Ashoka University Economics
Discussion Paper 62

Implementation in Large Population Games with Multiple Equilibria

June 2021

Sarvesh Bandhu, Indian Institute of Management, Bangalore
Ratul Lahkar, Ashoka University

https://ashoka.edu.in/economics-discussionpapers

Plot No. 2, Rajiv Gandhi Education City,
National Capital Region,
P.O.Rai, Sonepat, Haryana 131029

Implementation in Large Population Games with Multiple Equilibria

Sarvesh Bandhu* and Ratul Lahkar†

June 17, 2021

Abstract

Evolutionary implementation is a standard method of implementation in large population games. Such implementation may, however, be ineffective in certain situations. We consider one such situation where strategic complementarities generate multiple Nash equilibria. The planner constructs an externality adjusted game by adding the positive externalities in the game to the original payoffs. However, strategic complementarities render the Pareto inferior Nash equilibrium evolutionarily stable. The society, therefore, fails to converge to the efficient state of the model leading to the failure of evolutionary implementation. We provide a new solution to this problem of implementation in large population games with multiple equilibria using dominant strategy implementation. Our main result is that the efficient state can be implemented in strictly dominant strategy by applying Pigouvian pricing calculated on the basis of the distribution of reported types.

Keywords: Aggregative games; Potential games; Externalities; Implementation.

JEL classification: C72, C73, D62, D82.

*Economics and Social Sciences, Indian Institute of Management, Bangalore, Bannerghatta Main Road, Bangalore, Karnataka 560076, India, email: `sarvesh.bandhu@iimb.ac.in`.

†(corresponding author) Department of Economics, Ashoka University, Rajiv Gandhi Education City, Sonipat, Haryana 131029, India, email: `ratul.lahkar@ashoka.edu.in`.

1 Introduction

Pigouvian pricing is the most well known method of implementing a socially efficient outcome in the presence of externalities. However, when agents' type is private information, classical Pigouvian pricing will not work.[1] The mechanism design literature offers solutions to this problem. Perhaps the most well known such solution is the Vickrey–Clarke–Groves (VCG) direct mechanism in which truthful revelation of type becomes the dominant strategy for each agent (Vickrey [31], Clarke [5], Groves [11]). As agents reveal type truthfully, it becomes feasible for the planner to implement a transfer scheme like Pigouvian pricing that induces agents to internalize externalities.

An alternative to canonical models of implementation like the VCG mechanism that relies on truthful revelation is evolutionary implementation (Sandholm [28, 29]). This method is like Pigouvian pricing but with the crucial difference that the transfer price is calculated with respect to the prevailing social state rather than the optimal one. Using standard methods of evolutionary game theory, we can then identify situations in which the society will converge to the optimal state. Several authors have noted reasons to prefer an evolutionary approach to implementation, particularly when the number of agents is large. Unlike traditional methods like VCG, evolutionary implementation does not rely on truthful revelation as the planner only uses information about the current social state in calculating the current transfer. As a result, the planner does not need to collect information about types or compute optimal assignments based on reported types (Sandholm [29]). Nor are there concerns in this approach about revealing confidential information about types and the possibility of cheating by the bid taker and competing bidders (Rothkopf et al. [25], Rothkopf [26]).

Despite such suggested practical advantages of evolutionary implementation, it is possible that it may not work in certain situations. This paper analyzes one such situation which shows that despite the higher informational requirements, dominant strategy implementation will still succeed in achieving efficiency, even when evolutionary implementation fails. Evolutionary implementation has been applied to situations like congestion pricing (Sandholm [28, 29]), public goods, public bads and the tragedy of the commons (Lahkar and Mukherjee [17, 19]). A common feature of these models is the uniqueness of Nash equilibrium. This is true not just in the original game that models these applications but also the externality adjusted game obtained by adding externalities to the original payoff function of an agent. The key insight upon which evolutionary implementation rests, provided by Sandholm [28, 29], is that such an externality adjusted game is a potential game (Monderer and Shapley [21], Sandholm [27]). Moreover, the efficient state of the original game must be a Nash equilibrium of the externality adjusted game. The fact that this game is also a potential game then implies that if indeed it has a unique Nash equilibrium, evolutionary dynamics must converge globally to this equilibrium or, equivalently, the original efficient state.

The problem, however, arises if there are multiple equilibria in the externality adjusted game. In that case, there will not be global convergence to the efficient state. Instead, it is possible that

[1]This is because the classical Pigouvian price is calculated with reference to the social efficient outcome which the social planner cannot calculate without knowing the type distribution of agents.

the society converges to a Pareto inferior Nash equilibrium. Evolutionary implementation would breakdown in such a situation. We analyze such a situation and propose a solution to the problem. We show that despite multiple equilibria, a more conventional dominant strategy implementation approach does succeed in achieving coordination on the efficient state. This possibility of evolutionary implementation failing has also been noted by Sandholm [30]. The solution we offer, however, is very different from that in Sandholm [30].[2] Of course, dominant strategy implementation, by itself, is well understood in economic theory. But its application to achieve Pareto efficiency in large population games in an environment of multiple equilibria is the novel feature of this paper and its main contribution.

We model our problem as a large population aggregative game which are games in which the payoff of an agent depends upon individual strategy and the aggregate strategy level (Corchón [7]). The aggregative structure of the model enhances analytical tractability and enables precise identification of conditions causing multiple equilibria or success of dominant strategy over evolutionary implementation. Players are of different types and benefits to an agent depend positively on both individual strategy and the aggregate strategy level. This causes best responses in the model to be increasing with respect to the aggregate strategy level or, equivalently, generates strategic complementarities. As is generally the case in economics, strategic complementarities give rise to multiple equilibria. One such equilibrium is the outcome where all agents play the worst strategy which, in our strategy set, is the zero strategy. We show that this equilibrium is locally stable under evolutionary dynamics so that if the society gets trapped in this Pareto inferior equilibrium, it cannot escape it. In establishing this result, we use the fact that our model also constitutes a potential game. Strategic complementarities mean that some economic contexts in which our model may be relevant are education choice (McMahon [20]), technology choice (Katz and Shapiro [13]) and macroeconomic applications like search (Cooper [6]). It is also worth noting that due to strategic complementarities, externalities in our model are positive.

In contrast to this equilibrium where all agents play the zero strategy (or other possible Nash equilibria which does not play an important role in our analysis) is the Pareto efficient state. This is the state that maximizes the aggregate payoff in the society. It is this efficient state that the planner wishes to implement. Under evolutionary implementation, the planner seeks to do so by adding the current level of positive externalities to the original payoffs as a positive transfer price. As noted earlier, the original efficient state is a Nash equilibrium of this new externality adjusted game. But due to strategic complementarities, so is the Pareto inferior state where all agents play the zero strategy. Moreover, this zero equilibrium is also locally evolutionarily stable. Therefore, if the society starts in the vicinity of the zero equilibrium, it will gravitate towards it instead of the Pareto efficient state. Global convergence to social efficiency is, therefore, impossible in our model through evolutionary implementation. This highlights the stark possibility that evolutionary

[2]We should note that the ineffectiveness of evolutionary implementation we refer to is about the deterministic form of such implementation. The solution explored in Sandholm [30] is stochastic evolutionary implementation, which we describe in more detail towards the end of this section. It bears no relation to dominant strategy implementation.

implementation fails in our model. [3]

In contrast, dominant strategy implementation works regardless of multiplicity of equilibria. Direct mechanisms like VCG that rely on truthful revelation as a dominant strategy are, of course, very well known in the context of finite player implementation problems. Recently, this approach has also been extended to large population models (Lahkar and Mukherjee [19]). In fact, due to players being of measure zero, the large population results are stronger than the classical results of the VCG mechanism. In this setup, truthful revelation emerges as a strictly dominant strategy instead of being just weakly dominant. The intuition behind this approach is that of Pigouvian pricing but with respect to the reported type distribution. Based on reported type and type distribution, the planner assigns to each agent the efficient strategy and Pigouvian transfer. But due to truthful revelation of type being strictly dominant, the assigned strategies and transfers are those that coincide with the social optimum.

Our analysis, however, in no way suggests that evolutionary implementation is irrelevant. As the previous literature shows, evolutionary implementation is very effective when there is uniqueness of equilibrium; this includes important economic contexts like public goods provision and mitigating the tragedy of the commons. In fact, in such situations, evolutionary implementation would be informationally more parsimonious than dominant strategy implementation as the planner wouldn't need to know any type specific information. In fact, that is why, evolutionary implementation does not require truthful revelation. The fundamental point this paper makes is that there may be situations where evolutionary implementation is ineffective. Then, the planner will require alternatives like dominant strategy implementation despite its greater informational requirement and other concerns like confidentiality of private information that has been expressed in the literature. We should also note that while budget balance is easy to achieve in the dominant strategy implementation mechanism, it may fail individual rationality for high cost types of agents. But in that case, this would also be a weakness with the evolutionary implementation mechanism.

The method of evolutionary implementation we have considered here as well as in Sandholm [28, 29] and Lahkar and Mukherjee [17, 19] is deterministic evolutionary implementation. The social state changes under deterministic evolutionary dynamics. Previously, Sandholm [30] has also noted the possibility that deterministic evolutionary implementation may not work due to multiple equilibria. The solution proposed in that paper was stochastic evolutionary implementation. Evolution in this method is described by an ergodic Markov process. There are no absorbing states in this process. But the stationary distribution of the process in the externality adjusted game places nearly all the probability mass on that Nash equilibrium which is the efficient state of the original game.[4] Thus, even though the process keeps visiting all states, in the long run, it will be at the efficient state most of the time. This constitutes stochastic evolutionary implementation. Notice

[3] This is not to say that evolutionary implementation always fails. If the initial state is near the efficient state, we will have convergence to that state.

[4] In general, such stationary distributions concentrate their mass on the maximizer of the potential function in a potential game. In an externality adjusted game, the potential function is the aggregate payoff of the original game whose maximizer is the efficient state.

that in contrast to deterministic evolution, such convergence happens regardless of the number of Nash equilibrium. But it is also a feature of stochastic evolutionary methods that the time span required for convergence may be so long that it is doubtful how relevant it is in explaining social or economic phenomenon (Ellison [9]). Our method of resolving the problem with deterministic evolutionary implementation is completely different. Since dominant strategy implementation does not rely on any evolutionary process, the coordination it achieves is instantaneous. Hence, the question of delayed coordination associated with stochastic implementation does not arise.

The rest of the paper is as follows. Section 2 introduces the model and defines potential games. In Section 3, we characterize Nash equilibria and efficient states of the model and also establish local stability of the Pareto inferior equilibrium involving zero strategies. In Section 4, we describe evolutionary implementation and explain its possible failure in the model. Section 5 presents dominant strategy implementation as a solution and Section 6 concludes.

2 Model

We consider a society of a continuum of agents of mass 1. Each agent, therefore, is of measure zero. The society is divided into a finite set of *populations* or *types* $\mathcal{P} = \{1, 2, \cdots, n\}$. We denote the mass of population or type $p \in \mathcal{P}$ as $m_p \in (0, 1)$ with $\sum_{p \in \mathcal{P}} m_p = 1$. Hence, we refer to the distribution $m = (m_1, m_2, \cdots, m_n)$ as the type distribution. Every agent in the society has the common strategy set $\mathcal{S} = [0, \infty)$. We denote the state of a population p by the finite positive measure μ_p, with $\mu_p(B) \in (0, m_p)$ describing the mass of agents in population p playing strategies in $B \subseteq \mathcal{S}$. A social state is, therefore, $\mu = (\mu_1, \mu_2, \cdots, \mu_n)$. For our subsequent analysis, we will require the notion of the aggregate strategy level in the society. We define the aggregate strategy at the social state μ as

$$A(\mu) = \sum_{p \in \mathcal{P}} \int_{\mathcal{S}} x \mu_p(dx). \tag{1}$$

Since we have assumed that $\sum_{p \in \mathcal{P}} = 1$, it follows that $A(\mu) \in [0, \infty)$.

