The Existence and Stability of Equilibria in the Groves Ledyard Mechanism

Scott E. Page*
Troy Tassier†

September 4, 2003

Abstract

In this paper, we describe all interior and boundary equilibria of the Groves-Ledyard mechanism and test their stability. We provide closed form solutions for the equilibria discovered by Bergstrom, Simon, and Titus and show that these equilibria fail to exist when the punishment parameter is high. We further show that the Groves-Ledyard equilibrium is the unique stable interior equilibrium for all values of the punishment parameter. Interestingly, this stability rests on two dynamics each of which is unstable. We also show that all of the equilibria found by Bergstrom, Simon, and Titus are unstable. Finally, we locate an additional set of boundary equilibria of the same cardinality as the Bergstrom Simon and Titus equilibria and show that these are inefficient and stable. Yet, they also only exist for low values of the punishment parameter.

1 Introduction

In their seminal paper, Groves and Ledyard (1976) construct a balanced incentive compatible mechanism for an economy that includes both public and private goods. Their solution to the free rider problem remains a major contribution to the theory of public goods and to mechanism design. It demonstrates that it is possible to construct payoff functions that yield balanced and efficient outcomes. The Groves Ledyard
mechanism accomplishes these twin goals through an ingenious penalty function that induces agents to contribute to the public good by punishing them for deviating from the average contribution levels of the other agents in the economy. At the same time, it rewards agents if they deviate from the mean less than others do on average in such a way as to maintain balance.

In their original paper, Groves and Ledyard showed that all interior equilibria of their mechanism provided for efficient levels of the public goods. What they did not show was that the mechanism had a unique equilibrium, nor did they address the stability of the equilibrium they derived, though uniqueness does hold for quasi linear preferences, in which case the equilibrium is also stable. Following the publication of the Groves Ledyard mechanism, Bergstrom, Simon, and Titus (1983) showed that the Groves Ledyard Mechanism has multiple equilibria for a more general class of utility functions that allow the benefits from private goods to depend upon the level of the public good provided. Actually, this understates their result. They found that with \( N \) agents and a single public good, the Groves Ledyard mechanism has nearly \( N! \) over two equilibria. So with even modest numbers of agents, the number of equilibria can run into the billions. They, however, also undertook no analysis of the stability of these equilibria.

A close reading of these two papers reveals important open questions: First, given the symmetry among these Bergstrom, Simon, and Titus, equilibria – each subset of more than half the agents over-contributes while the remainder of agents under-contribute – these equilibria should all be stable or all be unstable. Second, for some values of \( N \), the total number of equilibria found jointly by Groves, Ledyard, Bergstrom, Simon and Titus is even. These two facts suggest that even more equilibria exist. Third, increasing the punishment parameter in the mechanism creates an incentive for agents to conform. In light of that fact, it seems unlikely that the
asymmetric equilibria found by Bergstrom, Simon, and Titus should exist with large punishments. Finally, even though Bergstrom, Simon, and Titus proved that the additional equilibria exist, they did not solve for them directly. They relied on a change of basis making it difficult to understand these equilibria intuitively.

In this paper, we fill those gaps. First, we explicitly solve for the Bergstrom, Simon, and Titus (hereafter BST) equilibria. We then identify an additional set of equilibria that lie at the boundary of the message space, or at negative and positive infinity in the case of unbounded message spaces. One of these equilibria exists for each BST equilibrium. This effectively doubles what was already an enormous number of equilibria. These equilibria, unlike the others, are not efficient. We further find that these boundary equilibria as well as the BST equilibria fail to exist when the punishment parameter is large. Thus, as intuition suggests, the BST equilibria (as well as the newly found boundary equilibria) fail to exist when punishments become severe. In such cases, the Groves Ledyard equilibrium becomes unique. We then state claims that imply that the Groves Ledyard equilibrium and the BST equilibria comprise all of the interior equilibria for the mechanism and offer two proofs as to why each produces the efficient level of the public good.

To test the stability of the equilibria, we use linear stability analysis. We find that the Groves Ledyard equilibrium is the unique stable interior equilibrium of the mechanism. The BST equilibria are all unstable. Interestingly, the stability of the Groves Ledyard equilibria rests on a combination of two unstable dynamics aggregating to form a stable dynamic. Moreover, these instabilities have intuitive interpretations as does the resulting aggregate stability. The boundary equilibria are also stable, as would be expected from degree theory, but as we remarked in the previous paragraph, none of them are efficient. At the end of this paper, we discuss how a budget constraint, like a high punishment parameter, can also wipe out the
BST equilibria and the boundary equilibria, making the Groves Ledyard equilibrium unique. Thus, not only does the Groves Ledyard mechanism have a unique equilibrium for quasi linear preferences, it also does for more general preferences given either a budget or borrowing constraint or a large punishment parameter.

As written, this paper can be read as moving us toward a complete understanding of one of the most important mechanisms. But, we also encourage a second reading. Many papers focus on symmetric equilibria, ignoring the possibility of asymmetric equilibria. And until recently, even fewer focused on the stability of equilibria.\footnote{Often it is the case, that when asymmetric equilibria arise, they become stable and the symmetric equilibrium becomes unstable (Brock and Durlauf 2001). Here, we find that the interior asymmetric equilibria are always unstable while the symmetric equilibrium remains stable independent of the existence of the other equilibria.} In the past decade, economists have placed greater emphasis on learning and the processes by which equilibria are attained. This warrants a revisiting of previous models that prove that an efficient equilibrium exists and even papers like the Groves Ledyard paper which proves that all interior equilibria are efficient. As we have come to learn, the existence of efficient equilibrium and the attainment of those equilibria are quite distinct. As is the case for this mechanism, there may also exist boundary equilibria, which in this case are inefficient. These boundary equilibria may matter if they are stable and if the dynamic generated by standard learning rules generate large basins of attraction (Page and Tassier 2003).

The remainder of this paper is organized in five parts. In the first part, we describe the Groves Ledyard mechanism and solve for the symmetric and BST equilibria in closed form. In the second part, we solve for the boundary equilibria. In part three we show that the Groves Ledyard equilibrium, the BST equilibrium, and our boundary equilibria comprise all of the equilibria for this mechanism. In the fourth section, we prove the instability of the BST equilibria and the stability of the other equilibria. And, in section five, we offer some conclusions.
2 The Groves Ledyard Mechanism

We consider a restriction of the Groves-Ledyard Mechanism to an economy with a single public good and a single private good. We index the agents by \( i \in \{1, \ldots, N\} \). Each agent has an initial wealth \( w_i \) and consumes a private good \( x_i \) and a public good \( y \). Utility is a function of both goods \([U_i(x_i(m), y(m))])\). Agent \( i \) pays a percentage of the public good \( \alpha_i \). In the Groves-Ledyard Mechanism, each agent announces a message, \( m_i \), which is \( 1/N \)th the amount of the public good that the agent would like to have produced. Agents pay a tax \( C_i(m) \) which is a function of the vector of messages. An agent’s budget constraint requires that \( x_i(m) = w_i - C_i(m) \). An equilibrium satisfies

\[
U_i(x_i(m), y(m)) \geq U(x(\hat{m}_i, m_{-i}), y(\hat{m}_i, m_{-i})) \quad \forall \hat{m}_i
\]

The Groves-Ledyard mechanism sets

\[
y(m) = \sum_{i=1}^{N} m_i
\]

Hereafter to simplify notation, we suppress \( y \) and use \( m \) to denote the sum of the \( m_i \)’s.

\[
C_i(m) = \alpha_i m + \frac{\gamma}{2} \left[ \frac{N-1}{N} (m_i - \bar{m})^2 - \sigma_i \right]
\]

The punishment parameter, \( \gamma \), plays an important role in the analysis. Increasing \( \gamma \) creates an incentive for conformity. For large enough \( \gamma \), all agents send the same message in equilibrium. This destroys the BST equilibria.

