Economics Bulletin

Volume 38, Issue 3

A note on stable cartels

Liang Mao
Shenzhen University

Abstract

Non-cooperative cartel formation games usually carry the assumption that cartel members will maximize their joint payoffs. Through an example, this note shows that this assumption is problematic because it imposes some unnecessary restrictions on cartel members' actions.

I would like to thank a reviewer for helpful comments.
Citation: Liang Mao, (2018) "A note on stable cartels", *Economics Bulletin*, Volume 38, Issue 3, pages 1338-1342
Contact: Liang Mao - maoliang@szu.edu.cn.
Submitted: February 14, 2018. **Published:** July 18, 2018.

1 Introduction

The non-cooperative coalition (cartel) formation model has been widely applied in many economic situations, such as collusion in oligopolistic markets (d'Aspremont et al., 1983; Diamantoudi, 2005), R&D joint ventures (Katz, 1986), customs unions (Yi, 1996), international environmental agreements (Carraro and Siniscalco, 1993; Barrett, 1994), and natural resources sharing (Miller and Nkuiya, 2016). In a typical application of this model, cooperation among players may potentially create a surplus. However, externalities and the lack of binding agreements may cause a free rider problem, which can hinder cooperation and lead to inefficient outcomes.

One possible method to overcome this problem is to form a cartel that regulates its members' actions. Players that voluntarily choose to be a member will form the cartel and sign a self-enforcing agreement. When payoffs are transferable, a common assumption about the agreement is that all cartel members should coordinate their actions to maximize their joint payoffs. We call this the MJP assumption, and call the resulting agreement the MJP agreement. This assumption is intuitive; otherwise, the cartel members are likely to renegotiate among themselves to replace the agreement with the MJP agreement so all may gain larger payoffs.

Nevertheless, one may still wonder whether the MJP assumption is indeed reasonable. For example, why does the MJP assumption require that members of all possible cartels, rather than only the stable ones, maximize their joint payoffs? After all, only stable cartels matter, since non-stable ones will not form. Therefore, it seems that the MJP assumption imposes too many restrictions on players' actions than are necessary.

To justify the MJP assumption, we should show that the MJP agreement can be optimal or that it is an equilibrium outcome when the agreement is endogenously determined. Many studies in the coalition formation literature discuss the endogenous determination of coalition agreements. In cooperative game settings where players can sign binding agreements, Shenoy (1979), Zhou (1994), Okada (1996), Ray and Vohra (1999), and Gomes (2005) highlight that the formation of a coalition and the allocation of payoffs in this coalition should be determined simultaneously and endogenously.[1] However, to the best of my knowledge, few studies in the literature show explicit evidence supporting or falsifying the MJP assumption in non-cooperative coalition formation models.

In this note, we present a simple open membership cartel formation model[2] where the MJP agreement will lead to a stable cartel in which all

[1]See Bloch (2003) for a review.

[2]Open membership means no player can be prevented from becoming a coalition mem-

members are willing to adopt a different agreement. This example shows that the MJP assumption is indeed problematic.

2 An example

Suppose a set $N = \{1, 2, \ldots, 5\}$ of homogeneous players may produce a public good. Let x_i denote player i's output. Player i's payoff is[3]

$$u_i = \sum_{k \in N} x_k - \frac{1}{300}\Big(\sum_{k \in N} x_k\Big)^2 - \frac{1}{2}x_i^2, \tag{1}$$

which depends on the total output of the good $\sum_{k \in N} x_k$ and i's individual cost $\frac{1}{2}x_i^2$.

Social welfare is the sum of all players' payoffs $\sum_{k \in N} u_k$, which is maximized when $x_i^* = 4.23$ for all i. However, each player's Nash equilibrium output is $x_i^0 = 0.97 < x_i^*$. The free rider problem causes this common social dilemma of insufficient public goods provision.

To overcome this problem, we can form a cartel to coordinate its members' actions. Consider a two-stage cartel formation game. In stage one, all players simultaneously decide whether or not to join the cartel. Those choosing to join become cartel members. In stage two, all members coordinate actions to maximize their joint payoffs (the MJP assumption), while simultaneously, non-members choose their own actions.

We can solve this game by backward induction. Suppose that the cartel formed in stage one is M, with cardinality $|M| = m$. In stage two, each non-member $j \notin M$ chooses output x_j to maximize u_j, while each member $i \in M$ chooses x_i to maximize $\sum_{k \in M} u_k$. In equilibrium, these lead to

$$x_j = \frac{150}{155 - m + m^2}, \quad j \notin M \tag{2}$$

$$x_i = \frac{150m}{155 - m + m^2}, \quad i \in M \tag{3}$$

if $|M| = m$.

We apply the stability concept introduced by d'Aspremont et al. (1983) to predict which cartel will form in stage one. When there are $|M| = m$ members, let $u^C(m)$ and $u^I(m)$ denote the payoffs for a member and a non-member, respectively. A cartel $M \notin \{\emptyset, N\}$ is said to be stable if $u^I(m) \geq$

---

ber.

[3]This payoff function is a special case of that used in Barrett (1994) and many others.

$u^C(m+1)$, and $u^C(m) \geq u^I(m-1)$. Further, N is stable if $u^C(n) \geq u^I(n-1)$, while $\emptyset$ is stable if $u^I(0) \geq u^C(1)$. A stable cartel is one in which no player has an incentive to unilaterally deviate from his or her participation decision.

