

DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES  
**CALIFORNIA INSTITUTE OF TECHNOLOGY**

PASADENA, CALIFORNIA 91125

Voting and Lottery Drafts as Efficient Public Goods Mechanisms

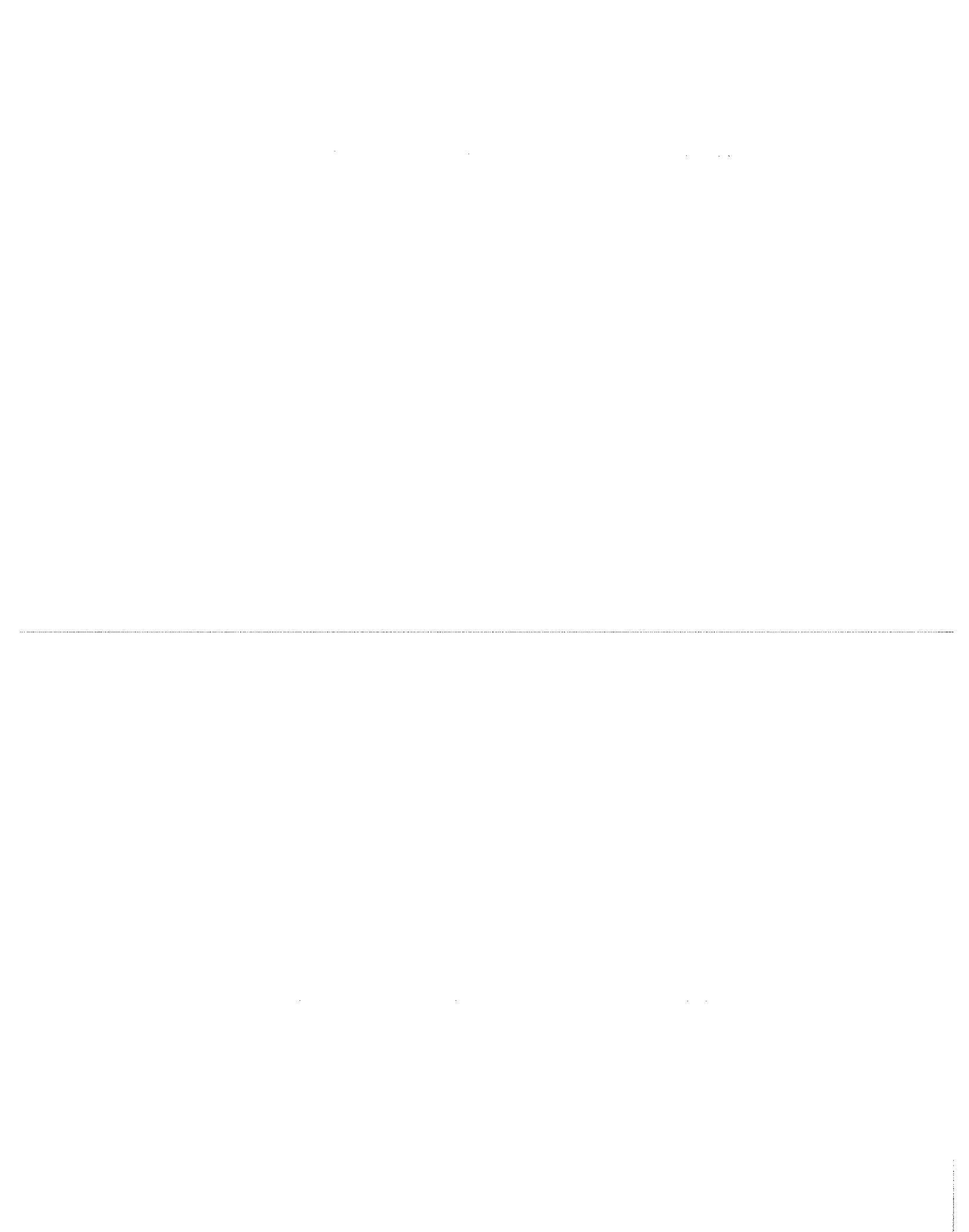
John O. Ledyard

Thomas R. Palfrey



**SOCIAL SCIENCE WORKING PAPER 717**

March 1990  
Revised September 1993



# Voting and Lottery Drafts as Efficient Public Goods Mechanisms

John O. Ledyard

Thomas R. Palfrey

## Abstract

This paper characterizes interim efficient mechanisms for public good production and cost allocation in a two-type environment with risk neutral, quasi-linear preferences and fixed size projects, where the distribution of the private good, as well as the public goods decision, affects social welfare. An efficient public good decision can always be accomplished by a majority voting scheme, where the number of “YES” votes required depends on the welfare weights in a simple way. The results are shown to have a natural geometry and an intuitive interpretation. We also extend these results to allow for restrictions on feasible transfer rules, ranging from the traditional unlimited transfers to the extreme case of no transfers.

For a range of welfare weights, an optimal scheme is a two-stage procedure which combines a voting stage with a second stage where an even-chance lottery is used to determine who pays. We call this the “lottery draft mechanism”. Since such a cost-sharing scheme does not require transfers, it follows that in many cases transfers are not necessary to achieve the optimal allocation. For other ranges of welfare weights the second stage is more complicated, but the voting stage remains the same. If transfers are completely infeasible, randomized voting rules may be optimal. The paper also provides a geometric characterization of the effects of voluntary participation constraints.

JEL Classification: 024, 026



# Voting and Lottery Drafts as Efficient Public Goods Mechanisms

John O. Ledyard and Thomas R. Palfrey\*

## 1 Introduction

Deciding whether or not to undertake and how to distribute the cost of a public project of fixed size is a classic problem in economics. This problem and variations on it have been at the center of attention of much of the work on optimal mechanism design under conditions of asymmetric information. Most of the work in the past has concentrated on the first question,<sup>1</sup> in the sense that optimality did not depend on the allocation of the private good. Utility was perfectly transferable through the private good, so cost distribution played a role only insofar as transfers provided a convenient way of relaxing incentive constraints.

The results for that case have been, for the most part, rather satisfying. If utility is transferable via a private good, then dominant strategy mechanisms exist which fit the bill<sup>2</sup> and which do not depend on prior information the players and the planner may have about the distribution of preferences for the public good. If the planner does have priors which are common knowledge among the players then there exist first best mechanisms if the players' preferences are statistically and functionally related in a particular way

---

\*This paper replaces an earlier version, entitled "On the Optimality of Lottery Drafts: Characterization of Interim Efficiency in a Public Goods Problem." We have benefited from comments by two referees and by seminar participants at Caltech, MIT, Northwestern, Texas, and Texas A&M, and the 1989 FRET Conference in San Diego. We especially wish to acknowledge useful conversations with Kim Border and Ed Green. Ledyard gratefully acknowledges the financial support of the Flight Projects Office of the Jet Propulsion Laboratory of NASA. Palfrey gratefully acknowledges the support of the National Science Foundation under grants SES-8718657 and SES-8815097.

<sup>1</sup>Exceptions include the work of Groves and Ledyard [1977b], Hurwicz [1979], and others, but their use of Nash equilibrium is controversial when applied in environments with incomplete and asymmetric information. Recently Jackson and Moulin [1992] have addressed distributional concerns, in a complete information framework.

<sup>2</sup>See Groves [1973], Green and Laffont [1979], and Groves and Ledyard [1977a] for descriptions of many of these results, and for extensive commentary on the restrictive nature of the assumptions needed for the main results. One serious limitation is the lack of budget balance.

(d'Aspremont and Gerard-Varet [1979,1982], d'Aspremont, Cremer, and Gerard-Varet [1990]).

All these results share two basic assumptions. First, the welfare properties of the final allocation depend only on the level of public good, and not on the distribution of cost shares. Second, there are no feasibility limitations on the transfer schemes. In particular, bankruptcy and other problems that could restrict the set of feasible sidepayments, are ignored.

In this paper, we relax both of these assumptions in the context of a Bayesian mechanism design problem with two preference-types, a 0-1 public good decision, and statistical independence. First, we characterize the set of all interim incentive efficient allocation rules. We show that the earlier limitation to a welfare function that depends only on the public good allocation is equivalent to solving the second-best allocation problem for exactly one set of (interim) welfare weights.<sup>3</sup> For any other choice of welfare weights, distributional considerations play a role, and the planner faces a tradeoff between using transfers to relax incentive constraints and using transfers to achieve distributional goals. We fully characterize the set of optimal allocation schemes for all welfare weights (Theorem 6). Second, we show that the public good allocation rule depends on the welfare weights in a very simple way: The more the planner wants to shift the payment burden towards individuals who value the public good more, the less often the public good is produced (see the last two paragraphs of Section 3.1).

Third, we investigate the implications of a wide class of restrictions on transfers, ranging from the traditional case of no limitations to the extreme case where transfers are impossible. We find systematic effects of these restrictions, and show that the qualitative features of these effects depend on the welfare weights. For example, in the traditional case, where welfare weights are chosen so that there are no distributional concerns, the first best solution can be achieved without resorting to transfers at all (see the commentary following corollary 9). More generally, however, transfers are needed to attain the second best solution. If transfer constraints are binding the solution is either first best or least cost in the sense that individuals who value the public good most are taxed to the limit (see Theorem 11).

Fourth, we show that the decision whether to produce the public good can be separated from the transfer rule, and, in the model we consider, can be accomplished by  $j^*$  majority voting scheme, where  $j^*$ , the number of "YES" votes required, depends on the welfare weights. The results are shown to have a natural geometric interpretation. Furthermore, for a wide range of welfare weights, an optimal scheme is a two-stage procedure, where first a vote is taken on whether to produce the public good. If the vote is to produce the public good, then a even-chance lottery is used to determine who pays. We call this the "lottery draft mechanism". Since such a cost-sharing scheme does not

---

<sup>3</sup>In asymmetric information models, these interim welfare weights depend on an individual's private information, or type. In our model, types are just preferences (marginal rates of substitution between the public and private good).

require transfers, it follows that in many cases transfers are not necessary to achieve the optimal allocation (see corollary 9).

Finally, we investigate the implications of imposing voluntary participation constraints. From the geometric characterization obtained earlier in the paper, it is easy to establish several properties of the participation-constrained solution. In particular, it provides a transparent proof of the results obtained by Mailath and Postlewaite [1990] and Rob [1989], that in large economies participation constraints imply that the public good will never be produced, unless it is common knowledge that a lottery draft mechanism does not violate participation constraints (see Section 4).

## 2 The Problem

### 2.1 Feasible Allocations, Preferences, and Information

A group of  $N$  people must decide whether or not to produce a discrete public good and who should pay for it.

A feasible allocation for this problem is any vector  $(x_1, \dots, x_N, y) \in X^N \times \{0, 1\}$  such that  $\sum_{i=1}^N x_i \geq Ky$ , where  $x_i$  is the transfer (tax) paid by agent  $i$ ,  $y$  is the amount of public good produced,  $K < N$  is a fixed constant, and  $X^i$  is the set of feasible transfers an individual may pay. Here,  $y = 1$  if the public good is produced,  $y = 0$  otherwise and  $x_i$  is the amount of private good that  $i$  contributes to the production of the public good. The set  $X^i$  of feasible transfers for  $i$  depends on  $w^i$ , the initial endowment, and on the possibilities for transfers of the private good. For this paper we assume  $w^i = 1$  for all  $i$ . We consider two types of transfer assumptions. First, we consider what happens if unlimited transfers are permissible and  $X^i = \mathbb{R}$ , the real line. Second, we consider limited transfer schemes where  $X^i \subset \mathbb{R}$ . Two special cases are: (1) no transfers at all,<sup>4</sup> where  $X^i = \{0, 1\}$ , and (2) consumption lower bounds where  $X^i = (-\infty, 1]$ , which restricts taxation to no more than the private good endowment. In all cases we will assume that  $\{0, 1\} \subseteq X$ .

Preferences of each individual depend upon whether or not the public good is produced and how much of the private good is available to be consumed by that individual. Individuals evaluate lotteries over outcomes using expected utility satisfying von Neuman Morgenstern axioms. In particular we assume if the probability that the project is built is  $q \in [0, 1]$ ,  $i$ 's endowment of the private good is  $w^i$  and  $i$ 's expected contribution is  $\alpha_i \in [0, 1]$  then  $i$ 's preferences can be represented by:

$$V^i = q - v^i \alpha^i.$$

Three facts should be noticed about this representation. First, as is standard, we assume that an individual's preferences are quasi-linear in the private good. Second, as is also

---

<sup>4</sup>This corresponds to an environment that has been studied extensively in laboratory experiments. See Palfrey and Rosenthal [1991a, 1991b] and the references cited therein.

standard, we assume individuals are risk-neutral. The first implies that each utility function is of the form  $g^i[h^i(y) + w^i - x^i]$ . The second implies that expected utility must be of the form  $A^i + r^i q + v^i(w^i - \alpha^i)$  where  $A^i, r^i$  and  $v^i$  are constants. The third fact is that we have normalized all utility functions so that  $r^i = 1$  and to eliminate  $(A^i + v^i w^i)$ . An alternative normalization, used by Clarke [1971] Groves [1973] and others, is  $v(q) = c^i q - \alpha^i$ . Since any linear transformation preserves preferences satisfying the axioms of expected utility, these represent the same preferences.

To simplify the problem, we restrict our attention in this paper to the two-type case<sup>5</sup> where  $V = \{v_1, v_2\}$  and  $0 < v_1 < v_2$ . We will call  $v_1$  the *low type* and  $v_2$  the *high type*. To make the problem interesting, we assume that  $1 \geq \frac{K}{N}v_1$ . Finally, we assume  $w^i = 1$  for all  $i$ .

We assume that the individual marginal rates of substitution,  $v_i$ , are private information, and that these are independently drawn from an identical distribution where  $\pi \in (0, 1)$  is the probability of a low type.

## 2.2 Mechanisms

How might we choose an allocation  $(x_1, \dots, x_n, y)$ ? Many natural possibilities exist, even if transfers are not possible. Perhaps the most familiar such mechanism is the *voluntary contribution mechanism*. Each individual voluntarily pays  $x_i$  and  $y$  is produced if and only if  $\sum_{i=1}^N x_i \geq K$ . Each  $i$  pays  $x_i$ . Another simple mechanism is a (*conditional*) *lottery draft* where  $y$  is produced with some probability,  $q$ , which does not depend on the information of the individuals. Then if  $y = 1$ ,  $K$  individuals are randomly selected and their  $x_i = w^i = 1$  while  $x_i = 0$  for the others. Both the voluntary contribution mechanism and the lottery draft are examples of mechanisms which do not use transfers of the private good between individuals. A somewhat different no transfer mechanism is the *least cost mechanism*. Each  $i$  is asked to report her marginal rate of substitution,  $v^i$ , of the public for the private good. If the public good is produced, these rates are ranked and the  $K$  lowest are required to pay  $x_i = w^i = 1$ ; the others pay nothing. If individuals correctly report their  $v^i$ , this might produce superior allocations to, say, the lottery draft. Unfortunately there is an incentive to misreport.

Economic theory suggests that using transfers can reduce the incentive to misreport in the last mechanism by compensating one type through taxes on the other type. Of course, if transfers are possible then we can improve things even more by transferring (arbitrarily) large amounts of the private good from the low  $v^i$  types to high  $v^i$  types — if  $v^i$  represents the social marginal cost of  $x_i$  — in which case a first best solution does not exist. Only if the welfare function is such that the marginal social cost of a transfer is the same for *both* types does a first-best solution exist, since, in this special case, no social gains or losses result from transfers. This is precisely what has been analyzed extensively by economists with a key contribution being made by d'Aspremont and Gerard-Varet

<sup>5</sup>For some results with continuous-types, see Ledyard and Palfrey [1989].

[1979]. In this special case, the *first-best solution* is any  $(x, y)$  satisfying  $y = 1$  if and only if  $\sum_{i=1}^N \frac{1}{v^i} \geq K$ . A mechanism provided by d'Aspremont and Gerard-Varet produces the first-best choice of  $y$  and provides incentives through the cost allocation  $x_1, \dots, x_n$ , for each  $i$  to correctly identify her  $v^i$ . To our knowledge no one has studied how to do this generally under conditions of asymmetric information when the social costs of transfers differ across types. We turn to that now.