To minimize notation, we restrict our attention to the case of agents with
identical wealth levels and preferences and assume that $\alpha_i = 1/N$ throughout. Even with identical agents the mechanism generates multiple equilibria. So we see no reason to make the analysis more complicated than necessary. We will take care to mention when this restriction matters substantively and when it just reduces notation. Our results do not appear to depend on the assumption of identical preferences and in some cases we show how the proofs can be extended.

Throughout the paper, we rely on a decomposition of $C_i(m)$ into two parts: the contribution $m\alpha_i = m/N$ and the punishment $T_i(\bar{m}) = \frac{\gamma}{2} \left[ \frac{N-1}{N} (m_i - \bar{m})^2 - \sigma_i \right]$ where $\bar{m}^i$ equals the average message sent by the $N - 1$ agents other than $i$ and $\sigma_i$ equals $\frac{1}{N-2}$ times the sum of the $(N - 1)$ terms of the form $(m_j - \bar{m}^i)^2$. Notice that $\sigma_i$ does not depend on $m_i$, so it is out of the agent’s control. Thus for incentive purposes all that matter are the tax share $m/N$ and the difference from the average.

### 2.1 Quasi-Linear Preferences

We begin with the case of quasi linear preferences (Chen and Tang 1998). Here, the Groves Ledyard Mechanism has a unique equilibrium. We can write the utility function of the $i$th agent as follows:

$$U_i(x_i, m) = x_i + B(m)$$

where $B(m)$ is a concave increasing function of $m$, the sum of the contributions to the public good. If we add the Groves-Ledyard tax into the budget constraint, we obtain the following equality.

$$x_i = w_i - \frac{m}{N} - \frac{\gamma}{2} \left[ \frac{N-1}{N} (m_i - \bar{m}^i)^2 - \sum_{j \neq i} \frac{1}{N-2} (m_j - \bar{m}^i)^2 \right]$$

Substituting this back into the utility function gives:
\[ U_i(x_i, m) = w_i - \frac{m}{N} - \frac{\gamma}{2} \left[ \frac{N-1}{N} (m_i - \bar{m})^2 - \sum_{j \neq i} \frac{1}{N-2} (m_j - \bar{m})^2 \right] + B(m) \]

The first order condition on \( m_i \) is as follows:

\[ -\frac{1}{N} - \gamma \frac{N-1}{N} (m_i - \bar{m}) + B'(m) = 0 \]

This system of \( N \) equations has a unique equilibrium. Suppose that the agents choose different \( m_i \)'s. Some of the messages will be above the average message of the others and some will be below the average of the other messages. Thus, \((m_i - \bar{m})\) will be positive for some \( i \) and negative for other \( i \). But this cannot hold in equilibrium since \((m_i - \bar{m}) = B'(m) - \frac{1}{N}\) for all \( i \). Since the \( m_i \)'s must all be the same it follows that that \( m_i = \frac{m}{N} \) for all \( i \), then each equation reduces to \( \frac{1}{N} = B'(m) \) which has a unique solution at \( m^* \) because of the concavity of \( B \). Thus, the equilibrium is unique.

### 2.2 More General Preferences

We now consider the more general preferences considered by BST. They show that the Groves Ledyard (hereafter GL) mechanism has many equilibria. BST rely on a change of basis to prove their result. Though this simplifies the math, it makes it much harder to maintain an intuitive feel for what is happening. We want to give some intuition for these equilibria and to test whether or not these equilibria are stable. Therefore, we will stick to the original notation but largely restrict ourselves to the case of identical agents.

BST consider a general class of utility functions. These utility functions allow for the amount of the public good to influence the value of the private good. As
an example, the value of beach front property (a private good) depends upon the cleanliness of the ocean and the air quality (public goods). This interaction is captured by the function $A(m)$, which is assumed to be concave, continuously differentiable, strictly positive, and strictly increasing. We can thus write the utility of the $i$th agent as

$$U_i(x_i, \bar{m}) = A(m)x_i + B_i(m)$$

The budget constraint requires that

$$x_i = w_i - m/N - T_i(\bar{m})$$

If we substitute the first of these two equations into the second we obtain

$$U_i(x_i, y) = A(m)w_i - A(m)m/N - A(m)T_i(\bar{m}) + B_i(m)$$

The first order condition with respect to $m_i$ can be written as follows:

$$A'(m)w_i - A'(m)m/N - A'(m)T_i(\bar{m}) - A(m)/N - A(m)\frac{N-1}{N}(m^*_i - \bar{m}^-) + B'_i(m) = 0$$

Assuming strictly concave preferences (this puts restrictions on $A(m)$), then there is a unique allocation which maximizes the sum of the agents’ utility functions. Call this $(x^*, y^*)$ and let $A^* = A(m^*)$ and $B^* = B(m^*)$.

### 2.3 Solving for the BST Equilibria

In the BST equilibria, the total amount of the public good provided is efficient. A minority of the agents sends the same low message and a majority sends the same high message. Let’s consider the example where we have $k$ people deviating below by
\( \epsilon \) and \((N - k)\) deviating above by \( \delta \).

Recall the first order condition for an equilibrium:

\[
A'(m)w_i - A'(m)m/N - A'(m)T_i(\bar{m}) - A(m)/N - \frac{N-1}{N}(m_i^* - m_i) + B_i'(m) = 0
\]

Since the total contribution to the public good is the same \( A(m) \) does not change. The agents who belong to the minority subset of size \( k \) who send the low message have a larger \( T_i \) because they differ from the mean by more than the other agents. We denote the punishment paid by the low message agents with \( T_l \) and the punishment paid by the high message agents with \( T_u \). Assuming all agents have the same wealth, we can write the first order conditions for the two types of agents as

\[
A'(m)w - A'(m)m/N - A'(m)T_l(m) - \frac{A(m)}{N} - A(m)\gamma \frac{N-1}{N}(m_l - \bar{m}_l) + B'(m) = 0
\]

\[
A'(m)w - A'(m)m/N - A'(m)T_u(m) - \frac{A(m)}{N} - A(m)\gamma \frac{N-1}{N}(m_u - \bar{m}_u) + B'(m) = 0
\]