We now define payoffs in our model. Consider an agent from population $p \in \mathcal{P}$ playing strategy $x \in \mathcal{S}$. We denote the payoff of such an agent at social state μ as $F_{x,p}(\mu)$ and define that payoff as

$$F_{x,p}(\mu) = x\pi(A(\mu)) - c_p(x). \tag{2}$$

We interpret $\pi : [0, \infty) \to \mathbf{R}_+$ as an aggregative function that captures the benefit that an agent derives from the aggregate strategy level in the society. The total benefit for the agent is the product of the agent's own strategy and this aggregative benefit. The function $c_p : \mathcal{S} \to \mathbf{R}_+$ captures the cost the agent incurs from using strategy x. While the aggregative benefit function is common for agents, the cost function differs across types but is the same for all agents of a particular type. An equivalent interpretation, therefore, is that the cost function determines the type of an agent. Once we subtract this cost from the total benefit, we obtain the agent's payoff (2). For reasons of conciseness, we denote the population game described by this payoff function as F. Since the payoff depends entirely upon the agent's own strategy and the aggregate strategy level (1), F is an aggregative game.[5] Of course, aggregative games can take other functional forms as well. But here, we will confine ourselves to the form (2). We now state the assumptions about the functions π and c_p in (2).

Assumption 2.1 *We assume the following about π and c_p in (2).*

1. *π and c_p, for all $p \in \mathcal{P}$, are smooth functions.*
2. *π is strictly increasing and strictly concave such that* $\lim_{\alpha \to \infty} \pi'(\alpha) = 0$.
3. *For all $p \in \mathcal{S}$, c_p is strictly increasing and strictly convex with its third derivative being non-negative.*
4. *$\alpha\pi'(\alpha)$ is non–decreasing and well–defined at $\alpha = 0$.*
5. *$\pi(\alpha) + \alpha\pi'(\alpha) < c'_p(0)$ at $\alpha = 0$ for all $p \in \mathcal{P}$.*
6. *For every possible type distribution $m = (m_1, \cdots, m_n)$, there exists at least one vector $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbf{R}^n_+$ such that if $\alpha = \sum_{p \in \mathcal{P}} m_p \alpha_p$, then $\alpha\pi(\alpha) - \sum_{p \in \mathcal{P}} m_p c_p(\alpha_p) > 0$.*

The first assumption is technical. In the next paragraph, we provide a more careful justification of why we require π to increasing. Strict concavity of π is a natural assumption and will have some bearing on ensuring a unique efficient state in our model. In the third assumption, strict convexity of c_p will generate a unique best response while the condition about the third derivative will ensure that the best response is finite. The fourth and fifth assumptions will ensure that the state where all agents play zero is a stable Nash equilibrium both in the original game (2) and the externality adjusted game we will define later. The last assumption is very mild and it is required only to ensure that efficient state doesn't become trivial.

Our most important assumption is that π is strictly increasing. As we will show in Section 3, this assumption will cause best responses to be upward sloping with respect to the aggregate strategy level. This is equivalent to the presence of strategic complementarities in the model which will be a major driving factor behind the main results on evolutionary versus dominant strategy implementation in this paper. We will also see in Section 4 that a strictly increasing π causes externalities to be positive which, again, plays an important role in our analysis. As mentioned in the Introduction, some important economic applications of our model may be education choice, technology choice and macroeconomic spillovers. In each case, x in (2) would be the relevant strategic variable like education or technology choice. Individual welfare is then increasing in both individual strategy and aggregate strategy, which may be interpreted as either strategic complementarities or positive externalities.

[5] Corchòn's [7] original definition of aggregative games is in the context of a finite number of players. Here, we are extending that notion to games with a continuum of strategies. See, for example, Lahkar and Mukherjee [17, 19] for other applications of such an extension.

2.1 Potential Games

Another important concept we use in this paper is that of potential games. (Monderer and Shapley [21], Sandholm [27]). These are games in which payoffs may be summarised using a real–valued function called the potential function. To define such games, we need some additional notation to deal with the technical complexities of a game with a continuous strategy set. Let $\mathcal{B}$ be the Borel $\sigma-$algebra on $\mathcal{S}$ and $\mathcal{M}(\mathcal{S})$ be the space of finite signed measures on $(\mathcal{S}, \mathcal{B})$. The space of finite measures that impose a total mass of $m > 0$ on $\mathcal{S}$ is then $\mathcal{M}_m^+(\mathcal{S}) \subset \mathcal{M}(\mathcal{S})$.[6] Thus, the set $\mathcal{M}_{m_p}^+(\mathcal{S}) \subset \mathcal{M}(\mathcal{S})$ is the set of states in population p, with $\mu_p \in \mathcal{M}_{m_p}^+(\mathcal{S})$ being such a population state. The set of states in the entire society is then $\Delta = \prod_{p=1}^n \mathcal{M}_{m_p}^+(\mathcal{S})$ with $\mu = (\mu_1, \cdots, \mu_n)$ being such a state.

We also require the notion of the Fréchet derivative, which is a generalization of the usual notion of the derivative to Banach spaces. Notice that the domain of the payoff function $F_{x,p}(\mu)$ in (2) is Δ. To define the Frechet derivative for $F_{x,p}(\mu)$, we extend its domain from Δ to $\mathscr{M} = \prod_{p=1}^n \mathcal{M}(\mathcal{S})$.[7] Consider now a direction $\zeta = (\zeta_1, \zeta_2, \cdots, \zeta_n) \in \mathscr{M}$. The Fréchet derivative $DF_{x,p}(\mu)\zeta$, represents the change in $F_{x,p}(\mu)$ when $\mu \in \mathscr{M}$ changes in the direction ζ. Let $f : \mathscr{M} \to \mathbf{R}$ be a Fréchet differentiable function and suppose there exists a function ∇f, called the gradient of f, on $\mathcal{S} \times \mathcal{P}$ such that

$$Df(\mu)\zeta = \sum_{p \in \mathcal{P}} \int_{\mathcal{S}} \nabla f(\mu)(x,p)\zeta_p(dx), \text{ for all } \zeta = (\zeta_1, \cdots, \zeta_n) \in \mathscr{M},$$

where $Df(\mu)\zeta$ is the Fréchet derivative of f at $\mu \in \mathscr{M}$ in the direction $\zeta = (\zeta_1, \cdots, \zeta_n) \in \mathscr{M}$. We then define a multipopulation potential game as follows (Lahkar and Mukherjee [17]).

Definition 2.2 *A population game F is a potential game if there exists a Fréchet differentiable (with respect to the variational norm) function $f : \mathscr{M} \to \mathbf{R}$ such that*

$$\nabla f(\mu) = F(\mu) \textit{ for all } \mu = (\mu_1, \cdots, \mu_n) \in \Delta$$

or, equivalently, $\nabla f(\mu)(x,p) = F_{x,p}(\mu)$ for all $(x,p) \in \mathcal{S} \times \mathcal{P}$. The function f is called the potential function of the game F.

Proposition 5.3 in Lahkar [16] shows that a large population continuous strategy model of Cournot competition with multiple types of cost functions is a potential game. The Cournot competition model, however, has the same functional form as the population game (2) we are considering in this paper. There is, of course, the distinction that in the earlier model, the aggregative benefit function is the downward sloping market inverse demand function whereas in the present model, the aggregative benefit function is increasing. This distinction will be important later in some respects. But the slope of the aggregative benefit function does not affect the way we establish

[6] Hence, $\mathcal{M}_1^+(\mathcal{S})$ is the space of probability measures on $\mathcal{S}$.

[7] To ensure that $\mathscr{M}$ is a Banach space, we impose the variational norm on it. See, for example, Appendix A.1.1 in Lahkar and Mukherjee [17] for more details.

these games to be potential games. Therefore, we obtain the following proposition, a result which will be useful in establishing stability of Nash equilibria in the next section.

Proposition 2.3 *The population game F defined by (2) is a potential game with potential function $f : \mathcal{M} \to \mathbf{R}$ defined by*

$$f(\mu) = \int_0^{A(\mu)} \pi(z)dz - \sum_{p \in \mathcal{P}} \int_{\mathcal{S}} c_p(x)\mu_p(dx). \tag{3}$$

Proof. See Proposition 5.3 in Lahkar [16]. ■

3 Nash Equilibrium and Efficient State

3.1 Nash Equilibrium

We now characterise Nash equilibria and efficient state of the population game defined in (2). First, we consider Nash equilibrium, which we formally define as follows.

Definition 3.1 *A Nash equilibrium of a multipopulation game F such as (2) is a social state $\mu^* = (\mu_1^*, \mu_2^*, \cdots, \mu_n^*) \in \Delta$ such that for all $x \in \mathcal{S}$, all $p \in \mathcal{P}$, if x lies in the support of μ_p^*, then $F_{x,p}(\mu^*) \geq F_{y,p}(\mu^*)$, for all $y \in \mathcal{S}$.*

We use the aggregative structure of F to derive its Nash equilibria. To simplify notation, we denote the aggregate strategy level $A(\mu)$ defined in (1) as $\alpha \in [0, \infty)$ and write the payoff function (2) as $x\pi(\alpha) - c_p(x)$. Due to our assumption that $c_p(x)$ is strictly convex, this function is strictly concave in x and, therefore, has a unique maximizer in $\mathcal{S} = [0, \infty)$. The non-negativity of third derivative of $c_p(\cdot)$ ensures that maximizer is finite. This maximizer is the unique best response of a type p agent to every social state μ such that $A(\mu) = \alpha$. We denote this best response as $b_p(\alpha)$ and note that is characterized by

$$b_p(\alpha) = \begin{cases} x^* \in (0, \infty) & \text{such that } \pi(\alpha) = c_p'(x^*) \text{ if } \pi(\alpha) > c_p'(0) \\ 0 & \text{if } \pi(\alpha) \leq c_p'(0). \end{cases} \tag{4}$$

The ease with which the best response can be characterised is due to the large population characteristic of the game. Since each agent is of measure zero, no agent can individually affect the aggregate strategy level α. Hence, finding the best response becomes a simple matter of differentiating the payoff function holding α constant. The following observations then arise from (4).

Observation 3.2 *The characterization of $b_p(\alpha)$ in (4) implies the following.*

1. *If $\alpha = 0$, then $b_p(\alpha) = 0$. This is because Assumption 2.1, parts 4 and 5, imply $\pi(0) < c_p'(0)$.*

2. *For each $p \in \mathcal{P}$, there exists $\bar{\alpha}_p > 0$ characterized by $\pi(\bar{\alpha}_p) = c_p'(0)$ such that for all $\alpha \in [0, \bar{\alpha}_p]$, $b_p(\alpha) = 0$. This again follows from the implication that $\pi(0) < c_p'(0)$. For all $\alpha > \bar{\alpha}_p$, $b_p(\alpha)$ is strictly increasing.*

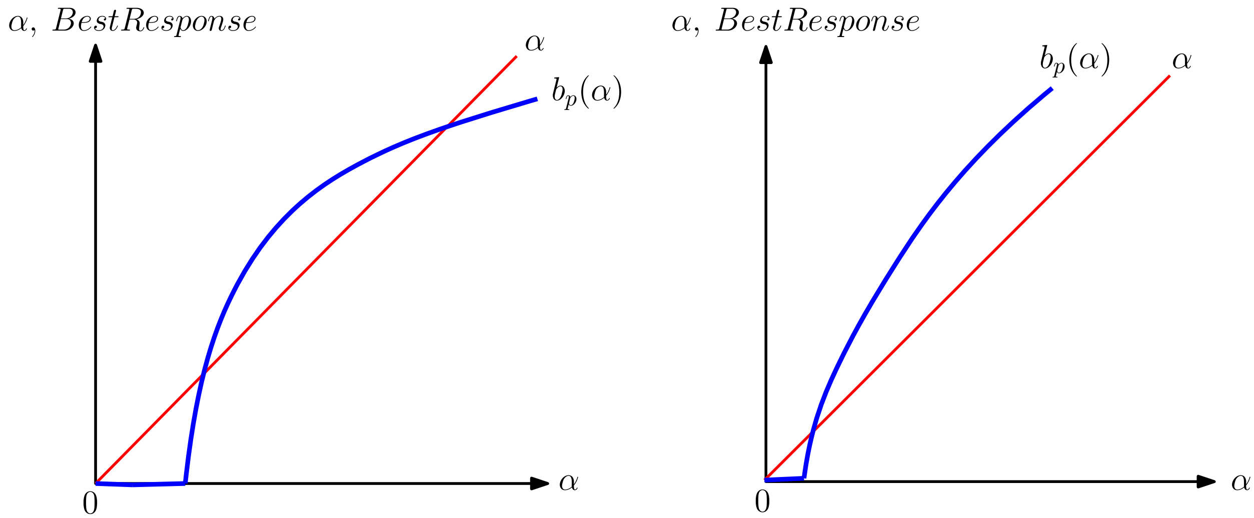

Figure 1: The blue curves represent $b_p(\alpha)$ and the red lines represent α. For each population, the best response $b_p(\alpha)$ is flat in the vicinity of $\alpha = 0$ and is then strictly increasing. The non–decreasing slope and concavity of $b_p(\cdot)$ are ensured by conditions assumed on π and c_p. We include α in the diagram because it is going to be useful later in characterising Nash equilibria.

