agent 1's IR failure can be avoided for large γ since the second term becomes positive. And $u_2(m^*)$ and $u_3(m^*)$ are positive as well. Therefore, to minimize the magnitude of the IR failure, we choose ρ to maximize $u_1^*(m^*)$ subject to the dynamic stability constraint of $\rho < 1/3$. Since $u_1(m^*)$ is increasing in ρ (when γ is large), and since $(\rho+3\rho^2)/(1+3\rho^2)$ converges to 1/2 as $\rho \to 1/3$, we have that $\sup u_1(m^*) = 0$. In other words, we can get arbitrarily close to eliminating the IR failure by letting $\gamma \to \infty$ and $\rho \to 1/3$.

Can the minimum utility be improved further by considering finite γ? For this economy the stability constraint can be computed as follows:

$$|\gamma(1+3\rho) + 3v_i''(y)| + |\gamma(1-3\rho) + 3v_i''(y)| < 2\gamma - 3v_i''(y) \;\; \forall i,$$

which reduces to $\{(\gamma,\rho) : \gamma > 12 \;\&\; \rho < \frac{1}{3} + \frac{1}{\gamma}\}$. We can obtain $\sup u_1(m^*) \approx 3.4$ at $\gamma = 12$ and $\rho = 5/12$. Thus, allowing for finite γ in the WGL mechanism completely eliminates the IR failure for this economy.

In these examples the minimal γ for stability is $\gamma = 12$, but for general economies this minimum can vary. And this defines the range over which γ should be chosen. Because of this complication, a general characterization of when IR can be improved with finite γ—and by how much—remains elusive.

VI. Concluding Thoughts

We have shown that it is possible to construct a PO mechanism that is simple, preserves stability, and improves IR in expectation over the GL mechanism. Although the message space is simple ($M_i = \mathbb{R}^1$), the tax function does involve more terms and, in that sense, is more complex. An open question is whether this sort of added complexity will hinder performance.

Furthermore, our results generally rely on small ρ (or, large γ relative to λ), which makes the mechanism act much more like a coordination game in which agents are severely penalized for deviating from others' announcements. Although this is stable in theory, it approaches a situation of multiplicity in which any message is approximately in equilibrium as long as all agents submit that message. It is possible this will effectively weaken the incentives and generate strategic uncertainty, undermining the equilibrium performance of the mechanism. Similar concerns were raised by Muench and Walker (1983) when considering the original GL mechanism with large γ, and Arifovic and Ledyard (2011) show that empirical convergence of the GL mechanism can actually become slower with very large values of γ.

Finally, the fact that players' taxes are heavily dependent on their neighbors' messages creates possible opportunities for collusion. Or animosity between neighbors. Future work could focus on anonymizing the relevant neighbors—or spreading the influence across many more neighbors—to increase the privacy of the mechanism. The mechanisms of Kim (1993), Chen (2002), and Healy and Mathevet (2012) accomplish this, but are not simple and not budget balanced out of equilibrium.[15] Many voting-based mechanisms preserve this sort of privacy, and perhaps this is why they are widely used despite their allocative inefficiencies, which in fact may become small for large economies (Ledyard and Palfrey, 2002).

References

Arifovic, J., Ledyard, J., 2011. A behavioral model for mechanism design: Individual evolutionary learning. Journal of Economic Behavior & Organization 78, 374–395. doi:10.1016/j.jebo.2011.01.021.

Bergstrom, T., Simon, C.P., Titus, C.J., 1983. Counting Groves-Ledyard Equilibria Via Degree Theory. Journal of Mathematical Economics 12, 167–184.

Bergstrom, T.C., Cornes, R.C., 1983. Independence of Allocative Efficiency from Distribution in the Theory of Public Goods. Econometrica 51, 1753–1765. doi:10.2307/1912115.

[15]Healy and Mathevet (2012) extend their mechanism to be budget balanced for all message profiles, but at considerable added complexity.

Chen, Y., 2002. A family of supermodular Nash mechanisms implementing Lindahl allocations. Economic Theory 19, 773–790.

Chen, Y., Plott, C.R., 1996. The Groves-Ledyard Mechanism: An Experimental Study of Institutional Design. Journal of Public Economics 59, 335–364.