Given that the sum of the deviations from the symmetric equilibrium add to zero, \( \delta = \frac{ke}{(N-k)} \). We first calculate the difference between the low message \( m_l \) and the average of the others, \( \bar{m}^*_l \). In the special case where \( \delta = \frac{ke}{(N-k)} \),

\[
(m_l - \bar{m}^*_l) = -\frac{\epsilon N}{(N-1)}
\]

For the agents who send higher than average messages the corresponding values are

\[
(m_u - \bar{m}^*_u) = \frac{\epsilon Nk}{(N-1)(N-k)}
\]
We next calculate the $\sigma$ term within $T_l$. For the agents with low messages

$$\sigma^l = \frac{1}{N-2} \left[ \frac{(k-1)(\epsilon + \delta)^2(N-k)^2}{(N-1)^2} + \frac{(N-k)(k-1)^2(\epsilon + \delta)^2}{(N-1)^2} \right]$$

which can be expanded as

$$\sigma^l = \frac{1}{N-2} \left[ (k-1) + (N-k) \right] \frac{(\epsilon + \delta)^2(N-j)(k-1)}{(N-1)^2}$$

which reduces to

$$\sigma^l = (\delta + \epsilon)^2 \left[ \frac{(k-1)(N-k)}{(N-1)(N-2)} \right]$$

For the agents with high messages a similar calculation gives

$$\sigma^u = (\delta + \epsilon)^2 \left[ \frac{(k)(N-k-1)}{(N-1)(N-2)} \right]$$

We can then calculate the punishment for both types of agents. For the agents who send the low message the punishment is:

$$\frac{\gamma}{2} (\delta + \epsilon)^2 \left[ \frac{(N-k)^2}{N(N-1)} - \frac{(k-1)(N-k)}{(N-1)(N-2)} \right]$$

Expanding we get

$$\frac{\gamma}{2} (\delta + \epsilon)^2 \left[ \frac{(N-2)(N-k)^2 - (k-1)(N-k)N}{N(N-1)(N-2)} \right]$$

which reduces to

$$\frac{\gamma}{2} (\delta + \epsilon)^2 (N-k) \left[ \frac{(N-1)(N-2k)}{N(N-1)(N-2)} \right]$$
which further reduces to

$$\frac{\gamma}{2}(\delta + \epsilon)^2 \left[ \frac{(N - k)(N - 2k)}{N(N - 2)} \right]$$

which is positive since $(N - 2k) > 0$. In the efficient case, this amount equals

$$\frac{\gamma}{2} \epsilon^2 \frac{N(N - 2k)}{(N - 2)(N - k)}$$

A similar calculation for the taxes paid by those who send the higher messages gives

$$\frac{\gamma}{2}(\delta + \epsilon)^2 \left[ \frac{k^2}{N(N - 1)} - \frac{k(N - k - 1)}{(N - 1)(N - 2)} \right]$$

Expanding we get

$$\frac{\gamma}{2}(\delta + \epsilon)^2 \left[ \frac{(N - 2)k^2 - Nk(N - k - 1)}{N(N - 1)(N - 2)} \right]$$

which reduces to

$$\frac{\gamma}{2}(\delta + \epsilon)^2 k \left[ \frac{(N - 1)(2k - N)}{N(N - 1)(N - 2)} \right]$$

which reduces to

$$\frac{\gamma}{2}(\delta + \epsilon)^2 \left[ \frac{k(2k - N)}{N(N - 2)} \right]$$

In the efficient case, this amount equals

$$-\frac{\gamma}{2} \epsilon^2 \frac{N(N - 2k)k}{(N - 2)(N - k)^2}$$
Notice that the sum of the taxes paid by the \( k \) agents sending the lower messages and the \((N - k)\) agents sending the higher messages equals zero, which is an artifact of the mechanism being balanced.

### 2.4 Characterization of Efficient Asymmetric Equilibria

We now restrict attention to messages that result in efficient levels of the public good. If the first order conditions hold then \( m^l, m^u, \) and \( m^* \) must satisfy the following two equations

\[
A'(m^*)w - A'(m^*)/N - A'(m^*)T^l - \frac{A(m^*)}{N} - A(m^*)\gamma \frac{N - 1}{N} (m^l - \bar{m}^l) + B'(m^*) = 0
\]

\[
A'(m^*)w - A'(m^*)/N - A'(m^*)T^u - \frac{A(m^*)}{N} - A(m^*)\gamma \frac{N - 1}{N} (m^u - \bar{m}^u) + B'(m^*) = 0
\]

Recall that in an efficient equilibrium, we have that

\[
A'(m^*)w - A'(m^*)/N - \frac{A(m^*)}{N} + B'(m^*) = 0
\]

Therefore, at any asymmetric efficient equilibrium it must be that

\[-A'(m^*)T^l - A(m^*)\gamma \frac{N - 1}{N} (m^l - \bar{m}^l) = 0\]

For the agents who send the low messages this reduces to

\[-A'(m)\gamma^2 \frac{N(N - 2k)}{(N - 2)(N - k)} = A(m)\gamma(N - 1)\frac{N}{N(N - 1)}\]

which can be rewritten
\[ \frac{N(N - 2k)}{2(N - 2)(N - k)} = \frac{A(m)}{A'(m)} \]

which can be simplified as

\[ \epsilon^{BST}(k) = \frac{2(N - 2)(N - k)A(m)}{A'(m)N(N - 2k)} \]

And for the agents who send the high message it reduces to

\[ A'(m)\frac{\gamma \epsilon^2}{2} \frac{N(2k - N)k}{(N - 2)(N - k)^2} = \frac{A(m)\gamma \epsilon kN(N - 1)}{N(N - 1)(N - k)} \]

Which can be reduced to

\[ \epsilon^{BST}(k) = \frac{2(N - 2)(N - k)A(m)}{A'(m)N(N - 2k)} \]

This characterizes the BST asymmetric equilibria. These are in agreement with the equilibria found in BST but they do not rely on their complicated change of basis.

2.5 Structure of the BST Equilibria

We now take a closer look at some characteristics of the BST equilibria. Note that the punishment parameter \( \gamma \) does not determine \( \epsilon \), ie, the location of these equilibria. \( \epsilon \) is only a function of \( N, k, \) and \( A(m) \). Yet, \( \gamma \) does play an important role for the BST equilibria.

To see the effect of \( \gamma \) most clearly let us decompose the utility function into two parts: the public good portion of portion of utility (PGU) which does not depend

\(^{2}\)Notice though that for \( k \) larger than \( \frac{N}{2} \) we getting the same equilibria that we obtain for \( k \) less than \( \frac{N}{2} \).
on $\gamma$ and the punishment portion of utility (PUNU) which does depend on $\gamma$.