Figure 1 shows some possible shapes of the best response function $b_p(\alpha)$ for two different populations. Notice that $b_p(\alpha)$ is generally upward sloping with respect to α, except for a possible flat stretch in the vicinity of $\alpha = 0$. Thus, there are strategic complementarities between individual strategy and the aggregate strategy level. As can be seen from (4), such complementarities arise from the assumption that π itself increases with α. This highlights the importance of this assumption as our subsequent results on Nash equilibria depends crucially on the upward sloping best responses.

Notice that $b_p(\alpha)$ is the best response for a type p agent at all social states μ such that $A(\mu) = \alpha$. Further, the best response $b_p(\alpha)$ is identical for every type p agents but will differ from agents of another type. If all agents of population p are playing the best response $b_p(\alpha)$ to μ, then the resulting population state is the monomorphic state $m_p\delta_{b_p(\alpha)}$.[8] If all agents in all populations are playing their respective best responses, then we denote the resulting social state as

$$B(\mu) = \left(m_1\delta_{b_1(\alpha)}, m_2\delta_{b_2(\alpha)}, \cdots, m_n\delta_{b_n(\alpha)}\right). \tag{5}$$

Applying (1) to (5), we obtain the aggregate strategy level at the social state $B(\mu)$ as

$$AB(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} x m_p \delta_{b_p(\alpha)} = \sum_{p\in\mathcal{P}} m_p b_p(\alpha). \tag{6}$$

The following proposition characterizes Nash equilibria in our model. A variant of this result also appears as Proposition 3.1 in Lahkar [16] in the context of indirect evolution of preferences in

[8] Here, δ_x is the Dirac measure putting probability 1 on x.

aggregative games. Since the proof is nearly identical, we do not reproduce it here and, instead, refer the reader to the earlier result.

Proposition 3.3 *Consider the equation*

$$\sum_{p \in \mathcal{P}} m_p b_p(\alpha) = \alpha. \tag{7}$$

The social state $\mu^* = \left(m_1 \delta_{b_1(\alpha^*)}, m_2 \delta_{b_2(\alpha^*)}, \cdots, m_n \delta_{b_n(\alpha^*)}\right)$ *is a Nash equilibrium of the population game* F *defined by (2) if and only if* α^* *is a solution of (7).*

Proof. See the proof of Proposition 3.1 in Lahkar [16]. ■

Proposition 3.3 greatly simplifies the characterisation of Nash equilibria in our aggregative game. Intuitively, the LHS of (7) is the aggregate best response function that arises when all agents in all populations are playing their best response. This result, therefore, shows that Nash equilibria in aggregative games can be characterised as fixed points of the aggregate best response function. More precisely, (7) implies that for α^* to be the aggregate strategy level at a Nash equilibrium, it must remain unchanged when all agents play a best response to it. The associated Nash equilibrium is then the social state at which every agent plays the unique best response to that aggregate strategy level. Hence, a Nash equilibrium consists entirely of monomorphic population states. We now obtain the following corollary.

Corollary 3.4 *Consider the population game* F *defined by (2). The social state*

$$\mu^0 = (m_1 \delta_0, m_2 \delta_0, \cdots, m_n \delta_0)$$

is a Nash equilibrium of F.

Proof. Recall from Observation 3.2(1) that if $\alpha = 0$, then $b_p(\alpha) = 0$ in F. Therefore, $\alpha = 0$ satisfies (7). Hence, by Proposition 3.3, $\mu = \left(m_1 \delta_{b_1(0)}, m_2 \delta_{b_2(0)}, \cdots, m_n \delta_{b_n(0)}\right)$ is a Nash equilibrium of F. But by Observation 3.2(1), $b_p(0) = 0$ for all $p \in \mathcal{P}$. The result, therefore, follows. ■

Corollary 3.4, therefore, implies that the state where all agents play the strategy 0 is a Nash equilibrium in our model. Clearly, this is the Pareto inferior Nash equilibrium in the game. The result arises because when the aggregate strategy level is zero, which can happen only when all agents play 0, then the best response for all types is also 0. Of course, that is not necessarily the only Nash equilibrium in the model. For example, we obtain Figure 2 by taking a convex combination of the two best response functions in Figure 1.[9] In this figure, there are three intersections. Thus, by Proposition 3.3, there are three Nash equilibria, each associated with one intersection of $\sum_p m_p b_p(\alpha)$ and α, namely, 0, $\hat{\alpha}$ and α^*. We note that such multiple equilibria are the direct consequence of

[9] Thus, in this example, there are two populations. Population 1's best response is the one in the left panel of Figure 1 and population 2's best response is the one in the right panel.

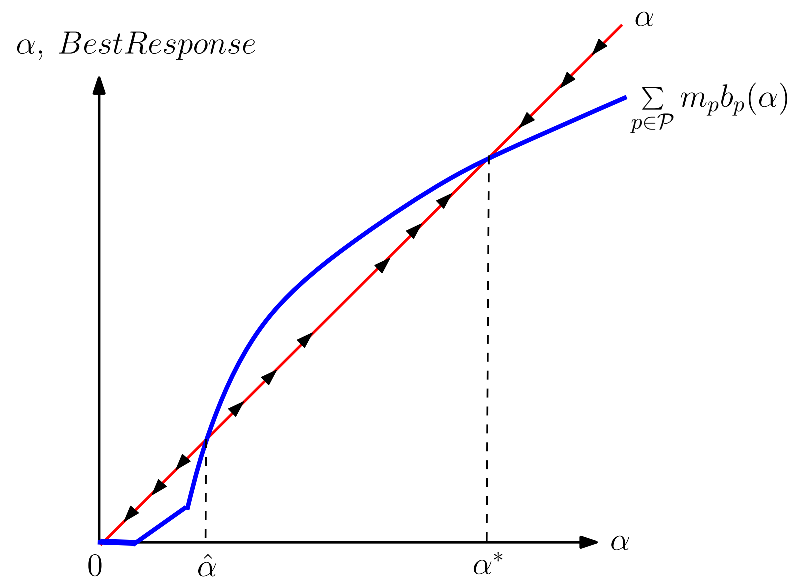

Figure 2: The blue curve represents $\sum_p m_p b_p(\alpha)$ and is a convex combination of the two best response functions in Figure 1. The red line represents α. The intersections 0, $\hat{\alpha}$ and α^{**} are the Nash equilibrium levels of aggregate strategy. The arrows represent the direction of convergence under the best response dynamic as discussed at the end of Section 3.1.1.

strategic complementarities in our model. Had best responses been declining, then there would have been only one intersection between $\sum_{p\in\mathcal{P}} m_p b_p(\alpha)$ and, hence, only one Nash equilibrium (by Proposition 3.3). This would have been the case had π in (2) been declining.

3.1.1 Potential Games and Evolutionary Stability

Recall from Proposition 2.3 that the aggregative game (2) is a potential game. Analytically, large population potential games are convenient because Nash equilibria in such games can be characterized by locally maximizing or minimizing the potential function (Sandholm [27]).[10] In principle, therefore, we could have characterized Nash equilibria in our model by using the potential function (3).

We, have, however chosen the alternative method of Proposition 3.3 for this purpose. The validity of this method is restricted to aggregative games but within this class of games, it allows a more transparent way of characterizing Nash equilibria by directly computing best responses. The potential game method may be applicable more generally than aggregative games but in the present case, would have been analytically less tractable. This is because unlike in, say, Lahkar [16], the potential function here is not concave due to the fact that the aggregative benefit function π is increasing. Without concavity, the potential function will have multiple local maximizers or minimizers associated with the multiple possible Nash equilibria in our model as can be seen in, for example, Figure 2. Computing all such local maximizers and minimizers would have been a

[10]See Cheung [2], Lahkar and Riedel [15] and Cheung and Lahkar [4] for extensions of the concept of potential games to games with continuous strategy sets.

cumbersome exercise, more so because of the abstract measure theoretic nature of the potential function.

Nevertheless, relating our model to potential games is still helpful. Potential games have attractive convergence properties under standard evolutionary dynamics. All such dynamics converge to a Nash equilibria in potential games (Sandholm [27], Cheung and Lahkar [4]). Indeed, it is due to such convergence properties that potential games constitute the basis of the whole concept of evolutionary implementation that we discuss in the next section. Dynamics that do converge to Nash equilibria in potential games include the replicator dynamic (Oechssler and Riedel [22, 23], Cheung [3]), the BNN dynamic (Hofbauer et al. [12]), the pairwise comparison dynamic (Cheung [2]), the logit dynamic (Lahkar and Riedel [15], Perkins and Leslie [24]) and the best response dynamic for aggregative games (Lahkar and Mukherjee [17]). We describe these dynamics in Appendix A.1. The potential function itself acts as the Lyapunov function along which solution trajectories of these dynamics ascend. Therefore, local maximizers of the potential function emerge as evolutionarily stable Nash equilibria while local minimizers are unstable (Sandholm [27], Cheung and Lahkar [4]).[11]

We apply this technique to show that the Nash equilibrium where all agents play the zero strategy is locally asymptotically stable. Recall from Corollary 3.4 that this is indeed a Nash equilibrium in our model. Showing local stability would require us to establish that potential function is locally maximized at this equilibrium. This, however, is not a trivial task because the potential function (3) is defined on an abstract measure theoretic space. We, therefore, construct a finite dimensional analogue of the potential function which we describe in the Appendix and which can be more easily analyzed. Using that function, we establish the following result on local stability of μ^0. Its proof is also in the Appendix.

Proposition 3.5 *The Nash equilibrium μ^0 characterized in Corollary 3.4 in which $\mu_p^0 = m_p\delta_0$ for every $p \in \mathcal{P}$ is a local maximizer of the potential function (3). Hence, this Nash equilibrium is locally asymptotically stable with respect to the weak topology under the replicator dynamic (25), the BNN dynamic (26), the pairwise comparison dynamic (27) and the best response dynamic (29) for aggregative games. For the logit dynamic (28), local asymptotic stability holds with respect to the logit equilibrium that comes arbitrarily close (with respect to the weak topology) to μ^0 as the perturbation parameter $\eta \to 0$ (see footnote 11).*

Proposition 3.5 implies that if the society is trapped in the Pareto inferior Nash equilibrium μ^0 in our model, then it will tend to remain there. Slight perturbations will be insufficient for the society to escape that equilibrium. This is easiest to see for the best response (BR) dynamic for

[11] Logit dynamic is generated through a perturbation of the best response using the Shannon entropy function (Lahkar and Riedel [15]). Rest points of this dynamic are not Nash equilibria but logit equilibria which approximate Nash equilibria as the perturbation parameter $\eta \to 0$. Therefore, convergence under logit equilibria happens to such approximate Nash equilibria. Under the replicator dynamic, it is well know that all social states that are entirely in monomorphic population states are rest points. This is true even if such a social state is not a Nash equilibrium. Therefore, under the replicator dynamic, convergence to Nash equilibrium is guaranteed only from the interior of the state space.

aggregative games which we describe in more detail in (29) in Appendix A.1. Under this dynamic, a population state μ_p changes according to $\dot{\mu}_p = m_p\delta_{b_p(\alpha)} - \mu_p$. The evolution of the aggregate strategy level is then given by

$$\begin{aligned}
\dot{\alpha} = \dot{A}(\mu) &= \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} x\dot{\mu}_p(dx) \\
&= \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} x\left(m_p\delta_{b_p(\alpha)} - \mu_p\right)(dx) \\
&= \sum_{p\in\mathcal{P}} m_p b_p(\alpha) - A(\mu) \\
&= \sum_{p\in\mathcal{P}} m_p b_p(\alpha) - \alpha. \qquad (8)
\end{aligned}$$

We refer to (8) as the aggregate best response (ABR) dynamic. Consider now an application of (8) to the situation illustrated in Figure 2 with the three intersections, 0, $\hat{\alpha}$ and α^*.[12] If we provide a dynamic interpretation to Figure 2, then the difference between the two curves in the figure represents the ABR dynamic (8). Each intersection is a rest point of the dynamic. Moreover, 0 and α^* are stable rest points and $\hat{\alpha}$ is unstable. Thus, from all aggregate strategy levels in $(0, \hat{\alpha})$, the ABR dynamic converges to 0. But the only way $A(\mu) = 0$ is if all agents play the 0 strategy. Therefore, equivalently, from all initial states μ such that $A(\mu) \in (0, \hat{\alpha})$, the BR dynamic (29) converges to μ^0. Hence, μ^0 is locally asymptotically stable under the BR dynamic. By a similar argument, μ^* where every agent in population p plays $b_p(\alpha^*)$ will also be locally asymptotically stable Nash equilibrium. On the other hand, $\hat{\mu}$ where every agent in population p plays $b_p(\hat{\alpha})$ is unstable under the BR dynamic.

