PN ABK-052
74913

INSTITUTE
FOR
POLICY
REFORM

Working Paper Series

The objective of the Institute for Policy Reform is to enhance the foundation for broad based economic growth in developing countries. Through its research, education and training activities the Institute will encourage active participation in the dialogue on policy reform, focusing on changes that stimulate and sustain economic development. At the core of these activities is the search for creative ideas that can be used to design constitutional, institutional and policy reforms. Research fellows and policy practitioners are engaged by IPR to expand the analytical core of the reform process. This includes all elements of comprehensive and customized reform packages, recognizing cultural, political, economic and environmental elements as crucial dimensions of societies.

1400 16th Street, NW / Suite 350
Washington, DC 20036
(202) 939 - 3450

This paper was prepared under a cooperative agreement between the Institute for Policy Reform (IPR) and Agency for International Development (USAID), Cooperative Agreement No. PDC# 0095-A-00-9079-00. Views expressed in this paper are those of the author and not necessarily those of IPR or USAID.

INSTITUTE
FOR
POLICY
REFORM

The Logic of Collective Action: Theory and Applications

Todd Sandler

Professor of Economics and Political Science
Iowa State University

and

Senior Research Fellow
Institute for Policy Reform

Draft: September 1991

This research illustrates the principles developed in "The Logic of Collective Action: A Retrospective View." These applications serve to illustrate the importance of joint products, the neutrality theorem, and the requirements of privileged groups. Moreover, the applications indicate the validity of the three basic themes of collective action. Theme 1 relates group size to the effectiveness of collective action. Theme 2 concerns the relationship that group composition has on group formation and burden sharing within collectives. Theme 3 involves the means for promoting more effective collective action. Applications concern military alliances, foreign aid, philanthropy, the growth of nations, labor unions, international responses to terrorism, the commons, strategic trade policy, international organizations and public inputs (infrastructure).

: Applications

Perhaps the greatest achievement of collective action theory has been its wide-ranging application to major fields of economics. In *The Logic*, Olson investigated a host of applications that included the behavior of oligopolies, labor unions, political lobbies, and farm collectives. The Marxian theory of state and class was also illuminated with the principles of collective action by Olson (1965, pp. 98-110). Rather than review and extend Olson's applications, I intend to strike out on my own and deal with a different and broader set of applications. Some overlap between applications exists, since labor unions and political lobbies are also studied here.

For the most part, I aim to keep the discussion free of technical details, since this chapter is meant to demonstrate that the analysis of chapters 2-3 is applicable to diverse problems of importance not only in economics, but also in other social sciences. Technical details are, however, given for a few select applications (e.g., the commons, public inputs). Space does not permit a full analysis of each of the applications presented. In consequence, choices must be made in terms of emphasis and coverage. To assist the reader's further study of these applications, key references are either given in the text or in footnotes. After a couple historical examples, the following applications are analyzed: military alliances, foreign aid, philanthropy, the growth of nations, labor unions, international responses to terrorism, the commons, strategic trade policy, principal-agent analysis, international organizations, rent-seeking behavior, cost considerations, and public inputs.

4.1 Historical Examples

Although collective-action problems are more prevalent today owing to technological advancements, population pressures, and resource demands, collective-action problems are not a new phenomenon. In controlling plagues, fighting pollution, overthrowing repressive regimes, or deterring a common enemy, nations have throughout history faced collective-action contingencies.

The Delian League of Classical Greece serves as an apt example of a military alliance between city-states, in which free riding was a problem.[1] The League was formed in 478 B.C., during a period in which Athens and Sparta almost single-handedly repulsed invasions from Persia. Many small Greek city states had refused to help fight the Persians, mainly because of the rational incentives to free ride on a defense arrangement with nearly purely public outputs. The geographic compactness of Greece meant that Athens and Sparta could not afford the danger of allowing the Persians to hold any major positions in the Aegean. Smaller cities also did not want to risk fighting on the losing side. Defense benefits were nonexcludable owing to geographical considerations. Moreover, the removal of the Persian threat provided nonrival benefits to the city states of the Aegean. The Delian League solved this public good problem in the Olsonian tradition with a simple selective incentive: coercion, including violence. Each member of the League was required to take a permanent oath of loyalty, which included making annual contributions to the treasury of the League in Delos. In addition to deterring further Persian expeditions, the League provided another public good: the suppression (largely through the Athenian navy) of piracy throughout the region and the consequent promotion of trade. Some cities (e.g., Eubeoa in 472 B.C.) were forced to join; a greater number attempted to secede once the Persians were defeated, but were forcibly prevented (e.g.,

Naxos in 467 B.C., Thasos in 462 B.C.).

As Athens and Sparta became enemies, the League gradually became an instrument of Athenian imperialism, providing (in the view of allied cities) more of a private than a public good (e.g., imposing Athenian settlers and democratic regimes, and moving the alliance's treasury to Athens), which increased the incentives for allies to defect. As disaffection mounted, Athens became increasingly harsh in maintaining the alliance. When the island of Melos refused to join the alliance in 415 B.C., Athenian forces executed the men and sold the women and children into slavery.

Thucydides' (1970) account of the Peloponnesian War abounds with reference to public good problems. Pericles, an Athenian leader, urged a strong stand against Sparta, on the grounds that Sparta and her allies had a major collective-action problem:

> . . . they cannot fight a war against a power unlike themselves, so long as they have no central deliberative authority to produce quick decisive action, when they all have equal votes, though they all come from different nationalities and everyone of these is concerned with its own interests--the usual result of which is that nothing gets done at all. . . . It never occurs to any of them that the apathy of one will damage the interest of all. Instead each state thinks that the responsibility for its future belongs to someone else, and so, while everyone has the same idea privately, no one notices that from a general point of view things are downhill (Thucydides, 1970, p. 93).

A second historical example is due to Adam Smith, who believed that Britain, like Athens, was exploited by free riders in times of military crisis. Wars with Spain (1739-48) and France (1756-63) cost £120 million (tripling the national debt), and were fought for the security of North American colonies that contributed neither revenue nor military force to the effort. "If any of the provinces of the British empire," Smith (1976, Vol. II, p. 486) concluded, "cannot be made to contribute towards the support of the whole empire, it is surely time that Britain should free herself from the

expense of defending those provinces." Britain, unlike Athens, could afford to withdraw from the provision of international public goods not directly connected with the maintenance of domestic sovereignty.

These two examples suggest that alliance public or collective goods are not unique to the contemporary world. Athens could not defend itself against the Persians, without simultaneously defending small free riders. Britain voluntarily provided a public defense good to its colonies, in the belief (mistaken, according to Adam Smith) that the empire provided indirect sidepayments which compensated for public good provision. The only aspect of historical alliances that may, at first, seem unique is the use of coercive selective incentives, but recent actions by the U.S. in fostering the alliance for Operation Desert Storm (e.g., the forgiveness of billions of dollars of Egyptian debt, the shipment of Patriot missiles to Israel, and veiled threat of future actions by the U.S. Congress regarding troop deployments) indicate that selective incentives are still used. Like modern alliances, historical alliances share defense activities with private, impurely public, and purely public benefits.

A host of 18th and 19th century alliances shared arsenals that yielded joint products of varying degrees of publicness. In an interesting paper, Thies (1987) examined two alliances (the Triple Alliance and the Triple Entente) begun prior to World War I and a number of post-World-War-I alliances that include the Anglo-Franco (1924-38), the Franco-Polish (1926-38), the Franco-Czech (1930-38), and the Franco-Belgian (1920-36). Since these alliances relied upon conventional armaments, defense benefits included deterrence as well as impurely public and private benefits (see discussion of defense benefits in Section 4.2). The existence of private and excludable

benefits (e.g., conventional protection in times of war) meant that benefits could be withheld through deployment decisions. Owing to the existence of these impurely public benefits, in very few instances did these historical alliances display the disproportionate burden sharing behavior, associated with the pure public good paradigm of defense (Thies, 1987). In a follow-up study, Conybeare and Sandler (1990) found little evidence of free riding in the Triple Alliance (Germany, Austria-Hungary, and Italy) and the Triple Entente (United Kingdom, France, and Russia) during 1880-1914. The sole exception was France and Russia, whose geographical position on either side of the enemy, allowed for some substitution of forces and, consequently, free riding between the two fronts. Unlike the Thies (1987) study which did not rely on statistical inferential tests, Conybeare and Sandler (1990) utilized two-stage least squares estimations to identify free riding.

As a final historical example of collective action, the containment of plagues and diseases serves as an appropriate example. Since plagues and diseases transgress national boundaries, one nation's public health problem becomes that of another through the migration of population, airborne viruses, and the transportation of pests. The bubonic plague, a bacterium-based disease carried by fleas from infected rats, decimated parts of Europe and Africa on a number of different occasions. The migration of infected rats in the holds of ships spread the disease far and wide. In early times prior to vaccines and other health measures, nations attempted to protect its own citizens and those of other nations by isolating themselves. John Forbes (1980) indicates that,

> The origins of international cooperation in public health are to be found in the response of nations to the threat of epidemic disease. Throughout history nations and communities adopted various prophylactic measures by which they attempted to isolate themselves against the

> importation of pestilence, plague, and disease. An early prophylactic measure was the cordon sanitaire enforced by a military blockage of an infected area. A less comprehensive form of population isolation--quarantine--was introduced in the fourteenth century when Venice adopted a forty-day (quaranta) isolation period that applied to incoming persons and goods. Between the fourteenth and nineteenth centuries quarantine measures were widely adopted throughout the world as nearly all countries instituted unilateral and largely uncoordinated regulations to control the international diffusion of disease (Forbes, 1980, p. 116).

Although a quarantine or cordon sanitaire conferred public benefits to other nations, independent national actions were taken since such measures also protected the nation's population against new exposure to the disease-bearing agents. That is, significant nation-specific or private benefits motivated these acts of containment. As Forbes (1980) noted, frequent and prolonged quarantines created other collective-action concerns as communication and commercial links were temporarily severed. A concern for these other collective-action problems led to new forms of cooperation, culminating in the creation of the World Health Organization (WHO) in 1948. WHO sought to foster the treatment, containment, eradication, and tracking of diseases that threatened the world community.

4.2 Military Alliances: The Case of NATO

Defense has long been viewed as a pure public good with nonexcludable and nonrival benefits. When defense is, however, shared between nations in an alliance, the pure publicness of defense expenditure may be doubted, since for some defense outputs the providing ally may be able to withhold benefits from allies; exclusion may be practiced. If the actual deployment of the weapon system affects the area receiving protection, then benefits may also be partially rival. In a recent contribution, Brian Goff and Robert Tollison

(1990) argued that defense benefits within a nation may impact and protect regions differently, thereby fostering impure public benefits even within a nation.

To distinguish purely public, impurely public, and private defense benefits, a series of articles developed a joint product analysis of military alliances.[2] The joint product model views an alliance arsenal as fulfilling at least three functions: (1) deterrence, (2) damage limitation or protection, and (3) private or nation-specific goals. Deterrence forestalls an attack by an opposing alliance by threatening a sufficiently costly punishment to any would-be aggressor. To be effective, deterrence must be based on a credible threat that is swift and automatic. If the threat does not appear to be automatic, then a would-be aggressor may discount the threat and act. Deterrence is nonrival, since a threat can protect additional nations and/or population without diminution in benefits to those already protected if the threat is sufficiently credible and if the promised punishment is sufficiently devastating. Deterrence will fail whenever there is serious doubts on behalf of those threatened that the threat will be executed, as Saddam Hussein's failure to withdraw from Kuwait sadly illustrates. Nonexcludability may or may not characterize deterrence. When a nation's people, property, territory or airspace must be compromised in any enemy attack directed against another ally, the nation cannot withhold retaliation on behalf of the ally and deterrence is nonexcludable. If an allied nation hosts the troops, citizens, or direct investment of another ally, then an enemy invasion on the host ally ensures collateral damage to the interests of the other ally; in consequence, this latter ally cannot withhold a retaliatory pledge. In other cases, geographical considerations cause

allies to face common threats that make deterrence nonexcludable. For example, a nuclear attack on the Republic of Ireland (Canada) would, due to fallout, misses, and wind direction, kill millions in the U.K. (U.S.). Obviously, the U.K. (U.S.) could not deny its neighbor deterrent protection.

Defense outputs are impurely public between allies when their benefits are either partially rival or else partially excludable by the providing nations. Conventional forces and arsenals yield both deterrence and damage-limiting protection (needed when deterrence fails and conflict begins). Although the deterrence benefits are purely public, damage-limiting protection are subject to consumption rivalry in the form of force thinning as a fixed arsenal is spread to defend a longer perimeter or a greater surface area. In other words, the deployment of conventional forces to limit damage does not afford equal protection to all allies. Moreover, increasing the concentration of troops along one ally's border may increase the vulnerability of another owing to rivalry. Since deployment decisions can exclude one or more allies, some conventional armaments possess benefits that are partially excludable. The possibility of exclusion and the presence of thinning or congestion costs imply that a club arrangement can be used, in part, to allocate defense within a conventional alliance. As such, an optimal size for the alliance could be determined so as to adjust cost sharing and thinning considerations, at the margin, with respect to an entrant. The presence of purely public deterrence means that some free riding will be present even under a club arrangement. If, however, the alliance primarily shares purely public deterrence, as <u>may</u> have been the case for the North Atlantic Treaty Organization (NATO) in the late 1940s and 1950s, then a club arrangement cannot be implemented. With purely public benefits, there is no need to restrict alliance membership.