To find the optimal mechanism for the public good problem one might begin by trying to identify the space of all possible mechanisms. Luckily that is not necessary. Using the Revelation Principle,<sup>6</sup> it can be shown that any allocation rule which can be attained as the Bayes-equilibrium of some mechanism can also be attained as the Bayes-equilibrium of a direct revelation mechanism: direct mechanisms which satisfy Incentive Compatibility Constraints.<sup>7</sup> Direct mechanisms (each of which uniquely defines our allocation rule) assign a probability measure over feasible allocations to every possible profile of types.

Now since  $\{v_1, v_2\}$  is the set of possible types, a feasible *direct mechanism* is

$$\mu : \{v_1, v_2\}^N \longrightarrow M(Z)$$

where  $Z = \{(x, y) \in X^N \times \{0, 1\} \mid \sum_{i=1}^N x_i \geq Ky\}$ , and  $M(Z)$  is the set of all probability measures on  $Z$ . Because lotteries between asymmetric mechanisms are possible and we are interested in symmetric welfare functions, we restrict attention to symmetric mechanisms whereby individuals are treated differently only if they have different  $v_i$ 's. This is justified by lemmas 1 and 2 in the appendix.

We can represent any direct symmetric mechanism by a collection of  $N + 1$  triples  $\left\{ \left\{ a_j^1, a_j^2, q_j \right\} \right\}_{j=0}^N$ , where

$a_j^1 =$  (expected) payment by a low type if there are exactly  $j$  low types and  $(N - j)$  high types.

$a_j^2 =$  (expected) payment by a high type if there are exactly  $j$  low types and  $(N - j)$  high types.

$q_j =$  probability the public good is produced if there are exactly  $j$  low types and  $(N - j)$  high types.

We represent mechanisms in this form by

$$\eta = (a^1, a^2, q) = [a_0^1, \dots, a_N^1, a_0^2, \dots, a_N^2, q_0, \dots, q_N].$$

$\eta$  is feasible if and only if there is a feasible symmetric direct mechanism  $\mu$  which yields  $\eta$ . It is shown in the appendix, lemma 11, that  $\eta$  is feasible iff, for  $k = 0, \dots, N$ ,

$$q_k \in [0, 1] \tag{1.1}$$

<sup>6</sup>See Gibbard [1973] or Harris and Townsend [1981].

<sup>7</sup>We do, nevertheless, return to other mechanisms below.

$$ka_k^1 + (N - k)a_k^2 \geq Kq_k \quad (1.2)$$

$$a_0^1 = 0, a_N^2 = 0 \quad (1.3)$$

$$\text{and } a_k^i \in A \equiv CO(X) \text{ for } i = 1, 2 \quad (1.4)$$

and, for  $k = 1, \dots, N - 1$ ,

$$a_k^1 \geq q_k \max \left\{ \underline{x}, \frac{K - (N - k)\bar{x}}{k} \right\} + (1 - q_k) \max \left\{ \underline{x}, \frac{-(N - k)\bar{x}}{k} \right\} \quad (1.5)$$

$$a_k^2 \geq q_k \max \left\{ \underline{x}, \frac{K - k\bar{x}}{N - k} \right\} + (1 - q_k) \max \left\{ \underline{x}, \frac{-k\bar{x}}{N - k} \right\} \quad (1.6)$$

where  $\underline{x} = \min\{x \in X\}$ ,  $\bar{x} = \max\{x \in X\}$ , and  $CO(X)$  is the convex hull of  $X$ ; that is,  $A = [\underline{x}, \bar{x}]$ .

There are three special cases of interest. If unlimited transfers are possible then  $A = X = (-\infty, +\infty)$  and (1.5) and (1.6) are never binding. If there are lower bounds on consumption then  $A = X = (-\infty, 1]$  and, again, (1.5) and (1.6) do not bind. If no transfers are possible,  $X = \{0, 1\}$ ,  $A = [0, 1]$  and (1.5) and (1.6) may come into play as  $a_k^1 \geq q_k \max \left\{ 0, \frac{K - (N - k)}{k} \right\}$  and  $a_k^2 \geq q_k \max \left\{ 0, \frac{K - k}{N - k} \right\}$ . The necessity for (1.5) and (1.6) arises because of the fact that  $y \in \{0, 1\}$ . If large subsidies or taxes are possible then lotteries and transfers can be used to convexify everything without (1.5) and (1.6). Limited transfers and discreteness of either or both  $X$  and feasible  $y$  requires additional care.

### 2.3 Feasible Reduced-form Allocation Rules:

An individual's interim<sup>8</sup> expected utility depends only on her type and two "reduced form" numbers representing that individual's expected transfer and the probability of public good production, given her type. Denoting these by  $(Q_1, \alpha_1)$  and  $(Q_2, \alpha_2)$  respectively for the two types, we have

$$Q_1(q) = \sum_{j=0}^{N-1} \binom{N-1}{j} \pi^j (1 - \pi)^{N-j-1} q_{j+1} \quad (2.1)$$

$$Q_2(q) = \sum_{j=0}^{N-1} \binom{N-1}{j} \pi^j (1 - \pi)^{N-j-1} q_j \quad (2.2)$$

$$\alpha_1(a) = \sum_{j=0}^{N-1} \binom{N-1}{j} \pi^j (1 - \pi)^{N-j-1} a_{j+1}^1 \quad (3.1)$$

$$\alpha_2(a) = \sum_{j=0}^{N-1} \binom{N-1}{j} \pi^j (1 - \pi)^{N-j-1} a_j^2 \quad (3.2)$$

<sup>8</sup>Interim means  $i$  knows her own  $v_i$  but not that of the others.

We will say  $(Q, \alpha)$  is a *feasible reduced form* if and only if there exists a feasible  $\eta = (a^1, a^2, q)$  such that (2) and (3) are true. We will say  $Q$  is *feasible* if and only if  $\exists q_k \in [0, 1] \forall k$  such that (2) is true. We illustrate reduced forms with three examples:

1. *Lindahl-optimal public good production* ( $Q^o$ ). According to this production rule, the public good is produced if and only if the sum of the marginal rates of substitution,  $\sum_{i=1}^n \frac{1}{v_i}$ , exceeds the marginal rate of transformation,  $K$ . Letting  $j \leq N$  denote the number of low cost types, this says that  $q = 1$  if and only if  $\frac{j}{v_1} + \frac{N-j}{v_2} \geq K$ , or  $q = 1$  if and only if  $j \geq j^o = \max \left\{ 0, \frac{v_1(v_2 K - N)}{v_2 - v_1} \right\}$ . Thus,  $(Q_1^o, Q_2^o)$  is given by:

$$Q_1^o = \sum_{j=j^o-1}^{N-1} \binom{N-1}{j} \pi^j (1-\pi)^{N-j-1}$$

$$Q_2^o = \sum_{j=j^o}^{N-1} \binom{N-1}{j} \pi^j (1-\pi)^{N-j-1}$$

2. *Conditional Lottery Draft* ( $Q^L, \alpha^L$ ). Under this rule the good is produced if and only if  $j \geq j^*$ , where  $j^*$  is some fixed number between 0 and  $n$ . If it is produced, taxes are raised by randomly “drafting”  $K$  individuals and taxing their entire endowment. Thus  $(Q_1^L, Q_2^L)$  is given just as in  $Q^o$ , except with  $j^*$  replacing  $j^o$ , and  $(\alpha_1^L, \alpha_2^L) = (\frac{K}{N} Q_1^L, \frac{K}{N} Q_2^L)$ .
3. *Conditional Least Cost Mechanism* ( $Q^{LC}, \alpha^{LC}$ ). In this case,  $Q^{LC} = Q^L$ , given  $j^*$ , but whenever possible  $v_1$ -types are taxed. The only time  $v_2$ -types are taxed is when the number of  $v_1$ -types is less than  $K$ , but greater than  $j^*$ . Thus:

$$\alpha_1^{LC} = \sum_{j=j^*-1}^{N-1} \binom{N-1}{j} \pi^j (1-\pi)^{N-j-1} \min \left\{ 1, \frac{K}{j} \right\}$$

$$\alpha_2^{LC} = \sum_{j=j^*}^{N-1} \binom{N-1}{j} \pi^j (1-\pi)^{N-j-1} \max \left\{ 0, \frac{K-j}{N-j} \right\}$$

Finally, observe that in all of the above examples, the budget balances. This is not required by feasibility. For example, the disastrous reduced form mechanism  $Q_1 = Q_2 = 0, \alpha_1 = \alpha_2 = 1$  is a feasible reduced form. However, as we will show below, “budget balance” is implied by interim incentive efficiency.

## 2.4 Incentive Compatible Allocation Rules:

Because interim utility depends only on  $Q, \alpha$  and  $v_i$ , we call  $V(Q, \alpha, v_i) = Q_i - \alpha_i v_i$  the *interim utility for type  $v_i$* . A reduced form allocation rule  $(Q, \alpha)$  is incentive compatible if and only if, for all  $v_i, v_j \in T, V(Q, \alpha, v_i) \geq Q_j - \alpha_j v_i$ . Observe that incentive compatibility of  $(Q, \alpha)$  is independent of the particular normalization of utility.

In addition to feasibility and incentive compatibility constraints, it is a common practice to impose some sort of requirement on the mechanism that participation be voluntary.<sup>9</sup> This may be interpreted as either a conservative judgement by the planner respecting prior “utility claims” of the individual, restrictions based on legal standards respecting property rights, or a lack of sufficient enforcement power by the planner to impose such allocations. In Section 3, we will solve the optimal mechanism problem in the absence of any such constraints. In Section 4 we will analyze exactly how these constraints affect the properties of the optimal mechanism.

## 2.5 Optimality

In this paper, we characterize the interim incentive efficient frontier corresponding to all type-dependent welfare weights that are symmetric across individuals (i.e. anonymous).<sup>10</sup> For our problem (see Holmstrom and Myerson [1983]), this means that for each  $\lambda \geq 0$  we look for reduced form allocation rules that maximize

$$\lambda\pi(Q_1 - v_1\alpha_1) + (1 - \pi)(Q_2 - v_2\alpha_2) \quad (E)$$

subject to incentive compatibility and feasibility. One might wonder why, instead of mapping out the (symmetric) interim efficiency frontier we do not simply find an *ex ante* optimal symmetric mechanism as in Myerson [1981] and Myerson and Satterthwaite [1983]. Actually, in our problem, the *ex ante* optimal allocation rule is simply a particular point on the interim efficiency frontier, corresponding to  $\lambda = 1$ . In fact, each of the other points on the frontier ( $\lambda \neq 1$ ) corresponds to an *ex ante* optimal allocation rule for a different normalization of the utility functions.<sup>11</sup> For example, the allocation rule that maximizes (E) subject to incentive compatibility and feasibility when  $\lambda = v_2/v_1$  corresponds to the *ex ante* incentive efficient rule in the case where utility functions are normalized as in Green-Laffont or d’Aspremont-Gerard-Varet ( $r^i q - \alpha^i$ ) instead of our normalization ( $q - v^i \alpha^i$ ). In this sense, one can view our approach of characterizing the

<sup>9</sup>Rob [1989] and Mailath and Postlewaite [1990] have recently analyzed these constraints in the context of public goods problems. Much of the initial work on optimal public goods mechanisms did not impose individual rationality constraints.

<sup>10</sup>We do not address the problem of optimal mechanisms with welfare weights that may depend on individual as well as type. Also, at least in the unlimited transfer case, optimal mechanisms may fail to exist with asymmetric weights, since optimality would call for arbitrarily large type-independent transfers between agents.

<sup>11</sup>See d’Aspremont and Gerard-Varet [1989] for a related discussion.

symmetric interim efficient frontier as equivalent to characterizing the set of allocation rules that are *ex ante* efficient for some normalization of the utility functions.<sup>12</sup>

## 2.6 Summary

Bringing together all the assumptions, comments, and observations above, we can summarize the problem of characterizing all interim-efficient public good production and cost allocation mechanisms. The main points follow.

1. We assume agents satisfy the expected utility hypothesis (they act as if they maximize expected utility) with risk-neutral, quasi-linear preferences.
2. The revelation principle and the hypothesis of Bayes-Nash equilibrium behavior allow us to restrict attention to feasible and incentive compatible direct mechanisms.
3. We restrict attention to symmetric mechanisms.
4. There are only 2 types.
5. We are interested in the class of symmetric interim efficient mechanisms for the normalization  $u^i = q - v^i \alpha^i$ . This is equivalent to the class of symmetric *ex ante* efficient mechanisms for all normalizations.

We can, therefore, solve our problem in two steps. **First**, for each  $\lambda > 0$ : choose the reduced form  $(Q, \alpha)$  to solve

$$(P) \quad \max \lambda \pi(Q_1 - v_1 \alpha_1) + (1 - \pi)(Q_2 - v_2 \alpha_2)$$

subject to:

$$Q_1 - v_1 \alpha_1 \geq Q_2 - v_1 \alpha_2 \tag{4.1}$$

$$Q_2 - v_2 \alpha_2 \geq Q_1 - v_2 \alpha_1 \tag{4.2}$$

$$(Q, \alpha) \text{ feasible} \tag{1)-(3)}$$

where (4.1)-(4.2) require  $(Q, \alpha)$  to be incentive compatible. **Second**, construct a feasible direct mechanism  $\mu$  which implements the optimal solution to  $(P)$ . We turn to the first of these tasks next.

---

<sup>12</sup>Implicit in this statement is that the choice of normalization is the same across agents. The implications of considering different normalizations for different individuals in the *ex ante* optimization problem are similar to the implications of nonanonymous type-dependent welfare weights in our *interim* optimization problem.

### 3 Characterization of Optimal Rules

In this section we characterize the class of interim-efficient, incentive-compatible mechanisms for public good provision and cost allocation under a variety of different assumptions about transferability of the private good. We do this by substituting a simpler but equivalent problem for  $(P)$ .

First, we rewrite (4.1) and (4.2) using two results. Incentive compatibility immediately implies that  $Q_1 \geq Q_2$  and  $\alpha_1 \geq \alpha_2$ . This is a standard result and is proved in Lemma 3 of the appendix. Second, in Lemmas 4, 5, and 6 we prove that any solution to  $(P)$  must satisfy  $\bar{\alpha} = \frac{K}{N}\bar{Q}$ , where  $\bar{\alpha} = \pi\alpha_1 + (1 - \pi)\alpha_2$  and  $\bar{Q} = \pi Q_1 + (1 - \pi)Q_2$ . That is, there is no waste. Production always takes place on the production possibilities frontier. This simplifies the analysis. Substituting  $\bar{\alpha} = \frac{K}{N}\bar{Q}$  into 4.1 and 4.2 gives:

$$\begin{aligned}\alpha_1 &\leq \frac{K}{N}\bar{Q} + \frac{1}{v_1}(1 - \pi)(Q_1 - Q_2) \equiv IC_U(Q) \\ \alpha_1 &\geq \frac{K}{N}\bar{Q} + \frac{1}{v_2}(1 - \pi)(Q_1 - Q_2) \equiv IC_L(Q).\end{aligned}$$

Notice that  $Q_1 \geq Q_2 \Rightarrow IC_U(Q) \geq IC_L(Q)$  which is also true with strict inequalities.

A second set of constraints relate to feasibility. The first such constraint is on  $Q = (Q_1, Q_2)$  and comes from (1.1), (2.1), and (2.2). We represent the set of feasible  $(Q_1, Q_2)$  in the following way. For any  $Q_1 \in [0, 1]$  define  $\underline{Q}_2(Q_1) = \min \{Q_2 | \exists q = (q_0, \dots, q_N)$  feasible such that  $Q_2 = \mathbf{Q}_2(q)$  and  $Q_1 = \mathbf{Q}_1(q)\}$ . This function is well defined since the set of feasible  $q$  is compact and  $\mathbf{Q}_1(q)$  and  $\mathbf{Q}_2(q)$ , as defined by (2) and (3), are both continuous in  $q$  and the set of  $q$  such that  $Q_1 = \mathbf{Q}_1(q)$  is nonempty for all  $Q_1 \in [0, 1]$ . Define

$$Q_F = \{(Q_1, Q_2) \mid 1 \geq Q_1 \geq Q_2 \geq 0 \text{ and } Q_2 \geq \underline{Q}_2(Q_1)\}.$$

It follows immediately from definitions that  $Q \in Q_F$  if and only if  $Q$  is a feasible reduced form and  $Q_1 \geq Q_2$ . The set  $Q_F$  is illustrated in Figure 1. We summarize the properties of  $Q_F$  in the following proposition:

**Proposition:**  $Q_F$  is a convex polygon, with a boundary defined by  $N + 1$  vertices that correspond to the  $j^*$  conditional mechanisms:

$$Q_1 = \sum_{j=j^*-1}^{N-1} \binom{N-1}{j} \pi^j (1 - \pi)^{N-j-1}$$

$$Q_2 = \sum_{j=j^*}^{N-1} \binom{N-1}{j} \pi^j (1 - \pi)^{N-j-1}$$

for  $j^* = 0, 1, \dots, N$ .