Notice that in establishing convergence under the BR dynamic, we did not have to use the potential game property of the underlying game.[13] This is because the dynamics of the BR dynamic for aggregative games can be fully captured using the simpler one–dimensional ABR dynamic (8) of the aggregate strategy level. Such a reduction is not possible for the other dynamics mentioned in Proposition 3.5. For them, we need to use the more sophisticated machinery of potential games. Nevertheless, the ABR dynamic provides some intuition behind the behavior of such dynamics as well.

3.2 Efficient State

To define an efficient state of a population game, we need to introduce its aggregate payoff. Given a population game F, the aggregate payoff at a social state $\mu \in \Delta$ is the function $\bar{F} : \Delta \to \mathbf{R}$ defined as

$$\bar{F}(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} F_{x,p}(\mu)\mu_p(dx). \qquad (9)$$

[12] The finiteness of the highest possible Nash equilibrium level of aggregate strategy, which in this case is α^*, is ensured by assumption that $\pi'(\alpha) \to 0$ as $\alpha \to \infty$.

[13] Hence, such arguments would have worked even if F had not been a potential game.

An efficient state of F is then defined as follows.

Definition 3.6 *An efficient state of a multipopulation game F such as (2) is a social state $\mu^{**} = (\mu_1^{**}, \mu_2^{**}, \cdots, \mu_n^{**}) \in \Delta$ that maximizes the aggregate payoff $\bar{F}$ as defined in (9).*

The aggregate payoff in the population game (2) is

$$\begin{aligned}\bar{F}(\mu) &= \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} F_{x,p}(\mu)\mu_p(dx) \\ &= \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} (x\pi(A(\mu)) - c_p(x))\,\mu_p(dx) \\ &= \pi(A(\mu)\sum_{p\in\mathcal{P}} \int_{\mathcal{S}} x\mu_p(dx) - \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} c_p(x)\mu_p(dx) \\ &= A(\mu)\pi(A(\mu)) - C(\mu), \end{aligned} \tag{10}$$

where we have used the notation $C(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} c_p(x)\mu_p(dx)$ to denote aggregate cost at the social state μ. In order to characterize an efficient state of the population game (2), we establish the following proposition. Note that the aggregate payoff function is defined over an abstract measure space. Hence, directly maximizing it to characterize an efficient state is difficult. This proposition, however, shows that due to the aggregative nature of our model and the strict convexity of the cost functions, any efficient state of F must be in monomorphic population states. Characterizing an efficient state then becomes the simple task of maximizing a function defined over $n-$tuple of real numbers. The formal proof is in the Appendix.

Proposition 3.7 *Suppose $\mu^{**} = (\mu_1^{**}, \mu_2^{**}, \cdots, \mu_n^{**})$ is an efficient state of F defined by (2). Then, for every $p \in \mathcal{P}$, μ_p^{**} is a monomorphic state. Hence, an efficient state of F can be characterized by maximizing the function $\bar{G} : \prod_{p\in\mathcal{P}} \mathcal{S} \to \mathbf{R}$ such that*

$$\bar{G}(\alpha_1, \alpha_2, \cdots, \alpha_n) = \pi\left(\sum_{p\in\mathcal{P}} m_p\alpha_p\right) \sum_{p\in\mathcal{P}} m_p\alpha_p - \sum_{p\in\mathcal{P}} m_p c_p(\alpha_p). \tag{11}$$

This function has a unique global maximizer, which we denote as $(\alpha_1^{**}, \alpha_2^{**}, \cdots, \alpha_n^{**})$, with α_p^{**} finite for every $p \in \mathcal{P}$. Further, the unique efficient state of F is $\mu^{**} = \left(m_1\delta_{\alpha_1^{**}}, m_2\delta_{\alpha_2^{**}}, \cdots, m_n\delta_{\alpha_n^{**}}\right)$.

Proposition 3.7, therefore, implies that if $(\alpha_1^{**}, \alpha_2^{**}, \cdots, \alpha_n^{**})$ is the global maximizer of $\bar{G}$, then at the efficient state of F, every agent of type p plays strategy α_p^{**}. Using (11), we can characterize α_p^{**} as

$$\pi\left(\sum_{q\in\mathcal{P}} m_q\alpha_q^{**}\right) + \pi'\left(\sum_{q\in\mathcal{P}} m_q\alpha_q^{**}\right) \sum_{q\in\mathcal{P}} m_q\alpha_q \leq c_p'(\alpha_p^{**}), \tag{12}$$

with strict inequality holding only if $\alpha_p^{**} = 0$.

Recall part 6 of Assumption 2.1. This assumption implies that $\alpha_p^{**} > 0$ for at least one $p \in \mathcal{P}$ so that $\mu^{**} \neq \mu^0$. In fact, typically, the efficient state will differ from any of the Nash equilibria of our model. Intuitively, this happens due to the presence of externalities, a notion we formalize in the next section. As we will show later, externalities in our model are positive. Standard microeconomics then implies that the aggregate strategy level at the efficient state will be higher than that at any of the Nash equilibria.[14]

4 Evolutionary Implementation

We now consider the problem of implementing the efficient state. For this purpose, we introduce a social planner. Formally, given any type distribution $m = (m_1, m_2, \cdots, m_n)$, the planner wishes to implement the efficient state μ^{**} as characterized in Proposition 3.7. Thus, in the conventional terminology of the mechanism design literature, the mapping $m \mapsto \mu^{**}$ is the *social choice function* the planner wishes to implement. We assume that the planner does not know the type distribution. But in this section, we will assume that the planner knows the aggregative benefit function π.

We first consider the method of evolutionary implementation (Sandholm [28, 29]).[15]. Evolutionary implementation relies on the notion of *variable externality pricing*. This is a transfer that equals the total externality that an agents creates at the current social state. Therefore, to describe evolutionary implementation, we first need to formalize the notion of externalities in a population game with a continuous strategy set.

This again requires the Fréchet derivative. Recall from Section 2.1 that the Fréchet derivative $DF_{y,q}(\mu)\zeta$, represents the change in $F_{y,q}(\mu)$ when $\mu \in \mathcal{M}$ changes in the direction ζ. Using the Fréchet derivative, we then define the *total externality* imposed by agents of type p who play strategy x at the social state μ as[16]

$$e_{x,p}(\mu) = \sum_{q \in \mathcal{P}} \int_{\mathcal{S}} DF_{y,q}(\mu)\mathcal{D}_{x,p}\mu_q(dy), \tag{13}$$

where $DF_{y,q}(\mu)$ is the Fréchet derivative of the payoff function $F_{y,q}(\mu)$ and $\mathcal{D}_{x,p} = \zeta \in \mathcal{M}$ such that $\zeta_p = \delta_x$ and $\zeta_k = 0$ for all $k \in \mathcal{P} \setminus p$. This measure ζ represents a situation in which the only change in μ is an increase in the mass of agents using strategy x in population p. Thus, $DF_{y,q}(\mu)\mathcal{D}_{x,p}$ is the change in the payoff of strategy $y-$users in population q when the mass of strategy $x-$users in population p changes. Summing up over all strategies $y \in \mathcal{S}$ and all populations $p \in \mathcal{P}$ then gives us the total externality (13).

Consider now a general aggregative game $F_{x,p}(\mu) = \beta_p(x, A(\mu))$. Proposition 4.1 in Lahkar and

[14] Somewhat informally, this is because the aggregate strategy level at any Nash equilibrium, $\sum_p m_p\alpha_p^*$, would be given by $\pi\left(\sum_p m_p\alpha_p^*\right) \leq c_p'(\alpha_p^*)$. But then, the aggregate strategy level $\sum_p m_p\alpha_p^{**}$ that solves (12) must be strictly higher because the LHS of (12) is strictly greater than $\pi\left(\sum_p m_p\alpha_p\right)$ due to the fact that $\pi'(\cdot) > 0$.

[15] This method has been extended to continuous strategy games by Lahkar and Mukherjee [17, 19]

[16] See Appendix A.1.1 in Lahkar and Mukherjee [17] for further details.

Mukherjee [19] then shows that the total externality in such a game is

$$e_{x,p}(\mu) = x \sum_{q \in \mathcal{P}} \int_{\mathcal{S}} \beta_{q,2}(y, A(\mu))\mu_q(dy) \tag{14}$$

where $\beta_{q,2}(y,\alpha) = \frac{\partial \beta_q(y,\alpha)}{\partial \alpha}$. In the population game (2), $\beta_p(x, A(\mu)) = x\pi(A(\mu)) - c_p(x)$. Hence, applying (14), we obtain the following observation.

Observation 4.1 *The total externality imposed by a type p agent playing $x \in \mathcal{S}$ in the population game F defined by (2) is*

$$e_{x,p}(\mu) = xA(\mu)\pi'(A(\mu)). \tag{15}$$

It is clear from (15) that due to our assumption that $\pi'(\alpha) > 0$, externalities in our model are positive. This is another important implication of this assumption in our model.

Variable externality pricing then entails the planner imposing a transfer $t_{x,p}(\mu) = e_{x,p}(\mu)$ upon type p agents who play strategy x at state μ. Thus, from (15), this transfer is

$$t_{x,p}(\mu) = e_{x,p}(\mu) = xA(\mu)\pi'(A(\mu)). \tag{16}$$

Notice that although in principle, the transfer can vary according to the type of an agent, this does not happen in the present case. This is because the aggregative benefit function π is common for all agents. Hence, the transfer (16) is the same for all agents who play strategy x irrespective of the type of the agent. Therefore, in applying this transfer, the planner does not need to know the type of individual agents. In fact, due to the aggregative nature of the game, the planner also does not need to know the type distribution or the current social state μ as long as he can observe the aggregate strategy level $A(\mu)$. It is certainly plausible to assume that the planner can indeed observe this variable as well as individual strategy levels x. The planner, however, does need to know π.

Intuitively, the transfer takes the form of a price $A(\mu)\pi'(A(\mu))$ for every unit of the strategy that the agent uses. Since this price varies with the social state μ, it is called a variable externality price. In the present model where externalities are positive, the transfer takes the form of a subsidy. This idea has similarities with the classical notion of Pigouvian pricing but with the crucial difference that a Pigouvian price is calculated with respect to the socially efficient state. Thus, in the present context, the Pigouvian price on an agent playing strategy x would have been $xA(\mu^{**})\pi'(A(\mu^{**})) = x\alpha^{**}\pi'(\alpha^{**})$ where, applying Proposition 3.7, $\alpha^{**} = \sum_{p \in \mathcal{P}} m_p \alpha_p^{**}$ is the efficient aggregate strategy level. Notice also that calculating the Pigouvian price would have required the planner to know the cost functions c_p while evolutionary implementation does not require this type specific information. This issue will be relevant in the next section on dominant strategy implementation.