Most importantly, free riding should be prevalent with an exploitation of the large ally by the small.

Conventional forces can also be used by allies to pursue country-specific private benefits (e.g., protecting coastal waters, engineering civil projects, thwarting terrorist threats, curbing domestic unrest, pursuing nationalistic goals, and providing disaster relief). Market principles can be applied to private defense goods. For instance, a nation cannot rely on a neighbor or ally for protection of, say, its ocean resources without entering a contractual arrangement in both parties' interests. As the share of private and/or excludable benefits from an alliance arsenal increases, market or clublike arrangements can be pursued more fully. In consequence, the extent of free riding is inversely related to the proportion of private defense outputs derived from the arsenal. Alliances that depend on conventional forces are more apt to share burdens according to the proportion of benefits received by the allies. In contrast, strategic nuclear forces are less likely to serve country-specific defense needs and to contain rivalrous benefits. Since such forces possess a small proportion of private and excludable benefits, these forces are most apt to be associated with the free-rider problem highlighted by Olson and Zeckhauser (1966) in their seminal study of NATO.

NATO was formed in 1949 as an offset to Soviet domination of Eastern Europe. The original alliance included Belgium, Canada, Denmark, France, Iceland, Italy, Luxembourg, the Netherlands, Norway, Portugal, the U.K., and the U.S. Greece and Turkey joined NATO in 1952; West Germany in 1955; and Spain in 1982. Thus, the alliance has grown to 16 allies.

Until the mid-1960s, NATO relied on its strategic arsenal to deter

Soviet's use of conventional forces in Western Europe, since NATO's conventional forces were outnumbered by Soviet tanks and ground troops. Moreover, the U.S. threat to retaliate on behalf of its NATO allies was credible during this period, inasmuch as the Soviet had not yet developed a retaliatory strike force. Consequently, the U.S. had little to fear from using its strategic arsenal, as shown by President Kennedy's aggressive stance during the Cuban Missile Crisis. Once the Soviet strike force had been deployed, the credibility of the U.S. retaliatory threat was diminished. The stationing of U.S. troops along the central front and at European airfields, U.S. investments in Europe, and a shared cultural heritage gave some credence to the commitment of the U.S. pledge. Nevertheless, some Europeans had their doubts.

By the early 1970s, many developments had altered the nature of the defense benefits derived by NATO and, hence, the possibility for free riding. Important events included the Soviet nuclear arsenal buildup, the deployment in the 1960s and 1970s of a small nuclear strategic deterrent by Britain and France, and the increased importance placed on conventional weapons. This last development took place throughout the last half of the 1960s and the early 1970s as the NATO alliance changed its emphasis from a strategy of mutual assured destruction deterrence (MAD), based on nuclear annihilation, to that of flexible response. The flexible response doctrine permits NATO to respond in different ways to a Warsaw Pact challenge; conventional forces, tactical nuclear forces, or strategic force may be used and, in the latter case, a missile exchange may be limited or complete. With the new doctrine, the European allies must be prepared to defend themselves against convention aggression in the European theater. Any ally that sits back and does not

increase its military preparedness could invite aggression, since an opposing alliance might have a better opportunity to gain an advantage in a conventional foray on that ally's soil.

In the 1980s, the Reagan buildup emphasized weapon procurement over operations and maintenance. The strategic weapons' share of the budget increased for the three nuclear allies--the U.S., France, and the U.K.--thereby increasing the relative share of the NATO budget going to weapons with a greater degree of publicness (Hansen, Murdoch, and Sandler, 1990).

Two popular burden sharing measures are now presented to take a retrospective view of NATO since 1955. In Table 4.1, an ally's share of total NATO military expenditures is calculated. From 1965 to 1970, the overwhelming defense burden was placed on the U.S., with the U.K., France, and West Germany sharing much of the residual. This pattern conforms well to the exploitation hypothesis that the large allies would shoulder the defense burden of the smaller owing to NATO's reliance on MAD and relatively pure public strategic weapons.

In Table 4.2, a second burden measure--the proportion of GDP spent on defense--shows much the same pattern for ten NATO allies: the U.S., the U.K., and France devoted by far the largest shares of GDP to defense. Six allies have been left out of the sample. Portugal, Luxembourg, and Iceland were excluded since their defense expenditures were so small. Spain only joined NATO in 1982. After the late 1970s, Greece and Turkey defense figures were unreliable. In recent years, both nations allocated about six percent of GDP to defense due, in part, to their territorial dispute.

By the mid-1970s, the doctrine of flexible response, with its emphasis on conventional armaments, had taken hold and had achieved the predicted

impact of reducing free riding and shifting more of the burden for NATO to the Europeans. In Table 4.1, the U.S. share of NATO defense spending dropped from an average value well over 70 percent to 58.3 percent. The rest of the allies, except Canada, were assuming a greater defense burden. In Table 4.2, the U.S., France, U.K., and West Germany devoted a smaller percentage of GDP to defense after 1975 as compared with the 1960s, while the smaller allies, except the Netherlands, maintained GDP percentages. This pattern of GDP burdens is behind the dramatic reallocation displayed in Table 4.1. Apparently, the new strategic doctrine accomplished what two decades of arm twisting could not--a shifting of defense burden from one side of the Atlantic to the other.

As expected, a reversal of burdens took place in the 1980s during the Reagan buildup. In Table 4.1, the U.S. share of the NATO burden increased. Of the three next largest allies, the two nuclear allies decreased their burdens less than the non-nuclear ally (West Germany). Similar patterns show up in Table 4.2.

Sandler and Forbes (1980, Table 2) provided a test of the exploitation hypothesis using NATO data for the 1960s to the mid-1970s. In particular, these authors computed the Kendall rank correlation between an ally's GDP and its percentage of GDP spent on defense. In keeping with the exploitation hypothesis, there is a significant correlation until 1966, thereby implying that the large allies shouldered the burden of NATO until about the start of flexible response.

The analysis of NATO indicates that the mix of pure public, impure public, and private benefits in a collective is not immutable. A change in strategic doctrine, the development of new technology (e.g., the deployment of

laser-guided munitions in the early 1970s), and alterations in diplomacy can alter the mix of public/private benefits in NATO and with it the way in which burdens are shared. In consequence, the development of star war technology, the integration of Western Europe by the year 1992, and the deployment of a new generation of strategic weapons by the medium-sized nuclear allies could have a profound effect on the way future burdens are shared in NATO.

Until now the literature on military alliances has treated the technology of public supply as either summation or weighted sum (see McGuire, 1990). The type of weapons shared and the kind of war anticipated may require a different type of technology of supply than those used to date. In the case of a conventional war, the weakest fortification along a front may ultimately determine the overall defense of the alliance. As such, Hirshleifer's (1983) weakest-link technology would apply. In the case of an alliance reliant on a strategic nuclear deterrent or a star war umbrella, the best-shot technology may be most appropriate. Weakest-link tends to attenuate the free-rider problem, while best-shot tends to exacerbate it. Future work should attempt to test for the best underlying technology for the NATO alliance during various stages of its history.

4.3 Foreign Aid

Foreign aid poses a collective-action problem for the international community. Many developing nations must rely on other nations to provide aid to finance public investment projects. By increasing the well-being of a recipient nation, foreign aid serves as an input that produces an output that is both nonexcludable and nonrival to all nations with an interest in the

recipient's well-being. That is, the recipient's well-being, produced in part by foreign aid, enters potential donors' utility functions. Moreover, once augmented, the well-being of the recipient nation is nonexcludable to all would-be donors. Foreign aid may also provide benefits to a donor that is private between nations but public within the donor nation. When the donor derives an advantage from its gift of aid and this advantage is not shared with others, private benefits may also motivate giving, much as private aspects stimulate allies' defense expenditures. First, we examine the collective-action problem associated with foreign aid when all potential donors view the economic advancement or well-being of another, less-fortunate nation as purely public. We next introduce private motivations along with the public motivation. The collective-action aspects of foreign aid is analogous to the case of charity or philanthropy within a nation.

If foreign aid is untied, aggregate aid to a recipient represents a fungible resource. In consequence, a summation technology of supply applies, since the recipient's well-being is dependent on the sum of aid received from others. Suboptimality in the supply of foreign aid is then to be expected. Suppose that all potential donors have the same tastes but different endowments. Further suppose that a recipient's well-being is a normal good in the utility function of would-be donors, then the demand for giving will be positively correlated to income. Wealthier nations have a greater desire to give, so that the richest nations will bear the burden of foreign aid. A clear asymmetry arises and Olson's exploitation hypothesis applies. Some potential donors may contribute nothing, relying instead on the foreign aid given by the wealthier nations. Suboptimal foreign aid levels indicate the need for some policy initiated at the supranational level. The manner used to

finance foreign aid at the supranational level becomes a crucial consideration owing to the neutrality theorem. If an international agency (e.g., U.S. Agency for International Development (AID), World Bank) supplements a recipient's foreign aid from revenues collected from donor nations, then foreign aid at the supranational level would crowd out foreign aid from donors on a dollar-for-dollar basis. This then implies that agencies must seek their funding from nondonor countries if they intend to increase foreign aid by their action. Any shock to the international system, such as sharp oil price rises, that redistribute income among the set of donor nations may yield no net change in foreign aid even if the gainers have a higher propensity to contribute. This neutrality result could be upset if the redistribution of income, no matter how it was engineered, altered the <u>set</u> of contributors. When foreign aid yields pure public benefits to the donors, proposed schemes to redistribute income to the underdeveloped world by taxing the resources of the wealthier nations may be self-defeating, since such taxes might reduce voluntary giving.

In many instances, foreign aid may yield both a purely public benefit to the set of donors and country-specific private benefits. Donor-specific private benefits may arise owing to a donor's relative location to the recipient. If, say, a recipient country is positioned strategically vis-à-vis a donor nation, aid-assisted growth may augment the recipient's political stability, which, in turn, fosters the donor's own security. More-distant donors may not receive any security benefits. Private benefits may arise from tying foreign aid to certain stipulations that are advantageous to the donor. For example, foreign aid may carry the condition that the donor can maintain military bases on the recipient's territory. When foreign aid possesses both

private and public benefits, the neutrality theorem does not necessarily apply and crowding out may not be a problem. A carefully engineered redistribution or tax scheme (see e.g., Andreoni, 1989; Sandler and Posnett, 1991) could increase overall giving if international agencies were to tap donors, who derive relatively greater private benefits from foreign aid, for resources.

There appears to be a clear analogy between charity and foreign aid. In the case of charity, private and public benefits are jointly produced by the act of giving (Posnett and Sandler, 1986; Andreoni, 1987). The extent of crowding out can be examined for foreign aid in much the same fashion as the charities (see Burton Abrams and Mark D. Schmitz, 1978; Charles Clotfelter, 1985). Evidence of dollar-for-dollar crowding out would lend great support for treating foreign aid as a pure public good.

4.4 Charities

Although foreign aid and charities share some important analogies as indicated above, there are a number of crucial differences. First, charities invariably involves large number of donors, whereas foreign aid at the nation-state level often concerns bilateral agreements. At times, multiple countries may come to the rescue of a country confronting famine (e.g., Ethiopia in the 1980s) or catastrophe, but, nevertheless, the number of participating countries typically is less than a hundred. In contrast, charities draw donors from the general population and, consequently, donors may number in the millions. Second, charities can garner funding from multiple levels of governments (e.g., state, local, and central) and can engage in commercial activities to supplement revenue sources. In recent years, there has been a

tremendous growth in non-profit institutions that raise revenues in the form of donations, fees, investment income, rents, sales, and government grants, so as to provide charitable outputs (Burton Weisbrod, 1988; Henry Hansman, 1980). In the U.S., 4.4 percent of national income originated from productive activities in the non-profit sector in 1985 (Weisbrod, 1988, p. 172). Third, charities have developed novel forms of institutional structures to circumvent the collective-action problem by making their activities more attractive to donors. Fourth, charities can often take advantage of favorable tax treatments to lower the effective price of giving.

The large number of contributors associated with many charities raises a conundrum: If charities are providing a pure public good in their philanthropic activities, and if a large number of donors are relied upon for contributions, then the group should be latent and, hence, not form. In fact, Robert Sugden (1982) provided an elegant demonstration that the pure public good theory of philanthropy is inconsistent with the existence of large fund-raising charities (large in terms of income and the number of donors), when taken together with available evidence concerning the income elasticity of charitable giving. Yet fund-raising charities exist. U.S. charities collect billions of dollars annually. In 1985, the 150,000 charities registered in England and Wales received a total income in excess of £12.5 billion or 4 percent of gross national product (Posnett, 1987). Clearly, the standard wisdom, drawn from conventional theories of pure public goods, does not apply to charities. But why?