From (1), (2), and (3), the condition for M to be stable is $m = 2$. Therefore, the payoff for a cartel member is $u^C(2) = 4.71$, the payoff for a non-member is $u^I(2) = 6.08$, and the social welfare is $2u^C(2) + 3u^I(2) = 27.67$.

In this example, the MJP assumption requires each member $i \in M$ to follow a specific agreement in stage two — to produce $x_i = \frac{150m}{155-m+m^2}$ if $|M| = m$. However, is this a reasonable assumption? Let us examine this cartel formation game with the following agreement: each member $i \in M$ should produce $x_i = 0.5m$ if $|M| = m$. As a reaction to this alternative agreement, each non-member's $j \notin M$ output is $x_j = \frac{300-m^2}{310-2m}$ if $|M| = m$. Thus, it is easy to verify that the condition for M to be stable is $m = 5$,[4] the payoff for each member is $u^C(5) = 8.85$, and the social welfare is $5u^C(5) = 44.27$.

This outcome shows that everyone (members, non-members, and the social planner who cares about social welfare) will agree to change the cartel agreement from $x_i = \frac{150m}{155-m+m^2}$ to $x_i = 0.5m$. The new agreement is better than the MJP agreement, regardless of the criterion used to evaluate it. Intuitively, this is because the new agreement provides a smaller incentive for players to free ride on other players' effort than the MJP agreement does[5]. Consequently, more players choose to join the cartel, which results in a more efficient outcome.

3 Discussion

We showed that the MJP assumption is problematic, despite seeming quite intuitive and being widely applied. At least in some situations, everyone has an incentive to replace the agreement based on the MJP assumption by one that provides a lower free-riding incentive. The problem with the former is that it imposes some unnecessary restrictions on members' actions. We should only care about players' payoffs in a stable cartel, rather than the payoffs in all possible cartels, as the MJP assumption requires.

Since agreements are not binding, some readers may wonder whether the non-MJP agreement $x_i = 0.5m$ is renegotiation-proof against the MJP

[4] We omit another stable cartel $M = \emptyset$, which is a trivial case where no player joins the cartel.

[5] Let $\Delta(m) = u^I(m-1) - u^C(m)$ characterizes the degree of a cartel member's free-riding incentive when $|M| = m$. The value of $\Delta(m)$ under agreement $x_i = \frac{150m}{155-m+m^2}$ is larger than that under agreement $x_i = 0.5m$ for all $m \geq 1$.

agreement $x_i = \frac{150m}{155-m+m^2}$. That is, once the cartel forms and all players receive their payoffs under $x_i = 0.5m$, will the members have incentives to switch this agreement to the MJP agreement? In fact, none of the members will choose to do so because their payoffs will either decrease from 8.85 to 4.71 (as a member), or decrease from 8.85 to 6.08 (as a non-member) otherwise.

A question is whether the new agreement ($x_i = 0.5m$) is "optimal". The point is that we need an explicit criterion to establish whether or not an agreement is "optimal". For future studies, a lesson we can learn from this note is that a cartel agreement should be endogenously determined, rather than exogenously given. In a non-cooperative coalition formation setting, some studies (Carraro et al., 2009; Köke and Lange, 2017; Mao, 2017) have already discussed endogenous agreements in some specific applications, but more work is needed in more general situations.

References

Barrett, S. (1994), "Self-enforcing International Environmental Agreements." *Oxford Economic Papers* **46**, 878–894.

Bloch, F. (2003), "Non-cooperative Models of Coalition Formation in Games with Spillovers" in *The Endogenous Formation of Economic Coalitions* by C. Carraro, ed., Edward Elgar, 35–79.

Carraro, C., C. Marchiori, and S. Oreffice (2009), "Endogenous Minimum Participation in International Environmental Treaties." *Environmental and Resource Economics* **42**, 411–425.

Carraro, C. and D. Siniscalco (1993), "Strategies for the International Protection of the Environment." *Journal of Public Economics* **52**, 309–328.

d'Aspremont, C., A. Jacquemin, J. J. Gabszewicz, and J. Weymark (1983), "On the Stability of Collusive Price Leadership." *Canadian Journal of Economics* **16**, 17–25.

Diamantoudi, E. (2005), "Stable Cartels Revisited." *Economic Theory* **26**, 907–921.

Gomes, A. (2005), "Multilateral Contracting with Externalities." *Econometrica* **73**, 1329–1350.

Katz, M. L. (1986), "An Analysis of Cooperative Research and Development." *The RAND Journal of Economics* **17**, 527–543.

Köke, S. and A. Lange (2017), "Negotiating Environmental Agreements under Ratification Constraints." *Journal of Environmental Economics and Management* **83**, 90–106.

Mao, L. (2017), "Designing International Environmental Agreements under Participation Uncertainty." Working paper.

Miller, S. and B. Nkuiya (2016), "Coalition Formation in Fisheries with Potential Regime Shift." *Journal of Environmental Economics and Management* **79**, 189–207.

Okada, A. (1996), "A Noncooperative Coalitional Bargaining Game with Random Proposers." *Games and Economic Behavior* **16**, 97–108.

Ray, D. and R. Vohra (1999), "A Theory of Endogenous Coalition Structures." *Games and Economic Behavior* **26**, 286–336.

Shenoy, Prakash P (1979), "On Coalition Formation: A Game-theoretical Approach." *International Journal of Game Theory* **8**, 133–164.

Yi, S.-S. (1996), "Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism is Good." *Journal of International Economics* **41**, 153–177.

Zhou, L. (1994), "A New Bargaining Set of an n-person Game and Endogenous Coalition Formation." *Games and Economic Behavior* **6**, 512–526.