The transfer policy (16) transforms the original game F defined by (2) into a new game $\hat{F}$ in which the payoff of a type p agent playing strategy x is

$$\begin{aligned}\hat{F}_{x,p}(\mu) &= F_{x,p}(\mu) + t_{x,p}(\mu) \\ &= x\pi(A(\mu)) - c_p(x) + xA(\mu)\pi'(A(\mu)) \\ &= x\left[\pi(A(\mu)) + A(\mu)\pi'(A(\mu))\right] - c_p(x). \end{aligned} \tag{17}$$

Evolutionary implementation relies on the crucial insight that an externality adjusted game like (17) is a potential game whose potential function is the aggregate payoff function (9) (Sandholm [28, 29]).[17] In fact, this results holds irrespective of whether the original game is a potential game of not. In our case, of course, the underlying game F is a potential game. But even if it had not been, once we add the variable externality price $t_{x,p}(\mu) = e_{x,p}(\mu)$ to the original payoff, we will obtain a potential game. This leads to the following proposition, which is an application of the more general result in Lahkar and Mukherjee [19] to our model.

Proposition 4.2 *The population game $\hat{F}$ defined by (17) is a potential game with potential function $\bar{F}$ defined by (9). Hence, the efficient state of the original game F, μ^{**}, characterized in Proposition 3.7 is a Nash equilibrium of $\hat{F}$.*

Proof. For the proof that $\hat{F}$ is a potential game with potential function $\bar{F}$, see Proposition 2.4 in Lahkar and Mukherjee [19]. Since $\bar{F}$ is the potential function of $\hat{F}$, the global maximizer of $\bar{F}$ must be a Nash equilibrium of $\hat{F}$. But the global maximizer of $\bar{F}$ is the efficient state μ^{**} of the original game (Proposition 3.7). ■

Had μ^{**} been the unique Nash equilibrium of $\hat{F}$, Proposition 4.2 would have been sufficient for evolutionary implementation. Since $\hat{F}$ is a potential game, all standard evolutionary dynamics would have converged globally to μ^{**}. This is indeed the case in the models of public good, public bad and the tragedy of the commons considered in Lahkar [17, 19]. The aggregate payoff function is those models are concave along which evolutionary dynamics in the externality adjusted game ascends to converge to the unique efficient state of F, μ^{**}. This then would constitute evolutionary implementation; global convergence to the efficient state once the planner creates the externality adjusted game by imposing a variable externality price.

In our present model, however, even though μ^{**} is the unique efficient state of F, it is not the only Nash equilibrium of $\hat{F}$. To see this, denote the unique best response of a type p agent at a state μ such that $A(\mu) = \alpha$ in $\hat{F}$ as $\hat{b}_p(\alpha)$. Note from (17) that this best response is characterized by

$$\hat{b}_p(\alpha) = \begin{cases} x^* \in (0,\infty) & \text{such that } \pi(\alpha) + \alpha\pi'(\alpha) = c_p'(x^*) \text{ if } \pi(\alpha) + \alpha\pi'(\alpha) > c_p'(0) \\ 0 & \text{if } \pi(\alpha) + \alpha\pi'(\alpha) \le c_p'(0). \end{cases} \tag{18}$$

[17] This result has been extended to continuous strategy games by, for example, Cheung [2] and Lahkar and Mukherjee [19].

Due to our assumption that $\pi(\alpha) + \alpha\pi'(\alpha) < c_p'(0)$ at $\alpha = 0$ (Assumption 2.1(5)) for all $p \in \mathcal{P}$, (18) implies $\hat{b}_p(0) = 0$ for all $p \in \mathcal{P}$. Therefore, $\sum_p m_p\hat{b}_p(0) = 0$. Hence, by Proposition 3.3, $\mu^0 = (m_1\delta_0, m_2\delta_0, \cdots, m_n\delta_0)$, the state where all agents play strategy 0, is also a Nash equilibrium of $\hat{F}$. Moreover, because the inequality $\pi(\alpha) + \alpha\pi'(\alpha) \leq c_p'(0)$ holds strictly at $\alpha = 0$, there must exist, for every $p \in \mathcal{P}$, $\hat{\alpha}_p$ such that $\hat{b}_p(\alpha) = 0$ for all $\alpha \in [0, \hat{\alpha}_p]$. This property of $\hat{b}_p(\alpha)$ is akin to Observation 3.2(2). Thus, the general shape of $\hat{b}_p(\alpha)$ is similar to that of $b_p(\alpha)$ as depicted in Figure 1 with a flat stretch on $[0, \hat{\alpha}_p]$ and strictly increasing for all $\alpha > \hat{\alpha}_p$. While the upward sloping character of $b_p(\alpha)$ arose due to $\pi(\alpha)$ being strictly increasing, the upward slope of $\hat{b}_p(\alpha)$ is due to $\pi(\alpha) + \alpha\pi'(\alpha)$ being strictly increasing (as implied by Assumption 2.1, parts 2 and 4).

We, therefore, have strategic complementarities even in the externality adjusted game $\hat{F}$. Besides 0, these complementarities generate other points of intersection between $\sum_p m_p\hat{b}_p(\alpha)$ and α (akin to Figure 2) and, hence, other Nash equilibria of $\hat{F}$ besides μ^0. One such Nash equilibrium is $\mu^{**} \neq \mu^0$ (by Proposition 3.7). Moreover, because $\hat{F}$ is a potential game (Proposition 4.2), an argument akin to Proposition 3.5 implies that μ^0 is locally asymptotically stable in $\hat{F}$ under standard evolutionary dynamics.[18] We summarize this conclusion in the following proposition.

Proposition 4.3 *The state $\mu^0 = (m_1\delta_0, m_2\delta_0, \cdots, m_n\delta_0)$ is a Nash equilibrium of the externality adjusted game $\hat{F}$ defined in (17). Further, μ^0 is a locally asymptotically stable Nash equilibrium (with respect to the weak topology) under the standard evolutionary dynamics; namely the replicator dynamic (25), the BNN dynamic (26), the pairwise comparison dynamic (27), the logit dynamic (28) and the best response dynamic (29) for aggregative games.*

Proof. Follows from Propositions 3.3, 3.5 and Proposition 4.2. ■

Proposition 4.3 demonstrates the problem with evolutionary implementation in our model. Suppose the society converges to the Pareto inferior Nash equilibrium μ^0 in the original game F under some standard evolutionary dynamic. Then, because this state is also asymptotically stable in $\hat{F}$, the imposition of the variable externality price (16) will fail to move the society away from μ^0 and towards the efficient state μ^{**} of F. Evolutionary implementation will, in that case, fail to restore efficiency. Of course, this is not to say that evolutionary implementation will always fail. If the initial state happens to be in the basin of attraction of the efficient state, then the social state in $\hat{F}$ will converge to μ^{**}. But there cannot be global convergence.

Ultimately, evolutionary implementation fails due to strategic complementarities, which is manifested in the upward sloping best response functions in our model. This generates a multiplicity of Nash equilibria, including μ^0. On the other hand, in the aggregative models considered in Lahkar and Mukherjee [17, 19], best responses are non–increasing which leads to a unique Nash equilibrium. In particular, the unique Nash equilibrium of the externality adjusted game is the efficient state of the original game and is, therefore, globally attracting. We also note that although our assumption that $\pi'(\alpha) > 0$ causes externalities to be positive, positive externalities alone need not imply failure

[18] Once again, the caveat mentioned in footnote 11 applies with respect to the logit dynamic and the replicator dynamic.

of evolutionary implementation. This is true, for example, in the model of public goods in Lahkar and Mukherjee [17] where externalities are positive but efficiency is globally implementable in an evolutionary sense.[19] Instead, it is the slope of the best responses that is the key factor.[20]

5 Dominant Strategy Implementation

Evolutionary implementation does not rely on inducing truthful revelation of type by agents. This is because in designing the transfer scheme (16), the planner does not need any information about the type specific cost functions c_p. We now consider a more classical form of implementation wherein the planner seeks to implement the socially efficient state by directly inducing truthful revelation. We will find that such a mechanism, which will be variant of Pigouvian pricing, will in fact elicit truthful revelation as a strictly dominant strategy. It will, therefore, implement social efficiency in dominant strategies. Moreover, it will succeed even in a situation where evolutionary implementation will fail.

As in the last section, we continue to assume that the planner knows π but not the type distribution m. In addition, however, we now need to assume that the planner knows the possible type specific cost functions $(c_1, c_2, \cdots, c_n)$ that an agent can have. Formally, $m \mapsto \mu^{**}$ continues to be the social choice function the planner wishes to implement. For this purpose, the planner invites reports of type from agents. Based on reports, which may or may not be truthful, the planner calculates the reported type distribution $\tilde{m} = (\tilde{m}_1, \tilde{m}_2, \cdots, \tilde{m}_n)$. Thus, $\tilde{m}_p$ is the proportion of agents who report that their cost function is c_p. It is possible that $\tilde{m}_p \neq m_p$, the true proportion of type p agents.

The planner now applies the reported type distribution $\tilde{m}$ to the $\bar{G}$ function defined in (11) and maximizes it. This calculation is feasible because we have assumed that the planner knows the relevant functions $(\pi; c_1, \cdots, c_n)$. Let this maximizer be $(\tilde{\alpha}_1^{**}, \tilde{\alpha}_2^{**}, \cdots, \tilde{\alpha}_n^{**})$. Denoting $\tilde{\alpha}^{**} = \sum_{q \in \mathcal{P}} \tilde{m}_q \tilde{\alpha}_q^{**}$, we can use (12) to characterize $\tilde{\alpha}_p^{**}$ as the unique solution to

$$\pi\left(\tilde{\alpha}^{**}\right) + \tilde{\alpha}^{**} \pi'\left(\tilde{\alpha}^{**}\right) \leq c_p'(\tilde{\alpha}_p^{**}), \tag{19}$$

with strict inequality only if $\tilde{\alpha}_p^{**} = 0$. The planner then assigns to an agent who reports type to be q the action $\tilde{\alpha}_q^{**}$ and the transfer $\tilde{\alpha}_q^{**} \tilde{\alpha}^{**} \pi'(\tilde{\alpha}^{**})$. The motivation behind postulating this transfer $\tilde{\alpha}_q^{**} \tilde{\alpha}^{**} \pi'(\tilde{\alpha}^{**})$ comes from (16). The planner applies Proposition 3.7 to calculate the efficient aggregate strategy level $\tilde{\alpha}^{**}$ that would have resulted had $\tilde{m}$ been the true type distribution. By (16), the transfer $\tilde{\alpha}_q^{**} \tilde{\alpha}^{**} \pi'(\tilde{\alpha}^{**})$ then equals the externality the agent would have generated had

[19] The payoff function in Lahkar and Mukherjee [17] is $F_{x,p}(\mu) = v_p(A(\mu)) - c_p(x)$, with $v_p(\alpha)$ strictly increasing and concave. Strategy 0 is strictly dominant and, hence, μ^0 is the unique Nash equilibrium. Externality is $e_{x,p}(\mu) = x \sum_{q \in \mathcal{P}} v_q'(\alpha)$. The externality adjusted game is $\hat{F}_{x,p}(\mu) = v_p(A(\mu)) + x \sum_{q \in \mathcal{P}} v_q'(A(\mu)) - c_p(x)$. Thus, best responses in this game depend only upon $\sum_{q \in \mathcal{P}} v_q'(\alpha)$ which is strictly declining due to concavity. Hence, best responses are also declining generating efficiency as the unique Nash equilibrium.

[20] Recall from the Introduction our justification of using aggregative games on the grounds that it enables precise identification of the conditions that drives our results. An example is the condition that $\pi(\alpha) + \alpha\pi'(\alpha)$ is strictly increasing, which gives rise to strategic complementarities in $\hat{F}$.

she played her type specific efficient strategy $\tilde{\alpha}_q^{**}$ at this hypothetical efficient aggregate strategy level $\tilde{\alpha}^{**}$.