These can be written as:

$$PGU = A(m^*)(w_i - \frac{m^*}{N}) + B(m^*)$$

and

$$PUNU = A(m^*)\left(\frac{\gamma}{2}\frac{N - 1}{N}(m_i - \bar{m}_i)^2 - \sum_{j \neq i} \frac{1}{N - 2}(m_j - \bar{m}_i)^2\right)$$

Figures 1 and 2 show the this decomposition of utility for the high and low agents for an example utility function with 3 agents at a BST equilibrium (2 agents are playing the high message and one is playing the low message.) Note first that both the public good portion of utility and the punishment portion of utility are at a maximum for the high agents. Thus total utility is also at a maximum. The decomposition of utility for the low agent is much different. For her the public good portion is also at a maximum but the punishment portion is at a minimum.

We can see this formally through the derivatives of the punishment portion of utility for the low agent. The first derivative is:

$$-A'(m^*)T^l - A(m^*)\gamma \frac{N - 1}{N}(m^l - \bar{m}^l)$$

And the second derivative is:

$$-A''(m^*)T^l - 2A'(m^*)\gamma \frac{N - 1}{N}(m^l - \bar{m}^l) - A(m^*)\gamma \frac{N - 1}{N}$$

Recall that $m^l - \bar{m}^l = -\frac{\epsilon N}{N - 1}$ and that $\epsilon = \frac{2(N - 2)(N - k)A(m)}{N(N - 2k)a'(m)}$. Substituting these into the 2nd derivative and rearranging yields:
Figure 1: Utility of the high message agents when $\gamma$ is low.

Figure 2: Utility of the low message agent when $\gamma$ is low.
Figure 3: Utility of the low message agent when $\gamma$ is high.

\[-A''(m^*)T^4 + \frac{\gamma A(m^*)}{N} \left[ \frac{4(N - 2)(N - k)}{(N - 2k)} - (N - 1) \right] \]

Note that the public good portion of utility does not depend on $\gamma$. Thus as $\gamma$ increases the PGU remains the same. But, for the low agent, the punishment portion of utility is decreasing linearly in $\gamma$. Thus as $\gamma$ increases the total utility at the BST equilibria is decreasing for the low agents in the punishment portion of utility. Thus for $\gamma$ sufficiently large, utility is minimized at the BST equilibria. In other words, the BST equilibria no longer exist for high $\gamma$. But this agreed with our initial intuition. If $\gamma$ is too high, the pressure to conform is sufficiently great to wipe out the asymmetric equilibria. This logic can be seen graphically. Using the same utility function as above we show the utility decomposition for the low agent with a high level of $\gamma$ in Figure 3.
2.6 Characterization of Interior Equilibria

In the next three claims we state that any interior equilibrium must be efficient and must have all agents sending one of either two messages, a high message or a low message. These claims imply that the symmetric equilibrium found by Groves and Ledyard and the asymmetric equilibria found by BST comprise all of the interior equilibria of the mechanism.

To prove these and later claims, we decompose the derivative of the utility function into two parts: the public good part and the punishment part. The public good part of marginal utility (PGMU) equals

\[ PGMU(\vec{m}) = A'(m)(w - \frac{m}{N}) - \frac{A(m)}{N} + B'(m) \]

The punishment part of marginal utility (PUNMU) equals

\[ PUNMU(\vec{m}) = -A'(m)T_i(\vec{m}) - A(m)\gamma \frac{N-1}{N}(m_i - \bar{m}) \]

We first state a claim that says that any interior equilibrium of the Groves Ledyard Mechanism is efficient, even when agents are not identical. Here we allow the wealth levels to change as well as preferences for the public good which we denote by \(B_i\).

**Claim 1** All interior equilibria of the Groves Ledyard mechanism provide for an efficient level of the public good.

pf. The amount of the public good is efficient if the sum of the PGMU’s for all of the agents equals zero. At an interior equilibrium the sum of each agent’s PGMU and PUNMU is zero; therefore, it suffices to show that the sum of the PUNMU’s equals zero at any equilibrium. We can write the sum of the PUNMU’s as follows

17
\[
\sum_{i=1}^{N} \left( -A'(m)T_i(\bar{m}) - \gamma \frac{N-1}{N} (m_i - \bar{m}^i) \right)
\]

The sum of the \((m_i - \bar{m}^i)\) terms is trivially zero. Each \(m_i\) is added once and each is subtracted \((N - 1)\) as part of an average of \((N - 1)\) messages. It therefore, suffices to show that the sum of the \(T_i(\bar{m})\)'s equal zero. Groves and Ledyard show this in their original paper. It is a condition of the mechanism being balanced.

The next claim states that with identical agents all agents sending the higher message send the same high message.

**Claim 2** Let \((m_1, m_2, \ldots, m_N)\) be an interior equilibrium of the Groves Ledyard Mechanism with identical agents. If \(m_i \geq \bar{m}^i\) and \(m_j \geq \bar{m}^j\), then \(m_i = m_j\).

pf. Set the public good part of marginal utility (PGMU) equal to \(X\), which may be positive, negative, or zero.

\[A'(m)\left( w - \frac{m}{N} \right) - \frac{A(m)}{N} + B'(m) = X\]

It follows that the punishment part of marginal utility (PUNMU) equals \(-X\).

\[-A'(m)T_i(\bar{m}) - A(m)\gamma \frac{N-1}{N} (m_i - \bar{m}^i) = -X\]

The PGMU’s are identical given that the agents have identical wealth and preferences. The PUNMU’s differ, we can write them as

\[-A'(m)T_i(\bar{m}) = -X + A(m)\gamma \frac{N-1}{N} (m_i - \bar{m}^i)\]
\[-A'(m)T_j(\bar{m}) = -X + A(m)\gamma \frac{N - 1}{N} (m_j - \bar{m}^j)\]

Suppose that \(m_i > m_j\). It follows then that \(\bar{m}^j > \bar{m}^i\) and therefore that \((m_i - \bar{m}^i) > (m_j - \bar{m}^j)\). By the equations above it follows that \(T_i(\bar{m}) < T_j(\bar{m})\). However, if we solve for \(T_i(\bar{m})\) and \(T_j(\bar{m})\) directly we see that \(T_i(\bar{m}) > T_j(\bar{m})\), a contradiction.

\[T_i = \frac{\gamma}{2} \left[ \frac{N - 1}{N} (m_i - \bar{m}^i)^2 - \frac{1}{N - 2} \sum_{\ell \neq i} (m_\ell - \bar{m}^i)^2 \right]\]

\[T_j = \frac{\gamma}{2} \left[ \frac{N - 1}{N} (m_j - \bar{m}^j)^2 - \frac{1}{N - 2} \sum_{\ell \neq j} (m_\ell - \bar{m}^j)^2 \right]\]

Recall that \((m_i - \bar{m}^i) > (m_j - \bar{m}^j)\). It therefore suffices to show that

\[\sum_{\ell \neq i} (m_\ell - \bar{m}^i)^2 < \sum_{\ell \neq j} (m_\ell - \bar{m}^j)^2\]

We will rewrite the left hand side of the inequality in order to show it is less than the right.

\[\sum_{l \neq i} (m_l - \bar{m}^i)^2 = \sum_{l \neq i} m_l^2 - 2 \sum_{l \neq i} m_l\bar{m}^i + (N - 1)\bar{m}^i^2\]

\[= \sum_{l \neq j} m_{l}^2 - 2 \sum_{l \neq j} m_l\bar{m}^i + (N - 1)\bar{m}^i^2 + m_j^2 - 2m_j\bar{m}^i - m_i^2 + 2m_i\bar{m}^i\]