Chen, Y., Tang, F.F., 1998. Learning and Incentive-Compatible Mechanisms for Public Goods Provision: An Experimantal Study. Journal of Political Economy 106, 633–662.

Conlisk, J., 1973. Quick Stability Checks and Matrix Norms. Economica 40, 402–409.

Groves, T., Ledyard, J.O., 1977. Optimal Allocation of Public Goods: A Solution to the 'Free-Rider' Problem. Econometrica 45, 783–809.

Groves, T., Ledyard, J.O., 1980. The Existence of Efficient and Incentive Compatible Equilibria With Public Goods. Econometrica 48, 1487–1506.

Healy, P.J., 2006. Learning dynamics for mechanism design: An experimental comparison of public goods mechanisms. Journal of Economic Theory 129, 114–149. doi:10.1016/j.jet.2005.03.002.

Healy, P.J., Jain, R., 2017. Generalized Groves–Ledyard mechanisms. Games and Economic Behavior 101, 204–217. doi:10.1016/j.geb.2016.01.005.

Healy, P.J., Mathevet, L., 2012. Designing stable mechanisms for economic environments. Theoretical Economics 7, 609–661. doi:10.3982/TE898.

Hurwicz, L., 1979a. On allocations attainable through Nash equilibria. Journal of Economic Theory 21, 140–165.

Hurwicz, L., 1979b. Outcome Functions Yielding Walrasian and Lindahl Allocations at Nash Equilibrium Points. Review of Economic Studies 46, 217–225.

Jordan, J.S., 1986. Instability in the Implementation of Walrasian Allocations. Journal of Economic Theory 39, 301–328.

Kim, T., 1987. Stability Problems in the Implementation of Lindahl Allocations. Ph.D. thesis. University of Minnesota.

Kim, T., 1993. A stable Nash mechanism implementing Lindahl allocations for quasilinear environments. Journal of Mathematical Economics 22, 359–371.

Kim, T., 1996. A Stable Nash Mechanism for Quasi-Additive Public-Good Environments. The Japanese Economic Review 47, 144–156.

Ledyard, J.O., Palfrey, T.R., 2002. The approximation of efficient public good mechanisms by simple voting schemes. Journal of Public Economics 83, 153–171. doi:10.1016/S0047-2727(00)00161-4.

Marchant, T., Mishra, D., 2015. Mechanism design with two alternatives in quasi-linear environments. Social Choice and Welfare 44, 433–455.

Maskin, E., 1999. Nash equilibrium and welfare optimality. Review of Economic Studies 66, 23–38.

Moulin, H., Shenker, S., 2001. Strategyproof sharing of submodular costs: Budget balance versus efficiency. Economic Theory 18, 511–533.

Muench, T., Walker, M., 1983. Are Groves-Ledyard Equilibria Attainable? Review of Economic Studies 50, 393–396.

Page, S.E., Tassier, T., 2004. Equilibrium Selection and Stability for the Groves Ledyard Mechanism. Journal of Public Economic Theory 6, 311–335.

Page, S.E., Tassier, T., 2010. A characterization of equilibria in the Groves–Ledyard mechanism. Journal of Mathematical Economics 46, 1229–1242. doi:10.1016/j.jmateco.2010.08.016.

Samuelson, P.A., 1954. The pure theory of public expenditure. Review of Economics and Statistics 36, 387–389.

Searle, S.R., 1979. On inverting circulant matrices. Linear Algebra and its Applications 25, 77–89. doi:10.1016/0024-3795(79)90007-7.

Tian, G., 1990. Completely Feasible and Continuous Implementation of the Lindahl Correspondence with a Message Spce of Minimal Dimension. Journal of Economic Theory 51, 443–452.

Van Essen, M., Lazzati, N., Walker, M., 2012. Out-of-equilibrium performance of three Lindahl mechanisms: Experimental evidence. Games and Economic Behavior 74, 366–381.

Van Essen, M.J., 2013. A simple supermodular mechanism that implements Lindahl allocations. Journal of Public Economic Theory 15, 363–377.

Walker, M., 1981. A Simple Incentive Compatible Scheme for Attaining Lindahl Allocations. Econometrica 49, 65–71.