*Proof:* By the definitions of  $Q_1$  and  $Q_2$  in equations (1) and (2), one can characterize  $\underline{Q}_2(Q_1)$  in terms of  $q_0, \dots, q_N$ . Specifically, for  $Q_1 \leq \pi^{N-1}$ , we get  $\underline{Q}_2(Q_1) = 0$ , since we can produce  $(Q_1, 0)$  in this region by setting  $q_0, \dots, q_{N-1}$  all equal to 0, and setting  $q_N = \frac{Q_1}{\pi^{N-1}}$ . For any  $Q_1 > \pi^{N-1}$ ,  $\underline{Q}_2(Q_1)$  must be strictly positive because  $q_i > 0$  for some  $i < N$ . For values of  $Q_1 > \pi^{N-1}$ ,  $\underline{Q}_2(Q_1)$  is constrained by setting  $q_N = 1$  and choosing  $q_1, \dots, q_{N-1}$  to minimize

$$Q_2(q) = \sum_{j=0}^{N-1} \binom{N-1}{j} \pi^j (1-\pi)^{N-j-1} q_j$$

subject to

$$Q_1(q) = \sum_{j=0}^{N-1} \binom{N-1}{j} \pi^j (1-\pi)^{N-j-1} q_{j+1}, \quad q_j \in [0, 1] \quad j = 1, \dots, N-1$$

Since  $\frac{\partial Q_2 / \partial q_j}{\partial Q_1 / \partial q_j} = \frac{(N-j)\pi}{j(1-\pi)}$  is decreasing in  $j$ , the solution to this linear minimization problem is given by a pair  $(j^*, \theta)$  such that  $q_i = 1$  for  $N \geq j > j^*$ ,  $q_i = 0$  for  $0 \leq j < j^*$  and  $q_i = \theta$  for  $j = j^*$ . The result follows immediately. ■

The other constraints involve joint restrictions on  $Q$  and  $\alpha$  that depend on the restrictions on transfers. Specifically, we show in Lemma 7 that there exist piecewise linear functions,  $L(Q)$  and  $U(Q)$  such that  $(Q, \alpha)$  is feasible and  $\bar{\alpha} = \frac{K}{N} \bar{Q}$  if and only if  $Q \in Q_F$  and  $L(Q) \leq \alpha_1 \leq U(Q)$ . This is a reduced-form representation of the constraints on transfers, and will depend on the possibilities for transfers (the set  $A^*$ ). For example, with unlimited transfers,  $L(Q) = -\infty$  and  $U(Q) = \infty$  for all  $Q$ , so the constraints are never binding. Things are more complicated in the **no bankruptcy** and **no transfers** cases, which are taken up later.

From above, we can replace (1)-(3) in  $(P)$ , and write a simplified reduced form problem  $(R)$  for each  $\lambda > 0$ , as: choose  $(Q, \alpha_1)$  to

$$(R) \quad \text{maximize } \pi Q_1 \left[ \lambda - \frac{K}{N} v_2 \right] + (1-\pi) Q_2 \left[ 1 - \frac{K}{N} v_2 \right] + \pi (v_2 - \lambda v_1) \alpha_1$$

subject to

$$\alpha_1 \leq IC_U(Q) \tag{4.1}$$

$$\alpha_1 \geq IC_L(Q) \tag{4.2}$$

$$Q \in Q_F \tag{5.1}$$

$$L(Q) \leq \alpha_1 \leq U(Q) \tag{5.2}$$

The constraints (5.1) and (5.2), replace (1) (2) and (3) in  $(P)$ . It is easy to show:

**Theorem 1**  $(Q, \alpha_1)$  solves (R) and  $\alpha_2 = \frac{1}{1-\pi}(\frac{K}{N}\bar{Q} - \pi\alpha_1)$  if and only if  $(Q, \alpha)$  solves (P).

*Proof:* Uses lemmas 3, 5, 6, and 7. ■

Thus the solutions to (R) for  $(Q, \alpha_1)$  as we vary  $\lambda \in [0, \infty)$  characterize all interim-efficient mechanisms. Since (5.2) depends on the set of feasible transfers, we split the characterization into two parts: unlimited and limited transfers.

### 3.1 Unconstrained Transfers of the Private Good

By unconstrained transfers, we mean that  $a_j^i$  can be any real number. That is  $A^* = \mathfrak{R}$  and  $(a, q)$  is feasible iff (1.1)-(1.4). Implicitly, this means that the private good is continuous rather than lumpy, and there are no lower or upper bounds on any individual's consumption.

In this case  $L(Q) = -\infty$  and  $U(Q) = \infty$ , so constraint (5.2) drops out. The solution to (R) can be split into three cases which depend on the value of  $\lambda$ , the welfare weight placed on low cost types.

**Theorem 2** If  $\lambda < \frac{v_2}{v_1}$ ,  $(Q, \alpha)$  is an optimal mechanism if and only if:

$$\alpha_1 = IC_U(Q)$$

(A)

$$\bar{\alpha} = \frac{K}{N}\bar{Q}$$

and  $Q$  maximizes

$$\pi Q_1 \left[ \lambda - \frac{K}{N}v_2 \right] + (1 - \pi)Q_2 \left[ 1 - \frac{K}{N}v_2 \right] + \pi(v_2 - \lambda v_1)IC_U(Q)$$

subject to  $Q \in Q_F$ .

*Proof:* Since  $v_2 > \lambda v_1$  we want  $\alpha_1$  to be as large as possible for any  $Q$ , subject to incentive compatibility. Therefore  $\alpha_1 = IC_U(Q)$ . Then (4.2) is not binding. Substitute for  $\alpha_1$  in the objective function and eliminate (5.2) in (R) to finish. ■

**Theorem 3** If  $\lambda > \frac{v_2}{v_1}$ , then  $(Q, \alpha)$  is an optimal mechanism if and only if:

$$\alpha_1 = IC_L(Q)$$

(A')

$$\bar{\alpha} = \frac{K}{N}\bar{Q}$$

and  $Q$  maximizes

$$\pi Q_1 \left[ \lambda - \frac{K}{N}v_2 \right] + (1 - \pi)Q_2 \left[ 1 - \frac{K}{N}v_2 \right] + \pi(v_2 - \lambda v_1)IC_L(Q)$$

subject to  $Q \in Q_F$ .

*Proof:* Same as proof of Theorem 1 but  $\alpha_1$  is now chosen as small as possible so (4.2) binds, and (4.1) is slack. ■

**Theorem 4** *If  $\lambda v_1 = v_2$ , then  $(Q, \alpha)$  is an optimal mechanism if and only if:*

$$\bar{\alpha} = \frac{K}{N}\bar{Q}$$

(A'')

$$IC_L(Q) \leq \alpha_1 \leq IC_U(Q)$$

and  $Q$  maximizes

$$\pi \left[ \frac{1}{v_1} - \frac{K}{N} \right] Q_1 + (1 - \pi) \left[ \frac{1}{v_2} - \frac{K}{N} \right] Q_2$$

subject to  $Q \in Q_F$ .

*Proof:* Substitution of  $\frac{v_2}{v_1}$  for  $\lambda$  in the objective function eliminates  $\alpha$  from the objective function, so the optimal  $Q$  is compatible with a range of  $\alpha$ 's. This results in a **first best** solution, since the incentive compatibility constraints can be satisfied with strict inequality. ■

Several observations can be made. With unlimited transfers (and a linear utility structure) the private good is used to transfer welfare subject to incentive compatibility constraints. For example, if  $\lambda v_1 < v_2$  and if incentives were not an issue, optimality would require an infinite transfer from low-cost types to high cost types, no matter what  $Q$  is. But then low types would pretend to be high types. That incentive to misrepresent is prevented by the incentive constraint (4.1). Thus, the private good is always transferred to the extent allowed by incentive constraints. This determines an optimal incentive compatible financing scheme for each possible  $Q$ . Then  $Q$  is chosen so the welfare from  $Q$  net of the loss from its financing is optimal. At this point, the only incentive constraint is  $Q_1 \geq Q_2$ . This means that the class of interim efficient mechanisms is simply represented by the boundary of  $Q_F$ . The only role incentive compatibility plays is to rule out mechanisms where  $Q_2 > Q_1$ . It is not until one wants to select a mechanism for a particular welfare weight,  $\lambda$ , on low types that the actual incentive constraints

become important. Then, they have a similar effect to that noticed by Myerson [1981]. They create a new “virtual utility” function but do not affect the feasible set.

We have the following summary Theorem,<sup>13</sup> which is illustrated in Figure 2.

**Theorem 5** *The set of interim efficient mechanisms for public good production and cost allocation with unlimited transfers is  $\{(Q, \alpha) | Q \in \text{lower boundary of } Q_F, \text{ and } (Q, \alpha) \text{ is incentive compatible}\}$ . For almost all  $\lambda$ , the interim efficient mechanism is a vertex of  $Q_F$ .*

This completes the first step in solving (P). The second step involves identifying direct (non-reduced form) mechanisms which yield the desired reduced form. The following result shows how to do this.

**Theorem 6** *For any parameters  $(K, N, \pi, v_1, v_2, \lambda)$  there exists a  $j^*$  such that it is interim efficient to produce the public good if and only if  $j \geq j^*$ .*

*Proof:* Recall that  $Q_F$  is a convex polygon, with a boundary consisting of  $N + 1$  vertices. Since the objective functions are linear in  $Q_1$  and  $Q_2$  and the boundary of  $Q_F$  is piecewise linear, the objective function achieves its maximum at a vertex for all maximization problems in Theorems 2, 3, and 4. ■

We can see that there are only two kinds of mechanisms that solve our problem. In the first kind, the good is always produced,  $Q_1 = Q_2 = 1$ , and can be financed by a lottery draft,  $a_1^i = a_2^i = \frac{K}{N}$ . Pooling of types is efficient and, in Theorem 6,  $j^* = 0$ . More precisely,

**Corollary 7** *If  $\lambda \leq \frac{v_2}{v_1}$ , then  $q_j = 1$ ,  $a_j^1 = \frac{K}{N}$ ,  $a_j^2 = \frac{K}{N} \forall j$  is an optimal solution to (P) if and only if*

$$\left[1 - \frac{K}{N}v_1\right] \geq \frac{(v_2 - v_1)}{\lambda\pi v_1 + (1 - \pi)v_2}$$

*Proof:* From Theorem 2, if  $\lambda \leq \frac{v_2}{v_1}$ , the objective function is

$$E(Q_1, Q_2) = \pi Q_1(\lambda - \frac{K}{N}v_2) + (1 - \pi)Q_2(1 - \frac{K}{N}v_2) + \pi(v_2 - \lambda v_1)IC_U(Q)$$

Notice that  $\partial E/\partial Q_1 = \pi(\frac{1}{v_1} - \frac{K}{N})(\lambda v_1 \pi + (1 - \pi)v_2) > 0$ . Therefore  $Q^* = (1, 1)$  is optimal if and only if  $\partial E/\partial Q_2 \geq 0$ . The result follows immediately since

$$\partial E/\partial Q_2 = (1 - \pi) \left[ \left( \frac{1}{v_1} - \frac{K}{N} \right) (\lambda v_1 \pi + (1 - \pi)v_2) - \frac{(v_2 - v_1)}{v_1} \right].$$

<sup>13</sup>We conjecture that this theorem is also true for a continuum of types. The only technical problem in proving this is characterizing  $Q_F$ . See Ledyard and Palfrey [1989].

**Corollary 8** *If  $\lambda > \frac{v_2}{v_1}$ , then  $q_j = 1$ ,  $a_j^1 = \frac{K}{N}$ ,  $a_j^2 = \frac{K}{N} \forall j$  is an optimal solution to (P) if and only if  $1 \geq \frac{K}{N}v_2$ .*

*Proof:* From Theorem 3, when  $\lambda \geq \frac{v_2}{v_1}$ , the objective function is

$$E = \pi Q_1 \left( \lambda - \frac{K}{N} v_2 \right) + (1 - \pi) Q_2 \left( 1 - \frac{K}{N} v_2 \right) + \pi (v_2 - \lambda v_1) IC_L(Q)$$

Now

$$\partial E / \partial Q_1 = \pi \left[ \left( \frac{1}{v_2} - \frac{K}{N} \right) (\lambda v_1 \pi + (1 - \pi) v_2) + \frac{\lambda}{v_2} (v_2 - v_1) \right]$$

and

$$\partial E / \partial Q_2 = (1 - \pi) \left[ \left( 1 - \frac{K}{N} v_2 \right) (\lambda v_1 \pi + (1 - \pi) v_2) \right] / v_2.$$

To show (if), notice that if  $1 > \frac{K}{N}v_2$  then  $\frac{\partial E}{\partial Q_1} > 0$  and  $\frac{\partial E}{\partial Q_2} \geq 0$ . To show (only if), observe that if  $1 < \frac{K}{N}v_2$  then  $\frac{\partial E}{\partial Q_2} < 0$ , so  $Q = (1, 1)$  cannot be optimal. ■

For comparison to Corollary 7, note that  $1 \geq \frac{K}{N}v_2$  iff  $\left[ 1 - \frac{K}{N}v_1 \right] \geq \frac{(v_2 - v_1)}{v_2}$ . An interesting implication of Corollaries 7 and 8 is that if  $1 < \frac{K}{N}v_2$  then  $Q^* = (1, 1)$  is never optimal. Lottery drafts with  $Q = 1$  are never optimal when it is ex post inefficient to produce the public good if everyone is a high-cost type.

In the second kind of mechanism, the good is not produced unless there are sufficiently many low cost types. This happens if it is less costly to induce revelation of preference by restricting production when there are too many high-types than by using the private good. This will be true if the (relative) welfare weight on the high cost type is large enough so that it is important to avoid lotteries (pooling) and to differentiate between types (separating). More precisely,

**Corollary 9** *If  $\lambda \leq \frac{v_2}{v_1}$  and  $\left[ 1 - \frac{K}{N}v_1 \right] < \frac{(v_2 - v_1)}{\lambda \pi v_1 + (1 - \pi)v_2}$ , then the optimal mechanism is to produce the public good if and only if the number of low types is larger than  $j^*$ , with  $j^*$  equal to the smallest  $j$  such that*

$$\frac{j}{N} \geq 1 - \frac{(1 - \frac{K}{N}v_1)(\lambda \pi v_1 + (1 - \pi)v_2)}{(v_2 - v_1)}.$$

*Proof:* From Figure 1 we want the smallest  $j$  such that the slope of the indifference lines of the objective function is greater than or equal to the slope of the constraint set, given by the boundary of  $Q_F$ . That is, we want the smallest  $j$  such that

$$\frac{-\pi \left[ 1 - \frac{K}{N}v_1 \right] [\lambda \pi v_1 + (1 - \pi)v_2]}{[1 - \pi] \left[ \left[ 1 - \frac{K}{N}v_1 \right] [\lambda \pi v_1 + (1 - \pi)v_2] + [v_1 - v_2] \right]} \geq \frac{(N - j)\pi}{j(1 - \pi)}$$

■

An analogous result holds if  $\lambda \geq \frac{v_2}{v_1}$ . In particular  $j^*$  is the smallest  $j$  such that

$$\frac{j}{N} \geq \frac{-(1 - \frac{K}{N}v_2)(\lambda\pi v_1 + (1 - \pi)v_2)}{\lambda(v_2 - v_1)}.$$

There are two interesting special cases. If  $\lambda = \frac{v_2}{v_1}$ , we have the transfer-neutral model, which turns out to be the knife-edge case that separates Theorem 2 and Theorem 3. That is, individual utilities are normalized to  $c_i q^i - \alpha^i$  so that every type's marginal utility of the private good is 1, and the welfare function is the expected value of  $c^i q - \alpha^i$ . In this case  $\lambda\pi v_1 + (1 - \pi)v_2 = v_2$ . To find the optimal mechanism as in Corollary 9, we want the minimum  $j \geq 0$  such that  $j \left[1 - \frac{K}{N}v_1\right] v_2 + (N - j) \left[\left[1 - \frac{K}{N}v_1\right] v_2 - (v_2 - v_1)\right] \geq 0$ . That is, we want the minimum  $j$  such that  $j \left[\frac{1}{v_1} - \frac{K}{N}\right] + (N - j) \left[\frac{1}{v_2} - \frac{K}{N} \geq 0\right]$  or  $j \left[c_1 - \frac{K}{N}\right] + (N - j) \left[c_2 - \frac{K}{N}\right] \geq 0$ . But *this is just the first best solution*: the good is produced if and only if  $j c_1 + (N - j)c_2 \geq K$ .