Let \(\Delta = \frac{m_i - m_j}{N - 1}\). The previous equation reduces to:
\[\sum_{l\neq j} m_l^2 - 2 \sum_{l\neq j} m_l \bar{m}^i + 2 \sum_{l\neq j} m_l \Delta - \sum_{l\neq j} m_l \Delta + (N-1) \bar{m}^i + m_j^2 - 2 m_j \bar{m}^i - m_i^2 + 2 m_i \bar{m}^i\]

\[
= \sum_{l\neq j} m_l^2 - 2 \sum_{l\neq j} m_l (\bar{m}^i - \Delta) - 2 \sum_{l\neq j} m_l \Delta + (N - 1) \bar{m}^i + m_j^2 - 2 m_j \bar{m}^i - m_i^2 + 2 m_i \bar{m}^i
\]

\[
= \sum_{l\neq j} (m_l - \bar{m}^i)^2 - 2 \Delta \sum_{l\neq j} m_l + (N - 1)(\bar{m}^i - \bar{m}^j)^2 + (m_j^2 - m_i^2) + 2 \bar{m}^i (m_i - m_j)
\]

Given that \(\sum_{l\neq j} (m_l - \bar{m}^j)^2\) appears in the expression, it suffices to show that the remaining terms are negative. We can rewrite these terms as:

\[-2 \Delta (N - 1)(\bar{m}^j - \bar{m}^i) - (N - 1)(\bar{m}^i - \bar{m}^j)^2 - (m_i^2 - m_j^2)\]

The first of these three terms is clearly negative. We can rewrite the last two terms as:

\[-(N - 1)(\frac{(m - m_j)^2}{(N - 1)^2} - \frac{(m - m_i)^2}{(N - 1)^2}) - (m_i^2 - m_j^2)\]

\[-2m \Delta - (m_j^2 + m_i^2)(1 - \frac{1}{N - 1})\]

This completes the proof.

This second claim appears to depend heavily on the identical agent assumption, but in fact this is not true. Without identical agents it would not be the case
all of the agents sending the higher messages choose the exact same messages. However, it would be the case that any asymmetric equilibrium messages higher than the GL equilibrium messages would generate the same PUNMU as the messages under the Groves-Ledyard equilibrium since the amount of the public good provided would be the same in each case. Therefore, all agents sending messages above their GL equilibrium message would be obtaining the same marginal punishment for doing so.

**Claim 3** Let \((m_1, m_2, \ldots, m_N)\) be an interior equilibrium of the Groves Ledyard Mechanism. If \(m_i \leq \bar{m}^1\) and \(m_j \leq \bar{m}^j\), then \(m_i = m_j\).

pf This proof follows directly from the previous claim.

Using these insights it is possible to construct a simpler direct proof for the efficiency of all interior equilibria in the case of identical agents.

**Corollary 1** All interior equilibria of the Groves Ledyard mechanism provide for an efficient level of the public good when agents are identical.

pf. Suppose that we have an equilibrium where the amount of the public good is not equal to \(m^*\). It follows that the marginal utility of the public good part of the first order condition cannot equal zero, \(PGMU(m^*) \neq 0\). From the previous claims we know that for any interior equilibrium, there must be a set of agents sending a low message and a set agents sending a high message. The punishment portion of the first order condition for each set of agents plus the public good portion must equal zero. From our characterization of the asymmetric equilibria we know that for the agents who send the low messages the punishment portion of the first order condition is given by

\[
-A'(m) \frac{\gamma e^2}{2} \frac{N(N-2k)}{(N-2)(N-k)} - \frac{A(m)\gamma e(N-1)N}{N(N-1)}
\]
and for the agents who send the high message the punishment portion of the first order condition is given by

\[ A'(m)\gamma^2 \frac{N(2k-N)k}{(N-2)(N-k)^2} + \frac{A(m)\gamma k N(N-1)}{N(N-1)(N-k)} \]

These two values have opposite sign. The first is less than 0 since \( k < N/2 \) and \( A'(m) \) is positive. For the second value the left hand term is negative since \( 2k - N < 0 \) but the right hand term is positive. Thus we need to show that:

\[ \frac{A(m)\gamma k N(N-1)}{N(N-1)(N-k)} > A'(m)\gamma^2 \frac{N(2k-N)k}{(N-2)(N-k)^2} \]

Recall that \( \epsilon = \frac{2(N-2)(N-k)A(m)}{A'(m)N(N-2k)} \). If we cancel terms and substitute for \( \epsilon \) we find this equation reduces to: \( \frac{(2k-N)}{(N-2k)} < 1 \) which is true since \( k < N/2 \). Since they must be equal, this can only occur if they equal zero, which implies that \( PGMU(m^*) = 0 \). This completes the proof.

Taken together these three claims imply that the only interior equilibria of the Groves Ledyard Mechanism are the symmetric equilibrium and the BST equilibria. Further all of these equilibria are efficient. We next consider the possibility of there being other equilibria if we bound the space of messages.

### 2.7 Boundary Equilibria

Finally we consider the case where messages are constrained from below. This creates the possibility of additional boundary equilibria. Specifically we restrict messages to the interval \([-D, \infty)\]. By bounding the messages from below we prevent infinite negative amounts of the public good. We will show that for small \( \gamma \) such equilibria always exist provided \( D \) is large enough. Therefore, in the absence of a lower bound,
this implies that there would be equilibria at negative infinity.

With this construction it is possible that agents sending higher messages may still want to send infinite positive messages. Here we no longer assume that the amount of the public good provided is efficient. As before we assume that \(k\) agents, where \(k < \frac{N}{2}\) send the message \(-D\), and the other \(N-k\) agents send messages \(m^+\) that are positive. Let \(m\) denote the amount of the public good provided so that \(-kD + (N-k)m^+ = m\) or \(m^+ = \frac{m+kD}{(N-k)}\). As above, we first calculate the difference between the low message \(-D\) and the average of the others, \(\bar{m}^D\). A straightforward calculations gives that \(\bar{m}^D = \frac{m+kD}{(N-1)}\). It follows that

\[
(-D - \bar{m}^D) = -\frac{m + ND}{(N-1)}
\]

For the agents who send higher than average messages the corresponding values are

\[
\bar{m}^+ = \frac{(N-k-1)m-kD}{(N-k)(N-1)}
\]

\[
(m^+ - \bar{m}^+) = \frac{km+NkD}{(N-k)(N-1)}
\]

We next calculate the \(\sigma\) term within \(C\). For the agents at the lower boundary

\[
\sigma^D = \frac{1}{N-2} \left[ (k-1) \frac{(m+ND)^2}{(N-1)^2} + (N-k) \frac{(ky+NkD)^2}{(N-k)^2(N-1)^2} \right]
\]