Appendix A. Proofs

Proof of Proposition 2

Recall that $\beta_i(m_{-i})$ is the best response function. We use the implicit function theorem approach to derive its derivatives. The first-order condition from utility maximization gives that

$$v'\left(\beta_i(m_{-i})+\sum_{j\neq i} m_j\right)-\alpha_i\kappa-\lambda A_i m-\gamma\frac{n-1}{n}\left(\beta_i(m_{-i})-\frac{1}{n-1}\sum_{j\neq i} m_j\right)=0$$

for every m_{-i}. Differentiating with respect to any m_j ($j\neq i$) gives

$$\frac{\partial\beta_i(m_{-i})}{m_j}=\frac{\gamma\left(1-n\rho A_{ij}\right)-n(-v_i''(y))}{\gamma(n-1)-nv_i''(y)}.$$

The mechanism is contractive if, for every i, $\sum_{j\neq i}|\partial\beta_i(m)/\partial m_j|<1$, which is equivalent to

$$\sum_{j\neq i}\frac{\left|\gamma\left(1-n\rho A_{ij}\right)-n(-v_i''(y))\right|}{\gamma(n-1)+n(-v_i''(y))}<1. \tag{11}$$

If $\rho<1/(n\max(A))$ and

$$\gamma>\frac{n\eta}{1-n\rho\max(A)}$$

then the terms inside the absolute-value signs are positive and the stability condition can be rewritten as

$$\sum_{j\neq i}\left(\gamma(1-n\rho A_{ij})-n(-v_i''(y))\right)<\gamma(n-1)+n(-v_i''(y)).$$

But

$$\begin{aligned}\sum_{j\neq i}\left(\gamma(1-n\rho A_{ij})-n(-v_i''(y))\right)&=\gamma(n-1)-n(n-1)(-v_i''(y))\\&<\gamma(n-1)+n(-v_i''(y)),\end{aligned}$$

so the stability condition is satisfied.

Proof of Proposition 4

WLOG assume agent 1 has the lowest type, so that $\delta_1\le 0$ and $\delta_j-\delta_1\ge 0$ for all j. If $\delta_1=0,\delta_j=0$ for all j and the problem is trivial. If $\delta_1<0$, we compute as $\gamma\to\infty$

$$\begin{aligned}u_1^{WGL}-u_1^{GL}&=W_1^{-1}\delta-\delta_1\\&=\sum_{j\neq 1}W_{1j}^{-1}(\delta_j-\delta_1),\end{aligned}\tag{12}$$

where the first equality follows lemma 5 and the second equality holds because $W \cdot \mathbb{1} = \mathbb{1}$ implies $W^{-1} \cdot \mathbb{1} = \mathbb{1}$ (its rows sum to one). By Searle (1979) we have

$$\begin{aligned} W_{1j}^{-1} &= \frac{1}{\sqrt{1+4\rho^2}} \left(\frac{z_1^{j-1}}{1-z_1^n} - \frac{z_2^{j-1}}{1-z_2^n} \right) \\ &= \frac{1}{\sqrt{1+4\rho^2}} \frac{z_1^{j-1} - z_2^{j-1} + (-1)^{j-1}(z_1^{n-j+1} - z_2^{n-j+1})}{(1-z_1^n)(1-z_2^n)} \end{aligned}$$

where

$$z_1 = \frac{-1+\sqrt{1+4\rho^2}}{2\rho}, \quad z_2 = \frac{-1-\sqrt{1+4\rho^2}}{2\rho}.$$

The last equality holds because $z_1 z_2 = -1$.

Suppose for a certain economy E such that $\sum_{j \neq 1} W_{1j}^{-1}(\delta_j - \delta_1) < 0$, we consider an alternative economy $\tilde{E}$ obtained by switching the identity of agent 2 and agent n. In equation (12) this is mathematically equivalent to keeping the economy at E but replacing ρ with $-\rho$, which has the effect of switching the signs of z_1 and z_2. For the new economy agent 1 is still the lowest type agent and we want to show that

$$\tilde{u}_1^{WGL} - \tilde{u}_1^{GL} = \sum_{j \neq 1} \tilde{W}_{1j}^{-1}(\delta_j - \delta_1) > 0$$

where

$$\tilde{W}_{1j}^{-1} = \frac{1}{\sqrt{1+4\rho^2}} \frac{(-z_1)^{j-1} - (-z_2)^{j-1} + (-1)^{j-1}\left[(-z_1)^{n-j+1} - (-z_2)^{n-j+1}\right]}{(1-(-z_1)^n)(1-(-z_2)^n)}.$$

Case 1. n is even.