With  $\lambda = \frac{v_2}{v_1}$ , the range of cost allocations which induce incentive compatibility is given in Theorem 4. A particularly simple mechanism is a lottery<sup>14</sup> where  $K$  of the  $N$  individuals are randomly selected to pay for the public good, whenever at least  $j^*$  individuals report that they are low-cost types. One can interpret this as a conditional lottery draft, that is preceded by a vote between producing and not producing the public good. The vote outcome is determined by a weighted majority rule where  $j^*$  of  $N$  must vote in favor, in order for production to arise. Low cost types vote for it and high cost types vote against it. This corresponds to  $\alpha_1 = \frac{K}{N}Q_1$ , and  $\alpha_2 = \frac{K}{N}Q_2$ , and does satisfy incentive compatibility.

A second interesting case occurs when  $\lambda = 1$  (covered by Theorem 2): the problem studied in the social dilemma literature.<sup>15</sup> In this case we find that, even if  $1 - \frac{K}{N}v_2 \geq 0$ , and therefore it is common knowledge that benefits exceed costs with probability 1, it may be optimal not to produce with probability 1. In fact if  $\lambda < \frac{v_2}{v_1}$  and  $(v_2 - v_1) > \left[1 - \frac{K}{N}v_1\right] [\pi v_1 + (1 - \pi)v_2]$  then  $Q_2 < Q_1 < 1$  in the optimal mechanism.<sup>16</sup> The reason is simple: The distribution of taxes is more important than optimal production.

As  $\lambda$  changes, the feasible set  $Q_F$  does not change but the slope of the indifference lines of the objective function does change. As  $\lambda$  increases, the indifference lines become steeper (see Lemma 8 in the appendix), so the expected output of the optimal mechanism (weakly) increases.<sup>17</sup> There is no discontinuity at  $\lambda v_1 = v_2$ .

<sup>14</sup>If the private good were divisible, each member could be taxed an amount  $K/N$  from their endowment whenever the public good is produced.

<sup>15</sup>See Van de Kragt, Orbell and Dawes [1983], Palfrey and Rosenthal [1984, 1988, 1990, 1991] and Rapoport [1985].

<sup>16</sup>To see that these conditions are not vacuous pick  $\pi$  large,  $v_1$  small, and  $v_2$  close to 1.

<sup>17</sup>It is easy to see a similar effect happens if  $K$  is decreased or  $N$  is increased or  $v_2$  decreases or  $r$  increases. Of course when  $N$  or  $\pi$  are increased,  $Q_F$  changes.

An increase in  $\lambda$  implies a higher welfare weight on low cost types, who like output relatively more than the high cost types. Only for  $\lambda = (v_2/v_1)$ , are there no welfare gains from redistribution of the private good. For  $\lambda > (v_2/v_1)$  the low cost types are relatively more important and, therefore, the efficient mechanism may provide an expected output higher than first best, resulting in overproduction of the public good. For  $\lambda < (v_2/v_1)$  the high cost types are relatively more important and, therefore, the efficient mechanism provides an expected output lower than first best, resulting in underproduction of the public good. Interim efficiency does not necessarily imply the first best solution because of the trade-off between the use of transfer payments for incentive compatibility and possible welfare gains from cost allocation.

### 3.2 Limited Transfers of the Private Good

In many applications it is illegal or impossible to fully compensate individuals who contribute to the production of the public good. In pure voluntary contributions situations no compensation is possible so  $X = \{0, 1\}$  ( $A = [0, 1]$ ); an agent can only contribute or not<sup>18</sup> (perhaps with some probability). A less restrictive situation arises when compensation is possible but taxation is limited by a no-bankruptcy condition so that no taxation is allowed which would leave an agent with negative amounts of the private good. Here  $X = (-\infty, 1]$ ; an agent cannot be taxed more than the initial endowment.

The impact of constraints on possible transfers can be easily seen by referring to the maximization problem ( $R$ ) in Section 3.0, particularly constraint (5.2). Under limited transfers  $L(Q) > -\infty$  and  $U(Q) < \infty$ , and it is possible that (5.2) will bind at the optimum. If  $\lambda v_1 < v_2$  then the solution to ( $R$ ) will involve making  $\alpha_1$  as large as possible, whatever  $Q \in Q_1$  is chosen. Thus  $\alpha_1$  will equal the smaller of  $U(Q)$  or  $\frac{K}{N}Q + \frac{1}{v_1}(1 - \pi)(Q_1 - Q_2)$ . An analogous condition holds if  $\lambda v_1 > v_2$ . Of course  $Q$  must be restricted so that an  $\alpha_1$  satisfying (4) and (5) exists but, as we will soon see, if  $Q \in Q_F$  we will not need to worry about that when lottery drafts are a feasible financing option, i.e. when  $[0, 1] \subseteq X$ . To formalize this intuition define:

$$H(Q) = \max \{L(Q), IC_L(Q)\}$$

and

$$G(Q) = \min \{U(Q), IC_U(Q)\}$$

and

$$S(Q, \lambda) = \begin{cases} H(Q) & \text{if } \lambda v_1 > v_2 \\ G(Q) & \text{if } \lambda v_1 < v_2. \end{cases}$$

When  $\lambda v_1 = v_2$ , let  $S(Q, \lambda)$  be any number such that<sup>19</sup>  $S \in [H(Q), G(Q)]$ . The

<sup>18</sup>If we assume divisibility ( $X = [0, 1]$ ) then an agent can contribute part of his endowment. Since  $A = [0, 1]$ , the formal results for the limited transfer case are the same whether or not divisibility is assumed. With divisibility, it would be possible to tax  $K/N$  of each individual's private good endowment.

<sup>19</sup>When  $Q \in Q_F$ ,  $\frac{K}{N}Q_1 \in [H(Q), G(Q)]$ . So if lottery drafts are feasible, then  $[H(Q), G(Q)] \neq \emptyset$  for all  $Q \in Q_F$ .

following theorem generalizes Theorems 2,3, and 4 to any limitations on transfers such that  $[0, 1] \subseteq A$ .

**Theorem 10** *Let  $[0, 1] \subseteq A$ .  $(Q, \alpha)$  is an interim efficient mechanism for  $\lambda$  (that is, it solves (R)) if and only if  $\alpha$  solves*

$$\alpha_1 = S(Q, \lambda)$$

(A')

$$\bar{\alpha} = \frac{K}{N} \bar{Q}$$

and  $Q$  solves

$$(C) \quad \max \pi Q_1 \left[ \lambda - \frac{K}{N} v_2 \right] + (1 - \pi) Q_2 \left[ 1 - \frac{K}{N} v_2 \right] + \pi(v_2 - \lambda v_1) S(Q, \lambda)$$

subject to

$$Q \in Q_F.$$

*Proof:* Case 1:  $\lambda v_1 > v_2$ . We need  $\alpha_1$  satisfying (4) and (5) and as small as possible. This means that whatever  $Q$  is chosen, we want  $\alpha_1 = H(Q)$ . Substitute this for  $\alpha_1$  in the objective function of (R).

We can find such an  $\alpha_1$  satisfying (4) and (5) for a given  $Q$  if and only if

$$IC_L(Q) \leq U(Q) \tag{6.1}$$

$$IC_U(Q) \geq L(Q) \tag{6.2}$$

By Lemma 9 and Lemma 10, (6.2) is true for all  $Q \in Q_F$ . By Lemma 10, if  $1 \leq \frac{K}{N} v_2$  then (6.1) is true for all  $Q \in Q_F$ . If  $1 > \frac{K}{N} v_2$ , we know from Corollary 8 that  $Q^* = (1, 1)$  is optimal if (5) is ignored. But (5) is satisfied at  $Q^* = (1, 1)$ .

Case 2:  $\lambda v_1 = v_2$ . Using the same analysis as that following Corollary 9 in Section 3.1, the first best solution is attainable with a lottery draft. (See footnote 15.)

Case 3:  $\lambda v_1 < v_2$ . Here we need  $\alpha_1$  as large as possible subject to (4) and (5). Let  $\alpha_1 = G(Q)$ . In order for such an  $\alpha_1$  to be feasible,  $Q$  must satisfy (6). By Lemmas 9 and 10, (6.2) is true for any  $Q \in Q_F$ . By Lemma 10, if  $1 \leq \frac{K}{N} v_2$  then (6.1) is true for  $Q \in Q_F$ . If  $1 > \frac{K}{N} v_2$ , suppose (6.1) is binding at the optimum to (R). Then  $IC_L(Q) = U(Q) < IC_U(Q)$  and  $Q_1 > Q_2$  since if  $Q_1 = Q_2$ , then  $\frac{K}{N} \bar{Q} = \frac{K}{N} Q_1 < U(Q)$  and (6.1) would not bind. Now consider a change from the optimal  $Q$  caused by increasing  $Q_2$  slightly. Since  $\alpha_1 = U(Q) < IC_U(Q)$ , the objective function changes by

$$\Delta Q_2 \cdot \left[ \left(1 - \frac{K}{N} v_2\right) (1 - \pi) + \pi(v_2 - \lambda v_1) \partial U / \partial Q_2 \right].$$

Since  $1 > \frac{K}{N} v_2$  and  $\partial U / \partial Q_2 \geq 0$ , an increase in  $Q_2$  increases the objective function. Further the left-hand side of (6.1) declines as  $Q_2$  increases and the right hand side increases as  $Q_2$  increases. Therefore (6.1) can not bind at an optimum. ■

Several questions now arise, including:

1. When is the interim-efficient mechanism for limited transfers the same as that for unlimited transfers?
2. Is the characterization in Theorem 6 still valid with limited transfers?

We already know when  $\lambda v_1 = v_2$ , as long as  $[0, 1] \subseteq A$ , the first-best output financed by a lottery-draft is an interim efficient mechanism. Limitations on transfers have no effect and the analysis of Section 3.1 remains valid. If  $\lambda v_1 \neq v_2$ , that is no longer true.

Suppose we were to solve (R) ignoring the transfer constraints (5.2). This is the unlimited transfer solution  $(Q^*, \alpha^*)$ . If  $\lambda v_1 < v_2$  and  $U(Q^*) \geq IC_U(Q^*)$  then  $Q^*$  also solves (R). If  $\lambda v_1 > v_2$  and  $L(Q^*) < IC_L(Q^*)$  then  $Q^*$  also solves (R). Therefore, one possibility is that transfer limitation will have no effect.

Next, define the (modified) *first-best solution* as that  $(Q, \alpha)$  which solves<sup>20</sup>

$$\max \pi Q_1 \left( \lambda - \frac{K}{N} v_2 \right) + (1 - \pi) Q_2 \left( 1 - \frac{K}{N} v_2 \right) + \pi (v_2 - \lambda v_1) \alpha_1$$

subject to  $Q \in Q_F$ ,  $\bar{\alpha} = \frac{K}{N} \bar{Q}$ , and  $L(Q) \leq \alpha_1 \leq U(Q)$ . Suppose  $\lambda v_1 < v_2$ ,  $Q^{**}$  is the modified first best solution, and  $U(Q^{**}) < IC_U(Q^{**})$ . Or alternatively  $\lambda v_1 > v_2$  and  $L(Q^{**}) \geq IC_L(Q^{**})$ . Then  $Q^{**}$  solves (R). This gives us a second class of possible solutions.

The intuition behind these two cases is straightforward. If, for example,  $\lambda v_1 < v_2$  then relatively more weight is placed on high cost types so that the optimal solution (without incentive constraints) makes  $\alpha_1$  as large as possible. If transfer constraints are tight and that transfer is incentive compatible then the modified first best solution is optimal and this is the end of the story. If, however, large transfers are possible then as  $\alpha_1$  is increased incentive-compatibility eventually bites. Once that happens, transfer limits become irrelevant. For  $\lambda v_1 < v_2$ , when  $IC_U(Q) < U(Q)$  the production  $Q$  must be financed subject to incentive compatibility.

There does remain, however, a third possibility<sup>21</sup> where, for  $\lambda v_1 < v_2$ ,

$$U(Q^*) < IC_U(Q^*)$$

---

<sup>20</sup>A true first best optimization would allow  $Q_1 < Q_2$ .

<sup>21</sup>A fourth possibility, where both

$$U(Q^*) \geq IC_U(Q^*)$$

and

$$U(Q^{**}) \leq IC_U(Q^{**})$$

with one strict inequality cannot happen. If it could, then

$$\begin{aligned} & \pi Q_1^* \left( \lambda - \frac{K}{N} v_2 \right) + (1 - \pi) Q_2^* \left( 1 - \frac{K}{N} v_2 \right) + \pi (v_2 - \lambda v_1) IC_U(Q^*) \\ & \geq \pi Q_1^{**} \left( \lambda - \frac{K}{N} v_2 \right) + (1 - \pi) Q_2^{**} \left( 1 - \frac{K}{N} v_2 \right) + \pi (v_2 - \lambda v_1) IC_U(Q^{**}) \end{aligned}$$

and

$$U(Q^{**}) > IC_U(Q^{**})$$

Consider the revised problems and solutions where  $\hat{Q}$  solves  $\max \pi Q_1 \left( \lambda - \frac{K}{N} v_2 \right) + (1 - \pi) Q_2 \left( 1 - \frac{K}{N} v_2 \right) + \pi(v_2 - \lambda v_1) IC_U(Q)$  subject to  $Q \in Q_F$  and  $IC_U(Q) \leq U(Q)$ , and where  $\tilde{Q}$  solves  $\max \pi Q_1 \left( \lambda - \frac{K}{N} v_2 \right) + (1 - \pi) Q_2 \left( 1 - \frac{K}{N} v_2 \right) + \pi(v_2 - \lambda v_1) U(Q)$  subject to  $U(Q) \leq IC_U(Q)$ . If either  $IC_U(\hat{Q}) < U(\hat{Q})$  or  $IC_U(\tilde{Q}) > U(\tilde{Q})$  then we are in a previous case. Therefore  $IC_U(\hat{Q}) = U(\hat{Q})$  and  $IC_U(\tilde{Q}) = U(\tilde{Q})$ . Thus,  $\hat{Q} = \tilde{Q}$ .

We can now answer our first question for  $\lambda v_1 < v_2$ .

**Theorem 11** (a) Let  $\hat{Q}$  solve

$$\max \pi Q_1 \left( \lambda - \frac{K}{N} v_2 \right) + (1 - \pi) Q_2 \left( 1 - \frac{K}{N} v_2 \right) + \pi(v_2 - \lambda v_1) IC_U(Q)$$

subject to  $Q \in Q_F$  and  $IC_U(Q) \leq U(Q)$ . If  $IC_U(\hat{Q}) < U(\hat{Q})$  then  $\hat{Q}$  is interim efficient and  $\hat{Q}$  is the same as if there were no limits on transfers.

(b) Let  $\tilde{Q}$  solve  $\max \pi Q_1 \left( \lambda - \frac{K}{N} v_2 \right) + (1 - \pi) Q_2 \left( 1 - \frac{K}{N} v_2 \right) + \pi(v_2 - \lambda v_1) U(Q)$  subject to  $Q \in Q_F$  and  $U(Q) \leq IC_U(Q)$ . If  $U(\tilde{Q}) < IC_U(\tilde{Q}) + \frac{1}{v_1} (1 - \pi) (\tilde{Q}_1 - \tilde{Q}_2)$  then  $\tilde{Q}$  is interim-efficient and  $\tilde{Q}$  is the first best solution (subject to  $Q_1 \geq Q_2$ ).