Which reduces to

\[
\sigma^D = \frac{1}{N-2} \frac{(m+ND)^2(N(k-1)+k)}{(N-1)^2(N-k)}
\]
We next calculate the $\sigma$ term within $T^D$. For the agents choosing the high message

$$\sigma^+ = \frac{1}{N-2} \left[ k \frac{(m + ND)^2}{(N-1)^2} + (N - k - 1) \frac{(ky + NkD)^2}{(N-k)^2(N-1)^2} \right]$$

Which reduces to

$$\sigma^+ = \frac{1}{N-2} \frac{(m + ND)^2 k (N(N - k) - k)}{(N-1)^2(N-k)^2}$$

We can then calculate the punishment for both types of agents. For the agents who send the message at the lower boundary it equals

$$\gamma \frac{(m + ND)^2[(N-1)^2(N-2k) + (2k+1)N]}{2 N(N-2)(N-k)(N-1)^2}$$

Notice that this is a positive number. And for the agents who send the high messages, we obtain

$$\gamma \frac{(m + ND)^2[-N^3 + 2k(N^2 - N + 1)]}{2 N(N-2)(N-k)^2(N-1)^2}$$

Notice that this amount is negative. Therefore, the punishment serves as a subsidy from those sending the low message to those sending the high message. Given that the total amount of the public good is not necessarily efficient, we have that

$$A'(m)(w - m/N) - \frac{A(m)}{N} + B'(m) = X$$

Therefore, the first order conditions for the agents who send the message at the lower boundary is
\[-\frac{A'(m)}{N} - A'(y)T^D(\bar{m}) - A(m)\gamma \frac{N-1}{N}(m^D - \bar{m}^D) + X = 0\]

This can be written as

\[-\frac{A'(m)}{N} - A'(y)T^D(\bar{m}) + A(m)\gamma \frac{m + ND}{N} + X = 0\]

and for the agents who send the high message we have

\[-\frac{A'(m)}{N} - A'(m)T^+(\bar{m}) - A(m)\gamma \frac{N-1}{N}(m^+ - \bar{m}^+) = 0\]

Which equals

\[-\frac{A'(m)}{N} - A'(m)T^+(\bar{m}) - A(m)\gamma \frac{km + NkD}{N(N - k)} = 0\]

In order for a boundary equilibria to exist three conditions must hold: First, MU for the low message agents must be negative (they want to decrease their message more but the boundary prevents them from doing so.). Second, for low message agents the payoff at the boundary must be higher than at any interior point. And third, the high message agents must not want to send infinite messages, (the message of the high agents must stop at some value.) The second and third conditions are trivially satisfied since for low \(\gamma\) the utility function is concave and marginal utility is decreasing in \(m\).

From above the first condition requires

\[-\frac{A'(m)}{N} - A'(y)T^D(\bar{m}) + A(m)\gamma \frac{m + ND}{N} < 0\]
For sufficiently small $\gamma$ the second and third terms are smaller in absolute terms than the first which is negative. Thus for small $\gamma$ the low agents may have negative marginal utility at the boundary. Combining this with the analysis of the BST equilibria above, we see that for sufficiently large $\gamma$ only the symmetric equilibrium exists. Thus one can always increase the punishment parameter of the mechanism to enforce an equal sharing of the cost of public good provision.

3 Stability of the Equilibria

In models of human actors, stability is perhaps best defined relative to classes of learning rules that impose dynamics on the system (Van Huyck, Cook, Battalio 1994). That said, linear stability analysis remains a powerful tool. Given the best response functions, linear stability analysis implicitly assumes that if an agent’s marginal utility is positive in $m_i$ at a point, then the agent increases its message and if the marginal utility is decreasing that the agent will decrease its message and that the amount of that increase is proportional to the marginal utility.

$$\dot{m}_i = \frac{\partial U_i}{\partial m_i}$$

In our analysis we again find it useful to decompose the utility into the public good component and the punishment component. This time, we focus on their contributions to marginal utility.

$$PGMU(\vec{m}) = A'(m)(w - \frac{1}{N}) - \frac{A(m)}{N} + B'(m)$$
\[ PUNMU(\bar{m}) = -A'(m)T_i(\bar{m}) - A(m)\gamma \frac{N - 1}{N} (m_i - \bar{m}') \]

We can then write the equation of motion as the sum of these two terms.

\[ \dot{m}_i = PGMU_i(\bar{m}) + PUNMU(\bar{m}) \]

As we shall show, peculiarities of the Groves-Ledyard mechanism make this decomposition extremely useful in understanding the dynamics of the system. We first show that symmetric equilibrium is stable and then show that the BST equilibria are unstable given these equations of motion.

### 3.1 Stability of the Symmetric Equilibrium

The decomposition of the equation of motion into a public good portion and a punishment portion highlights something remarkable about the Groves-Ledyard Mechanism at the symmetric equilibrium. The punishment contribution to utility creates an unstable dynamic as does the public good contribution to utility. But together, they create a stable dynamic. The dynamics created by the public good portion of the payoffs allow for drift in messages so long as the amount of the public good provided remains efficient. The punishment portion of the payoffs creates an incentive for the agents to send the same message, whether or not it is efficient. Therefore, the punishment portion allows drifts from efficiency. The force toward efficiency created by the public good part overpowers the drift away allowed by the punishment part. And the force toward symmetry by the punishment part overpowers the drift away from symmetry allowed by the public good part. Thus, at the symmetric equilibrium,
unstable incentives plus unstable incentives yield stable incentives.

We show matrices as though there are only three players but using notation from the $N$ player case, so the proof is general. If we compute the Jacobian for the public good contribution to utility we get the following form

\[
\begin{bmatrix}
-\theta & -\theta & -\theta \\
-\theta & -\theta & -\theta \\
-\theta & -\theta & -\theta \\
\end{bmatrix}
\]

where $-\theta$ equals $A''(m)(w - \frac{1}{N}) - \frac{A'(m)}{N} + B''(m)$. The reason that every entry in the Jacobian has the same value is that the marginal effect of an increase in any agent’s message is the same for all players since the costs are split evenly. This value is negative at an efficient equilibrium. As we mentioned, considered in isolation, this is not a stable system. In the case of $N$ agents, $N - 1$ of the eigenvalues are 0 and the other has value $-N\theta$. The fact that the non negative eigenvalues are not strictly positive means that the system has other equilibria in the neighborhood of this point. In fact, any set of messages that sum to the efficient amount of the public good is stable with respect to the public good part. If one agent increases its message by $\epsilon$ and another decreases its message by $\epsilon$ then this new set of messages is stable for this portion of the equations of motion.