When j is odd, we have

$$W_{1j}^{-1} = \tilde{W}_{1j}^{-1} = \frac{1}{\sqrt{1+4\rho^2}} \frac{z_1^{j-1} - z_2^{j-1} + z_1^{n-j+1} - z_2^{n-j+1}}{(1-z_1^n)(1-z_2^n)} > 0.$$

The inequality holds because for all $|\rho| < \frac{1}{n}$ we have $|z_1| < 1$ and $|z_2| > 1$. This implies

$$\sum_{j \neq 1, j \text{ odd}} W_{1j}^{-1}(\delta_j - \delta_1) > 0.$$

When j is even, we have

$$\sum_{j \neq 1, j \text{ even}} W_{1j}^{-1}(\delta_j - \delta_1) = \sum_{j \neq 1, j \text{ even}} \frac{1}{\sqrt{1+4\rho^2}} \frac{z_1^{j-1} - z_2^{j-1} - z_1^{n-j+1} + z_2^{n-j+1}}{(1-z_1^n)(1-z_2^n)} (\delta_j - \delta_1) < 0.$$

The inequality holds because by assumption $\delta_j - \delta_1 > 0$, $\sum_{j \neq 1} W_{1j}^{-1}(\delta_j - \delta_1) < 0$ and $\sum_{j \neq 1, j \text{ odd}} W_{1j}^{-1}(\delta_j - \delta_1) > 0$.

Meanwhile we have $\tilde{W}_{1j}^{-1} = -W_{1j}^{-1} > 0$ for any j that is even, so that

$$\begin{aligned}
\sum_{j\neq 1} \tilde{W}_{1j}^{-1}(\delta_j - \delta_1) &= \sum_{j\neq 1, j \text{ odd}} \tilde{W}_{1j}^{-1}(\delta_j - \delta_1) + \sum_{j\neq 1, j \text{ even}} \tilde{W}_{1j}^{-1}(\delta_j - \delta_1) \\
&= \sum_{j\neq 1, j \text{ odd}} W_{1j}^{-1}(\delta_j - \delta_1) - \sum_{j\neq 1, j \text{ even}} W_{1j}^{-1}(\delta_j - \delta_1) \\
&> 0,
\end{aligned}$$

giving the result.

Case 2. n is odd.

When j is odd, we have

$$W_{1j}^{-1} = \frac{1}{\sqrt{1+4\rho^2}} \frac{z_1^{j-1} - z_2^{j-1} + z_1^{n-j+1} - z_2^{n-j+1}}{(1-z_1^n)(1-z_2^n)}.$$

When j is even, we have

$$W_{1j}^{-1} = \frac{1}{\sqrt{1+4\rho^2}} \frac{z_1^{j-1} - z_2^{j-1} - z_1^{n-j+1} + z_2^{n-j+1}}{(1-z_1^n)(1-z_2^n)}.$$