Apparently, the answer lies both in a failure of charities to provide a pure public good and in the charity's ability to design an institutional structure that manages to circumvent standard free-riding concerns. The pure

public good theory of philanthropy rests on three primary assumptions: (1) That the output of charitable activity produces benefits which are non-rival and which may, in principle, be enjoyed equally by contributor and noncontributor alike. (2) That each individual holds a zero (Nash) conjecture regarding the effect of his own contribution on the contributions of others. (3) That individuals behave in such a way as to maximize utility. In choosing among possible assumption violations, Sugden (1982) preferred candidate is that of utility maximization:

> If one interprets 'utility' in the classical Benthamite way, as a psychological experience of pleasure, it is not a matter of logical necessity that an individual should seek to maximize his own utility. For example, a public-spirited act utilitarian might instead seek to maximize the sum of utilities for all people in society. Another person might be a rule utilitarian following those rules that, if generally followed, would maximize the sum of utilities. A third person might act on the Kantian principle of following those rules that he could will to be general laws (Sugden, 1982, p. 349).

Most other researchers have turned to the violation of assumption 1--the pure publicness assumption--for an explanation of observed behavior concerning charities. As such, charities are viewed as providing an activity that has pure public and private outputs. In other words, people are motivated, in part, to contribute owing to excludable benefits, not available from the general contributions of others.

Table 4.3, taken from Posnett and Sandler (1986, p. 219), indicates income sources for eight of the largest charities in Britain in 1983. Fees and subscriptions constitute sizable income sources for the National Trust, Dr. Barnado's, and the Salvation Army. In the case of the National Trust, fees are collected to permit visits to the trust's sites--a clear private benefit is conferred to visitors. Dr. Barnado's, for example, receives fees from local authorities in respect to children placed in its care, and uses

surpluses to finance its care of other children. Direct trading, involving the sale of private goods, is a significant revenue source for Oxfam; in total, it accounted for 39.1 percent of its income. Voluntary contributions, in the form of donations and legacies, account for approximately 50 percent of the income of the charities listed in Table 4.3. Weisbrod (1988, Table C.4, p. 197) indicated that nonprofits in the U.S. depended on service fees and other incomes for approximately 50 percent of their receipts in 1980. Nonprofits, such as education and research institutions, received almost 80 percent of their income from sources other than private giving or the government.

Charities are able to compete against the for-profit sector in the sale of private goods owing to cost advantages. These may arise from favorable tax treatment that exempt direct trading from corporation taxes when certain requirements are met (see Posnett and Sandler, 1986, pp. 215-16). Moreover, charities can reduce costs through the use of voluntary labor and exemptions to minimum wage laws. In some situations, charities have monopoly power, as in the sale of visitation rights to historical sites by the National Trust.

The presence of these private benefits means that alternative funding sources, as provided by private trading or government grants, need not crowd out private donations in a dollar-for-dollar fashion, since the neutrality theorem is not applicable. In a study of British charities, Posnett and Sandler (1989) found little or no evidence of any crowding out. This result is consistent with results reported by Abrams and Schmitz (1978) and Clotfelter (1985) in which crowding out was nowhere near a dollar-for-dollar basis. Moreover, autonomous income (rents, investment income and fees) were a <u>net</u> addition to income with no crowding out (Posnett and Sandler, 1986, Table

4).

The design of institutional structures to deal with collective-action aspects of charity fund-raising has been recognized since *The Logic*, where Olson (1965, pp. 62-63) mentioned the use of federated structures to limit group size. For example, the United Fund and Easter Seals organize much of its fund-raising drives at a local level, where peer group pressures are more effective. Furthermore, in the case of the United Fund, revenues are used to support local charities so that benefits have the greatest potential impact on contributors. When organized at the local level, individuals may be more willing to contribute owing to a system of rewards (e.g., status or prestige) and punishments (e.g., stigma), valued by a sufficiently large number of individuals to make the group viable. This type of induced cooperation is more likely in a relatively small group, in which the costs of detecting noncooperative behavior are low and the value of rewards is high owing to close social interaction.

Since donors cannot control how their money is used by the charity, an important asymmetric information problem exists between uniformed donors and well-informed organizations. This problem is especially acute for organizations that depend on voluntary contributions from a large number of donors, since donors may have little incentive to become knowledgeable about activities of the charities that they support, especially if their contributions are modest. To limit transaction costs and to assure donors that funds go to stated purposes, large public fund-raising organizations limit the degree of trustee discretion. In contrast, charities that depend on a few donors who may themselves serve on the board of trustees are expected to have a high degree of trustee discretion. Thus, sources of income may help

shape a charity's organizational form in terms of trustee discretion. A study by Posnett and Sandler (1988) provide evidence consistent with this hypothesis.

4.5 Growth of Nations

In 1982, Olson published ***The Rise and Decline of Nations*** in which he spelled out the macroeconomic implications of ***The Logic***. Prior to the publication of ***The Rise and Decline of Nations***, most of the principles of collective action had been applied to microeconomic issues. Olson (1982) argued that a nation's political stability allowed for the emergence of (mostly small) special interest groups that vied for a nation's income and resources. In stable nations, the number of such groups grew with time. The agenda of these "distributional" coalitions was to redistribute income to its members. Since most of the coalitions were small, they were not concerned with the social and transaction costs that their actions imposed on the economy. Olson reasoned that a small coalition would experience only a tiny fraction of the efficiency loss associated with the excess burden caused by their redistributive policies. If, say, the coalition consisted of one millionth of the population, then on average only this fraction of the efficiency loss would be experienced by the coalition. If, however, the coalition's gain from the redistribution exceeds this fraction-weighted loss, then the coalition will continue their redistributive activity. Societies that accumulate relative large number of these redistributional coalitions would be especially prone to losses of efficiency that would impede growth. In contrast, large-sized coalition groups are more apt to curb inefficient

activities, since their share of the deadweight is greater and, consequently, is likely to outweigh redistributional benefits at a lower activity level.

In mature societies, significant social costs could arise as coalitions create excess burdens from redistributive tax policies, slow the adoption of new technologies, or inhibit the reallocation of resources to new growth-promoting sectors. A nation's growth was predicted by Olson (1982) to suffer from such coalitions. Since these coalitions increase as a nation ages, younger nations (states) were predicted to have better growth performance than older nations (states). This provocative hypothesis as well as the opening quote of the book show that many macroeconomic issues may depend, in part, on collective-action considerations.

There have been numerous attempts to test Olson's hypothesis; noteworthy examples include Kwang Choi (1983), Olson (1982), Peter Murrell (1983) and Richard Vedder and Lowell Gallaway (1986). Most tests involve regressions in which the dependent variable is some growth measure (e.g., growth in per capita income) and the independent variables include an age indicator for the state or nation. The studies cited above have found a significant negative relationship between growth and the nation's age, consistent with Olson's hypothesis. Many of the tests are, however, very simplistic and do not consider alternative hypotheses that could explain growth differences between nations.

Vedder and Galloway (1986) added in a host of additional independent variables that include tax measures, union membership, public assistance measures, and economic activity indices (e.g., agriculture and mineral). By including more factors, this study attempted to ascertain whether other alternatives could explain the personal income growth in 48 U.S. states during

1970-82. Their results strongly support Olson's hypothesis: the state's age and its membership in distributional coalitions were negative determinants of growth. The inclusion of distributional coalitions per se is important in lending support to Olson's hypothesis. Future econometric studies must examine the influence of the extent and behavior of interest groups over time on the growth of income.

Surely more work on the aggregate impact of special-interest groups on economic growth is needed. The influence of special interest groups in developing nations is of particular interest, since wasteful activities could siphon off much-needed resources from supporting infrastructure projects.

4.6 Labor Unions

Olson (1965) devoted chapter III of ***The Logic*** to depicting labor unions as a collective-action problem, in which union members share the collective goods of higher wages and other nonpecuniary fringe benefits associated with a union shop. Olson reasoned that either membership must be compulsory or else selective incentives must be provided to motivate laborers to join. If the union cannot motivate laborers, then workers have incentives to free ride on union accomplishments. Union action to establish a closed shop, whereby only union members can be hired, is an attempt to make union-provided collective actions excludable.

In two contributions, Alison Booth (1984, 1985) has re-examined whether workers may be motivated to join even without compulsory actions. Following George Akerlof's (1980) theory of social custom, Booth (1985) showed that codes of behavior may motivate workers to join unions even in the absence of

an enforcement mechanism. Booth (1985) made the union wage, w, both nonrival and nonexcludable to all workers in the industry, but viewed reputation, r, as an excludable benefit only acquired by members. Reputation is an increasing concave function of the proportion of labor force, M, belonging to the union. This membership proportion varies between 0 (no members) and 1 (everyone in the industry is a member). Union members receive utility,

(4.1) $$U^{in} = U^{in}[r(M),\ w - s],$$

in which superscript "in" denotes a member and s is the membership fee,

(4.2) $$s = a + b/M,$$

which include a fixed component, a, and a variable component b/M that falls with membership size. Nonunion workers receive

(4.3) $$U^{out} = U^{out}(0,\ w),$$

in which superscript "out" depicts a nonunion worker. Since reputation can only be acquired by union membership, reputation for nonunion members is 0. An individual is motivated to join whenever net benefits from membership exists, so that

(4.4) $$U^{in}[r(M),\ w - a - b/M] - U^{out}(0,\ w) > 0.$$

Booth (1985, pp. 257-260) demonstrated, via the intermediate value theorem, that stable equilibria can exist at the corners, where M = 0, 1, as long as everyone is identical and the difference depicted in (4.4) equals 0 for some $M \in (0, 1)$. The stable equilibrium at M = 1 violates the free-rider tendency since everyone chooses to join. Contrary to Booth's assertion, this equilibrium is very much in keeping with Olson's view of collective-action problems, inasmuch as reputation is, in essence, a selective incentive produced jointly, owing to social custom, with the purely public wage

component. Without this excludable, jointly produced union benefit, membership would, indeed, be zero as Olson and Booth recognize. Jointly produced outputs can motivate, provided that they are excludable. These motivating benefits may, however, be rival, nonrival, or partially rival. Booth's analysis is reminiscent of club theory.

4.7 International Responses to Terrorism

Since the late 1960s, the international community has experienced an ever-increasing threat of terrorism. Terrorism is the premeditated use, or threat of use, of extra-normal violence or force to gain a political objective through intimidation or fear. Although terrorist motivations vary widely between groups (e.g., nationalism, separatism, nihilism, issue-specific concern), terrorist tactics are similar and include hostage taking, bombings, assassinations and hoaxes. Transnational terrorism concerns terrorist activities involving terrorists or government participants from two or more nations. Incidents originating in one nation and terminating in another, like a skyjacking, are transnational, as are incidents involving the demands made of an agent in a nation other than the one hosting the incident. Terrorist events that include victims, terrorists, or the institutions of two or more nations are considered transnational. These transnational terrorist incidents can be associated with numerous collective-action problems between countries owing to externalities.

If a set of nations are targeted by a state sponsoring terrorism (e.g., Iraq, Iran, Syria), then any retaliatory response on the part of one nation would confer nonexcludable and nonrival benefits to the other <u>targeted</u>

nations. Situations may arise in with the net benefits from retaliation to the retaliator are negative (i.e., retaliatory costs exceed the nation's derived benefits) if the nation is the sole retaliator. If, moreover, net benefits are positive when two or more nations retaliate, then a prisoner's dilemma game, like those in Figure 2.1 and 2.3, applies and no nation may go it alone. Unless the payoffs from the status quo of not retaliating is sufficiently negative, nations may not expend a Pareto-optimal level of retaliatory actions. When, however, a state sponsor selectively targets some nation's people and property (e.g., the people and property of the U.S. and Israel), benefits derived from retaliation may be sufficiently great to warrant an unilateral response even though the full costs are shouldered by the sole retaliator. In the case of the Libyan raid on April 15, 1986, the Reagan administration, obviously viewed these net benefits as positive. For state-sponsored terrorism, the retaliation dilemma raises many of the issues (e.g., group asymmetry, suboptimality) that Olson mentioned in *The Logic*.

Dwight Lee (1988) raised the further possibility that some nations may hurt the collective interests of the group in ways worse than free riding. Lee (1988) puts forward the notion of "paid riding", whereby an agent "sells" the public good of terrorism deterrence that the efforts of others attempt to create. Paid riding occurs when a country offers terrorists a sanctuary if the terrorists promise restraint on the country's own soil. The paid-rider option dominates the free-rider option of doing nothing. By selling or reducing the public good, the paid rider reduces the incentive of the retaliator to act, since the level of deterrence that its actions would achieve is partially undone by the paid rider. Thus, the retaliator's net benefits are even less than with free riders and may imply that the group is

no longer privileged.