In terms of notation, the planner constructs a direct mechanism $\phi : (q, \tilde{m}) \mapsto \big(\tilde{\alpha}_q^*, \tilde{\alpha}_q^* \tilde{\alpha}^{**} \pi'(\tilde{\alpha}^*)\big)$. Thus, ϕ takes the reported type q of an agent and the reported type distribution $\tilde{m}$ and assigns the strategy and transfer $\big(\tilde{\alpha}_q^*, \tilde{\alpha}_q^* \tilde{\alpha}^{**} \pi'(\tilde{\alpha}^*)\big)$ to that agent.[21] Intuitively, ϕ is simply a Pigouvian pricing mechanism but which calculates the Pigouvian price with respect to the reported type distribution. As we had discussed in Section 4, Pigouvian price equals the externality generated by an agent at the efficient state. Here, of course, the planner does not know the efficient state ex ante but calculates the Pigouvian price at the efficient state that the reported type distribution induces. The resulting payoff of a type p agent who announces type to be q when the reported type distribution is $\tilde{m}$ in this mechanism ϕ is

$$\begin{aligned}\phi_p(q;\tilde{m}) &= \tilde{\alpha}_q^{**}\pi(\tilde{\alpha}^{**}) - c_p(\tilde{\alpha}_q^{**}) + \tilde{\alpha}_q^* \tilde{\alpha}^{**}\pi'(\tilde{\alpha}^{**}) \\ &= \tilde{\alpha}_q^{**}\left[\pi(\tilde{\alpha}^{**}) + \tilde{\alpha}^{**}\pi'(\tilde{\alpha}^{**})\right] - c_p(\tilde{\alpha}_q^{**}). \end{aligned} \tag{20}$$

This is the externality adjusted payoff (17) calculated at the strategy $\tilde{\alpha}_q^{**}$ and transfer $\tilde{\alpha}_q^* \tilde{\alpha}^{**} \pi'(\tilde{\alpha}^*)$ assigned to this type p agent. We now show that this mechanism renders truthful revelation of type incentive compatible in strictly dominant strategies and, therefore, implements the efficient state of F. The proof of the result relies on the fact that each agent is of measure zero due to which individual announcements cannot affect aggregate variables.

Proposition 5.1 *For any type distribution m, the direct mechanism ϕ described by (20) implements the efficient state μ^{**} of F as characterized in Proposition 3.7. Further, ϕ satisfies incentive compatibility in strictly dominant strategies.*

Proof. We need to show that truthful revelation of type is strictly dominant for an agent in the direct mechanism ϕ. Due to the large population nature of the problem, any individual agent's action cannot have any impact on the aggregate strategy level $\tilde{\alpha}^{**}$ in (20). Therefore, the problem for a type p agent in (20) can be equivalently stated as

$$\max_{x\in[0,\infty)} x\left[\pi(\tilde{\alpha}^{**}) + \tilde{\alpha}^{**}\pi'(\tilde{\alpha}^{**})\right] - c_p(x). \tag{21}$$

If we show that the unique maximizer of (21) is $\tilde{\alpha}_p^{**}$, then that would imply that truthfully revealing type to be p for a type p agent is the strictly dominant strategy in ϕ. Clearly, this maximizer x^{**} is characterized by $\pi(\tilde{\alpha}^{**}) + \tilde{\alpha}^{**}\pi'(\tilde{\alpha}^{**}) \le c_p'(x^{**})$, with strict inequality only if $x^{**} = 0$. But by (19), the unique solution to this equation is $\tilde{\alpha}_p^{**}$. Hence, all agents report type truthfully as a strictly dominant strategy so that $\tilde{m} = m$ and $\tilde{\alpha}_p^{**} = \alpha_p^{**}$ as characterized in Proposition 3.7. ■

Formally, (21) is the externality adjusted payoff (17) at any social state at which the aggregate strategy level is $\tilde{\alpha}^{**}$. The unique best response (18) of a type p agent is then $\hat{b}_p(\tilde{\alpha}^{**}) = \tilde{\alpha}_p^{**}$.

[21] We can, equivalently, also view $(q, \tilde{m}) \mapsto \big(\tilde{\alpha}_q^*, \tilde{\alpha}_q^* \tilde{\alpha}^{**} \pi'(\tilde{\alpha}^*)\big)$ as the output function of the mechanism ϕ.

Therefore, every agent reports type truthfully as a strictly dominant strategy so that the reported type distribution is the actual distribution m. Therefore, every agent of every type p is assigned the socially efficient strategy α_p^{**} as characterized in Proposition 3.7. The resulting social state is the efficient state μ^{**}. Once implemented, and using the notation in (20), an agent of type p ends up receiving the payoff

$$\phi_p(p;m) = \alpha_p^{**}\left[\pi(\alpha^{**}) + \alpha^{**}\pi'(\alpha^{**})\right] - c_p(\alpha_p^{**}). \tag{22}$$

where α_p^{**} is the efficient strategy level as characterized in Proposition 3.7 and $\alpha^{**} = \sum_{q\in\mathcal{P}} m_q\alpha_q^{**}$. The transfer that each agent ends up receiving is the true Pigouvian price calculated at the true efficient state.[22]

This is dominant strategy implementation. Notice that it succeeds in implementing social efficiency irrespective of the initial social state. Thus, as we have noted, if the society is initially near the Pareto inferior Nash equilibrium μ^0, then evolutionary implementation will fail in achieving efficiency. But dominant strategy implementation succeeds even in this situation. This is because unlike evolutionary implementation, it does not seek a gradual transition towards the efficient state. Instead, the planner consciously creates incentives for truthful revelation thereby achieving instantaneous coordination on the efficient state. This renders the initial state irrelevant. Similar to evolutionary implementation, dominant strategy implementation also does not impose any requirement upon the planner to know the type of agents or even the type distribution. Agents voluntarily disclose type and, hence, the type distribution in equilibrium. Unlike in evolutionary implementation where the planner needs to continuously update the variable transfer price, dominant strategy implementation only requires the planner to intervene once in inviting reports about type.

But in one respect, dominant strategy implementation does have a greater informational requirement. In addition to knowing the aggregative benefit function π, it also needs the planner to know the possible cost functions $\{c_1, c_2, \cdots, c_n\}$ as without such knowledge, the planner will not be able to calculate $(\tilde{\alpha}_1^{**}, \tilde{\alpha}_2^{**}, \cdots, \tilde{\alpha}_n^{**})$. But evolutionary implementation only required the planner to know π and observe x and $A(\mu)$. The type specific information about cost functions was not needed. In that respect, evolutionary implementation is informationally more parsimonious than dominant strategy implementation.

Along with incentive compatibility, it is also desirable that a mechanism satisfy individual rationality and strict budget balance. Individual rationality implies that all agents receive a non–negative equilibrium payoff so that no one has to be coerced to participate in the mechanism. Strict budget balance means the planner is left with neither a surplus nor a deficit once the transfers have been made. It is obvious that the mechanism ϕ satisfies individual rationality. With π being strictly increasing, it is clear from (22) that every agent receives a subsidy $\alpha_p^{**}\pi'(\alpha^{**})$ in the dominant strategy equilibrium of ϕ. But then, it is also equally obvious that strict budget balance will not

[22] As in footnote 20, the aggregative structure of our model enables us to identify precisely the condition behind dominant strategy implementation. This condition is that (19) has a unique solution, which arises directly from the fact that we have modelled our problem as an aggregative game.

be satisfied. With every agent being given a subsidy, the planner is left with a deficit which is a more serious problem than that of a surplus.

There is, of course, an easy way to restore budget balance. Recall from (20) that if an agent reports type to be q and the reported type distribution is $\tilde{m}$, the agent receives transfer $\tilde{\alpha}_q{}^{**}\tilde{\alpha}^{**}\pi'(\tilde{\alpha}^{**})$. Therefore, the total transfer made when the reported type distribution is $\tilde{m}$ is $\tilde{\alpha}^{**}\pi'(\tilde{\alpha}^{**})\sum_{p\in\mathcal{P}} m_p\tilde{\alpha}_p^{**} = \left(\tilde{\alpha}^{**}\right)^2\pi'(\tilde{\alpha}^{**})$. We, therefore, introduce a new mechanism

$$\phi^B : (q, \tilde{m}) \to \left(\tilde{\alpha}_q^{**}, \tilde{\alpha}_q{}^{**}\tilde{\alpha}^{**}\pi'(\tilde{\alpha}^{**}) - \left(\tilde{\alpha}^{**}\right)^2\pi'(\tilde{\alpha}^{**})\right). \tag{23}$$

Thus, this new mechanism takes reported type and reported type distribution as inputs and assigns the same action $\tilde{\alpha}_q^{**}$ as the earlier mechanism ϕ but a different transfer $\tilde{\alpha}_q{}^{**}\tilde{\alpha}^{**}\pi'(\tilde{\alpha}^{**}) - \left(\tilde{\alpha}^{**}\right)^2\pi'(\tilde{\alpha}^{**})$, which differs from the earlier transfer only by the constant $\left(\tilde{\alpha}^{**}\right)^2\pi'(\tilde{\alpha}^{**})$.

Clearly, ϕ^B is strategically equivalent to ϕ and, therefore, by Proposition 5.1, satisfies incentive compatibility in strictly dominant strategies. Unlike ϕ, ϕ^B satisfies budget balance because by construction of the constant term $\left(\tilde{\alpha}^{**}\right)^2\pi'(\tilde{\alpha}^{**})$, the total transfer made in this mechanism is always zero irrespective of the reported type distribution. However, ϕ^B will not satisfy individual rationality. It may be possible that agents whose cost of effort is particularly high receive a strictly negative payoff in equilibrium. We show this using a counterexample in Appendix A.2.

Having said this, when we compare dominant strategy implementation with evolutionary implementation, this failure of individual rationality should not be construed as a weakness of only dominant strategy implementation. Had we replaced the variable externality price $t_{x,p}(\mu) = xA(\mu)\pi'(A(\mu))$ with, as in (23), the strategically equivalent transfer $xA(\mu)\pi'(A(\mu)) - A(\mu)^2\pi'(A(\mu))$, evolutionary implementation too would have failed individual rationality at the efficient state upon convergence to that state.

The approach of this section has followed that of Lahkar and Mukherjee [18] which applied dominant strategy implementation to a large population public goods game. This section shows that this method is more generally applicable. The approach has clear similarities to the classical VCG mechanism in its reliance on Pigouvian pricing. Our main result is, however, stronger as we obtain truthful revelation to be strictly dominant whereas it is only weakly dominant in the VCG mechanism. There is, however, one key difference between the results of this section and those of Lahkar and Mukherjee [18]. In the earlier paper, the mechanism akin to ϕ^B designed for budget balance also satisfies individual rationality. Here, however, it is possible that ϕ^B may fail individual rationality. In that respect, our conclusions are similar to the classical dAGV mechanism (Arrow [1], d'Aspremont and Gérard–Varet [8]) which also may fail individual rationality. But the dAGV mechanism only implements truthful revelation as a Bayesian Nash equilibrium whereas we obtain the stronger form of implementation in strictly dominant strategies. Of course, as we have shown with the mechanism ϕ, we can construct a dominant strategy mechanism which is individually rational. But in that case, we need to forgo budget balance. We note that the incompatibility among these axioms is consistent with classical results in the mechanism design literature (Green and Laffont [10]).[23]

6 Conclusion

We have considered two alternative means of implementation in large population games in this paper, evolutionary implementation and dominant strategy implementation. Our focus has been on an aggregative game that satisfies strategic complementarities, i.e. best responses are increasing with respect to the aggregate strategy level. This creates multiple Nash equilibria. For evolutionary implementation, the planner creates a new game by adding the positive externalities an agent generates to the original payoff of the agent. This externality adjusted game is a potential game. Hence, evolutionary dynamics converge to Nash equilibria. The original efficient state is a Nash equilibrium of the externality adjusted game. But so are other Pareto inferior states like the one where all agents play the zero strategy. Hence, instead of converging to the efficient state, evolutionary implementation may converge to some other Pareto inferior state. Thus, evolutionary implementation may fail. In contrast, dominant strategy implementation succeeds because the planner deliberately creates incentive for truthful revelation of types, thereby ensuring instantaneous coordination on the efficient state. The paper, therefore, has provided a new approach towards implementing social efficiency in large population games in the presence of multiple equilibria.

Our exercise raises certain interesting research questions. One is on the scope of dominant strategy implementation in large population models. The present analysis as well as that in Lahkar and Mukherjee [18] on this topic has been confined to aggregative games with a continuous strategy set. These features play a crucial role in our main result that truthful revelation is a strictly dominant strategy (Proposition 5.1). This result arises from the uniqueness of best response in our model for which, we require both the aggregative structure of the model and the continuous strategy set. If we relax these features and consider non-aggregative population games or population games with a finite strategy set, we may not be able to obtain strict dominance. Weak dominance of truth telling may be the best we can hope for.