Note that

$$\begin{aligned}
\sum_{j\neq 1} W_{1j}^{-1}(\delta_j - \delta_1) &= \sum_{j\neq 1, j \text{ odd}} W_{1j}^{-1}(\delta_j - \delta_1) + \sum_{j\neq 1, j \text{ even}} W_{1j}^{-1}(\delta_j - \delta_1) \\
&= \frac{1}{\sqrt{1+4\rho^2}} \frac{1}{(1-z_1^n)(1-z_2^n)} \cdot \\
&\quad \Bigg[\sum_{j\neq 1, j \text{ odd}} (z_1^{j-1} - z_2^{j-1})(\delta_j - \delta_1) + \sum_{j\neq 1, j \text{ even}} (z_1^{j-1} - z_2^{j-1})(\delta_j - \delta_1) \\
&\quad + \sum_{j\neq 1, j \text{ odd}} (z_1^{n-j+1} - z_2^{n-j+1})(\delta_j - \delta_1) - \sum_{j\neq 1, j \text{ even}} (z_1^{n-j+1} - z_2^{n-j+1})(\delta_j - \delta_1) \Bigg] \\
&= \frac{1}{\sqrt{1+4\rho^2}} \frac{1}{(1-z_1^n)(1-z_2^n)} \cdot \\
&\quad \Bigg[\sum_{j\neq 1, j \text{ even}} (z_1^{j} - z_2^{j})(\delta_{j+1} - \delta_1) + \sum_{j\neq 1, j \text{ even}} (z_1^{n-j} - z_2^{n-j})(\delta_{n-j+1} - \delta_1) \\
&\quad + \sum_{j\neq 1, j \text{ even}} (z_1^{n-j} - z_2^{n-j})(\delta_{j+1} - \delta_1) - \sum_{j\neq 1, j \text{ even}} (z_1^{j} - z_2^{j})(\delta_{n-j+1} - \delta_1) \Bigg] \\
&= \frac{1}{\sqrt{1+4\rho^2}} \frac{1}{(1-z_1^n)(1-z_2^n)} \cdot \\
&\quad \Bigg[\sum_{j\neq 1, j \text{ even}} (z_1^{j} - z_2^{j})(\delta_{j+1} - \delta_{n-j+1}) + \sum_{j\neq 1, j \text{ even}} (z_1^{n-j} - z_2^{n-j})(\delta_{n-j+1} + \delta_{j+1} - 2\delta_1) \Bigg].
\end{aligned}$$

where the third equality obtains by replacing each j by $j-1$ when j is odd and replacing each $j-1$ by $n-j$ when j is even, and the first term is positive

$$\frac{\sum_{j\neq 1, j \text{ even}}(z_1^{n-j}-z_2^{n-j})(\delta_{n-j+1}+\delta_{j+1}-2\delta_1)}{\sqrt{1+4\rho^2}(1-z_1^n)(1-z_2^n)}>0$$

because $z_1>0, z_2<0$, n is odd, j is even and that δ_1 is the lowest type. If $\sum_{j\neq 1} W_{1j}^{-1}(\delta_j-\delta_1)<0$, it must be that the second half is negative

$$\frac{\sum_{j\neq 1, j \text{ even}}(z_1^{j}-z_2^{j})(\delta_{j+1}-\delta_{n-j+1})}{\sqrt{1+4\rho^2}(1-z_1^n)(1-z_2^n)}<0.$$

If we replace ρ by $-\rho$,

$$\begin{aligned}\sum_{j\neq 1}\tilde{W}_{1j}^{-1}(\delta_j-\delta_1)=&\frac{1}{\sqrt{1+4\rho^2}}\frac{1}{-(1-z_1^n)(1-z_2^n)}\cdot\\ &\left[\sum_{j\neq 1, j \text{ even}}(z_1^{j}-z_2^{j})(\delta_{j+1}-\delta_{n-j+1})-\sum_{j\neq 1, j \text{ even}}(z_1^{n-j}-z_2^{n-j})(\delta_{n-j+1}+\delta_{j+1}-2\delta_1)\right]\\ =&\frac{1}{\sqrt{1+4\rho^2}}\frac{1}{(1-z_1^n)(1-z_2^n)}\cdot\\ &\left[-\sum_{j\neq 1, j \text{ even}}(z_1^{j}-z_2^{j})(\delta_{j+1}-\delta_{n-j+1})+\sum_{j\neq 1, j \text{ even}}(z_1^{n-j}-z_2^{n-j})(\delta_{n-j+1}+\delta_{j+1}-2\delta_1)\right]\\ >&0,\end{aligned}$$

which is what we wish to prove. Note that the first equality holds because n is odd, $z_1z_2=-1$, and

$$\begin{aligned}[1-(-z_1)^n][1-(-z_2)^n]&=(1+z_1^n)(1+z_2^n)\\ &=1+z_1^n+z_2^n+z_1^nz_2^n\\ &=z_1^n+z_2^n\\ &=-1+z_1^n+z_2^n-z_1^nz_2^n\\ &=-(1-z_1^n)(1-z_2^n).\end{aligned}$$