A collective-action problem may also characterize nations' deterrence expenditure decisions to curb attacks when confronting terrorists that operate on two or more nations' soil. If each nation decides its deterrence independently, then each may allocate too many resources to inducing terrorists to switch where they stage their events. This follows because the nations do not account for the negative externality (by inducing terrorists to operate on another nation's soil) that these deterrence choices create for others. If nations were to share intelligence concerning the group's true preferences for attacking alternative targets, then the overdeterrence outcome would be aggravated as nations are better able to calculate what it takes to make the terrorists go elsewhere. Piecemeal policy, in which intelligence but not deterrence decisions are shared, may make everyone worse off. Both intelligence sharing and deterrence decision pose collective-action concerns. Neither decision can be handled in isolation since each is interdependent. The problem of second best applies. A grand strategy for coordinating policy on both fronts is required to reach Pareto optimality.

4.8 Problems of the Commons

An apt example of a collective-action problem is the commons in which a scarce resource is owned collectively by a set of agents. Common property examples include fisheries, oil pools, hunting grounds, deep-sea mineral beds, orbital bands in geostationary space, and radio-wave frequencies. The analysis of the commons demonstrates the overexploitation of the scarce fixed resource, in with the average product of the variable input, not its marginal

product, is equated to the input's rental rate when access is free and the set of exploiters large. Overexploitation is relevant to a commons in both a static, steady-state framework and a dynamic analysis. To maintain simplicity, we present a static analysis.

Consider a set of exploiters consisting of a given number of profit-maximizing firms, each having free access to an exogenously fixed common property resource. We refer to this resource as a fishing ground. Each firm combines the common property resource with a single private input, fishing vessel, to produce an output of fish. With the size of the commons fixed, the total output or catch, C, depends solely on the size of the total fishing fleet, R, which represents effort on behalf of exploiters. The production function is

(4.5) $C = F(R)$,

which is increasing and strictly concave (i.e., $F'(R) > 0$ and $F''(R) < 0$). Other sources of supply for fish are assumed so that the price of catch is exogenously fixed at unity. Vessels are also assumed to be sold on competitive markets but at a price of p per vessel.

To illustrate a commons problem with many of the Olsonian themes, we further assume a "pure" homogeneous commons where fish are evenly distributed throughout the fishing grounds, so that each vessel hauls in the same catch. This assumption allows the catch, c, of each firm to equal its share of total effort times the total catch:

(4.6) $c = [r/R]F(R)$

$= [r/(r + \bar{R})]F(r + \bar{R})$,

in which r is the firm's number of vessels and $\bar{R}$ is the aggregate fleet for the other firms in the commons. Obviously, the total fleet in the commons, R,

equals $r + \bar{R}$. Each exploiter's benefit equals its value of catch, which, in turn, is a fraction of the total value derived from the commons. This representation is analogous to the Olson depiction of a collective-action problem.

The Pareto-optimal solution for the commons is found by choosing the aggregate fleet size that maximizes total profit, π, for the set of exploiters:

(4.7) $\pi(R) = F(R) - pR.$

The optimizing fleet size, R^*, for the commons is uniquely determined by the first-order conditions, $F'(R) = p$, which is independent of the distribution of vessels among the exploiters. Profit maximization is equivalent to maximizing the sum of producer and consumer surplus in this stylized problem. The fleet's marginal product must be equated to the variable input's price. This is the allocation of effort resulting from competitive exploitation in the presence of well-defined property rights.

The Nash equilibrium corresponds to the profit-maximizing choice of r on the part of the individual firms; i.e.,

$$\max_r \{[r/(r + \bar{R})]F(r + \bar{R}) - pr\},$$

in which $\bar{R}$ is an exogenous parameter owing to Nash behavior. The associated first-order conditions can be written as

(4.8) $p = (r/R)F'(R) + (\bar{R}/R)[F(R)/R]$

with some manipulations. At a symmetric equilibrium involving n exploiters, we have $r/R = 1/n$ and $\bar{R}/R = (n - 1)/n$, which means that (4.8) can be expressed as

(4.9) $p = (1/n)F'(R) + [(n - 1)/n]F(R)/R.$

In (4.9), the price of a vessel is equated to a weighted sum of its marginal and average product. If $n = 1$, price equals marginal product and Pareto optimality is obtained; if, instead, $n \to \infty$, price converges to average product and profits approach zero. In the latter case, the full tragedy of the commons is experienced.

A comparison of the Nash fleet, R^N, for the commons that satisfies (4.8) with the Pareto-optimal fleet, R^*, shows an overexploitation in which $R^N > R^*$ when $n > 1$. Since R^* is independent of the number of exploiters, while R^N increases as n increases reaching the R corresponding to $p = F(R)/R$ in the limit, the ratio R^N/R^* becomes larger as n rises (Cornes and Sandler, 1986, pp. 128-131). In consequence, inefficiency does worsen as n increases, consistent with Theme 1 of *The Logic*. If the size of the exploiters is unequal, then a clear asymmetry can be shown in which effort and, hence, exploitation are positively correlated to fleet size in a pure commons. Small firms are crowded out of the commons by larger firms with mightier fleets. In consequence, the small is exploited by the large--the reverse of the usual exploitation theme.

Equation (4.9) indicates that the Nash equilibrium lies in-between the Pareto-optimal solution and the zero-rent equilibrium, R^O, associated with unlimited access so that $R^O > R^N > R^*$. As such, the Nash equilibrium depicted above can be viewed as that of a limited-access commons if exploitation is prior to the point at which $p = F(R)/R$. Profits need not be zero in the Nash equilibrium. With long-run considerations, entry will be pushed until profits, net of entry fees, are zero.

In an interesting series of experiments, Walker, Gardner and Ostrom (1990) tested the above theory of the commons. Each subject was allowed to

allocate tokens to a private good market and a commons market. In the latter, the subject's return declined with increased allocations by others. A summation technology characterized the payoff function and a player's payoff depended on his share of the total token allocated to the commons. Group size was maintained at eight; thus, the influence of group size was not investigated. These experiments showed that the Nash equilibrium best characterized the subjects.

If the commons were less pure, so that fish were not evenly distributed, allowing some fishing spots to be more productive than others, then firms may derive private benefits by positioning their fleets over these bountiful areas. The presence of these firm-specific private benefits can serve to lessen inefficiency. Symmetry of effort may also be lost if firms differ according to technology. Although *The Logic*'s characterization of collective-action problems is clearly analogous to the pure commons problem above, less stylized models may yield results at odds with the Olsonian themes. As in the case of public goods, assurance and other game structures may be more appropriate for some commons problems.

Until now, the output is assumed to be sold in a competitive market so that the price of fish has been treated as a constant normalized at a value of 1. An interesting second-best problem arises when the output of the commons is sold in an imperfectly competitive market in which the price of fish, P, is a function of the total catch so that the representative firm's profit is

(4.10) $\pi = \{[P(C)C/R - p]\ r\}$,

in a pure commons. Since demand inelasticity due to monopoly power leads to overconservation (i.e., a lessening of the tragedy of the commons), while an increase in the number of exploiting firms typically leads to an

underconservation, a finite number of firms for a commons can be found corresponding to a social or Pareto optimum (Cornes, Charles Mason, and Sandler, 1986). One distortion can serve to offset or ameliorate another, thereby limiting the need for outside intervention to correct for the collective-action failure. At a symmetric equilibrium, the Nash equilibrium associated with (4.10) would yield a Pareto optimum for the following number of exploiters:[4]

(4.11) $n^* = 1 + \epsilon_C/[(\epsilon_C - 1)\epsilon_D]$,

where ϵ_C is the elasticity of input productivity [i.e., $\epsilon_C = RF'(R)/C$] and ϵ_D is the price elasticity of market demand. Since the ϵ_C lies between zero and one, n^* equals or exceeds one.

Equation (4.11) depicts a number of possibilities. If market demand is perfectly elastic, then (4.11) implies the conventional wisdom: a single firm must exploit the commons to achieve a social optimum. For a given elasticity of input productivity, (4.11) indicates that the more inelastic is market demand, the greater should be the number of firms in the commons if a social optimum is to be achieved. This follows because the increase in market imperfection leads to a greater degree of conservation as firms restrict output, thereby taking advantage of buyer's unresponsiveness to price. Hence, more exploiters can be admitted to the commons, since the resulting increase in exploitation is needed to offset the conservation associated with monopoly power. Holding market demand elasticity fixed allows us to focus on the influence of input-side distortions by varying ϵ_C. In (4.11), as ϵ_C nears one, n^* approaches infinity, implying that free access is desirable. If ϵ_C = 1, then average and marginal products are equal; hence, equating average product to the real rental rate leads to no input-size distortion. As

diminishing returns increase and ϵ_C approaches zero, the commons problem intensifies and n^* must be restricted.

Sandler and Sterbenz (1990) have extended the pure commons model to cases where harvest uncertainty is present so that effort, in terms of the variable input (vessels), may be associated with different catch levels owing to random stock considerations (e.g., uncertain migratory patterns). To capture stock uncertainty, the aggregate catch is now

(4.12) $C = Z(R)\bar{X}_0 v,$

where $\bar{X}_0$ is the mean initial stock and v is a nonnegative random variable with a mean of one. With an uncertain stock, the representative firm now solves the problem,

$$\max_{(r)} \int_0^{\beta} U(\pi^*)g(v)dv,$$

in which $\pi^* = (r/R)Z(R)\bar{X}_0 v - pr$, $g(v)$ is the probability density function. Expected utility corresponds to the above integral. If the first-order condition for the above problem is investigated, uncertainty is seen to reduce the exploitation of the commons for a fixed number of risk-averse firms when compared with certainty (Sandler and Sterbenz, 1990, p. 159). As a result of risk aversion, uncertainty is undesirable for the firms. In this pure commons model, total uncertainty is fixed for the commons; but the uncertainty faced by each exploiter depends on its <u>share</u> of the exploitation efforts in the commons. By reducing its fleet, a risk-averse firm is therefore attempting to limit its own uncertainty. In so doing, the collective-action problem known as the tragedy of the commons is ameliorated. This amelioration also characterizes the long-run equilibrium as well as some stylized noncompetitive cases with risk-neutral and risk-averse exploiters.

The underdeveloped world appears especially prone to commons problems. A relevant example is the gathering of firewood for fuel, which has laid waste to forests throughout the developing world. When monsoons come, the loss of the forests means widespread flooding, famine, and pestilence. Within many developing nations, the property rights to the forests are allocated on a free access basis. In other cases, the forests that protect a lowland nation (e.g., Bangladesh) are found in upland nations with no incentives to consider the hardships that their exploitation can cause another nation. Commons problems are more prevalent in the developing world, since the means for assigning and enforcing property rights are at a more primitive stage than in the developed world. The law of the sea treaty, for example, gave the rights to the seabed resources within 200 miles to coastal nations in an effort to overcome the commons problem by making $n = 1$. Developing countries are apt to suffer under this arrangement for two reasons. Firstly, these nations do not have the resources or technology to enforce their 200-mile rights, thus leaving them vulnerable to exploiters from within and outside the country. Secondly, the wealth of the deep seabeds, located beyond 200 miles, can be taken by those developed countries with the technology to do so. Similar problems arise with respect to polar and outer-space resources.

Some of the most-pressing problems confronting mankind (e.g., ozone depletion, carbon dioxide buildup, and acid rain) arise because the atmosphere and air sheds are open-access commons. The buildup of carbon dioxide is especially difficult to deal with, since some nations stand to gain as the earth heats up. These potential gainers have perverse incentives that make global warming a potential windfall. The best chance for mankind to deal with this problem is now when uncertainty with respect to gainers and losers is

still present. Once gainers can be identified, they have a strong incentive to work against collective interests.

Institutional arrangements are, in some instances, needed to control overexploitation in the commons (see Ostrom, 1990). The use of contracting has been examined as a means for controlling overexploitation (Ronald Johnson and Gary Libecap, 1982; Libecap and Steven N. Wiggins, 1984). For fisheries, Johnson and Libecap (1982) have shown that contracts are difficult to formulate and enforce when heterogeneity characterizes the exploiters. In other words, pure commons with symmetric players are easier to control with contracts. This result is in keeping with Theme 2 of *The Logic* concerning asymmetries. For oil fields in Oklahoma and Texas, Libecap and Wiggins (1982) demonstrated empirically that the concentration of ownership facilitated contractual arrangements in agreement with Theme 1 of *The Logic*.

Sanctions can also be used to limit exploitation. Since these sanctions are themselves a collective benefit to the exploiter, their funding and provision pose additional difficulties. In addition, the certainty of enforcement is important. If, for example, exploiters do not believe that their activities can be monitored sufficiently, the threat of sanctions may not have much effect. Thus, the certainty of enforcement is a crucial concern. The imposition of sanctions could alter the game structure from that of prisoner's dilemma to that of "fully privileged".

In a renewable resource commons, the problem of extinctions is ever-present. Uncontrolled exploitation could sufficiently deplete populations so that the species becomes extinction, as in the case of the blue whale. In some instances, as extinction is approached, the payoffs for overexploitation may turn so negative that the dominant strategy may become "to cooperate".

Thus, the game structure may change from that of prisoner's dilemma to that of assurance, provided that players are properly informed about the resource status. The latter is, however, sometimes not known until it is too late.