Next we consider the Jacobian associated with the punishment portion of payoffs. Recall that the punishment portion of marginal utility at the symmetric equilibrium are zero. It can be shown that the Jacobian is given by
This matrix has the form
\[-A(m)\gamma \frac{N-1}{N} \quad A(m)\gamma \frac{1}{N} \quad A(m)\gamma \frac{1}{N}\]

\[A(m)\gamma \frac{1}{N} \quad -A(m)\gamma \frac{N-1}{N} \quad A(m)\gamma \frac{1}{N}\]

\[A(m)\gamma \frac{1}{N} \quad A(m)\gamma \frac{1}{N} \quad -A(m)\gamma \frac{N-1}{N}\]

This matrix has the form
\[-(N-1)\omega \quad \omega \quad \omega \]

\[\omega \quad -(N-1)\omega \quad \omega \]

\[\omega \quad \omega \quad -(N-1)\omega \]

This matrix has \((N - 1)\) eigenvalues equal to \(-N\omega\) and one eigenvalue equal to 0. Therefore, this dynamic like the previous dynamic is also not stable, but neither does it explode. If all of the agents increase or decrease their messages by a common amount, the new messages are stable. Because this deviation has to be coordinated among all \(N\) agents, this dynamic has only one eigenvalue equal to zero. In the previous dynamic, almost any deviation will lead to a new equilibrium, that is why almost all of the eigenvalues are zero.

These two unstable dynamics combine to create stability. The public good portion keeps the agents at an efficient level of the public good but lets them vary in who contributes. The punishment portion keeps them at the identical contribution levels but the combined level can stray from equilibrium. When we add the two
together, we get stability. The first dynamic forces efficiency. The second forces symmetry. We can see this mathematically by adding the two Jacobians to get the combined Jacobian at the symmetric equilibrium.

\[
\begin{pmatrix}
-(N-1)\omega - \theta & \omega - \theta & \omega - \theta \\
\omega - \theta & -(N-1)\omega - \theta & \omega - \theta \\
\omega - \theta & \omega - \theta & -(N-1)\omega - \theta
\end{pmatrix}
\]

This matrix has \((N-1)\) eigenvalues of \(-N\theta\) and one eigenvalue of \(-N\omega\). Therefore, it is stable. Note that the assumption of identical agents may appear to be playing a huge role here but it is not. If we give each agent a unique \(B_i(m)\), then each row of the public good matrix gets multiplied by a unique constant, \(\alpha_i > 0\). \(N-1\) of the eigenvalues are still zero and the other eigenvalue equals \(-N\sum_{i=1}^{N} \alpha_i \theta\). This nonzero eigenvalue replaces \(-N\theta\) as an eigenvalue of the combined Jacobian and the other eigenvalues remain unchanged.

4 The Instability of the BST equilibria

We now perform the same analysis for the BST equilibria. Here, the calculations are more involved because for each subset of size \(k < N/2\) that deviates we get a distinct set of dynamics. Since this creates a potentially infinite set of systems, we consider the case where \(k\) equals one in full and show that it is unstable. We then sketch a proof for why that same logic applies for any \(k\).

As in the symmetric case, we decompose the equations of motion into two parts: the public good part and the punishment part, and then combine them to
form the Jacobian for the system of equations. The Jacobian for the public goods
portion considered alone is the same as for the symmetric equilibrium and takes the
form

\[-\theta \ -\theta \ -\theta \]

\[-\theta \ -\theta \ -\theta \]

\[-\theta \ -\theta \ -\theta \]

This is not a stable system. In the case of \(N\) agents, \(N - 1\) of the eigenvalues
are 0 and the other has value \(-N\theta\). Calculating the Jacobian for the punishment
portion of the dynamics at the asymmetric equilibrium is cumbersome. Recall that
\(T^i(\vec{m})\) is the punishment paid by the agents who send the lower message. In this case,
that is just one agent. The equation of motion for the message sent by that agent is
given in the third row. With some effort it can be shown that the Jacobian takes the
following form:

\[
\begin{array}{ccc}
\alpha - (N - 1)\beta & \alpha + \beta & \alpha - \beta \\
\alpha + \beta & \alpha - (N - 1)\beta & \alpha - \beta \\
-(N - 1)\alpha + \beta & -(N - 1)\alpha + \beta & -(N - 1)\alpha + (N - 1)\beta \\
\end{array}
\]

where \(\alpha = A''(m)T'(\vec{m})/(N - 1)\) and \(\beta = \gamma A(m)\) It can be show that his matrix has
one eigenvalue of 0, one equal to \(\beta\) and \((N - 2)\) equal to \(-N\beta\). So it is unstable.
When we combine the Jacobian for the public goods contribution to the dynamics
and the punishment portion of the dynamics we get a matrix of the following form
\[
\begin{array}{ccc}
\alpha - (N - 1)\beta - \theta & \alpha + \beta - \theta & \alpha - \beta - \theta \\
\alpha + \beta - \theta & \alpha - (N - 1)\beta - \theta & \alpha - \beta - \theta \\
-(N - 1)\alpha + \beta - \theta & -(N - 1)\alpha + \beta - \theta & -(N - 1)\alpha + (N - 1)\beta - \theta \\
\end{array}
\]

This matrix has one eigenvalue equal to \(-N\theta\), \((N - 2)\) eigenvalues equal to \(-N\beta\), and one eigenvalue equal to \(\beta\). Therefore, the system is not stable. As before, symmetry of agents does not play a large a role in the dynamics. If each agent has a unique value for the public good, a unique \(B_i\), then as in the symmetric case, only the \(-N\theta\) eigenvalue is affected and the change is only in magnitude not in sign.

This proves that the BST equilibria in which one agent sends a low message and the rest send high messages are unstable, but it does not prove the general case. Moreover, the calculation does not always provide any intuition behind why a system is stable or unstable. Though, in the symmetric case, we found that by decomposing the dynamical system into two parts, we could uncover the causes of stability.

The logic driving the instability of the BST equilibria relies on the decomposition of marginal utility into a public good contribution (PGMU) and the punishment contribution (PUNMU) to marginal utility. At \(m - \epsilon\), PUNMU equals

\[-A'(m^*)T^l(m) - A(m^*)\gamma\frac{N - 1}{N}(m^l - \bar{m}^l)\]

This can be rewritten as

\[\gamma[\epsilon A(m^*) - \epsilon^2 A'(m^*)\frac{N(N - 2k)}{2(N - 2)(N - k)}\]

32
Similarly, PUNMU at \( m + \frac{k\epsilon}{N-k} \) equals

\[
-A'(m^*)T^u - A(m^*)\gamma \frac{N-1}{N}(m^u - \bar{m}^u)
\]

This can be rewritten as

\[
\gamma[-\epsilon A(m^*) - \frac{k}{(N-k)} + \epsilon^2 A'(m^*)] \frac{N(N-2k)k}{2(N-2)(N-k)^2}
\]

For \( \epsilon = 0 \) PUNMU at \( m - \epsilon \) and PUNMU at \( m + \frac{k\epsilon}{N-k} \) become the same equation which trivially equals zero. By assumption \( A(m^*) > 0 \) and \( A'(m^*) < 0 \); therefore, the punishment portion of the FONC for the agents sending the low message is positive for small values of \( \epsilon \). It eventually hits zero and thereafter is negative. The value for which this expression hits zero is the \( \epsilon \) which gives the BST equilibria. Recall that this is denoted \( \epsilon^{BST} \). Similarly, the punishment portion of the FONC for the agents sending the higher message equals 0 at \( \epsilon = 0 \) and then remains negative until \( \epsilon^{BST} \) and thereafter it is strictly positive.