Proof of Lemma 8

We first establish that the two forms of A_1 we propose lead to the same W_{ii}. Recall that by Searle (1979) we have

$$W_{ii}^{-1}=W_{11}^{-1}=\frac{1}{\sqrt{1+4\rho^2}}\left(\frac{1}{1-z_1^n}-\frac{1}{1-z_2^n}\right)$$

where

$$z_1 = \frac{-1+\sqrt{1+4\rho^2}}{2\rho}, \quad z_2 = \frac{-1-\sqrt{1+4\rho^2}}{2\rho}.$$

Let

$$f(\rho) = \frac{1}{\sqrt{1+4\rho^2}}, \quad g(\rho) = \frac{1}{1-z_1^n} - \frac{1}{1-z_2^n},$$

It is immediate that $f(\rho) = f(-\rho)$, so we only need to check $g(\rho) = g(-\rho)$. When n is even, it holds naturally as $z_h^2(\rho) = z_h^2(-\rho)$ for $h = 1,2$. When n is odd, we have:

$$\begin{aligned}
g(\rho) - g(-\rho) &= \left(\frac{1}{1-z_1^n} - \frac{1}{1-z_2^n}\right) - \left(\frac{1}{1+z_1^n} - \frac{1}{1+z_2^n}\right) \\
&= \frac{z_1^n - z_2^n}{(1-z_1^n)(1-z_2^n)} + \frac{z_1^n - z_2^n}{(1+z_1^n)(1+z_2^n)} \\
&= \frac{2(z_1^n - z_2^n)(1+z_1^n z_2^n)}{(1-z_1^n)(1-z_2^n)(1+z_1^n)(1+z_2^n)} \\
&= 0.
\end{aligned}$$

The last equality holds because $z_1 z_2 = \frac{1-(1+4\rho^2)}{4\rho^2} = -1$.

Now we establish monotonicity. We have

$$\begin{aligned}
f'(\rho) &= -4\rho(1+4\rho^2)^{-3/2} = \frac{-4\rho}{1+4\rho^2} f(\rho), \\
g'(\rho) &= \frac{n z_1^{n-1} z_1'(\rho)}{(1-z_1^n)^2} - \frac{n z_2^{n-1} z_2'(\rho)}{(1-z_2^n)^2} = \frac{\big(1-f(\rho)\big) n z_1^{n-1}}{2\rho^2(1-z_1^n)^2} - \frac{\big(1+f(\rho)\big) n z_2^{n-1}}{2\rho^2(1-z_2^n)^2}.
\end{aligned}$$

Substitute these two terms into $\frac{d}{d\rho}W_{11}^{-1}$, we obtain

$$\begin{aligned}
\frac{d}{d\rho}W_{11}^{-1} =& f'(\rho)g(\rho)+g'(\rho)f(\rho)\\
=& f(\rho)\left[\frac{-4\rho}{1+4\rho^2}g(\rho)+g'(\rho)\right]\\
=& f(\rho)\left[\frac{-4\rho}{1+4\rho^2}\left(\frac{1}{1-z_1^n}-\frac{1}{1-z_2^n}\right)+\frac{\big(1-f(\rho)\big)nz_1^{n-1}}{2\rho^2(1-z_1^n)^2}-\frac{\big(1+f(\rho)\big)nz_2^{n-1}}{2\rho^2(1-z_2^n)^2}\right]\\
=& f(\rho)\left[\frac{-4\rho}{1+4\rho^2}\frac{z_1^n-z_2^n}{(1-z_1^n)(1-z_2^n)}+\frac{\big(1-f(\rho)\big)nz_1^{n-1}(1-z_2^n)^2-\big(1+f(\rho)\big)nz_2^{n-1}(1-z_1^n)^2}{2\rho^2(1-z_1^n)^2(1-z_2^n)^2}\right]\\
=& f(\rho)\left[\frac{-4\rho}{1+4\rho^2}\frac{z_1^n-z_2^n}{(1-z_1^n)(1-z_2^n)}+\frac{2\rho f(\rho)nz_1^n(1-z_2^n)^2+2\rho f(\rho)nz_2^n(1-z_1^n)^2}{2\rho^2(1-z_1^n)^2(1-z_2^n)^2}\right]\\
=& \frac{f(\rho)^2\left[-4\rho^2f(\rho)(z_1^n-z_2^n)(1-z_1^n)(1-z_2^n)+nz_1^n(1-z_2^n)^2+nz_2^n(1-z_1^n)^2\right]}{\rho(1-z_1^n)^2(1-z_2^n)^2}\\
=& \frac{f(\rho)^2\left[-4\rho^2f(\rho)(z_1^n-z_2^n)(1-z_1^n)(1-z_2^n)+n(1-z_2^n)(z_1^n-(-1)^n)+n(1-z_1^n)(z_2^n-(-1)^n)\right]}{\rho(1-z_1^n)^2(1-z_2^n)^2}
\end{aligned}$$