4.9 Strategic Trade Policy

Collective-action problems are also relevant in the study of international trade. If perfect competition and nonincreasing returns to scale prevail so that the terms of trade cannot be influenced by a nation's actions, then the theory of comparative advantage indicates that social welfare can be maximized under a regime of free trade. Although it is in the collective interest of all nations to promote free trade, a defector may achieve individual gains at the expense of the collective by imposing protectionist policies (e.g., tariffs, quotas). Since each nation may experience a payoff matrix in which the imposition of a tariff is a dominant strategy, a classic prisoner's dilemma may apply and distortionary tariffs may characterize the entire trading community. The study of an optimal tariff involves the choice of the Nash equilibrium tariff, whereby the externality imposed on others is ignored. More sophisticated players may anticipate one another's strategies leading to nonzero conjectures. Asymmetries in information may support leader-follower behavior, whereby the follower moves first and the leader then chooses its tariff using the follower's Nash reaction path as its relevant constraint. Distributional aspects can be severely influenced by the strategic assumption.

In the last couple decades, trade economists have extended the theory to include increasing returns to scale and imperfect competition in their study

of free trade (Paul Krugman, 1987). The latter qualification is especially significant in terms of collective action, since imperfect competition raises a whole new set of justifications, whereby protectionist policies can yield nation-specific benefits at the expense of the international community. Strategic trade policy "holds that government policy can tilt the terms of oligopolistic competition to shift excess returns from foreign to domestic firms" (Krugman, 1987, p. 134). Industries that generate knowledge externalities (e.g., computer, aerospace) that can benefit the entire economy of the protectionist are the prime candidates for subsidies and/or other distortions that enable these industries to gain a foothold in the international market. Krugman (1987, pp. 135-37) demonstrated with a simple 2 x 2 matrix how, in the absence of a subsidy, competition between Airbus and Boeing would yield a coordination game with two pure-strategy Nash equilibria, in which one of the two firms controls the market. If, however, a sufficient European subsidy is given to Airbus, then there is only one pure-strategy Nash equilibrium with Airbus in control. Trade restrictionist policies, whether based on strategic considerations or otherwise, will lead to retaliation and could plunge the international community into another era of protectionism where all suffer.

The problem with this new analysis of strategic trade policy concerns its use of an atemporal model to analyze an intertemporal problem -- the adoption of new technologies. The atemporal model does not allow reputation costs, associated with repeated games, to be included. Although clear short-term benefits can come from strategic trade policies as put forward by Krugman, long-term costs may outweigh these gains as trading partners retaliate and impose other restrictions. A second problem with strategic

trade policy is its partial equilibrium nature with its focus on a single activity. Other activities may be affected through forward and backward linkages to the protection proposed; second-best considerations are relevant.

4.10 Principal-Agent Analysis

In recent years, economists[5] have developed agency theory to design an optimal contractual arrangement between a principal and an agent when asymmetric information prevails. In a firm, the principal is the owner or manager, while the agents are the workers; in a government, the principal is the electorate, while the agents are the elected officials; and in an union, the principal is the union membership, while the agents are union leadership. Asymmetric information is germane when a principal can view the final outcome (e.g., an output level, a provision level of the public good) but is unable to observe the agent's actual action. This situation would prevail when an exogenous risk factor intervenes in the output or provision process so that the agent's effort is no longer uniquely tied to each outcome. Rather, a distribution of outcomes is associated with each effort level. Information is one-sided since the agent, and not the principal, knows the agent's true effort. This asymmetric information leads to a moral hazard problem, insofar as the agent may take advantage of the principal's ignorance as an excuse to supply suboptimal levels of effort. Agent-principal difficulties can be minimized if the principal can design a contract or payment schedule for agent efforts that induces the agent to supply a fuller effort level. An entire schedule relating effort to payment must be chosen to maximize the principal's welfare subject to incentive-compatible and individually rational constraints.

Incentive compatible constraints account for the agent's optimizing responses (in terms of first-order conditions) derived in the first step of the procedure, while an individually rational constraint ensures that agents' welfare is no less with the contract than in alternative employment opportunities. Such contracts often involve the sharing of risk between the principal and the agent, unless the latter is extremely risk averse. If a first-best result cannot be achieved by the contract, then agency costs arise in terms of losses associated with second-best contracts.

Although there are many general representations for the agent-principal problem, the analysis is sufficiently complex for even single-agent, single-principal problems to preclude exact or closed-form solutions for the optimal payment schedule. Unlike the standard optimization problem in which a scalar value is chosen for a single level problem, principal-agent analysis involves the choice of an entire function (i.e., a payment schedule) for a multi-level problem. At least two levels are required, since optimizing the principal's welfare must be consistent with the first level involving the agent's welfare maximization.

At least two relevant aspects of the principal-agent analysis apply to the theory of collective action. The first involves the number of agents, while the second concerns the use of nonmarket institutions to correct collective-action failures. In regards to the first, Bengt Holmstrom (1982) and Eric Rasmusen (1987) have noted a free-rider problem common in teams when individual effort is unobservable. The difficulty that Holmstrom recognized is a classic collective-action problem that limits individual efforts, induced by payment schedules. With symmetric players, each shirker loses pay equal to only 1/n (where n is the team size) of his reduced effort, but gains in the

efforts of others. For linear payment schemes, individual effort decreases in proportion to the team size. The larger the team, the more inefficient individual effort. Thus, the first theme of ***The Logic*** is fully supported (Cauley and Sandler, 1991). In fact, the problem displayed by Holmstrom (1982) is isomorphic to the commons problem of section 4.8. When the team confronts a fixed amount of uncertainty, an increase in team size may have positive influences that offset, to some extent, free riding. If multiple risk-averse agents are involved, an increase in team size makes the risk premium for agents fall by a factor of $(1/n)^2$ owing to risk spreading (Cauley and Sandler, 1991). Since an increase in team size is a double-edged sword owing to negative free-riding influences and positive risk-spreading influences, an optimal-sized collective could be determined to balance these forces at the margin.

To circumvent the free-rider problem, nonlinear payment schedules are proposed that, like preference-revelation mechanisms, make each agent face through the use of side payments the social choice problem of the team. This can be accomplished by withholding rewards to each agent unless Pareto-optimizing levels of output are achieved, so that the existence of even a single shirker would cause harm to all team members (Holmstrom, 1982, p. 327). Many variations of this theme are possible (Rasmusen, 1987).

A second aspect of the principal-agent problem that concerns collective action has to do with Theme 3 of ***The Logic***, in which mechanisms are designed to overcome suboptimality. Public policy itself poses a principal-agent problem. Since the agent (e.g., elected official, bureaucrat) has more information regarding their efforts and the random states of nature than the principal for whom they serve, informational asymmetries may be present.

Thus, the provision of a pure public good by a government may not improve greatly on efficiency unless an incentive mechanism is put into place to motivate elected officials. Since this mechanism is costly, transaction costs in the form of agency costs much be considered when determining whether a nonmarket structure should be instituted to provide the collective good.

A more basic information problem confronts these nonmarket institutions that may still exist even though the principal is properly motivated by an incentive scheme. That is, the principal must still ascertain the preferences of his/her constituency so as to determine the provision level of the collective good. Hence, a properly functioning nonmarket institutional structure requires well-informed agents who are sufficiently motivated to satisfy the (revealed) preferences of the constituency. This is a tall order.

4.11 International Organizations

Thus far, two international organizations--military alliances and the World Health Organization--have been mentioned as a means for confronting collective-action problems. The alliance provides deterrence, damage-limiting protection, and other benefits to the allies, whereas WHO provides disease containment and eradication for its member nations. Many other international organizations have come into existence owing to collective-action considerations. A few examples include the United Nations (U.N.), INTELSAT, the European Common Market, Interpol, pollution pacts, and the European Space Agency. The U.N. tries to promote world peace (a pure public good), while pursuing a host of other humanitarian goals (e.g., curbing world hunger). INTELSAT links over 80 nations in an external communication network made up of

geostationary satellites that carry over two-thirds of all transoceanic messages. The European Common Market promotes free trade among member nations, while affording protective barriers to nonmember products. The promotion of trade within the trading bloc provides a pure public good in terms of efficiency gains to the members. Interpol is a transnational police linkage that facilitates the sharing of intelligence with respect to criminal activity of a transnational character (e.g., terrorism, drug running). Pollution pacts, such as the Baltic Pact, give forth a pure public good to its members in the form of a cleaner environment (see, e.g., Michael Hoel, 1991). Finally, the European Space Agency allows the countries of Western Europe to pool resources so as to achieve sufficient scale economies to develop launch capabilities, space colonization, and space explorations.

When the analysis is at the supranational level, participants may include nation-states, whose interests are represented by policymakers. In international organizations, participants may include both governments and private corporations as in the case of INTELSAT. Clearly, the principal-agent considerations, raised above, abound in international organizations owing to the additional supranational level. Unless policymakers are given the proper incentive schedule, they may not pursue the goals of their constituent principals, which are often the nation's population. At the agent level, cooperation must be fostered by the institutional arrangement so that net gains can be achieved beyond the autarchic noncooperative equilibrium.

Transaction costs are a crucial consideration for the supranational structures since sizable expenditures of fixed costs in terms of communication, infrastructure, and administration are required. Variable transaction costs arise as the structure or institution is utilized to decide

policies. Unless a sufficient number of policies are enacted, a nonconvexity in the transaction set might exist so that the international organization cannot achieve sufficient benefits to match or exceed transaction costs. Since many inputs (e.g., communication networks, infrastructure, administrative staff) can be used for a host of different activities, these inputs are public in the sense that they enter the production function of several policy outputs simultaneously. The presence of public inputs is a sufficient condition (William Baumol, John Panzar, and Robert Willig, 1988, pp. 75-77) for economies of scope, which exists when

(4.13) $C(q_1, q_2) < C(q_1, 0) + C(0, q_2)$

holds. In (4.13), the cost, $C(q_1, q_2)$, of providing both q_1 of activity 1 and q_2 of activity of activity 2 in a single institution is cheaper than the total cost of providing the two activities in separate institutions.

If economies of scope truly characterize international organizations, then these institutions should provide multiple collective goods. This, indeed, appears to be the case. The NATO alliance, for example, not only provides deterrence, but also polices elicit drug trade, improves highway safety throughout Europe, and facilitates scientific research. The U.N. engages in a wide range of peace-promoting, educational, philanthropic, and scientific pursuits. Common markets not only fosters free trade, but promotes technology development and transfer among members. The pursuit of these multiple benefits may be an important means for circumventing nonconvexities.

The appearance of economies of scope makes it more difficult to determine an optimal group size when club goods are shared by these institutions. This follows because economies of scale with respect to the membership size for one activity may have little relationship to economies of

scope when multiple outputs with diverse congestion functions are involved. Thus, the introduction of economies of scope requires us to redefine the optimal membership when excludable collective goods are shared by international organizations or other types of clubs.

A final issue concerns the nature of the supranational linkage: Should participating nations be tightly or loosely joined? In a tight link, the participating nations sacrifice its autonomy to the collective, while in a loose link, the participants do their own thing with little consultation or coordination with their counterparts. A structure like NATO, in which less than one percent of its expenditure is commonly funded, is a loose structure (Sandler and Forbes, 1980, Table 1, p. 432). Moreover, NATO allies use an unanimous voting rule that also serves to keep the structure loose. Collectives that share activities with a large proportion of excludable benefits, as may be the case in NATO since the late 1960s, need not coordinate activities closely to take advantage of efficiency gains. Surely the nature of the collective good as well as transaction costs are behind the optimal architecture of an international organization.

4.12 Rent-Seeking Behavior

Another example of collective action is the activity of rent seeking, in which agents or a collective expend resources to obtain a return that results in no net addition of output to society. Robert Tollison (1982, p. 578) states that, "Rent seeking is the expenditure of scarce resources to capture an artificially created transfer." Distributional coalitions, such as those behind Olson's *Rise and Decline of Nations*, are rent seekers since they spend

resources in trying to obtain a greater piece of the economic pie without adding to the size of the pie through their activities. As such, rent seeking denotes an economically wasteful activity and is an example where the pursuit collective gains can lead to losses at an aggregate level.

The notion of rent seeking was first put forward by Gordon Tullock (1967) when he argued that the waste of monopoly included more than the traditional deadweight triangular loss. Figure 4.1 illustrates Tullock's analysis for the monopoly case. In Figure 4.1, a linear demand curve and its marginal revenue curve (dashed curve MR) are displayed. If marginal cost is horizontal and if costs are independent of market structure, then the monopoly charges P_M and sells Q_M, while the perfect competitive industry charges P_C and produces Q_C. The deadweight loss is equal to the area of triangle ABC, whereas monopoly profit is the area of rectangle P_CP_MBA. Tullock considers this profit area as a potential source of additional social loss through rent seeking. Potential entrants or competitors may induce the monopolist to expend resources up to area P_CP_MBA to protect its rent. When competitive rent seeking occurs, society loses not only ABC, but also P_CP_MBA. Tullock (1980) has shown that when groups or individuals compete for a rent in a noncompetitive tournament arrangement that each will carry its rent seeking activity until marginal costs equal marginal return from rent seeking. In the Tullock model, marginal adjustments in bids are allowed, since the likelihood of winning equals the rent seeker's expenditure on the tournament as a proportion of the total expended by all rent seekers. Tullock (1980) demonstrated that, in total, rent seeking captures only half of the rectangle when just two players compete.