(c) If  $IC_U(\hat{Q}) = U(\hat{Q})$  in (a) and if  $IC_L(\tilde{Q}) = U(\tilde{Q})$  in (b), then  $\hat{Q} = \tilde{Q}$  is the interim efficient for  $\lambda$  and may be neither first best nor identical to the unlimited transfer solution.

(d)  $IC_U(\hat{Q}) < U(\hat{Q})$  and  $IC_L(\tilde{Q}) > U(\tilde{Q})$  not possible.

*Proof:* See above. ■

Part (a) of the theorem characterizes when the unlimited transfer solution remains unaffected by restrictions on transfers. For example, if a mechanism financed by lottery-draft is interim-efficient with unlimited transfers it remains so even with tight restrictions on transfers if  $[0, 1] \subseteq A$ . Part (b) identifies when the solution is first best, in the restricted sense given the transfer technology. Part (c) identifies when the solution differs from the unlimited transfer case in a substantial way.

Also,

$$\begin{aligned} & \pi Q_1^{**} \left( \lambda - \frac{K}{N} v_2 \right) + (1 - \pi) Q_2^{**} \left( 1 - \frac{K}{N} v_2 \right) + \pi(v_2 - \lambda v_1) U(Q^{**}) \\ & \geq \pi Q_1^* \left( \lambda - \frac{K}{N} v_2 \right) + (1 - \pi) Q_2^* \left( 1 - \frac{K}{N} v_2 \right) + \pi(v_2 - \lambda v_1) U(Q^*). \end{aligned}$$

Thus

$$IC_U(Q^*) - U(Q^*) \geq IC_U(Q^{**}) - U(Q^{**}) \geq 0$$

with one strict, which is a contradiction.

To answer the second question about the characterization in Theorem 6 we need to explore case (c) of the Theorem above more deeply. In particular, if  $\lambda v_1 < v_2$ ,  $IC_U(\hat{Q}) = U(\hat{Q})$ ,  $IC_U(\hat{Q}) = U(\hat{Q})$ , and  $\hat{Q} = \hat{Q}$  then can it be that  $\hat{Q} \in \text{Interior}(Q_F)$ ? The answer is no.

To see this, suppose  $\hat{Q}$  does not lie on the right boundary of  $Q_F$ . Then there exists a point in  $Q_F$ ,  $\hat{Q}' = (\hat{Q}'_1, \hat{Q}'_2)$  such that  $\hat{Q}'_1 > \hat{Q}_1$  and  $U(\hat{Q}') \geq U(\hat{Q})$  (Lemma 12). If  $[\hat{Q}', U(\hat{Q}')] satisfies the incentive constraint for low cost types ( $U(\hat{Q}') \leq IC_U(\hat{Q}')$ ) we are done since this will contradict  $\hat{Q}$  maximizing  $\pi Q_1 \left( \lambda - \frac{K}{N} v_2 \right) + (1 - \pi) Q_2 \left( 1 - \frac{K}{N} v_2 \right) + \pi(v_2 - \lambda v_1)U(Q)$  subject to  $U(Q) \leq IC_U(Q)$ . If  $[\hat{Q}', U(\hat{Q}')] violates incentive compatibility for low cost types, then compare the mechanisms  $[\hat{Q}', \hat{\alpha}']$  with  $\hat{\alpha}'_1 = U(\hat{Q}')$  versus  $[\hat{Q}, \hat{\alpha}]$  with  $\hat{\alpha}_1 = U(\hat{Q})$ . Since  $[\hat{Q}', \hat{\alpha}']$  violates incentive compatibility for low types and  $[\hat{Q}, \hat{\alpha}]$  satisfies that same constraint, it must that  $\hat{\alpha}'_2 < \hat{\alpha}_2$ . This implies that there exists a way to finance  $\hat{Q}'$  so that  $\hat{\alpha}'_2 = \hat{\alpha}_2$ , in which case  $(\hat{Q}, \hat{\alpha}')$  is incentive compatible. But since  $\hat{Q}' > \hat{Q}$  this means that  $\hat{\alpha}'_1 > U(\hat{Q})$ . Therefore, we compare two mechanisms  $\hat{Q}$  and  $\hat{Q}'$ , with$$

- (1)  $\hat{Q}_2 = \hat{Q}'_2$
- (2)  $\hat{\alpha}_2 = \hat{\alpha}'_2$
- (3)  $\hat{Q}_1 < \hat{Q}'_1$
- (4)  $\hat{\alpha}_1 < \hat{\alpha}'_1$

where  $(\hat{Q}, \hat{\alpha})$  satisfies (4.1) strictly and (4.2) is slack, and in  $(\hat{Q}', \hat{\alpha}')$  (4.1) is slack. Since  $\lambda v_1 < v_2$ ,  $(\hat{Q}', \hat{\alpha}')$  generates higher expected welfare than  $(\hat{Q}, \hat{\alpha})$ . Furthermore, for all  $0 < \epsilon < 1$ ,  $(\hat{Q}_\epsilon, \hat{\alpha}_\epsilon) = \epsilon(\hat{Q}, \hat{\alpha}) + (1 - \epsilon)(\hat{Q}', \hat{\alpha}')$  generates higher expected welfare than  $(\hat{Q}, \hat{\alpha})$ , and will satisfy (4.1) and (4.2) if  $\epsilon$  is close to 1.

**Theorem 12** *If  $[0, 1] \subseteq A^*$ , then for any parameters  $(K, N, \pi, v_1, v_2, \lambda)$  there exists  $j^*$  such that it is interim efficient to produce the public good iff the number of low types ( $v = v_1$ ) is  $\geq j^*$ .*

## 4 Extensions

### 4.1 Large Populations

We consider a sequence of economies indexed by the number of agents in the population. As  $N$  increases, the total amount of private good needed to produce the public good is denoted  $K(N) < N$ . We assume  $\limsup_{N \rightarrow \infty} \frac{K(N)}{N} = \liminf_{N \rightarrow \infty} \frac{K(N)}{N} \equiv k$  and assume that  $0 < k < 1$ .

There are three parts of the optimization problem that can change as  $N$  grows: the objective function, the feasible set, and the incentive constraints. We first consider the effects on the feasible set.

Because the right hand boundary of  $Q_F$  is piecewise linear for every  $N$ , the maximum difference  $Q_1 - Q_2$  is achieved at one of the vertices. This difference at a vertex is easy to compute. From equations (2) and (3) and the characterization of  $Q_F$  before Theorem 1 we can characterize the sequence of feasible sets  $\{Q_F^N\}$ . At each vertex except  $(0, 0)$ , we have, for some  $m \in \{0, \dots, N-1\}$

$$Q_1 = \sum_{j=m-1}^{N-1} \binom{N-1}{j} \pi^j (1-\pi)^{N-j-1}$$

$$Q_2 = \sum_{j=m}^{N-1} \binom{N-1}{j} \pi^j (1-\pi)^{N-j-1}$$

Recall that vertex  $m$  corresponds to a mechanism whereby the public good is produced if and only if there are at least  $m$  low-cost-types. The next theorem follows immediately.

**Theorem 13** *Fix  $\pi \in (0, 1)$ . For any  $\varepsilon > 0$ ,  $\exists N^*$  such that for all  $N \geq N^*$*

$$(Q_1, Q_2) \in Q_F^N \Rightarrow |Q_1 - Q_2| < \varepsilon.$$

*Proof:* From above, for every  $N$ ,

$$Q_1 - Q_2 \leq \max_{m=1, \dots, N-1} \binom{N-1}{m-1} \pi^{m-1} (1-\pi)^{N-m}$$

It is well known that the right hand side, which is the likelihood of the most likely number of successes out of  $N$  independent Bernoulli trials, converges to 0, as  $N$  gets larger. ■

Therefore, in the limit both types (interim) expect the same public good production, i.e.  $Q_F$  converges to the set of all convex combinations between full production ( $Q = 1$ ) and no production ( $Q = 0$ ). It is important to note that this result depends only on the feasibility of allowing output to depend on the number of low types. *The result does not even depend on  $K(N)$ .* It says nothing about whether the free rider problem gets better or worse in  $N$ , since incentive constraints play no role.

By the linearity of the objective function, the interim efficient output will be an extreme point. If  $(\lambda\pi + 1 - \pi) < k[\lambda\pi v_1 + (1 - \pi)v_2]$  then, in the limit, the interim efficient mechanism will have zero production. If  $(\lambda\pi + 1 - \pi) > k[\lambda\pi v_1 + (1 - \pi)v_2]$  then, in the limit, the interim efficient mechanism will have full production. This is essentially a law of large numbers result. The proportion of  $v_1$ -types converges to  $\pi$  almost surely

in  $N$ . Consequently, in the limit economy the optimal  $Q$  is known to be either 0 or 1, depending on the direction of the above inequality.<sup>22</sup> Since  $Q_1 - Q_2$  converges to 0 in the limit, incentive compatibility then implies that, in the limit, for any feasible  $Q$ , both types must have the same expected transfer. Optimality then implies that either  $\alpha_1 = \alpha_2 = 0$  or  $\alpha_1 - \alpha_2 = k$ , depending on whether  $Q$  is 0 or 1. Thus, incentive compatibility and feasibility imply that in the limit only lottery drafts (or equal payments if the endowments are divisible) are possible.<sup>23</sup>

## 4.2 Individual Rationality

A natural constraint to impose on the choice of a mechanism is that, given the rules  $(Q, \alpha)$  and given a player's observation of type,  $v_i$ , the player should agree to play the game. This leads to what are called *interim individual rationality constraints*. One interim individual rationality constraint is  $Q_2 \geq v_2 \alpha_2$ . By incentive compatibility, whenever this constraint is satisfied, low-cost types  $v_1$  will also automatically satisfy interim individual rationality, since  $Q_1 - v_1 \alpha_1 \geq Q_2 - v_1 \alpha_2 \geq Q_2 \geq 0$ . More specifically, since the incentive compatibility constraint, for  $\bar{\alpha} = \frac{K}{N} \bar{Q}$ , is  $IC_L(Q) \leq \alpha_1 \leq IC_U(Q)$ , the interim individual rationality constraint for high-cost types may be written:

$$Q_2 - v_2 \left( \frac{1}{1-\pi} \right) \left( \frac{K}{N} \bar{Q} - \pi \alpha_1 \right) \geq 0$$

or

$$\alpha_1 \geq -\frac{1}{v_2} Q_2 \frac{1-\pi}{\pi} + \frac{K}{N} \left( \frac{1}{\pi} \bar{Q} \right).$$

It is possible to find  $\alpha_1$  satisfying all of this if and only if

$$(IR) \quad Q_1 \left( \frac{1}{v_1} - \frac{K}{N} \right) + Q_2 \left[ \left[ \frac{1}{v_2} - \frac{K}{N} \right] - \pi \left[ \frac{1}{v_1} - \frac{K}{N} \right] \right] \geq 0.$$

Thus as long as  $\frac{K}{N} v_2 < 1$ , incentive compatibility guarantees that  $(IR)$  is not binding since  $Q_1 \geq Q_2 \Rightarrow (IR)$ . Notice that  $\frac{K}{N} v_2 < 1$  means that it would be ex post efficient to produce even if all  $N$  players had high costs.

Geometrically  $(IR)$  defines a line through the origin in  $(Q_1, Q_2)$  space, with the  $(IR)$  constraint requiring  $Q$  to lie to the lower right of this line. If  $\frac{K}{N} v_2 < 1$ , then the line has slope greater than 1 and lies above  $Q_F$ . If  $\frac{K}{N} v_2 > 1$ , then the line has slope less than 1 and cuts through  $Q_F$ . If  $IR$  cuts the boundary of  $Q_F$  above the unconstrained solution,  $Q^*$ , it is not binding. If  $IR$  cuts the boundary below  $Q^*$ , it is binding and results in *lower* production (i.e.  $Q_{IR}^* < Q^*$ ). See Figure 3. In no case will  $IR$  lead to greater production than  $Q^*$ . Notice that if the slope of  $IR$  is less than 1, the only mechanism that treats both types the same ( $Q_1 = Q_2$ ) involves zero production. Therefore in

<sup>22</sup>In the borderline case where  $k = \frac{\lambda\pi+1-\pi}{\lambda\pi v_1+(1-\pi)v_2}$  then in the limit every  $Q$  is optimal.

<sup>23</sup>More precisely, only mechanisms that generate the same interim expected utilities as a lottery draft are possible in the limit.

large populations, individual rationality constraints drive the optimal solution to zero production ( $Q_1 = Q_2 = 0$ ) whenever  $\lim_{N \rightarrow \infty} \frac{K(N)}{N} v_2 > 1$ .

This provides a simple geometric interpretation for the findings in Mailath and Postlewaite [1990]<sup>24</sup> and Rob [1989]: in Bayesian public goods environments with independent types and some other minor restrictions (that rule out cases similar to  $\lim_{N \rightarrow \infty} \frac{K(N)}{N} v_2 \leq 1$ ), the probability the public good will be produced goes to zero in the number of agents if interim individual rationality is imposed. Feasibility implies that, for any mechanism  $Q_1 \approx Q_2$  when  $n$  is large. By incentive compatibility,  $Q_1 \approx Q_2$  implies  $\alpha_1 \approx \alpha_2$ . By individual rationality,  $1 > kv_2$  implies  $Q_1 \approx Q_2 = 0$  is the only possible mechanism.

## 5 Concluding Remarks

The central finding of the paper is that interim efficient public goods provision is always of a particularly simple form: Given a welfare criterion  $\lambda$ , the public good should be produced if the fraction of low cost types in the population is greater than or equal to  $\frac{j^*(\lambda)}{N}$ . The result continues to hold up when there are restricted transfers, even if transfers are entirely ruled out.<sup>25</sup>

Furthermore,  $j^*(\lambda)$  is decreasing in  $\lambda$ , the welfare weight on low-cost types. This has a intuitive interpretation. There is a natural benchmark,  $\lambda = \frac{v_2}{v_1}$ , where total welfare is neutral with respect to transfers and the public good is produced if and only if the sum of the benefits exceed its total production cost. Call this “efficient production.” When  $\lambda$  increases relative to this, there is a welfare gain by shifting the burden of the cost onto the high cost types, and by over-producing relative to efficient production (since low types favor production). When  $\lambda$  decreases, there is a welfare gain to shifting the burden onto low-cost types and (when  $v_2$  is sufficiently high) to reducing production of the public good. In many ways the latter case ( $\lambda < \frac{v_2}{v_1}$ ) is a useful case to focus on. It corresponds to the common welfare criterion that “those who benefit the most should pay the most.” Our result establishes that doing so entails a production inefficiency in the form of restricted output of the public good.

A second set of findings relates to practical aspects of mechanism design. This is easily seen in the case of a very simple optimal mechanism when  $\lambda = \frac{v_2}{v_1}$ . In that case the optimal solution can be implemented by simply having a vote between production and non-production. If the number voting in favor exceeds the critical level,  $j^*(\frac{v_2}{v_1})$ , corresponding to production efficiency, then the good is financed by randomly selecting  $K$  of  $N$  individuals to share the cost equally, a (conditional) **lottery draft** mechanism. Incentive compatibility implies that an equilibrium of this simple mechanism is for all

<sup>24</sup>This is also closely related to an earlier result in Guth and Hellwig [1986, Prop. 4.5].

<sup>25</sup>When there are restrictions on transfers, then the good may also be produced with some probability  $0 < q < 1$  when the fraction of low-cost types exactly equals  $\frac{j^*(\lambda)}{N}$ .

low types to vote “yes” and all high types to vote “no.”<sup>26</sup>

Only slightly more complicated mechanisms are needed for other  $\lambda \neq \frac{v_2}{v_1}$  and in the case of transfer limitations. A  $\frac{j^*}{N}$  majority rule can still be used to determine whether the public good is to be produced, but a random tie-breaking rule when the number of Yes votes exactly equals  $j^*$  may be needed when there are restricted transfers. In addition, lottery drafts are not generally optimal financing schemes unless  $j^* = 0$ ,  $\lambda = \frac{v_2}{v_1}$ , or  $j^* \geq K$ . In the last of these cases, then lottery draft financing may be optimal when the  $U(Q)$  transfer constraint is binding. In that case the lottery selects  $K$  out of the  $j^*$  yes voters. The optimality of these simple mechanisms hinges on our assumption of only two types. More complicated mechanisms will typically be required to achieve optimality with more types, at least in small populations.