where the fifth equality is because $2\rho f(\rho)z_1=\frac{2\rho}{\sqrt{1+4\rho^2}}\frac{-1+\sqrt{1+4\rho^2}}{2\rho}=1-\frac{1}{\sqrt{1+4\rho^2}}=1-f(\rho)$ and, similarly, $2\rho f(\rho)z_2=-1-f(\rho)$. The sixth equality holds because $\frac{-4\rho}{1+4\rho^2}=-4\rho f(\rho)^2$. The last equality holds because $z_1z_2=-1$. To show that this term is negative, we further simplify the expression by considering two cases: n is odd and n is even.

Case 1. When n is odd,

$$\begin{aligned}
\frac{d}{d\rho}W_{11}^{-1} =& \frac{f(\rho)^2\left[4\rho^2f(\rho)(z_1^n-z_2^n)(z_1^n+z_2^n)+n(1-z_2^n)(z_1^n+1)+n(1-z_1^n)(z_2^n+1)\right]}{\rho(1-z_1^n)^2(1-z_2^n)^2}\\
=& \frac{f(\rho)^2\left[4\rho^2f(\rho)(z_1^n-z_2^n)(z_1^n+z_2^n)+n(1-z_2^n+z_1^n-z_1^nz_2^n)+n(1-z_1^n+z_2^n-z_1^nz_2^n)\right]}{\rho(1-z_1^n)^2(1-z_2^n)^2}\\
=& \frac{4f(\rho)^2\left[\rho^2f(\rho)(z_1^n-z_2^n)(z_1^n+z_2^n)+n\right]}{\rho(1-z_1^n)^2(1-z_2^n)^2}.
\end{aligned}$$

The first equality holds because n is odd. The third equality holds because $z_1z_2=-1$. This result implies that the sign of $\frac{d}{d\rho}W_{11}^{-1}$ is the same as that of $\rho^2f(\rho)(z_1^n-z_2^n)(z_1^n+$

$z_2^n)+n$. We argue that this term is negative, because for any ρ we consider in this paper

$$\begin{aligned}
\rho^2 f(\rho)(z_1^n - z_2^n)(z_1^n + z_2^n) + n =& \rho^2 f(\rho)(z_1^{2n} - z_2^{2n}) + n \\
\leq& \rho^2 f(\rho)(z_1^2 - z_2^2) + 1 \\
=& \frac{\rho^2}{\sqrt{1+4\rho^2}} \left(\frac{1+2\rho^2 - \sqrt{1+4\rho^2}}{2\rho^2} - \frac{1+2\rho^2 + \sqrt{1+4\rho^2}}{2\rho^2} \right) + 1 \\
=& 0.
\end{aligned}$$

where the inequality follows from the next lemma.

Lemma 9. $\rho^2 f(\rho)(z_1^{2n} - z_2^{2n}) + n$ decreases with n for any $\rho < \frac{1}{3}$.