Eli Applebaum and Eliakim Katz (1986) have extended Tullock's tournament

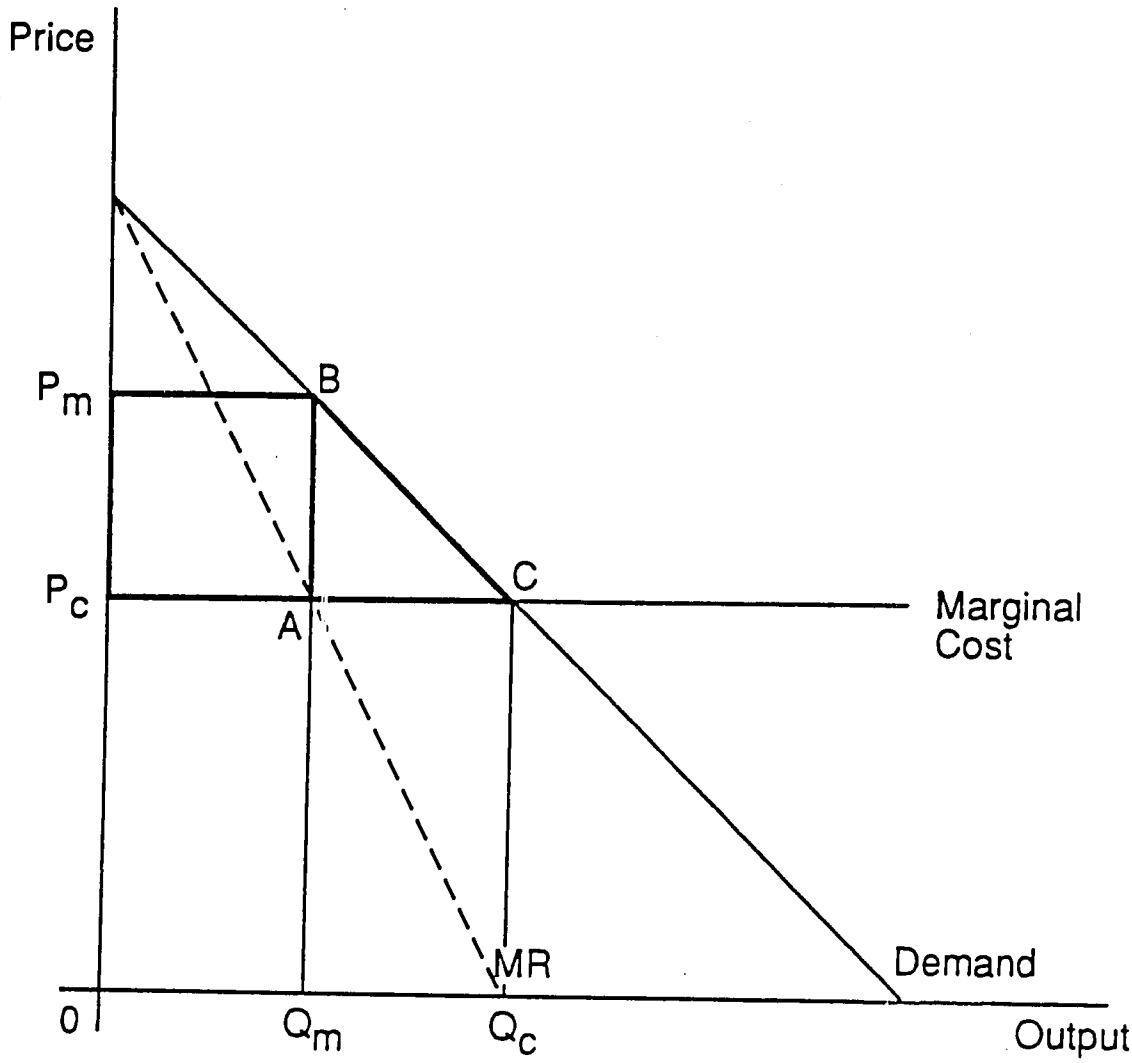

Figure 4.1: Rent seeking for monopoly

result to include not only the prize, but also the deadweight loss from rent seeking. Among other results, these authors demonstrated that total rent seeking increases with the size of the rent (prize) and the deadweight loss, and decreases with the number of competing groups when a symmetric equilibrium is assumed. Tullock (1980) has examined situations in which there is overexploitation of the rent-seeking rectangle. A classic collective-action problem then results, since everyone could be better off by not rent seeking.

Katz, Nitzan and Jacob Rosenberg (1990) altered Tullock's (1980) model to the case where two groups of differing sizes pursue rents by lobbying a government for public goods. Group 1 with n members lobbies for public good x that benefits <u>each</u> member by αR dollars with $0 < \alpha < 1$; group 2 with m members lobbies for public good y that benefits each member also by αR dollars. For each group, once the good is provided, the per-person ex post benefit is independent of group size. Furthermore, benefits are not based on a summation technology. Following Tullock, the likelihood of success (ϕ) for groups 1 and 2 are

$$(4.14) \quad \phi_1 = \sum_{i=1}^{n} x_i / \left(\sum_{i=1}^{n} x_i + \sum_{i=1}^{m} y_i \right)$$

$$\phi_2 = 1 - \phi_1 = \sum_{i=1}^{m} y_i / \left(\sum_{i=1}^{n} x_i + \sum_{i=1}^{m} y_i \right).$$

where x_i is the contribution of the ith member of group 1 and y_i is the contribution of the ith member of group 2. With risk neutrality, the expected utility of individual i in group 1 is

$$(4.15) \quad U_i = \phi_1(\alpha R - x_i) + \phi_2(-x_i),$$

which is equivalent to

$$(4.16) \quad U_i = \phi_1 \alpha R - x_i.$$

Choosing x_i for each member of group 1 gives the following first-order

condition at a symmetric equilibrium where $x_i = x$ for every i and $y_j = y$ for every j:

(4.17) $my/(nx + my)^2 = (\alpha R)^{-1}$.

By an analogous set of steps, the first-order condition for a group 2 member is

(4.18) $nx/(nx + my)^2 = (\alpha R)^{-1}$.

If the two sides of (4.17) and (4.18) are added together and the result simplified, we get an expression for total rent seeking,

(4.19) $nx + my = \alpha R/2$,

for the symmetric Nash equilibrium. This finding is reminiscent of Tullock's (1980) result. In (4.19), total rent seeking is only 50 percent of the per-person rent-seeking benefits.

Equation (4.19) is neutral to the distribution of individuals between the groups. This result is in contrast to Theme 1 of *The Logic* and arises owing to the technology of public supply. To pursue this finding, we follow the methodology of Katz, Nitzan, and Rosenberg (1990, pp. 52-53), and divide (4.17) by (4.18) to yield

(4.20) $my = nx$,

or that total rent seeking is equal between groups so that success probabilities are one half. For each group, (4.20), (4.18), and (4.17) implies

(4.21) $x = \alpha R/4n$

(4.22) $y = \alpha R/4m$.

By these equations, we see that an increase in group size decreases individual rent-seeking efforts since $\alpha x/\alpha n = -R/4n^2$ and $\alpha y/\alpha m = -R/4m^2$. Owing to

symmetry, total contributions decrease by $\alpha R/4n$ and $\alpha R/4m$ in the two groups, respectively. This decrease is exactly offset by the entrant's contribution as seen in (4.21) and (4.22); hence, group size is immaterial to the equilibrium. Obviously, other technologies of supply than the one considered could make group size a more important determinant of relative rent-seeking activities between groups. Group heterogeneity may also affect the role played by relative group size. Thresholds or discrete aspects of group support could also alter the influence of group size in competing rent-seeking collectives.

A second potential instance of rent seeking is depicted in Figure 4.2 where an externality associated with the output is corrected with the use of a quota or output standard of q^*. Marginal private cost is depicted by curve MPC, while marginal external cost is denoted by curve MEC. For each level of output, both a private and an external cost are incurred. The latter is experienced by third parties in the population. To find the marginal social cost (MSC) curve, we sum the MPC and MEC curves vertically at each output level. The demand curve for the outputs is schedule D. With no government intervention, output is at q^o, and MSB is less than MSC. The imposition of a quota at level q^* where MSC = MSB assists producers and externally affected third parties at the expense of consumers. In Figure 4.2, consumers lose area P^oP^*ac of consumer surplus. Producers gain the difference between area P^oP^*ae and area bec of producer surplus, while third parties gain (or avoid) area badc. All together, the quota leads to a deadweight loss of area adc.

Since producers can actually gain from a quota, they might lobby the government for this kind of intervention as opposed to a per-unit Pigouvian tax of ab at output level q^*. With such a tax, producers lose area PP^ocb,

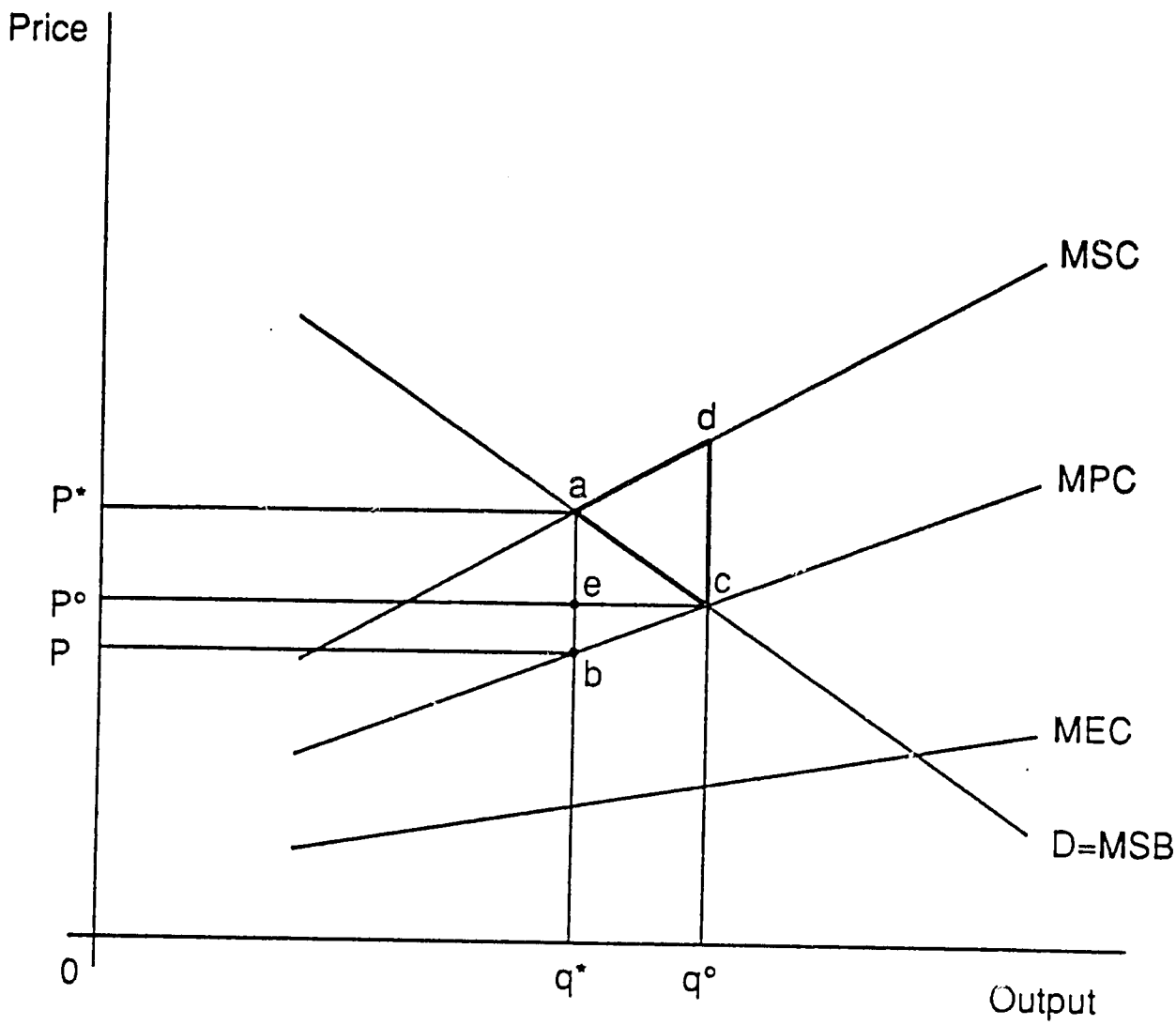

Figure 4.2: Rent seeking external correction

while consumers and third parties welfare is unchanged as compared with the standard. In contrast, the government gains PP^*ab in tax revenues. Although the deadweight loss is still area adc so that efficiency considerations have not changed, distributional changes will induce rent-seeking activities on the part of producers. These activities could add to the inefficiency by as much as area PP^*ab. Such lobbying efforts could explain why standards are indeed more prevalent than taxes.