The second question is the reverse of the one we have considered here. Whether there are situations where dominant strategy implementation fails but evolutionary implementation succeeds. It may not be possible to provide a broad answer to this question unless we generalize dominant strategy implementation to all large population games. But at least in aggregative games with a continuous strategy set, the answer will be no as long as there is a unique best response to every social state. In such cases, through an argument similar to Proposition 5.1, truthful revelation will be strictly dominant.

[23] Green and Laffont [10] have shown that in finite player models there is no budget-balanced mechanism which can implement the efficient outcome in dominant strategy. Our result, in fact, shows that we can overcome this impossibility by considering a model of large population. But this may be at the cost of sacrificing individual rationality.

A Appendix

For the proof of Proposition 3.5, we introduce the function $g : \prod_{p\in\mathcal{P}}[0,\infty) \to \mathbf{R}$ defined as

$$g(\alpha_1, \alpha_2, \cdots, \alpha_n) = \int_0^{\sum_{p\in\mathcal{P}} m_p\alpha_p} \pi(z)dz - \sum_{p\in\mathcal{P}} m_p c_p(\alpha_p). \tag{24}$$

The following lemma establishes two important properties of g which establishes its relationship to the potential function (3). Due to this relationship, we interpret g as a finite dimensional analogue of the potential function.[24]

Lemma A.1 *Consider the potential function f defined by (3) and its finite dimensional analogue g defined by (24).*

1. *Suppose $\mu = (m_1\delta_{\alpha_1}, m_2\delta_{\alpha_2}, \cdots, m_n\delta_{\alpha_n})$. Then $f(\mu) = g(\alpha_1, \alpha_2, \cdots, \alpha_n)$.*
2. *Suppose for any $p \in \mathcal{P}$, μ_p is polymorphic with $\int_{\mathcal{S}} x\mu_p(dx) = m_p\alpha_p$. Then, $g(\alpha_1, \alpha_2, \cdots, \alpha_n) > f(\mu)$.*

Proof.

1. If $\mu = (m_1\delta_{\alpha_1}, m_2\delta_{\alpha_2}, \cdots, m_n\delta_{\alpha_n})$, then $A(\mu) = \sum_p m_p\alpha_p$ and $\int_{\mathcal{S}} c_p(x)\mu_p(dx) = m_p c_p(\alpha_p)$. Therefore, $f(\mu) = g(\alpha_1, \alpha_2, \cdots, \alpha_n)$.
2. If μ_p is polymorphic with $\int_{\mathcal{S}} x\mu_p(dx) = m_p\alpha_p$, then $A(\mu) = \sum_p m_p\alpha_p$. Hence, $\int_0^{A(\mu)} \pi(z)dz = \int_0^{\sum_{p\in\mathcal{P}} m_p\alpha_p} \pi(z)dz$. On the other hand, due to the strict convexity of c_p, $m_p c_p(\alpha_p) < \int_{\mathcal{S}} c_p(x)\mu_p(dx)$. Therefore, $g(\alpha_1, \alpha_2, \cdots, \alpha_n) > f(\mu)$. ■

Proof of Proposition 3.5: We know from Lemma A.1(1) that $f(\mu^0) = g(0, 0, \cdots, 0)$, where f and g are as defined in (3) and (24) respectively. We first show g is strictly declining at $(0, 0, \cdots, 0)$. For this, denote $\alpha = \sum_p m_p\alpha_p$ and note from (24) that $\frac{\partial g}{\partial \alpha_p} = m_p\big[\pi(\alpha) - c_p'(\alpha_p)\big]$. Therefore, at $(0, 0, \cdots, 0)$, $\frac{\partial g}{\partial \alpha_p} = m_p\big[\pi(0) - c_p'(0)\big] < 0$ for all $p \in \mathcal{P}$ by (4).

Now, consider $\tilde{\mu}$ close to μ^0 under the weak topology (topology induced by convergence in distribution) so that $A(\tilde{\mu}) \approx 0$. First, assume $\tilde{\mu}$ consists of monomorphic population states. Thus, every agent in population p plays $\tilde{\alpha}_p$ such that $\sum_p m_p\tilde{\alpha}_p = A(\tilde{\mu}) = \tilde{\alpha}$. For $\tilde{\mu}$ sufficiently close to μ^0, each $\tilde{\alpha}_p$ will also be sufficiently close to 0. Therefore, the fact that $\frac{\partial g}{\partial \alpha_p} < 0$ at $(0, 0, \cdots, 0)$ for all p implies $g(0, 0, \cdots, 0) > g(\tilde{\alpha}_1, \tilde{\alpha}_2, \cdots, \tilde{\alpha}_n)$. By Lemma A.1(1), this then implies $f(\mu^0) = g(0, 0, \cdots, 0) > g(\tilde{\alpha}_1, \tilde{\alpha}_2, \cdots, \tilde{\alpha}_n) = f(\tilde{\mu})$.

Next, suppose $\tilde{\mu}$ is such that at least one population state is polymorphic. Suppose in all monomorphic states, players play $\tilde{\alpha}_p$ while for all polymorphic states, consider $\tilde{\alpha}_p = \frac{\int_{\mathcal{S}} x\tilde{\mu}_p(dx)}{m_p}$. For $\tilde{\mu}$ sufficiently close to μ^0, all such $\tilde{\alpha}_p$ will also be sufficiently close to 0. Therefore, by the argument

[24] Such a function is called a quasi–potential function in Lahkar [14] and Cheung and Lahkar [4]. The function (24) is a more general version as it is defined for multiple populations.

in the earlier paragraph, $g(0,0,\cdots,0) > g(\tilde{\alpha}_1,\tilde{\alpha}_2,\cdots,\tilde{\alpha}_n)$. By Lemma A.1(2), this then implies $f(\mu^0) = g(0,0,\cdots,0) > g(\tilde{\alpha}_1,\tilde{\alpha}_2,\cdots,\tilde{\alpha}_n) > f(\tilde{\mu})$.

Hence, μ^0 is a local maximizer of f. The conclusion about the local stability of μ^0 under the standard evolutionary dynamics then follows from the papers on the dynamics cited earlier. ■

Proof of Proposition 3.7: Suppose μ_q^{**} is not monomorphic for some $q \in \mathcal{P}$. Define $a(\mu_q) = \int_{\mathcal{S}} x\mu_q(dx)$ as the aggregate strategy level in population q. Let $a(\mu_q^{**}) = \alpha_q$. Consider a new social state $\hat{\mu}$ such that $\hat{\mu}_p = \mu_p^{**}$ for all $p \neq q$ and $\hat{\mu}_q = m_q\delta_{\frac{\alpha_q}{m_q}}$. Thus, in constructing $\hat{\mu}$, we have replaced the non–monomorphic population state μ_q^{**} with a monomorphic state with the same population level aggregate strategy.

Consider $\bar{F}(\mu) = A(\mu)\pi(A(\mu)) - C(\mu)$ as derived in (10). We show that $\bar{F}(\hat{\mu}) > \bar{F}(\mu^{**})$, thereby obtaining a contradiction. Since $A(\mu^{**}) = A(\hat{\mu})$, clearly $A(\mu^{**})\pi(A(\mu^{**})) = A(\hat{\mu})\pi(A(\hat{\mu}))$. The desired result will be established if we show $C(\hat{\mu}) < C(\hat{\mu}^{**})$, i.e.

$$\sum_{p\in\mathcal{P}} \int_{\mathcal{S}} c_p(x)\hat{\mu}_p(dx) < \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} c_p(x)\mu_p^{**}(dx).$$

Again, because $\hat{\mu}_p = \mu_p^{**}$ for all $p \neq q$, this inequality will hold if $\int_{\mathcal{S}} c_q(x)\hat{\mu}_q(dx) < \int_{\mathcal{S}} c_q(x)\hat{\mu}_q(dx)$. But this follows from the strict convexity of c. Thus, if μ^{**} is an efficient state of F, it must be in monomorphic population states.

For the second part of the result, suppose $\mu = (m_1\delta_{\alpha_1}, m_2\delta_{\alpha_2},\cdots,m_n\delta_{\alpha_n})$ is a social state in monomorphic social states. Then, $A(\mu) = \sum_{p\in\mathcal{P}}\int_{\mathcal{S}} x\mu_p(dx) = \sum_{p\in\mathcal{P}}\int_{\mathcal{S}} xm_p\delta_{\alpha_p}(dx) = \sum_{p\in\mathcal{P}} m_p\alpha_p$. Further, $C(\mu) = \sum_{p\in\mathcal{P}}\int_{\mathcal{S}} c_p(x)\mu_p(dx) = \sum_{p\in\mathcal{P}}\int_{\mathcal{S}} c_p(x)m_p\delta_{\alpha_p}(dx) = \sum_{p\in\mathcal{P}} m_pc_p(\alpha_p)$. Comparing (10) and (11), we then conclude that at all such social states of the form

$$\mu = (m_1\delta_{\alpha_1}, m_2\delta_{\alpha_2},\cdots,m_n\delta_{\alpha_n}),$$

$\bar{F}(\mu) = \bar{G}(\alpha_1,\alpha_2,\cdots,\alpha_n)$.

It, therefore, follows that it suffices to focus on $\bar{G}$ to characterize the efficient state of F. If $(\alpha_1^{**},\alpha_2^{**},\cdots,\alpha_n^{**})$ is the global maximizer of $\bar{G}$, then $\mu^{**} = \left(m_1\delta_{\alpha_1^{**}}, m_2\delta_{\alpha_2^{**}},\cdots,m_n\delta_{\alpha_n^{**}}\right)$ must be the unique efficient state of F. Uniqueness of the global maximizer of $\bar{G}$ follows from the strict convexity of the cost functions c_p. This global maximizer is characterized by a solution to (12). Any such solution must be finite due to our assumptions that the third derivative of c_p is non–negative for all $p \in \mathcal{P}$ and limit of $\pi'(\cdot)$ being zero at infinity (Assumptions 2.1(2) and (3)). ■

A.1 Standard Evolutionary Dynamics

Consider a population game F with strategy set $\mathcal{S}$ and let the payoff of a population p agent playing strategy x at social state μ be $F_{x,p}(\mu)$. Denote the average payoff in population p μ as $\bar{F}_p(\mu) = \frac{1}{m_p}\int_{\mathcal{S}} F_{x,p}(\mu)\mu_p(dx)$ and define the excess payoff of a strategy x in population p as

$F_{x,p}(\mu) - \bar{F}_p(\mu)$. This notion of the excess payoff will be required in the definitions of the replicator dynamic and the BNN dynamic.

In addition, for the logit dynamic, define the probability measure $L_{\eta,p}(\mu)$ on $\mathcal{S}$, known as the logit choice measure, as $L_{\eta,p}(\mu)(B) = \int_B \frac{\exp(\eta^{-1}F_{x,p}(\mu))}{\int_{\mathcal{S}} \exp(\eta^{-1}F_{y,p}(\mu))dy} dx$, $B \subseteq \mathcal{S}$, $\eta > 0$. Here, η is a perturbation parameter. Intuitively, $L_{\eta,p}(\mu)$ is an approximation of the best response in the sense that for η small, it puts most of the probability mass on the set of best responses to μ. It is generated when agents best respond to a perturbed version of payoffs, where the perturbation depends upon η (Lahkar and Riedel [15]).