Proof. We directly check the first order derivative:[16]

$$\begin{aligned}
\frac{d}{dn}\left[\rho^2 f(\rho)(z_1^{2n} - z_2^{2n}) + n\right] =& \rho^2 f(\rho)\left(z_1^{2n}\ln(z_1^2) - z_2^{2n}\ln(z_2^2)\right) + 1 \\
=& \rho^2 f(\rho) z_1^{2n}\ln(z_1^2) - \rho^2 f(\rho) z_2^{2n}\ln(z_2^2) + 1 \\
<& -\rho^2 f(\rho) z_2^2 \ln(z_2^2) + 1 \\
=& -\frac{\rho^2}{\sqrt{1+4\rho^2}} \frac{1+2\rho^2+\sqrt{1+4\rho^2}}{2\rho^2} \ln\frac{1+2\rho^2+\sqrt{1+4\rho^2}}{2\rho^2} + 1 \\
=& -\frac{1+2\rho^2+\sqrt{1+4\rho^2}}{2\sqrt{1+4\rho^2}} \ln\frac{1+2\rho^2+\sqrt{1+4\rho^2}}{2\rho^2} + 1 \\
<& -\frac{1+2\rho^2+\sqrt{1+4\rho^2}}{2\sqrt{1+4\rho^2}} 2.3895 + 1 \\
<& -2.38 + 1 \\
<& 0.
\end{aligned}$$

where the first inequality holds because 1) $z_1^2 < 1$ and, hence, $\rho^2 f(\rho)\ln(z_1^2) < 0$, and 2) $z_2^2 > 1$ which implies that $-\rho^2 f(\rho) z_2^{2n}\ln(z_2^2)$ decreases in n. We obtain the second inequality by letting $\rho = \frac{1}{3}$ which is the least upper bound of feasible ρ that satisfies dynamic stability. The inequality holds because $\frac{1+2\rho^2+\sqrt{1+4\rho^2}}{2\rho^2}$ decreases in ρ, which can be easily verified by its first order derivative. We have the third inequality because $\frac{1+2\rho^2+\sqrt{1+4\rho^2}}{2\sqrt{1+4\rho^2}}$ increases in ρ and, hence, we obtain the upper bound of its negation by letting $\rho = 0$ within this term. □

[16] Note that n is an integer larger than or equal to 3 in our original problem but here it suffices to show the monotonicity by treating it as a continuous variable.

Case 2. When n is even,

$$\begin{aligned}\frac{d}{d\rho}W_{11}^{-1} =&\frac{f(\rho)^2\left[-4\rho^2 f(\rho)(z_1^n - z_2^n)(2 - z_1^n - z_2^n) + n(1 - z_2^n)(z_1^n - 1) + n(1 - z_1^n)(z_2^n - 1)\right]}{\rho(1 - z_1^n)^2(1 - z_2^n)^2}\\ =&\frac{f(\rho)^2\left[-4\rho^2 f(\rho)(z_1^n - z_2^n)(2 - z_1^n - z_2^n) - n(2 - z_1^n - z_2^n) - n(2 - z_1^n - z_2^n)\right]}{\rho(1 - z_1^n)^2(1 - z_2^n)^2}\\ =&\frac{2f(\rho)^2(-2 + z_1^n + z_2^n)\left[2\rho^2 f(\rho)(z_1^n - z_2^n) + n\right]}{\rho(1 - z_1^n)^2(1 - z_2^n)^2}.\end{aligned}$$

The first inequality holds because n is even and the second is because $(z_1z_2)^n = (-1)^n = 1$. In this case the sign of $\frac{d}{d\rho}W_{11}^{-1}$ is the same as that of $2\rho^2 f(\rho)(z_1^n - z_2^n) + n$. This is because when n is even,

$$\begin{aligned}-2 + z_1^n + z_2^n =& -2 + |z_1|^n + |z_2|^n\\ =& -2 + |z_1|^n + \frac{1}{|z_1|^n}\\ >&0.\end{aligned}$$

We can now prove that

$$\begin{aligned}2\rho^2 f(\rho)(z_1^n - z_2^n) + n =&2\rho^2 f(\rho)(z_1^{2m} - z_2^{2m}) + 2m\\ <&0.\end{aligned}$$

The equality is obtained by letting $n = 2m$ which is feasible since n is even. The inequality follows directly from case 1. Note that although we assume that n is odd in case 1 to get the formula of $\frac{d}{d\rho}W_{11}^{-1}$, the proof of the inequality does not.