4.13 Economies of Scale, Economies of Scope, and Collective Action

This subsection is concerned with cost considerations that can promote collective action. In chapter 2, we discovered by way of examples that the technology of collective supply, institutional structures, and the strategic assumption are important in determining the underlying game structure. The prisoner's dilemma does not need to characterize the game structure for collective-action problems. We now demonstrate that cost considerations can influence the underlying game structure.

Economies of scale that involve a decrease in average cost can give a game structure in which mutual cooperation is among the Nash equilibria, as illustrated by the following example. Suppose that two players can each contribute one unit of the public good. Suppose further that if two units are produced in the same production process, then each unit costs $4. If, however, a single unit is produced, then its cost is $6. Scale economies are present since per-unit cost decreases when production is done cooperatively in a single plant. If <u>each</u> unit produced yields $5 in benefits to contributor and noncontributor alike, then the top 2 x 2 matrix in Figure 4.3 indicates

FIGURE 4.3
POTENTIAL EFFECTS OF ECONOMIES OF SCALE
AND ECONOMIES OF SCOPE ON GAME STRUCTURE
2 x 2 GAMES

		2: Cooperate	2: Defect
1	Cooperate	a 6, 6	b −1, 5
	Defect	c 5, −1	d 0, 0

ECONOMIES OF SCALE

		2: Cooperate	2: Defect
1	Cooperate	a 4, 4	b −1, 5
	Defect	c 5, −1	d 0, 0

SINGLE ACTIVITY

		2: Cooperate	2: Defect
1	Cooperate	a 12, 12	b 2, 10
	Defect	c 10, 2	d 0, 0

DOUBLE ACTIVITY

ECONOMIES OF SCOPE

the net payoffs to the two players, with player 1's payoff listed first. When both cooperate in contributing and production, they each gain 10 (= 2 x 5) at a cost of 4 for a net payoff of 6. If only one player contributes, then the contributor gains a net payoff of −1, since benefits of 5 come at a cost of 6 owing to the absence scale economies at one unit of output. The noncontributor can then free ride on the contributor for a gain of 5 without any cost. The payoff structure in the matrix is somewhat reminiscent of the Assurance 1 game of Figure 2.1, since the Nash equilibria are the same in pure strategies.

There is no dominant strategy. In pure strategies, Nash equilibria correspond to cells a and d where both players cooperate or defect. Unlike the prisoner's dilemma, defection is not a dominant strategy. If, moreover, player 1 (2) cooperates, it is then in player 2's (1's) interest to cooperate whenever scale economies reduce per-unit cost below the per-unit benefits. As in the assurance game, contracts are self-enforcing; if player 1 cooperates, then player 2 gains by cooperating. In an n-person game, these economies of scale can reduce the incentives for free riding as group size increases and per-unit cost falls further. Many clubs find their motivation in these economies of scale. In essence, exclusion is practiced with respect to cost savings, since an agent must contribute to gain the cost saving. Economies of scale implies decreasing average cost which, in turn, implies strictly subadditive costs:[6]

$$(4.23) \quad C\left(\sum_{i=1}^{n} q^i\right) < \sum_{i=1}^{n} C(q^i),$$

where n is the number of contributors, q^i is the number of units contributed by the ith agent, and $C(\bullet)$ denotes costs. Equation (4.23) indicates that the cost of producing the total number of units in one facility is less than the

sum of the costs of producing q^i in separate facilities. In the above equation, cost subadditivity applies since $C(2) < 2C(1)$.

In many public good situations including the provision of infrastructure, scale economies exist and costs are subadditive, thereby making collective action advantageous. This advantage may occur within a nation as a project serves multiple jurisdictions, or between nations as a project serves many nations. If the output of a production process is based on volume, while costs are based on surface area, then scale economies are likely. Scale economies may be obtained without collective production provided that both parties buy from the same producer whose average costs are declining.

The presence of scale economies as a motivation for collective action is well known. We merely emphasize that these economies may alter the game from that of the prisoner's dilemma to one more in keeping with collective action.

The number of activities undertaken by the collective may also be an important determinant of the feasibility of collective action. To illustrate, we consider a case of economies of scope as depicted in (4.13), where the costs of providing two activities is cheaper than the sum of the stand-alone costs of providing them separately. Suppose that the costs of providing <u>each</u> unit of collective activity one or two is \$6, while the costs of providing a unit of <u>both</u> activities is \$8. Further suppose that each unit of either activity gives <u>each</u> of the two players \$5 in benefits. In the bottom row of Figure 4.3, the left-hand matrix indicates the net payoffs for the provision of either activity alone. A classic case of the prisoner's dilemma applies for either activity provided alone, since per person benefit is less than per person costs. In consequence, neither activity is undertaken.

If, however, both activities are provided, then the bottom right-hand matrix of Figure 4.3 applies. For cell a, both players provide a unit of both activities at a per-person cost of $8 and a per-person benefit of $20. Gross benefits are $20 since each player gains 5 from each of the 2 units of the 2 activities; net payoffs are consequently $12. The other payoffs are computed similarly. In cell b, for example, player 1 cooperates by providing a unit of both activities at a cost of 8 with benefits of 10 to both players. With both activities provided, the existence of economies of scope makes the game fully privileged with cooperation as the dominant strategy. Thus, the presence of economies of scope may mean that more than one activity must be done by the group to elicit collective action. The latter may only be desirable for some combinations of activities.

4.14 Public Inputs and Infrastructure

An input is public or collective when it enters two or more firms' production functions.[7] If the input is nonexcludable and perfectly indivisible between firms, then the input is purely public. Public inputs may enter two or more firms' production functions (e.g., industry-specific knowledge or research findings), or they may enter two or more production functions with respect to different industries or products (e.g., highways are used for a number of different products). The latter situation may give rise to economies of scope (Baumol, Panzar, and Willig, 1988, pp. 76-79) and may justify a firm or an organization providing multiple products.

The existence of public inputs is quite prevalent in the real world and is behind a host of collective-action problems. Scientific findings, for

example, are not final consumption goods, but intermediate inputs used to produce final commodities. Similarly, an arsenal, held by an alliance or a nation, is an intermediate input that produces security, which is a final consumption good. These final goods may themselves be private or public. Thus, a rich set of possibilities exists. In the case of security, the final commodity is public between citizens of the country. Infrastructure may be used to yield output for firms, in which the final commodity (i.e., the firms' outputs) is private between the users of the public input.

Many possible models can be constructed. We present a simple representation due to Kaizuka (1965), Sandmo (1972), and McMillan (1979). Suppose that a single primary factor, labor, is utilized by n firms to produce a private consumption good, q. The ith firm produces q^i according to the following well-behaved twice-continuously differentiable production function:

$$(4.24) \qquad q^i = f_i(L_i, r), \qquad i = 1, \ldots, n,$$

where L_i is the labor used by firm i and r is a public intermediate good. The production function is increasing in both inputs and, moreover, is assumed to be strictly concave. The public input is itself produced by labor under constant returns to scale:

$$(4.25) \qquad r = f_r(L_r).$$

In (4.25), production of r is positively related to the input of labor (L_r), and the average productivity of labor (r/L_r) equals the marginal productivity (df_r/dL_r) owing to constant returns to scale. The model is completed by the labor distribution constraint,

$$(4.26) \qquad L = \sum_{i=1}^{n} L_i + L_r,$$

which indicates that the fixed supply of labor, L, is allocated between the production of the private consumption good and the production of the public input. In the former case, production processes (one for each firm) may draw

labor. Although r enters each firm's production function, a single production process is relevant, since, once produced, r is utilized by all n firms owing to the publicness assumption.

The Pareto-optimal requirement for efficient input supply is derived by maximizing any firm's output subject to the constancy of the other n − 1 firms' outputs and to the labor distribution constraint. In the Appendix, we derive the Pareto-optimal condition:

$$(4.27) \qquad \sum_{i=1}^{n} \frac{(\partial f_i/\partial r)}{(\partial f_i/\partial L_i)} = \frac{1}{df_r/dL_r},$$

which states that the <u>sum</u> of the ratios of marginal productivities between the public input and the primary input <u>equals</u> the reciprocal of the marginal productivity of labor in producing the public input. In (4.27), the left-hand side depicts the marginal benefits derived by the entire industry from an additional unit of the public input, while the right-hand side denotes the marginal costs, in terms of additional labor, associated with an additional unit of the public input. Since the industry consists of n firms, the marginal benefits must be aggregated over the entire industry. These marginal benefits correspond to the marginal productivities of r, normalized by the marginal productivity of L_i in producing good q^i. As such, (4.27) is reminiscent of the Samuelsonian sum of MRSs = MRT condition.

The analysis can be complicated in many ways. First, the input r may be public both within and between industries. This complication by itself would show up as a change in the summation indices, so that all affected firms' marginal productivity ratios are included. A double summation over firms within an industry and over industries may be required. Second, the number of public inputs may be increased beyond one. This extension would require a set

of conditions like (4.27), one for each public input (Manning, Markusen, McMillan, 1985, p. 235). Third, additional primary inputs could be included. In this case, each pair of primary inputs would require their ratio of marginal products to equal the input price ratio for all firms using both inputs. Fourth, congestion may characterize the public input, so that the utilization of the input reduces its level. This can be modeled by making r depend on primary inputs and some measure of public input utilization. When congestion is present, group size becomes an important variable to determine along with input allocation. Toll arrangements for users of the public input can "internalize" or adjust for crowding costs through user charges. In other words, a club arrangement can be instituted. Fifth, the public input may be used to produce a public consumption good. Publicness would then involve firms at the input stage and consumers at the output stage. Market failures may consequently be double tiered.

Financing the public input may prove especially problemmatic. We first examine financing possibilities for a single pure public input case as modeled above. Following McMillan (1979) and Sandmo (1972), we sketch a tax-based scheme under the best possible scenario, in which each firm pays a Lindahl-type-tax price based on its derived marginal productivity from the public input. The ith firm would attempt to maximize profits, π_i:

(4.28) $$\pi_i = p_i f_i(L_i, r) - wL_i - t_i,$$

where t_i is its tax share and p_i is the price of the private consumption output. If the public input industry is assumed to have zero profit wherein tax revenue just covers expenditures, we have

(4.29) $$\sum_{i=1}^{n} t_i = wL_r,$$

in which w is the wage rate for labor. With constant returns characterizing

$f_r(\bullet)$, the proper per-unit tax price is

$$t_i^*/r^* = p_i(\partial f_i/\partial r), \qquad i = 1, \ldots, n. \tag{4.30}$$

In (4.30), the ith firm pays a per-unit public input tax price equal to its marginal value product associated with the public input. As shown in the Appendix, this tax price meets the efficiency condition in (4.27).

Although a marginal productivity pricing scheme is a theoretical possibility, it is unlikely to be feasible in large industries unless a central authority can compute the public input's marginal productivities for each of the firms. The profit function in (4.28) can be rewritten as

$$\pi_i = \pi_i(L_i, t_i, t_i + \sum_{j \neq i} t_j), \tag{4.31}$$

with the use of (4.29) (see McMillan, 1979, p. 89). The Nash equilibrium choice of t_i, associated with maximizing profits, would achieve the following set of equations:

$$\frac{\partial f_i/\partial r}{\partial f_i/\partial L_i} = \frac{1}{df_r/dL_r}, \qquad i = 1, \ldots, n, \tag{4.32}$$

in which each firm ignores the benefits that its tax contribution to the public input confers on others [see equation (4.27)]. In consequence, too little of the public input is provided--a standard collective-action result.

There are, however, some reasons to be more hopeful with respect to public inputs in contrast to public goods. Public inputs may, at times, involve a small number of firms. In such cases, merger may occur owing to scale economies and/or scope economies resulting from the public input. If, say, three firms are affected by a public input, then a three-way merger motivates the conglomerate firm to account for all relevant marginal productivities associated with resources allocated to input r when making allocation decisions.

Since firms in a public input situation must interact repeatedly, a repeated game scenario is particularly relevant (McMillan, 1979, pp. 91-97). If the game's endpoint is uncertain and/or if the player's time horizon is infinitely long-lived, then a cooperative solution may be an appropriate Nash equilibrium, provided that discount rates are not too high.

In essence, the public input problem is a generalization of the commons problem. For the latter, a fixed (often nature-determined) input is public to a set of exploiters, who do not have to pay for the input's use. Moreover, the fixed input is subject to congestion which reduces the input's productivity. In the more general case considered here, the public input need not be fixed and may be augmented through production. For public inputs, crowding may, but need not, occur. If the number of firms is sufficiently small and/or interactions are repeated indefinitely, cooperation may evolve for public inputs as in the commons. The analogy with the commons is important, because of the wide range of institutional structures that have developed to manage the latter (see, especially, Ostrom 1990). Many of these institutional arrangements could work for public inputs.