This observation confirms our intuition for why the symmetric equilibrium is stable. If we perturb the stable equilibrium, then the agents sending the low message have a positive PUNMU, therefore, they announce higher messages and the agents sending the higher message have a negative PUNMU and therefore, send lower message, modulo the effect of the amount of the public good.

At the BST equilibria if the agents are not quite far enough apart (if the low agents are too high and the high agents are too low) then the agents sending the lower message will have a positive PUNMU and will increase their message. The agents sending the high message will have a negative PUNMU and they decrease their message, moving them away from the BST and toward the symmetric equilibrium.
To formally show the local instability of the BST equilibria, we must show that given a BST equilibrium $m^q$ for any $\rho > 0$ the neighborhood of radius $\rho$ around $m^q$ contains a point, $\tilde{m}$ such that beginning from $\tilde{m}$, the dynamical system will not converge to $m^q$. Choose an arbitrary BST equilibrium where $k$ agents send the message $m^* - \epsilon^{BST}(k)$ and $N - k$ agents send the message $m^* + \delta^{BST}(k)$ such that the allocation of the public good is efficient. Call this $m_k^{BST}(k)$. Given a $\rho > 0$ choose $\epsilon < \epsilon^{BST}(k)$ such that if $k$ agents send the message $m^* - \epsilon$ and $N - k$ agents send the message $m^* - \frac{(N-k)\epsilon}{k}$, the messages lie in the neighborhood of radius $\rho$ around $m_k^{BST}(k)$. The agents sending the lower message are sending a message greater than the equilibrium low message and the agents sending the higher message are sending a message lower than the equilibrium high message.

It suffices to show that the agents sending the lower message will increase their message and the agents sending the higher message will decrease their message. This will move the set of messages further from the BST equilibrium. These conditions are sufficient using the following logic: If the messages at each instance of time continue to provide for an efficient amount of the public good then the same argument applies: at each moment in time, the messages will be even further from the BST equilibrium. Alternatively, suppose the dynamics could lead to either over or under provision of the public good at some time $t_1$. Without loss of generality assume over provision of the public good. In this case the public good portion of the FONC will become negative. This will cause the agents sending the higher message to reduce their messages by even more. The agents sending the lower messages will have less incentive to increase their messages and may even have an incentive to decrease their messages if the over provision becomes too severe. Therefore, at some time $t_2 > t_1$ one of two things must happen. Either the low and high messages will converge (in which case the system will converge to the symmetric equilibrium) or the amount of the public good will again be
efficient. If the latter occurs, since the agents sending the higher messages have been decreasing their message, the new messages are further from the BST equilibrium at time $t_2$ then they were at time $t_1$.

Thus, we need only show that the derivative of the utility function for the agents sending the lower message with respect to their message is positive and that the opposite is true of the derivative of the utility function for the agents sending the higher message. Since the amount of the public good is efficient, the value PGMU equals zero. It suffices then to show that PUNMU is positive at $m - \epsilon$ and negative at $m + \frac{ke}{N-k}$. Recall that

$$
\epsilon^{BST}(k) = \frac{2(N - 2)(N - k)A(m*)}{A'(m*)N(N - 2k)}
$$

Consider the following deviation $\epsilon = (1 - \phi)\epsilon^{BST}(k)$, where $\phi > 0$ is arbitrarily small. PUNMU at $m - \epsilon$ equals

$$
\gamma \left[ \epsilon A(m*) - \epsilon^2 A'(m*) - \frac{N(N - 2k)}{2(N - 2)(N - k)} \right]
$$

substituting in the value for $\epsilon$ yields

$$
\gamma \left[ \frac{(1 - \phi)2(N - 2)(N - k)A(m*)^2}{A'(m*)N(N - 2k)} - \frac{2(1 - \phi)^2(N - 2)(N - k)A(m*)}{A'(m*)N(N - 2k)} \right]
$$

which reduces to

$$
\gamma \left[ ((1 - \phi) - (1 - \phi)^2) \frac{2(N - 2)(N - k)A(m*)^2}{A'(m*)N(N - 2k)} \right]
$$

which is strictly positive. Therefore, the agents at $m - \epsilon$ increase their message. To show that the agents at $m + \frac{ke}{N-k}$ decrease their message, we make the following
similar calculation. Recall from above that PUNMU at \( m + \frac{k\epsilon}{N-k} \) equals

\[
\gamma \left[ -\epsilon A(m^*) \frac{k}{(N-k)} + \epsilon^2 A'(m^*) \frac{N(N-2k)k}{2(N-2)(N-k)^2} \right]
\]

It follows that PUNMU at \( \epsilon \) equals

\[
\gamma \left[ -\frac{2(1-\phi)k(N-2)A(m^*)^2}{A'(m^*)N(N-2k)} + \frac{2(1-\phi)^2k(N-2)A(m^*)^2}{A'(m^*)N(N-2k)} \right]
\]

which can be simplified as

\[
\gamma \left[ (-1-\phi) + (1-\phi)^2 \right] \frac{2(k(N-2)A(m^*)^2)}{A'(m^*)N(N-2k)}
\]

which is negative, therefore, the BST equilibria are locally unstable.

5 Conclusion

In this paper, we have fully characterized all interior equilibria of the Groves Ledyard mechanism and located new equilibria at the boundary. We have also shown that the equilibria found by Bergstrom, Simon, and Titus, may fail to exist for high values of the punishment parameter and that even if they do exist they are not stable. Moreover, the Groves Ledyard equilibrium is stable and achieves stability by combining two unstable dynamics. The boundary equilibria, which unlike the other equilibria are non efficient, are also stable.

The fact that increasing the punishment parameter can make the Groves Ledyard equilibrium unique, as well as stable, suggests that this paper resurrects the mechanism from the implicit critique of Bergstrom, Simon, and Titus. Their result implied that the mechanism created a massive coordination or learning problem in
which agents would have to select from among one of possibly billions of equilibria. In that the agents that send the high message get higher utility in those equilibria, this could create a complicated system. However, the fact that a large punishment parameter eradicates this problem does not mean it is an ideal solution. Ramping up $\gamma$ implies that all agents give the same amount. Thus, the mechanism reduces to a tax. This tax is not imposed by the rule of law but by the law of incentives.

All however, is not lost. The explicit solutions for the BST equilibria show that the messages for the agents sending the low messages are large and negative for most parameter values that we tested. These equilibria are unstable. So, if a budget constraint prevents agents from sending messages that negative, the agents would increase their messages and head back to the Groves Ledyard equilibrium. Therefore, budget constraints mitigate the need to make the punishment parameter too large. It need not be the case that the BST equilibria no longer exist, only that the budget constraint prevents agents from finding these equilibria and then heading off to the boundary equilibria. After all, since the BST equilibria are unstable, the real issue is preventing the boundary equilibria from arising. Our analysis suggests that careful selection of a punishment parameter and a budget or borrowing constraint can force the system to the Groves Ledyard equilibrium.

References