This last kind of lottery is a special case of a more general type of financing that is optimal with limited transfers when the  $U(Q)$  transfer constraint is binding. We refer to this as “least cost,” and it involves always producing the public good in a way that taxes high-cost types as little as possible, and low-cost types as much as possible. Theorem 11 establishes that if the transfer constraints are binding, then the solution is either **first-best** (i.e. incentive constraints are not binding) or **least cost**.

An important feature of all of these simple 2-stage mechanisms is that the stages separate the production decision (the voting stage) from the cost allocation decision (the lottery, or transfer, stage). In order for this separation to work, it may be important that the planners have confiscatory power. If  $v_2$  is large (or if  $\alpha_1$  or  $\alpha_2$  is large), then some players might not wish to make the transfers required in the second stage, or may not be willing to partake in the lottery. These *ex post* individual rationality problems are severe. In fact the last section of the paper shows the comparatively mild constraint of **interim** individual rationality generally leads to gross inefficiency in large populations. While we do not work out the details of the stronger *ex post* participation constraints on mechanisms, it is clear that a similar, negative result for large populations will emerge since *ex post* constraints are stronger than interim constraints.

One extension would be to allow for continuous rather than discrete production technologies and continuous levels of contribution. As long as we restrict attention to *linear* technologies the main results remain unchanged.

A second direction would be to allow for continuous types. As noted earlier, the characterization of optimal mechanisms relies on the restriction to two types. Elsewhere (Ledyard and Palfrey [1989]) we have begun to investigate the continuous type case when a regularity condition on the distribution of types is satisfied. That paper derives analogous versions of Corollary 7 and Corollary 8. When the continuous version of the inequality in Corollary 7 or Corollary 8 fails to hold, “feasible implementation” constraints on the reduced form corresponding to  $Q_F$ ,  $U(Q)$  and  $L(Q)$  play a role. Preliminary work

---

<sup>26</sup>Observe how much simpler this is compared with the very complicated schemes based on Clarke-taxes. (See, for example, Laffont and Maskin [1982] p. 76.)

on this indicates that those feasibility constraints are closely related to the conditions found in the auction literature (Matthews [1984], Maskin and Riley [1984], and Border [1991]).

Finally, a more difficult set of issues arise if we drop the assumption of quasi-linear preferences. The existence of income effects, arising from non-separability or from risk aversion due to nonlinearities in the valuation of the private good, poses formidable technical difficulties in the analysis.

## Appendix: Some Lemmata

We establish a few simple but useful properties about the set of feasible and incentive compatible mechanisms.

**Lemma 1:** Let  $(Q, \alpha) = \{(Q^i, \alpha^i)\}_{i=1}^N$  and  $(Q', \alpha') = \{(Q'^i, \alpha'^i)\}_{i=1}^N$  be two feasible and incentive compatible (but not necessarily symmetric) mechanisms. Then  $p \in [0, 1] \Rightarrow [pQ + (1 - p)Q', p\alpha + (1 - p)\alpha']$  is also feasible and incentive compatible.

*Proof:* Incentive compatibility of the new mechanism follows from the linearity of utility in  $v_i, \alpha_i$  and  $Q_i$ . Feasibility follows because one can produce the new allocation rule as a compound lottery in the following way. First, everyone reports their types; then the planner uses a public randomizing device to choose  $(Q, \alpha)$  with probability  $p$  and  $(Q', \alpha')$  with probability  $(1 - p)$ . ■

**Lemma 2:** Let  $P = \{(Q_i, \alpha_i)\}_{i=1}^N$  be feasible and let

$$\omega = \sum_{i=1}^N \frac{1}{N} \left[ \lambda \pi V^i(Q_1^i, \alpha_1^i, v_1) + (1 - \pi) V^i(Q_2^i, \alpha_2^i, v_2) \right].$$

There exists a feasible symmetric mechanism  $\hat{P} = (\hat{Q}, \hat{\alpha})$  such that  $\pi \lambda (\hat{Q}_1 - v_1 \hat{\alpha}_1) + (1 - \pi) (\hat{Q}_2 - v_2 \hat{\alpha}_2) = \omega$ . Furthermore, if  $P$  is incentive compatible then  $\hat{P}$  is incentive compatible.

*Proof:* By the symmetry of the problem (i.e. the form of everyone's utility function is the same and everyone has symmetric priors)  $\{Q^{\rho(i)}, \alpha^{\rho(i)}\}_{i=1}^N$  is feasible and incentive compatible for every  $\rho : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$  that is one-to-one (i.e. every permutation operator). By the convexity property established in Lemma 3, every mechanism in the convex hull of the set of all such permutations of  $P$  is feasible and incentive compatible. Let  $\hat{P}$  be the equally weighted [by  $\frac{1}{N!}$ ] convex combination of the permutations. Then  $\hat{Q}^i = \frac{1}{N} \sum_{i=1}^N Q^i$  for all  $i$  and  $\hat{\alpha}^i = \frac{1}{N} \sum \alpha^i$  for all  $i$ , so  $\hat{P}$  is symmetric both mechanisms generate the same ex ante social value. ■

In view of Lemma 1 and Lemma 2, we can restrict attention to symmetric mechanisms. The next lemma is standard for incentive compatible mechanisms.

**Lemma 3:** Let  $v_1 < v_2$ . If  $(Q, \alpha)$  is incentive compatible then:

$$Q_1 \geq Q_2 \tag{A.1}$$

$$\alpha_1 \geq \alpha_2 \tag{A.2}$$

$$V(Q, \alpha, v_1) \geq V(Q, \alpha, v_2) \tag{A.3}$$

*Proof:* By incentive compatibility we have:

$$V(Q, \alpha, v_1) \geq Q_2 - v_1\alpha_2$$

$$V(Q, \alpha, v_2) \geq Q_1 - v_2\alpha_1$$

Since  $v_1 < v_2$ , we get

$$Q_1 - v_1\alpha_1 \geq Q_2 - v_1\alpha_2 \geq Q_2 - v_2\alpha_2 \geq Q_1 - v_2\alpha_1$$

Therefore, (A.3) follows immediately; (A.4) follows from

$$[Q_1 - v_1\alpha_1] - [Q_1 - v_2\alpha_1] \geq [Q_2 - v_1\alpha_2] - [Q_2 - v_2\alpha_2].$$

(A.1) then follows from (A.4) and from  $Q_1 \geq Q_2 + v_1[\alpha_1 - \alpha_2]$ . We need to show that efficiency requires balanced budgets. ■

**Lemma 4:** If  $(Q, \alpha)$  is feasible then

$$\bar{\alpha} \geq \frac{K}{N}\bar{Q}$$

*Proof:* Suppose not. Then contributions are on average insufficient to produce the output on the right hand side of the inequality. This means that for some  $j$ ,  $ja_j^1 + (N-j)a_j^2 < Kq_j$ , which violates feasibility. ■

**Lemma 5:** Suppose  $(a^1, a^2, q)$  is feasible. Then  $ja_j^1 + (N-j)a_j^2 = Kq_j$  for all  $j$  if and only if  $\bar{\alpha} = \frac{K}{N}\bar{Q}$ .

*Proof:* Follows immediately from the feasibility constraint that  $ja_i^1 + (N-j)a_j^2 \geq Kq_j$  and from equations (2) and (3). ■

**Lemma 6:** If  $(\alpha, Q)$  solves  $(P)$  then  $\bar{\alpha} = \frac{K}{N}\bar{Q}$ , where  $\bar{\alpha} = \pi\alpha_1 + (1-\pi)\alpha_2$  and  $\bar{Q} = \pi Q_1 + (1-\pi)Q_2$ . That is interim efficient mechanisms do not waste resources.

*Proof:* Suppose  $(\alpha, Q)$  solves  $(P)$  but that  $\bar{\alpha} > \frac{K}{N}\bar{Q}$ . By Lemmas 4 and 5, there is some  $k' \in \{0, \dots, N\}$  such that

$$k'a_{k'}^1 + (N-k')a_{k'}^2 > kq_{k'}.$$

We will show this implies  $(Q, \alpha)$  cannot solve  $(P)$ . We consider four cases.

*Case 1:* At the optimum  $(Q, \alpha)$ ,  $Q_1 - v_1\alpha_1 > Q_2 - v_1\alpha_2$  and  $Q_2 - v_2\alpha_2 > Q_1 - v_2\alpha_1$ . Then at the optimum neither (4.1) nor (4.2) is binding. Therefore, increase  $q_{k'}$  a little or lower one of  $a_{k'}^1$ , or  $a_{k'}^2$  a little. This is not possible only if  $q_{k'} = 1$ ,  $a_{k'}^1 = \max\left\{a, \frac{K - (N-k)\bar{\alpha}}{k'}\right\}$

and  $a_{k'}^2 = \max\left\{\underline{a}, \frac{K-k\bar{a}}{k'}\right\}$ . If  $a_{k'}^1 = \max\left\{\underline{a}, \frac{K-(N-k)\bar{a}}{k'}\right\}$  then  $a_{k'}^2 = \bar{a} > \underline{a}$ . Therefore,  $a_{k'}^2 = \frac{K-k'\bar{a}}{N-k'}$  and  $a_{k'}^1 = \bar{a}$ . Thus,

$$k'a_{k'}^1 + (N-k')a_{k'}^2 = K - (N-k')\bar{a} + K - k'\bar{a} = 2K - N\bar{a}$$

and

$$k'a_{k'}^1 + (N-k')a_{k'}^2 = k'\bar{a} + (N-k')\bar{a} = N\bar{a}.$$

Therefore  $N\bar{a} = K$ . But  $N\bar{a} > K$ . If  $a_{k'}^1 = \underline{a}$  then  $a_{k'}^2 = \underline{a}$  and

$$kq_{k'} < k'a_{k'}^1 + (N-k')a_{k'}^1 = N\underline{a} \leq 0.$$

But  $q_{k'} \geq 0$ , which is a contradiction.

*Case 2:* At the optimum,  $Q_1 - \alpha_1 v_1 = Q_2 - \alpha_2 v_1$  and  $Q_2 - \alpha_2 v_2 = Q_1 - \alpha_1 v_2$ . Thus  $Q_1 = Q_2 = \bar{Q}$  and  $\alpha_1 = \alpha_2 = \bar{\alpha}$ . Suppose that  $\bar{\alpha} > \frac{K}{N}\bar{Q}$ . Let  $\hat{\alpha}_1 = \hat{\alpha}_2 = \frac{K}{N}\bar{Q}$ .  $(Q, \hat{\alpha})$  is feasible since one can let  $a_k^2 = a_k^1 = \frac{K}{N}q_n \forall k$ . Then

$$\begin{aligned} \hat{E} &= \lambda\pi[Q_1 - v_1\hat{\alpha}_1] + (1-\pi)[Q_2 - v_2\hat{\alpha}_2] = \\ &(\lambda\pi + (1-\pi))\bar{Q} - (\lambda\pi v_1 + (1-\pi)v_2)\bar{Q}\frac{K}{N} > \\ &(\lambda\pi + (1-\pi))\bar{Q} - (\lambda\pi v_1 + (1-\pi)v_2)\bar{\alpha}. \end{aligned}$$

Therefore  $(Q, \alpha)$  was not optimal.

*Case 3:* At the optimum,  $(Q_1 - Q_2) < v_2(\alpha_1 - \alpha_2)$  and  $(Q_1 - Q_2) = v_1(\alpha_1 - \alpha_2)$ . A decrease in  $\alpha_1$ , keeping  $Q_1, Q_2, \alpha_2$  constant will improve  $E$  and satisfy (4) if feasible. If  $k' > 0$  and  $a_{k'}^1 > q_k \max\left\{\underline{x}, \frac{K-(N-k)\bar{x}}{k}\right\} + (1-q_k) \max\left\{\underline{x}, \frac{-(N-k)\bar{x}}{k}\right\}$  then such a change is feasible by reducing  $a_{k'}^1$ . If  $k' = 0$  or  $a_{k'}^1 = q_k \max\left\{\underline{x}, \frac{K-(N-k)\bar{x}}{k}\right\} + (1-q_k) \max\left\{\underline{x}, \frac{-(N-k)\bar{x}}{k}\right\}$ , then  $a_{k'}^2 > q_k \max\left\{\underline{x}, \frac{K-k\bar{x}}{N-k}\right\} + (1-q_k) \max\left\{\underline{x}, \frac{-k\bar{x}}{N-k}\right\}$  so we need to show that one can feasibly reduce  $a_{k'}^2$ , and for some  $k'' \neq k'$  ( $k'' < N$ ) increase  $a_{k''}^2$  in such a way that  $\alpha_2$  remains unchanged, and at the same time reduce  $a_{k''}^1$ . Such a  $k''$  exists unless, for all  $k \neq k'$  ( $k < N$ ) either

$$(1) \quad a_k^1 = q_k \max\left\{\underline{x}, \frac{K-(N-k)\bar{x}}{k}\right\} + (1-q_k) \max\left\{\underline{x}, \frac{-(N-k)\bar{x}}{k}\right\}, \quad k > 0$$

$$\text{or } (2) \quad a_k^2 = q_k \min\left\{\bar{x}, \frac{K-kx}{N-k}\right\} + (1-q_k) \min\left\{\bar{x}, \frac{-kx}{N-k}\right\}$$

$$\text{or } (3) \quad k = 0, \quad a_{k'}^1 = q_k \max\left\{\underline{x}, \frac{K-(N-k)\bar{x}}{k}\right\} + (1-q_k) \max\left\{\underline{x}, \frac{-(N-k)\bar{x}}{k}\right\}$$

If (1), then  $a_k^1 \leq q_k \frac{K}{k}$ , so  $(N-k)a_k^2 \geq Kq_k$ . If (2) or (3), then  $a_k^1 \leq a_k^2$ . Therefore,  $Na_k^2 \geq ka_k^1 + (N-k)a_k^2 \geq Kq_k$  for  $k = 0, 1, \dots, N-1$  and  $Na_k^2 > k'q_{k'}$  which implies  $\alpha_2 > \frac{K}{N}Q_2$ . But  $v_1(\alpha_1 - \alpha_2) = (Q_1 - Q_2) < v_2(\alpha_1 - \alpha_2)$  implies  $\alpha_1 > \alpha_2$ . Therefore

$\alpha_1 > \alpha_2 > \frac{K}{N}Q_2$ . If  $\alpha_1 \geq \frac{K}{N}Q_1$  then let  $\hat{\alpha}_1 = \frac{K}{N}Q_1$ , and  $\hat{\alpha}_2 = \frac{K}{N}Q_2$ . Then  $(Q, \hat{\alpha})$  yields a higher value for  $E$  than  $(Q, \alpha)$  and we have a contradiction as long as  $(Q, \alpha)$  is incentive compatible. If  $(Q, \hat{\alpha})$  is not incentive compatible, then for some  $\epsilon$  close to 0,  $(Q, \hat{\alpha}_\epsilon)$  is incentive compatible, where  $\hat{\alpha}_\epsilon = \epsilon\hat{\alpha} + (1 - \epsilon)\alpha$ . Since  $(Q, \hat{\alpha}_\epsilon)$  yields a higher value of  $E$  than  $(Q, \alpha)$ , we have a contradiction. Thus  $\frac{K}{N}Q_1 \geq \alpha_1 > \alpha_2 > \frac{K}{N}Q_2$ . Thus

$$\frac{K}{N}(Q_1 - Q_2) > \alpha_1 - \alpha_2 \quad \text{or} \quad v_1 \frac{K}{N}(Q_1 - Q_2) > v_1(\alpha_1 - \alpha_2)$$

But  $1 > \frac{K}{N}v_1$ . Therefore  $(Q_1 - Q_2) > v_1(\alpha_1 - \alpha_2)$ , a contradiction.