An important institutional solution, short of merger, is contracting between firms. If firms face repeated interactions among one another with respect to a public input, a contract by "convention" may result, in which each firm adheres to cooperative behavior (see Hardin, 1982, for an excellent analysis of contract by convention). Such conventions are feasible when either the game structure does not have a dominant (single-shot) defection strategy or else a cooperative Nash equilibrium exists in the repeated game framework. For public inputs, resolution by contract is most likely when the number of firms is small. If, however, a central authority representing the

firms' interest (e.g., a producer association) comes into existence, contracting is possible even in a large number scenario. An interesting real-world example is generic advertising in industries characterized by numerous firms. Generic advertising is a public input since it increases the profit of all firms in the industry and is produced itself by primary factors.[8] The beef, pork, milk, and orange-growers industries engage in generic advertising to increase sales and profits of member firms. In these four cases, a producer association taxes it firms to finance the advertising campaign. Obviously, it is in no single firms' interest owing to product homogeneity and, consequently, its small share of gains to engage in its own advertising campaign.

The technology of supply can come into play in at least two different ways--through the production function for the private consumption good, $f_i(\bullet)$, and through the production function for the public input, $f_r(\bullet)$. Since the commons is not necessarily produced by primary inputs, only the former is likely to be relevant in the study of the commons. Henderson (1974), Manning, Markusen, and McMillan (1985), and Martin, Zacharias, and Lange (1990) have analyzed the importance of the form of $f_i(\bullet)$. To illustrate a couple important technologies, we now assume that each q_i is produced with a vector of m primary inputs, denoted by $x^i = (x^i_1, \ldots, x^i_m)$, and a single public input so that

$$q^i = f_i(x^i, r). \tag{4.33}$$

An important form for (4.33) is

$$q^i = f_i(x^i, r) = g(r)\bar{f}_i(x^i), \tag{4.34}$$

where $f_i(\bullet)$ is multiplicatively separable in the two inputs and $f_i(\bullet)$ and $\bar{f}_i(\bullet)$ are assumed linearly homogeneous in only the primary inputs. Thus, a

doubling of the components of x^i doubles output for a fixed level of the public input. Increasing returns to scale are present since doubling all inputs including the public input would consequently more than double output.

The production technology in (4.34) leads to a financing problem when a Lindahl-pricing scheme is utilized, since insufficient revenues are collected (Manning, Markusen, and McMillan, 1985, pp. 236-37). Problems of self-financing are always associated with increasing returns to scale since marginal-cost-based financing, such as Lindahl pricing, does not collect enough owing to the marginal cost being less than average cost. In fact, returns to scale occurs when the ratio of marginal cost to average cost is less than one. This problem can be circumvented in a second-best manner by adding a full-financing constraint requiring the sum of tax revenues to cover the resource costs of the public input. The resulting tax scheme is termed Ramsey pricing and is dependent on output shares of the firms and elasticity of supply considerations.

Another possible form for (4.33) is

$$q^i = f_i[k_i(x_1^i, r), \ldots, k_m(x_m^i, r)], \tag{4.35}$$

in which the public input augments the factor productivity of each primary input so that $\partial k_j/\partial r > 0$ but $\partial k_j/\partial r \neq \partial k_s/\partial r$ so that the augmentation need not be neutral or equivalent between inputs as in (4.34). Obviously, numerous different scenarios are possible and each can have different implications for self financing and incentives for collective action.

4.15 Concluding Remarks

This chapter has attempted to indicate the rich array of applications

that involve the theory of collective action. Each of the subsections could itself be expanded into a chapter in its own rights, since each application has a significant body of literature with many related topics of interest. Our intent has been to introduce the reader to the large number of applications available. The interested reader could gain a more in-depth knowledge by referring to the readings listed. Numerous other applications, not considered, exist including the study of revolutionary movements, the control of pollution, the pursuit of scientific discovery, the adoption of advanced technologies, and the modification of weather. Hence, the applications here are merely representative. Further applications are considered in Chapters 5 and 6, where empirical studies and modern-day contingencies are investigated.

APPENDIX

Derivation of Optimality Condition (4.27)

The Lagrangian is

(1) $\mathcal{L} = f_1[L_1, f_r(L_r)] + \sum_{j \neq 1}^{n} \lambda_j [f_j[L_j, f_r(L_r)] - \bar{q}_j] + \sigma(L - \sum_{i=1}^{n} L_i - L_r)$,

where $\bar{q}_j$ is fixed level for the jth firm's output of the private consumption good. Maximizing (1) with respect to L_1, L_j, and L_r gives the following first-order conditions:

(2) L_1: $f_{11} - \sigma = 0$

(3) L_j: $\lambda_j f_{jj} - \sigma = 0$

(4) r: $f_{1r}(df_r/dL_r) + \sum_{j \neq 1} \lambda_j f_{jr}(df_r/dL_r) - \sigma = 0$,

where $f_{11} = \partial f_1/\partial L_1$, $f_{1r} = \partial f_1/\partial L_r$, $f_{jj} = \partial f_j/\partial L_j$, and $f_{jr} = \partial f_j/\partial L_r$. First, we use (2)-(3) to express σ in terms of f_{11} and $\lambda_j f_{jj}$. Second, both sides of (4) are divided by σ and df_r/dL_r to yield

(5) $f_{1r}/\sigma + \sum_{j \neq 1} \lambda_j f_{jr}/\sigma = 1/(df_r/dL_r)$.

Using the expressions for σ, derived with (2)-(3), in (5), we immediately transform (5) into (4.27).

Full Financing of the Efficiency Condition in (4.27)

We next intend to show that

(6) $t_i^*/r^* = p_i(\partial f_i/\partial r)$,

for i = 1,..., n implies (4.27). We first sum both sides of (6) over the firms to yield

(7) $\sum_{i=1}^{n} t_i^*/r^* = \sum_{i=1}^{n} p_i(\partial f_i/\partial r)$.

If labor is used efficiently in producing q^i, then

(8) $w = p_i f_{ii}$, $\qquad i = 1, \ldots, n$,

which follows from a maximization of profits in (4.28) with respect to L_i while treating r and t_i as parameters. Rewriting (8) as $p_i = w/f_{ii}$ and substituting for p_i in (7), we get

(9) $\Sigma t_i^*/(wr) = \Sigma f_{ir}/f_{ii}$,

where indices of summation are suppressed. By the zero profit condition of (4.29), equation (9) becomes

(10) $L_r/r = \Sigma f_{ir}/f_{ii}$,

which is (4.27) since $L_r/r = (df_r/dL_r)^{-1}$ by constant returns.

FOOTNOTES

1. The first two historical examples were developed by John A. C. Conybeare and appeared in an earlier draft of a paper dealing with the Triple Alliance and Triple Entente, two pre-World-War-I alliances (see Conybeare and Sandler, 1989, 1990).

2. See Sandler and Cauley (1975), Sandler (1977), Murdoch and Sandler (1982, 1984a), and Sandler and Forbes (1980). Also see Boyer (1989), Gonzales and Mehay (1990, 1991), and McGuire (1990).

3. See, e.g., Cornes and Sandler (1981, 1984a), Andreoni (1987, 1989, 1990), Posnett and Sandler (1986, 1988, 1989), and Steinberg (1986, 1987).

4. See Cornes, Mason, and Sandler (1986, pp. 642-45) for a derivation.

5. See, for example, S. A. Ross (1973), Joseph Stiglitz (1974), and Glenn MacDonald (1984).

6. If we let $q = \sum_{i=1}^{n} q^i$, then $q > q^i > 0$ if $(q^1, \ldots, q^n)$ is a nontrivial division of q. By decreasing average cost, we have

$$C(q)/q < C(q^i)/q^i$$

so that

$$(q^i/q)C(q) < C(q^i).$$

Summing both sides, we get

$$\sum_{i=1}^{n} C(q^i) > \sum_{i=1}^{n} (q^i/q)C(q) = C(q).$$

7. Public inputs are examined by Harry de Gorter and David Zilberman (1990), Theodore Groves and Martin Loeb (1975), J. V. Henderson (1974), Arye Hillman (1978), Keimei Kaizuku (1965), Kohli Ulrich (1985), Richard Manning, James Markusen, and John McMillan (1985), Robert Martin, Thomas Zacharias, and Mark Lange (1990), McMillan (1979b), Herbert Mohring (1970), Takashi Negishi (1973), and Agnar Sandmo (1972, 1973).

8. See the recent treatment in Mark Stegeman (1991) for competitive industries.

Table 4.1
NATO Defense Burdens*
(in percentages by country for various years)

	1955	1960	1965	1970	1975	1980	1985	1987
U.S.	77.1	73.2	71.2	74.5	58.3	56.2	62.5	69.9
Canada	3.0	2.6	2.2	1.9	1.7	1.8	1.9	2.0
Belgium	.6	.6	.7	.7	1.4	1.5	1.1	.8
Denmark	.3	.3	.4	.4	.6	.6	.5	.4
France	5.3	6.3	6.3	5.8	9.1	10.3	8.6	7.4
West Germany	2.7	5.4	7.0	5.9	10.6	10.4	8.3	5.7
Greece	.2	.3	.3	.5	1.0	.9	.9	.6
Italy	1.4	2.1	2.5	2.4	3.2	3.7	3.4	3.5
Luxembourg	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
Netherlands	.7	.9	1.1	1.1	2.0	2.1	1.6	1.3
Norway	.3	.3	.4	.4	.6	.7	.6	.6
Portugal	.1	.2	.4	.4	.5	.3	.2	.2
Turkey	.3	.5	.6	.6	1.2	1.0	1.0	.7
United Kingdom	7.8	7.2	7.0	5.6	9.8	10.4	9.3	6.9

*Military expenditures divided by total NATO military expenditures multiplied by 100. Military expenditures are expressed in 1980 prices using GDP price deflators for each country. Converted to U.S. dollars using 1980 exchange rates. Spain was left out of the sample since it joined NATO after 1980; Iceland was left out since it has no military expenditure. (Columns may not sum to 100 due to rounding errors).

Source: Stockholm International Peace Research Institute. *World Armaments and Disarmament: SIPRI Yearbook* (various years).

Table 4.2
Military Expenditures as a Percentage of GDP by Country for Various Years

	United States	France	West Germany	United Kingdom	Belgium	Nether-lands	Denmark	Norway	Italy	Canada
1961	9.18	6.30	3.97	6.29	3.26	4.56	2.61	3.01	2.90	4.25
1963	8.84	5.64	5.21	6.16	3.42	4.47	3.04	3.59	3.10	3.66
1965	7.56	5.19	4.33	5.87	3.16	4.00	2.83	3.75	3.10	2.94
1967	9.51	5.11	3.61	5.67	3.13	3.95	2.66	3.51	2.91	2.90
1969	8.69	4.52	3.61	4.94	2.95	3.62	2.46	3.60	2.53	2.34
1971	6.85	4.00	3.39	4.87	2.83	3.27	2.44	3.39	2.70	2.22
1973	5.83	3.80	3.48	4.75	2.75	3.10	2.04	3.13	2.60	1.92
1975	5.75	3.85	3.66	4.86	3.02	3.29	2.44	3.39	2.48	1.86
1977	5.14	3.83	3.36	4.69	3.14	3.37	2.27	3.10	2.39	1.89
1979	5.12	3.88	3.26	4.41	3.26	3.20	2.32	3.09	2.39	1.73
1981	5.83	4.10	3.39	4.71	3.44	3.20	2.53	2.89	2.11	1.73
1983	6.50	4.12	3.37	5.16	3.24	3.19	2.45	3.08	2.27	2.11
1985	6.65	3.97	3.20	5.13	2.98	3.10	2.15	3.09	2.28	2.13
1987	6.45	3.92	3.06	4.70	2.92	3.03	2.10	3.19	2.22	2.04

Sources: Military Expenditures from Stockholm International Peace Research Institute, *World Armaments and Disarmament: SIPRI Yearbook* (various years).

GDP from *International Monetary Fund Financial Statistics Yearbook*

Table 4.3

Major Fund-Raising Charities in Great Britain: Total Income by Source 1983

Charity	Fund Raising & Donations	Legacies	Fees and Subscriptions	Rent and Investment	Trading	Grants from Statutory Authorities	Grants from other Charities	Other Income
	(Percent of Total Income)							
National Trust	--	17.2	--	33.5	26.7	3.8	6.1	6.3
Dr. Barnardo's	19.1	20.7	45.9	7.0	7.3	--	--	--
Salvation Army	14.7	21.9	39.7	21.1	--	0.4	2.0	0.2
Imperial Cancer Research Fund	14.4	60.0	--	24.1	0.7	--	--	0.3
Oxfam	32.1	6.8	--	1.4	39.1	10.3	9.3	1.0
Cancer Research Campaign	29.5	62.5	--	7.4	0.7	--	--	--
Royal National Lifeboat Inst.	29.8	46.9	6.4	4.7	4.6	--	--	7.6
Save the Children Fund	--	68.1	--	--	3.8	2.9	22.4	--
Percent of Aggregate Income of all Charities Listed	24.3	23.9	21.0	14.6	6.8	4.2	2.7	2.5

Source: Annual Accounts. Table compiled by John Posnett, Univerity of York, U.K.