We now define the replicator dynamic, the BNN dynamic, the pairwise comparison dynamic and the logit dynamic respectively in F as follows.

$$\dot{\mu}_p(B) = \int_B \left(F_{x,p}(\mu) - \bar{F}_p(\mu)\right) \mu_p(dx), \tag{25}$$

$$\dot{\mu}_p(B) = m_p \int_B \left[F_{x,p}(\mu) - \bar{F}_p(\mu)\right]_+ dx - \mu_p(B) \int_{\mathcal{S}} \left[F_{y,p}(\mu) - \bar{F}_p(\mu)\right]_+ dy, \tag{26}$$

$$\dot{\mu}_p(B) = \int_{\mathcal{S}} \int_B \left[F_{x,p}(\mu) - F_{y,p}(\mu)\right]_+ dx\mu_p(dy) - \int_{\mathcal{S}} \int_B \left[F_{y,p}(\mu) - F_{x,p}(\mu)\right]_+ \mu_p(dx)dy, \tag{27}$$

$$\dot{\mu}_p(B) = m_p L_{\eta,p}(\mu)(B) - \mu_p(B), \text{ where } \eta > 0. \tag{28}$$

In each case, $\dot{\mu}_p(B)$ is the direction and magnitude of change in the mass of agents in population p who are playing strategies in $B \subseteq \mathcal{S}$. Under the replicator dynamic (25), the mass of agents playing strategies in B increases if the aggregate excess payoff of such strategies is positive. The BNN dynamic (26) involves agents adopting strategy x with probability proportional to the positive part of the excess payoff $F_{x,p}(\mu) - \bar{F}_p(\mu)$ of that strategy (note that $[a-b]_+ = \max(a-b, 0)$). Under the pairwise comparison dynamic, agents abandon strategy y and adopt strategy x with probability proportional to $[F_{x,p}(\mu) - F_{y,p}(\mu)]_+$. In the logit dynamic, the social state μ moves towards the logit choice measure $L_{\eta,p}(\mu)$.

The four dynamics (25)–(28) are well defined for all population games. The best response dynamic for aggregative games, as the name suggests, is valid only in such aggregative games where every social state μ generates a unique best response. In such a game, denote the aggregate strategy level $A(\mu)$ at μ as α. The best response dynamic is then a ODE in which the direction and magnitude of change in the population state μ is

$$\dot{\mu}_p = m_p \delta_{b_p(\alpha)} - \mu_p. \tag{29}$$

In more general games, the best response may not be uniquely defined or, due to the continuous structure of the strategy set, may not even exist. This would create technical difficulties in defining the dynamic.

A.2 Counterexample

We present an example that shows that the mechanism ϕ^B defined in (23) may violate individual rationality. The example will also illustrate the other results we have derived throughout the paper. Suppose there are two types of agents. Their respective cost functions are

$$\begin{aligned} c_1(x) &= 2x + 2x^2 \\ c_2(x) &= 8x + 6x^2. \end{aligned} \tag{30}$$

Notice $c_1'(0) = 2 > 0$ and $c_2'(0) = 8 > 0$. Let the population masses be $m_1 = \frac{1}{10}$ and $m_2 = \frac{9}{10}$.

We assume that the aggregative benefit function is $\pi(\alpha) = 12\sqrt{\alpha}$. Notice $\alpha\pi'(\alpha) = 6\sqrt{\alpha}$ which is strictly increasing in α and equals 0 at $\alpha = 0$. The payoffs of the two types are

$$\begin{aligned} F_{x,1}(\mu) &= 12x\sqrt{A(\mu)} - (2x + 2x^2) \\ F_{x,2}(\mu) &= 12x\sqrt{A(\mu)} - (8x + 6x^2). \end{aligned} \tag{31}$$

Applying (15), we obtain the positive externality in this example to be

$$e_{x,p}(\mu) = 6x\sqrt{A(\mu)}. \tag{32}$$

Using (4) and writing $A(\mu) = \alpha$, we obtain the best responses of the two types to be respectively

$$\begin{aligned} b_1(\alpha) &= \max\left\{0, \frac{1}{2}(-1 + 6\sqrt{\alpha})\right\} \\ b_2(\alpha) &= \max\left\{0, \frac{1}{3}(-2 + 3\sqrt{\alpha})\right\}. \end{aligned} \tag{33}$$

We apply Proposition 3.3 and solve $\sum_p m_p b_p(\alpha) = \alpha$. Using (33), we can show that this equation has a unique solution, $\alpha^0 = 0$. Therefore, this example has a unique Nash equilibrium μ^0 in which every agent plays strategy 0. Moreover, due to the potential game property, this unique equilibrium must be globally asymptotically stable under all the standard dynamics mentioned in Proposition 3.5.

To find the efficient state of the example, we apply Proposition 3.7. Thus, maximizing the associated $\bar{G}$ function in (11), we obtain the efficient strategy levels of the two types at the efficient state μ^{**} to be

$$\alpha_1^{**} = 5.35, \quad \alpha_2^{**} = 1.2833. \tag{34}$$

The associated aggregate strategy level at the efficient state is $\alpha^{**} = \sum_p m_p \alpha_p^{**} = 1.69$.

Adding $t_{x,p}(\mu) = e_{x,p}(\mu) = 6x\sqrt{A(\mu)}$ to the original payoffs (31), we obtain the externality adjusted payoffs

$$\hat{F}_{x,1}(\mu) = 18x\sqrt{A(\mu)} - (2x + 2x^2)$$

$$\hat{F}_{x,2}(\mu) = 12x\sqrt{A(\mu)} - (8x + 6x^2). \tag{35}$$

The best responses in this game are

$$\begin{aligned}\hat{b}_1(\alpha) &= \max\left\{0, \frac{1}{2}(-1 + 9\sqrt{\alpha})\right\} \\ \hat{b}_2(\alpha) &= \max\left\{0, \frac{1}{6}(-4 + 9\sqrt{\alpha})\right\}.\end{aligned} \tag{36}$$

Applying Proposition 3.3 to $\hat{F}$, we can calculate its Nash equilibria. There are five such equilibria; μ^0, μ^1, μ^2, μ^3 and μ^4 with associated aggregate strategy levels

$$\alpha^0 = 0, \quad \alpha^1 = \frac{1}{25}, \quad \alpha^2 = \frac{1}{16}, \quad \alpha^3 = \frac{1}{4}, \quad \alpha^4 = \frac{169}{100}. \tag{37}$$

Two of these equilibria are of particular interest; μ^0 where every agent plays 0 and μ^4. The latter is, in fact, the efficient state of the original game as can be seen by applying α^4 to (36). The type specific strategy levels of the two types at this equilibrium of $\hat{F}$ turn out to the efficient strategy levels (34) in the original game. Thus, $\mu^4 = \mu^{**}$.[25]

By Proposition 4.3, μ^0 is locally asymptotically stable in $\hat{F}$ under standard evolutionary dynamics.[26] Therefore, evolutionary implementation is ineffective in implementing efficiency in this example. In the original game, the society globally converges to μ^0. Hence, if the planner applies variable externality pricing, the local asymptotic stability of μ^0 in $\hat{F}$ means the society will not be able to escape this Pareto inefficient state.

Dominant strategy implementation will, however, succeed in establishing social efficiency. Truthful revelation becomes strictly dominant in the direct mechanism ϕ defined by (20) or, equivalently, in the mechanism ϕ^B defined by (23). Moreover, ϕ^B satisfies strict budget balance. But, it will not satisfy individual rationality. To see this, note that because there is truthful revelation (so that reported distribution is m itself) and the efficient state gets implemented, equilibrium payoffs in ϕ^B for types 1 and 2 are (recall $\pi(\alpha) = 12\sqrt{\alpha}$)

$$\begin{aligned}\phi^B(1; m) &= \alpha_1^{**}\left[\pi(\alpha^{**}) + \alpha^{**}\pi'(\alpha^{**})\right] - (\alpha^{**})^2\pi'(\alpha^{**}) - c_1(\alpha_1^{**}) \\ &= 18\alpha_1^{**}\sqrt{\alpha^{**}} - 6\alpha^{**}\sqrt{\alpha^{**}} - c_1(\alpha_1^{**}). \\ \phi^B(2; m) &= \alpha_2^{**}\left[\pi(\alpha^{**}) + \alpha^{**}\pi'(\alpha^{**})\right] - (\alpha^{**})^2\pi'(\alpha^{**}) - c_2(\alpha_2^{**}). \\ &= 18\alpha_2^{**}\sqrt{\alpha^{**}} - 6\alpha^{**}\sqrt{\alpha^{**}} - c_2(\alpha_2^{**}).\end{aligned} \tag{38}$$

Applying (30) and (34) to (38), we obtain $\phi^B(1; m) = 44.0629$ and $\phi^B(2; m) = -3.3003$. Thus, individual rationality is violated for type 2.

[25] Indeed, $\alpha^4 = 1.69$, the efficient aggregate strategy level $\sum_p m_p \alpha_p^{**}$ in the original game.

[26] In addition, at least under the BR dynamic (29), we can rigorously argue that μ^2 and μ^4 are also asymptotically stable.

References

[1] Arrow K (1979) The Property Rights Doctrine and Demand Revelation Under Incomplete Information. in Economics and Human Welfare, ed. by M. Boskin. New York: Academic Press.

[2] Cheung MW (2014) Pairwise comparison dynamics for games with continuous strategy space. J. Econ. Theory 153:344–375.

[3] Cheung MW (2016) Imitative Dynamics for Games with Continuous Strategy Space. Games Econ. Behav. 99:206–223.

[4] Cheung MW, Lahkar R (2018) Nonatomic Potential Games: the Continuous Strategy Case. Games Econ. Behav. 108:341–362.

[5] Clarke E (1971) Multi-part pricing of public goods. Public Choice 11:17–23.

[6] Cooper RJ (1999) Coordination Games: Complementarities and Macroeconomics, Cambridge University Press, Cambridge.

[7] Corchón L (1994) Comparative statics for aggregative games the strong concavity case. Math. Soc. Sci. 28:151–165.

[8] d'Aspremont C, Gérard-Varet L–A (1979) Incentives and incomplete information. J. Public Econ. 11: 25-45.

[9] Ellison G (1993) Learning, local interaction, and coordination. Econometrica 61:1047-–1071.

[10] Green J, Laffont J-J (1979) Incentives in public decision-making. Elsevier North-Holland

[11] Groves T (1973) Incentives in teams. Econometrica 41:617–631.

[12] Hofbauer J, Oechssler J, Riedel F (2009) Brown–von Neumann–Nash dynamics: the continuous strategy case. Games Econ. Behav. 65:406–429.

[13] Katz ML, Shapiro C (1986) Technology Adoption in the Presence of Network Externalities. J. Polit. Econ. 94:822–841

[14] Lahkar R (2017) Large population aggregative potential games. Dyn Games Appl. 7:443–467.

[15] Lahkar R, Riedel F (2015) The logit dynamic for games with continuous strategy sets. Games Econ. Behav. 91:268–282.

[16] Lahkar R (2019) Elimination of non-individualistic preferences in large population aggregative games. J. Math. Econ. 84:150–165.

[17] Lahkar R, Mukherjee S (2019) Evolutionary Implementation in a Public Goods Game. J. Econ. Theory 181:423–460.

[18] Lahkar R, Mukherjee S (2020) Dominant Strategy Implementation in a Large Population Public Goods Game. Econ. Lett. 197:109616.

[19] Lahkar R, Mukherjee S (2021) Evolutionary Implementation in Aggregative Games. Math. Soc. Sci. 109:137–151.

[20] McMahon WW (1999) Education and Development: Measuring the Social Benefits. Oxford University Press, New York.

[21] Monderer D, Shapley L (1996) Potential games. Games Econ. Behav. 14:124–143.

[22] Oechssler J, Riedel F (2001) Evolutionary dynamics on infinite strategy spaces. Econ. Theory 17:141–162.

[23] Oechssler J, Riedel F (2002) On the dynamic foundation of evolutionary stability in continuous models. J. Econ. Theory 107:223–252.

[24] Perkins S, Leslie D (2014) Stochastic fictitious play with continuous action sets. J. Econ. Theory 152:179–213.

[25] Rothkopf MH, Teisberg TJ, Kahn EP (1990) Why are Vickrey auctions rare? J. Political Econ. 98:94–109.

[26] Rothkopf MH (2007) Thirteen reasons why the Vickrey-Clarke-Groves process is not practical. Oper. Res. 55:191–197.

[27] Sandholm WH (2001) Potential games with continuous player sets. J. Econ. Theory 97:81–108.

[28] Sandholm WH (2002) Evolutionary Implementation and Congestion Pricing. Rev. Econ. Stud 69:667–689.

[29] Sandholm WH (2005) Negative Externalities and Evolutionary Implementation. Rev. Econ. Stud 72:885–915.

[30] Sandholm WH (2007) Pigouvian Pricing and Stochastic Evolutionary Implementation. J. Econ. Theory 132:367–382.

[31] Vickrey W (1961) Counterspeculation, auctions, and competitive sealed tenders. J. Finance 16:8–37.