*Case 4:* At the optimum  $(Q_1 - Q_2) = v_2(\alpha_1 - \alpha_2)$  and  $(Q_1 - Q_2) > v_1(\alpha_1 - \alpha_2)$ . A decrease in  $\alpha_2$ , keeping  $Q$  and  $\alpha_1$  fixed will improve  $E$  and satisfy (4) if feasible. If  $k' < N$  and  $a_{k'}^2 > q_k \max \left\{ \underline{x}, \frac{K - k\bar{x}}{N - k} \right\} + (1 - q_k) \max \left\{ \underline{x}, \frac{-k\bar{x}}{N - k} \right\}$  then such a change is feasible by reducing  $a_{k'}^2$ . If  $k' = N$  or  $a_{k'}^2 = q_k \max \left\{ \underline{x}, \frac{K - k\bar{x}}{N - k} \right\} + (1 - q_k) \max \left\{ \underline{x}, \frac{-k\bar{x}}{N - k} \right\}$  then  $a_{k'}^1 > q_k \max \left\{ \underline{x}, \frac{K - (N - k)\bar{x}}{k} \right\} + (1 - q_k) \max \left\{ \underline{x}, \frac{-(N - k)\bar{x}}{k} \right\}$ . So we need to show that one can feasibly lower  $a_{k'}^1$  and, for some  $k'' \neq k'$  ( $k'' > 0$ ), increase  $a_{k''}^1$  in such a way that  $\alpha_1$  remains unchanged, and at the same time reduce  $a_{k''}^2$ . Such a  $k''$  exists, unless for all  $k \neq k'$  ( $k > 0$ ) either:

$$\begin{aligned} (1) \quad & a_k^2 = q_k \max \left\{ \underline{x}, \frac{K - k\bar{x}}{N - k} \right\} + (1 - q_k) \max \left\{ \underline{x}, \frac{-k\bar{x}}{N - k} \right\}, \quad k < N \\ \text{or } (2) \quad & a_k^1 = q_k \min \left\{ \bar{x}, \frac{K - (N - k)\underline{x}}{k} \right\} + (1 - q_k) \min \left\{ \bar{x}, \frac{-(N - k)}{k} \right\} \\ \text{or } (3) \quad & k = N, a_{k'}^1 = q_k \max \left\{ \underline{x}, \frac{K - (N - k)\bar{x}}{k} \right\} + (1 - q_k) \max \left\{ \underline{x}, \frac{-(N - k)\bar{x}}{k} \right\}. \end{aligned}$$

If (1), then  $a_k^2 \geq \frac{K}{N - k}$  so  $ka_k^1 \geq Kq_k$ . If (2) or (3), then  $a_k^2 \leq a_k^1$ . Following a similar argument to Case 3, we get

$$\begin{aligned} \frac{K}{N}(Q_1 - Q_2) &< \alpha_1 - \alpha_2 \quad \text{or} \\ v_2 \frac{K}{N}(Q_1 - Q_2) &< v_2(\alpha_1 - \alpha_2) \end{aligned}$$

If  $1 \leq v_2 \frac{K}{N}$  then  $(Q_1 - Q_2) < v_2(\alpha_1 - \alpha_2)$ , a contradiction. If  $1 > \frac{K}{N}v_2$  then  $Q_2 = Q_1$  at the optimum. But then  $\alpha_1 = \alpha_2$ , a contradiction. QED. ■

Finally we provide a very useful lemma about feasible and efficient mechanisms with limited transfers which allows us to work entirely in the space of reduced-form mechanisms. First, given a feasible  $Q = (Q_1, Q_2)$ , we define two values,  $U(Q)$  and  $L(Q)$ , which correspond to the maximum and minimum expected payment of a low type.

$U(Q) = \max \alpha_1$  subject to (2), (3.1) and for  $k = 0, \dots, N$ ,  $q_k \in [0, 1]$  and

$$(*) \quad a_k^1 \leq q_k \min \left\{ \bar{x}, \frac{K - (N - k)\underline{x}}{k} \right\} + (1 - q_k) \min \left\{ \bar{x}, \frac{-(N - k)\underline{x}}{k} \right\}.$$

$L(Q) = \min \alpha_1$  subject to (2), (3.1) and for  $k = 0, \dots, N$ ,  $q_k \in [0, 1]$  and

$$(**) \quad a_k^1 \leq q_k \max \left\{ \underline{x}, \frac{K - (N - k)\bar{x}}{k} \right\} + (1 - q_k) \max \left\{ \underline{x}, \frac{-(N - k)\bar{x}}{k} \right\}.$$

**Lemma 7:** Suppose  $\bar{\alpha} = \frac{K}{N}\bar{Q}$ .  $(Q, \alpha)$  is feasible if and only if  $Q$  is feasible and  $L(Q) \leq \alpha_1 \leq U(Q)$ .

*Proof:* (only if) Since  $(Q, \alpha)$  is feasible,  $\exists$  feasible  $(q, a^1, a^2)$  such that  $ka_k^1 + (N - k)a_k^2 = Kq_k$  for  $k = 0, \dots, N$ . Since  $\bar{\alpha} = \frac{K}{N}\bar{Q}$ ,  $a_k^1 = \frac{Kq_k - (N - k)a_k^2}{k}$  for all  $k$ . Since  $(a, q)$  is feasible (1.5) and (1.6) are true. (1.6) implies  $(**)$  while (1.5) implies  $(*)$ . Thus,  $L(Q) \leq \alpha_1 \leq U(Q)$ .

(if) Given  $Q$  feasible, we know by definition that  $(Q, \alpha^U)$  and  $(Q, \alpha^L)$  are feasible where  $\alpha_1^U = U(Q)$ ,  $\alpha_1^L = L(Q)$ ,  $\alpha_2^U = \left[ \frac{K}{N}\bar{Q} - \pi\alpha_1^U \right] \left( \frac{1}{1 - \pi} \right)$ , and  $\alpha_2^L = \left[ \frac{K}{N}\bar{Q} - \pi\alpha_1^L \right] \left( \frac{1}{1 - \pi} \right)$ . That is, there are feasible  $[q^U, a^{1U}, a^{2U}]$  which with  $(Q, \alpha^U)$  satisfy (2) and (3). Similarly, there are feasible  $[q^L, a^{1L}, a^{2L}]$  which with  $(Q, \alpha^L)$  satisfy (2) and (3). Thus if  $Q$  is feasible and  $L(Q) \leq \alpha_1 \leq U(Q)$  and  $\bar{\alpha} = \frac{K}{N}\bar{Q}$  then there is a  $\mu \in [0, 1]$  such that  $\alpha_1 = \mu L(Q) + (1 - \mu)U(Q)$ . Therefore,  $\mu[q^U, a^{1U}, a^{2U}] + (1 - \mu)[q^L, a^{1L}, a^{2L}]$  is feasible and, together with  $(Q, \alpha)$ , satisfies (2) and (3). ■

**Lemma 8:** Let  $S$  be the slope of the indifference lines of the objective function for the maximization problems in Theorems 2, 3, 4.  $\partial S / \partial \lambda \geq 0$  or  $Q^* = 1$ .

*Proof:*  $\lambda \leq v_2/v_1$  implies

$$S = -\frac{\partial E / \partial Q_1}{\partial E / \partial Q_2} = \frac{\pi \left( \frac{1}{v_1} - \frac{K}{N} \right) (\lambda v_1 \pi + (1 - \pi) v_2)}{-(1 - \pi) \left[ \left( \frac{1}{v_1} - \frac{K}{N} \right) (\lambda v_1 \pi + (1 - \pi) v_2) - \frac{(v_2 - v_1)}{v_1} \right]}$$

where the denominator is  $> 0$ . Differentiation and simplification give:

$$\text{sgn}(\partial S / \partial \lambda) = \text{sgn} \left[ \frac{(v_2 - v_1)}{v_1} \right] > 0.$$

For  $\lambda > \frac{v_2}{v_1}$ , if  $1 \geq \frac{K}{N}v_2$  then  $Q^* = 1$  and we are done. For  $1 < \frac{K}{N}v_2$ ,

$$S = \frac{\pi \left[ \left( \frac{1}{v_2} - \frac{K}{N} \right) (\lambda v_1 \pi + (1 - \pi) v_2) + \frac{\lambda}{v_2} (v_2 - v_1) \right]}{-(1 - \pi) \left( \frac{1}{v_2} - \frac{K}{N} \right) (\lambda v_1 \pi + (1 - \pi) v_2)}.$$

Differentiation and simplification give:

$$\text{Sgn} \left( \frac{\partial S}{\partial \lambda} \right) = \text{Sgn} \left[ - \left( \frac{1}{v_2} - \frac{K}{N} \right) (v_2 - v_1)(1 - \pi) \right] > 0.$$

■

**Lemma 9:** If  $[0, 1] \in A^*$  and  $1 > \frac{K}{N}v_2$ , then for all  $Q \in Q_F$ ,

$$L(Q) \leq \frac{K}{N}Q_1 \leq IC_L(Q) \leq IC_U(Q)$$

*Proof:* The last inequality follows from  $v_1 \leq v_2$  and  $Q_1 \geq Q_2$ , since  $Q \in Q_F$ . Now at the solution to  $L(Q)$ ,

$$a_k^1 = q_k \max \left\{ \underline{x}, \frac{K - (N - k)}{k} \right\} + (1 - q_k) \max \left\{ \underline{x}, \frac{-(N - k)}{k} \right\} \leq \frac{K}{N}q_k.$$

Thus  $L(Q) \leq \frac{K}{N}Q_1$ . Finally  $\frac{K}{N}Q_1 \leq IC_L(Q)$  if and only if  $0 \leq \left( \frac{1}{v_2} - \frac{K}{N} \right) (1 - \pi)(Q_1 - Q_2)$  if and only if  $1 \geq \frac{K}{N}v_2$ . ■

**Lemma 10:** If  $[0, 1] \in A^*$  and  $1 > \frac{K}{N}v_2$ , then for all  $Q \in Q_F$ ,

$$U(Q) \geq \frac{K}{N}Q_1 \geq IC_L(Q)$$

and

$$L(Q) \leq \frac{K}{N}Q_1 \leq IC_U(Q).$$

*Proof:* Since  $[0, 1] \in A^*$ ,  $a_k^1 > \frac{K}{N}q_k$  is always feasible. Thus  $U(Q) \geq \frac{K}{N}Q_1 \geq L(Q)$ . The rest follows from  $Q_1 \geq Q_2$ ,  $v_1 \leq v_2$ ,  $\frac{K}{N}v_1 \leq r \leq \frac{K}{N}v_2$ . ■

**Lemma 11:** Let  $\eta = (a, q) = [a_0^1, \dots, a_N^1, a_0^2, \dots, a_N^2, q_0, \dots, q_N]$  as in Section 2.2.  $\eta$  is feasible if and only if

$$q_k \in [0, 1] \tag{1.1}$$

$$ka_k^1 + (N - k)a_k^2 \geq Kq_k \tag{1.2}$$

$$a_0^1 = 0, a_N^2 = 0 \tag{1.3}$$

$$\text{and } a_k^i \in A^* = CO(X) \text{ for } i = 1, 2 \tag{1.4}$$

and, for  $k = 1, \dots, N - 1$ ,

$$a_k^1 \geq q_k \max \left\{ \underline{x}, \frac{K - (N - k)\bar{x}}{k} \right\} + (1 - q_k) \max \left\{ \underline{x}, \frac{-(N - k)\bar{x}}{k} \right\} \tag{1.5}$$

$$a_k^2 \geq q_k \max \left\{ \underline{x}, \frac{K - k\bar{x}}{N - k} \right\} + (1 - q_k) \max \left\{ \underline{x}, \frac{-k\bar{x}}{N - k} \right\} \tag{1.6}$$

*Proof:* We need to show; there is a feasible symmetric direct mechanism  $\mu$  yielding  $\eta = (a^1, a^2, q)$  if and only if (1.1)-(1.6). The set of feasible symmetric direct mechanisms when there are two types is  $S = \{\mu : \{0, \dots, N\} \rightarrow M(Z_1)\}$  where  $\mu(k)$  is the measure  $Z$ , the feasible  $(x, y)$ , when there are exactly  $k$   $v^i$  equal to  $v_1$ .  $\mu \in S$  yields  $\eta$  if and only if  $\mu(k)$  yields  $(a_k^1, a_k^2, q_k)$  for  $k = 0, \dots, N$ .

**(only if)** Suppose  $\mu \in S$  yields  $\eta$ . Then  $q_k = \mu(k)(\{z \in Z | y = 1\})$ , and  $a_k^j = \mathcal{E}(x^i | v^i = v_j, \mu(k))$ , the expected value of  $x^i$  under the measure  $\mu(k)$  conditional on  $v^i = v_j$ . Let  $\hat{x}_{tk}^j = \mathcal{E}(x^i | v^i = v_j, \mu(k), y = t)$ , for  $t = 0, 1$ , the conditional expected value of  $x^i$  for  $\mu(k)$  conditional on  $v^i = v_j$  and  $y = t$ . Then  $a_k^j = q_k \hat{x}_{1k}^j + (1 - q_k) \hat{x}_{0k}^j$ . It is now easy to see that (1.1) (1.3) and (1.4) are true. (1.2) follows from the fact that  $k \hat{x}_{0k}^1 + (N - k) \hat{x}_{0k}^2 \geq 0$  and  $k \hat{x}_{1k}^1 + (N - k) \hat{x}_{1k}^2 \geq K$ . It also follows from these facts and  $co(X) = [\underline{x}, \bar{x}]$ , that

$$\min \left\{ \bar{x}, \frac{kt - (N - k)\underline{x}}{k} \right\} \leq \hat{x}_t^j \leq \max \left\{ \underline{x}, \frac{kt - (N - k)\bar{x}}{k} \right\}.$$

(1.5) and (1.6) follow from these.

**(if)** (Given  $(a, q)$  satisfying (1.1)-(1.6), we need to identify an appropriate  $\mu \in S$ . Given (1.5) and (1.6) we can choose  $\hat{x}_{tk}^j \in co(X)$  for  $j = 1, 2$  and  $t = 0, 1$  such that

$$a_k^j = q_k \hat{x}_{1k}^j + (1 - q_k) \hat{x}_{0k}^j$$

and

$$k \hat{x}_{tk}^1 (N - k) \hat{x}_{tk}^2 \geq tK.$$

Now let  $\hat{\mu}_t(k) \in M(X^N)$  be the measure on  $\{(x_1, \dots, x_N)\}$  such that the probability is one that, for all  $i$ ,  $x^i = \hat{x}_{tk}^j$  if  $v^i = v_j$ . Then let  $\mu(k)$  be the measure where  $\mu((x, y)) = q_k \hat{\mu}_1(k)(x)$  if  $y = 1$  and  $\mu((x, y)) = (1 - q_k) \hat{\mu}_0(k)(x)$  if  $y = 0$ .  $\blacksquare$

To see why (1.5) and (1.6) are necessary consider an example:  $X = \{0, 1\}$ ,  $K = 3$ ,  $N = 10$ ,  $k = 2$ . For  $y = 0$ , the only feasible  $x^i = 0$  for  $v^i = v_1$  and  $v^i = v_2$ . Therefore,  $\hat{x}_{jk}^i = 0$ . For  $y = 1$ , the largest  $\hat{x}_{0k}^1$  is 1 and the smallest is 0. The corresponding values of  $\hat{x}_{0k}^2$  are  $\frac{1}{8}$  and  $\frac{3}{8}$  respectively. Now (1.5) is  $a_k^1 \geq q_k \max\{0, \frac{3-(8)}{2}\} = 0$  and (1.6) is  $a_k^2 \geq q_k \max\{0, \frac{1}{8}\} = q_k \frac{1}{8}$ . If we were to try to find a feasible direct mechanism to implement  $q_k = \frac{2}{3}$ ,  $a_k^1 = 1$ ,  $a_k^2 = 0$  which satisfies  $ka_k^1 + (N - k)a_k^2 = q_k K$  we would not be able to do it even though (1.1) to (1.4) are satisfied. Because  $a_k^1 = 1$ ,  $\hat{x}_{jk}^1 = 1$  for  $j = 0, 1$ . Therefore it must be true that  $\hat{x}_{0k}^2 = -\frac{1}{8}$  but  $x^2$  must be either 0 or 1. Constraints (1.5) and (1.6) protect against this.

**Lemma 12:** Suppose  $A^* = [0, 1]$ . Let  $Q \in Q_F$  and  $\underline{Q}_2(Q_1) < Q_2 < Q_1 < 1$ . Then there exists  $Q'_1$  such that:

- (a)  $Q'_1 > Q_1$
- (b)  $Q' = (Q'_1, Q_2) \in Q_F$
- (c)  $U(Q') \geq U(Q)$

*Proof:* Let  $\{(q^k, a_1^k, a_2^k)\}_{k=0}^N$  be a mechanism for which the reduced form outputs are  $Q = (Q_1, Q_2)$  and the reduced form taxes are  $\alpha = (\alpha_1, \alpha_2)$  with  $\alpha_1 = U(Q)$ , and  $\bar{\alpha} = \frac{K}{N}\bar{Q}$ . First note that since  $Q$  is in the interior of  $Q_F$ , there exist  $k_1, k_2$  with  $k_1 > k_2 > 0$  such that  $q^{k_2} \in (0, 1]$  and  $q^{k_1} < 1$ . This follows from the characterization of  $Q_F$  given in the proof of Theorem 6. Furthermore, since  $\alpha_1 = U(Q)$ , we must have  $a_1^k = q_k \min\{\frac{K}{k}, 1\}$  for  $k = 1, \dots, N$ . Consider an increase of  $q^{k_1}$  to  $q'^{k_1} = q^{k_2} + \Delta^1$ , and a corresponding decrease of  $q^{k_2}$  to  $q'^{k_2} = q^{k_2} - \Delta^2$ , such that  $Q_2$  remains fixed, so:

$$\Delta^1 \binom{N-1}{k_1} \pi^{k_1} (1-\pi)^{N-k_1-1} = \Delta^2 \binom{N-1}{k_1} \pi^{k_2} (1-\pi)^{N-k_2-1}$$

This results in a new value  $Q'_1$ , given by

$$Q'_1 = Q_1 + \Delta^1 \binom{N-1}{k_1} \pi^{k_1-1} (1-\pi)^{N-k_1} - \Delta^2 \binom{N-1}{k_1} \pi^{k_2-1} (1-\pi)^{N-k_2} > Q_1$$

To establish (c), observe that it is feasible to finance  $(Q'_1, Q'_2)$  exactly as  $(Q_1, Q_2)$  was financed, except:

$$\begin{aligned} a_1'^{k_1} &= a_1^{k_1} + \Delta^1 \min\left\{\frac{K}{k_1}, 1\right\} \\ a_1'^{k_2} &= a_1^{k_2} - \Delta^2 \min\left\{\frac{K}{k_2}, 1\right\} \end{aligned}$$

Using a similar argument as before, it follows from  $k_1 > k_2$  that:

$$\begin{aligned} \alpha'_1 &= \alpha_1 + \Delta^1 \min\left\{\frac{K}{k_1}, 1\right\} \binom{N-1}{k_1} \pi^{k_1-1} (1-\pi)^{N-k_1} \\ &\quad - \Delta^2 \min\left\{\frac{K}{k_2}, 1\right\} \binom{N-1}{k_1} \pi^{k_2-1} (1-\pi)^{N-k_2} > \alpha_1. \end{aligned}$$

Therefore  $U(Q') > U(Q)$ . ■

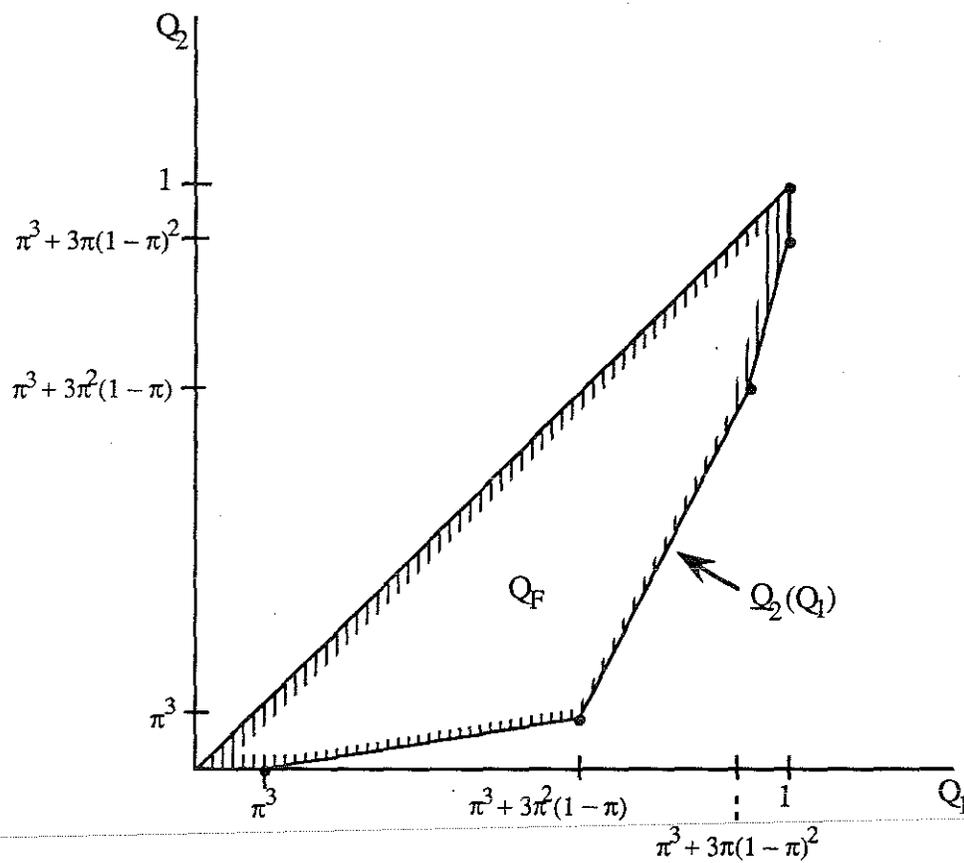


Figure 1. The set  $Q_F$  when  $N = 4$ .

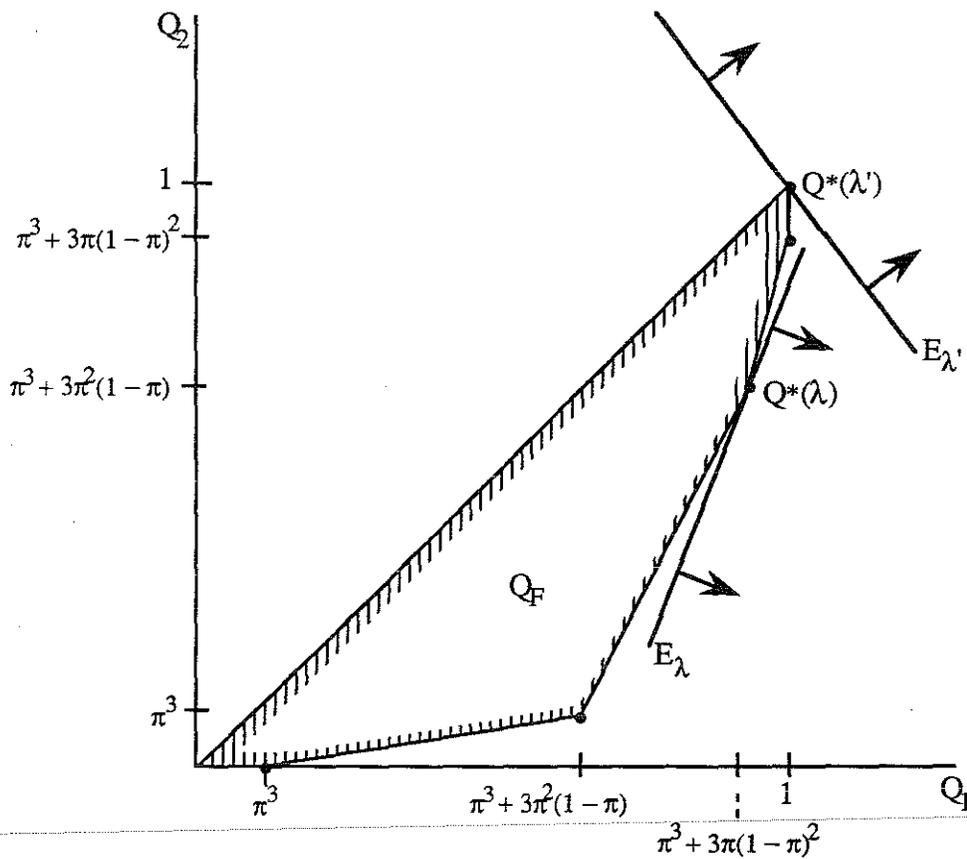


Figure 2. Optimal reduced form public good production for two welfare weights,  $\lambda < \lambda'$ . Optimal levels are vertices marked  $Q^*(\lambda)$  and  $Q^*(\lambda')$ , respectively. Level surfaces of the objective function,  $E(Q_1, Q_2)$ , are marked  $E_\lambda$  and  $E_{\lambda'}$ , respectively, with arrows indicating the direction of the gradient vector.

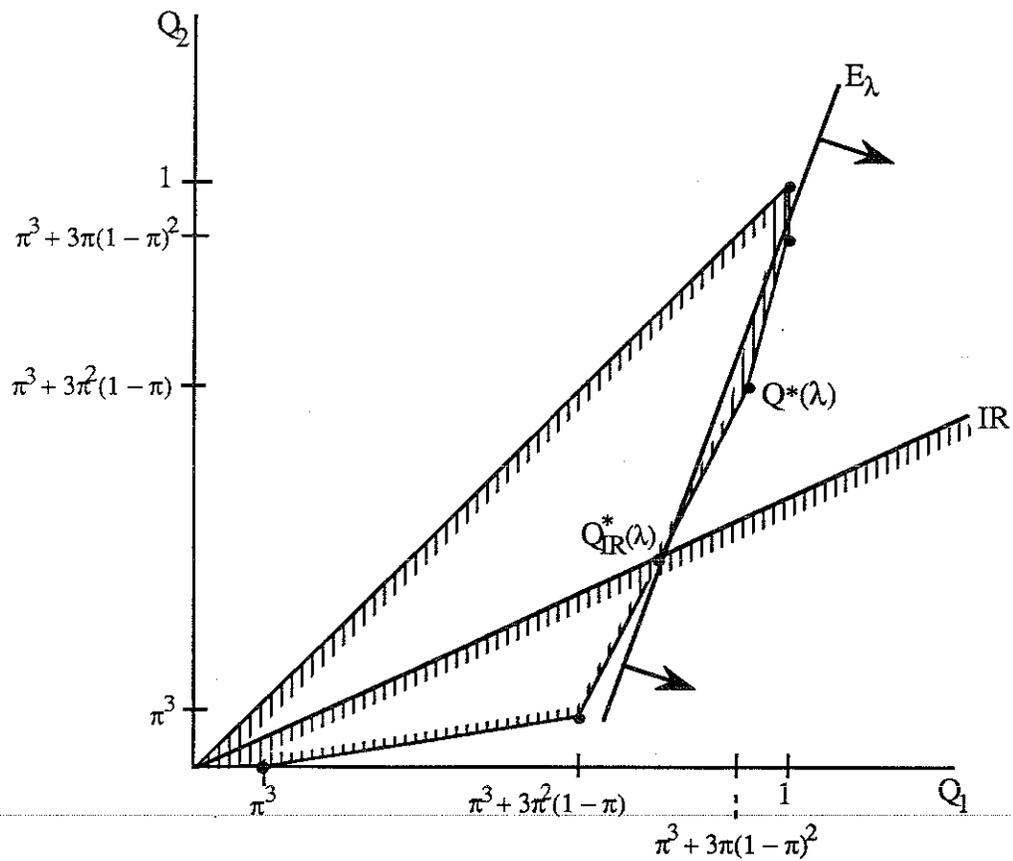


Figure 3. Effect of individual rationality constraint, IR. The unconstrained solution,  $Q^*(\lambda) > Q_{IR}^*(\lambda)$ , the IR-constrained solution. When IR constraint is binding, solution lies on boundary, but not typically on a vertex.

## References

- Border, K., "Implementation of Reduced Form Auctions: A Geometric Approach," *Econometrica* 59, 1175-1187, 1991.
- Clarke, E., "Multipart Pricing of Public Goods," *Public Choice*, 11, 1971, 17-33.
- d'Aspremont, C., Cremer, J., and Gerard-Varet, L.-A., "Incentives and the Existence of Pareto Optimal Revelation Mechanisms," *Journal of Economic Theory*, 51, 1990, 233-54.
- d'Aspremont, C. and Gerard-Varet, L.-A., "Incentives and Incomplete Information," *Journal of Public Economics*, 11, 1979, 25-45.
- d'Aspremont, C. and Gerard-Varet, L.-A., "Bayesian Incentive Compatible Beliefs," *Journal of Mathematical Economics*, 10, 1982, 25-45.
- d'Aspremont, C. and Gerard-Varet, L.-A., "Utilitarian Fundamentalism and Limited Information," CORE Discussion Paper #8935, Université Catholique de Louvain, 1989.
- Gibbard, A., "Manipulation of Voting Schemes," *Econometrica*, 41, 1973, 587-602.
- Green, J. and Laffont, J.-J., *Incentives in Public Decision Making*, Amsterdam, North-Holland, 1979.
- Groves, T., "Incentives in Teams," *Econometrica*, 41, 1973, 617-31.
- Groves, T. and Ledyard, J., "Some Limitations of Demand Revealing Processes," *Public Choice*, 29(2), 1977a, 107-24.
- Groves, T. and Ledyard, J., "Optimal Allocation of Public Goods: A Solution to the 'Free Rider' Problem," *Econometrica*, 45, 1977b, 783-809.
- Guth, W. and Hellwig, M. "The Private Supply of a Public Good," *Journal of Economics*, Suppl. 5, 121-59, 1986.
- Harris, M. and Townsend, R., "Resource Allocation with Asymmetric Information," *Econometrica*, 49, 1981, 33-64.
- Holmstrom, B. and Myerson, R., "Efficient and Durable Decision Rules," *Econometrica*, 51, 1983, 1799-819.
- Hurwicz, L., "Outcome Functions Yielding Walrasian and Lindahl Allocations at Nash Equilibrium Points," *Review of Economic Studies*, 1979, 217-25.
- Jackson, M. and H. Moulin, "Implementing a Public Project and Distributing Its Cost," *Journal of Economic Theory*, 57:125-40, 1992.

- Laffont, J.-J. and Maskin, E., "The Theory of Incentives: An Overview," in *Advances in Economic Theory*, W. Hildebrand, ed., Cambridge, England: Cambridge University Press, 1982.
- Ledyard, J. and Palfrey, T., "Mechanism Design Without Sidepayments: An Application to Public Goods," mimeo, California Institute of Technology, 1989.
- Mailath, G. and Postlewaite, A., "Asymmetric Information Bargaining Problems with Many Agents," *Review of Economic Studies*, 57:351-67, 1990.
- Maskin, E. and Riley, J., "Optimal Auctions with Risk Averse Buyers," *Econometrica*, 52, 1984, 1473-518.
- Matthews, S., "On the Implementability of Reduced Form Auctions," *Econometrica*, 52, 1984, 1519-22.
- Myerson, R., "Optimal Auction Design," *Mathematics of Operations Research*, 6, 1981, 58-73.
- Myerson, R., and Satterthwaite, M., "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 28, 1983, 265-81.
- Palfrey, T. and Rosenthal, H., "Participation and Provision of Discrete Public Goods: A Strategic Analysis," *Journal of Public Economics*, 24, 1984, 171-93.
- 
- Palfrey, T. and Rosenthal, H., "Private Incentives and Social Dilemmas," *Journal of Public Economics*, 35, 1988, 309-32.
- Palfrey, T. and Rosenthal, H., "Testing Game-Theoretic Models of Free-Riding: New Evidence on Probability Bias and Learning," in *Laboratory Research in Political Economy*, (T. Palfrey ed.), University of Michigan Press: Ann Arbor, 1991a, 239-68.
- Palfrey, T. and Rosenthal, H., "Testing for Effects of Cheap Talk in a Public Goods Game with Private Information," *Games and Economic Behavior*, 3, 1991b, pp. 183-220.
- Rapoport, A., "Public Goods and the MCS Experimental Paradigm," *American Political Science Review*, 70, 1985, 148-55.
- Rob, R., "Pollution Claim Settlements with Private Information," *Journal of Economic Theory*, 47, 1989, 307-33.
- Van de Kragt, A., Orbell, J., and Dawes, R., "The Minimal Contributing Set as a Solution to Public Goods Problems," *American Political Science Review*, 77, 1983, 112-21.

