Contestable Market And Cost Structure

Masahiro Koreeda

I. Introduction

Contestable market and Theory of Industry Structure by W. Baumol, J. Panzer, R. Willig provides a new approach for the multiproduct cost structure, competition, and market performance. Their main aim is to offer a product more directly useful to the policymaker than the more abstract mathematical writings.

There are at least three basic ideas presented here. First the crucial feature of a contestable market is its vulnerability to hit-and-run entry. Even a very transient profit opportunity need not be neglected by a potential entrant. They exclude the whole sunk cost of fixed capital. This means that all capital is salable or reusable, setting aside existence of its well-organized secondhand market. If all capital is so without loss other than that corresponding to normal user cost and depreciation, then any risk of entry is eliminated. Thus, contestable market may share an attribute with perfect competition.

Second, they introduce useful cost concepts for the multiproduct firm. These include economies of scope, product-specific return to scale, multiproduct scale economies, ray scale economies, and subadditive cost functions.

Third, they consider not only the notion of natural monopoly in single product case but also in multiproduct case, and then represent that
multiproduct monopoly can be financially viable if its cost function satisfies certain conditions. Traditional theory tells us that monopoly may result because of (1) the control by single firm of essential input, (2) patents on trademark, or (3) the exclusive right to sell. Monopoly firm caused by these reasons will obtain profits more than normal rate of profit. But a contestable market never offers more than a normal rate of profit to monopoly firm. This enables monopoly firm to have welfare properties.

The purpose of this paper is to review the theory of the contestable market and investigate strictly some important results provided by Boumol et al.

II. Natural Monopoly

We begin by analysing single product cost structure and then define natural monopoly.

All firms have the same cost function, $C(\cdot)$. Let $y^i$ be the output of firm $i$, and let $m$ be the number of firms. $Q(\cdot)$ is industry demand function, and $p$ is the price of good.

**Definition 1**: Strict Subadditivity A cost function $C(y)$ is strictly subadditive at $y$ if for any positive quantities of outputs $y^1, \cdots, y^m, y^j \neq y, j = 1, \cdots, m$, such that

$$\sum_{j=1}^{m} y^j = y,$$

we have

$$C(y) < \sum_{j=1}^{m} C(y^j).$$

**Proposition 1**: Strict subadditivity implies increasing return to scale.

**Proof.** Let $C$ be homogeneous function of degree $k$. Then $C(\lambda y) = \lambda^k C(y)$ for
\( \lambda > 0 \). If average cost is decreasing with \( y \), we have, using Euler’s theorem,

\[
\frac{d}{dy} \frac{C(y)}{y} = \frac{yC'(y) - C(y)}{y^2} < 0 \iff yC'(y) < C(y) \iff k < 1
\]

We can see that return to scale is increasing if and only if \( k < 1 \). If for \( \lambda > 1 \) \( C(\lambda y) < \lambda C(y) \), it is obvious that \( C(\lambda y)/\lambda < C(y) \) for \( \lambda > 1 \). This means that average cost is increasing with \( y \). Hence this completes the proof.

Then we have the following.

**Proposition 2.** Let \( C \) be a function from \([0, \infty)\) to \([0, \infty)\). Then \( C \) is strictly subadditive if \( C \) is strictly concave 2) and \( C(0) = 0 \).

**Proof.** Since \( C \) is strictly concave, for \( x = 0 \), any \( y \in (0, \infty) \), \( \lambda > 0 \), \( \mu > 0 \), \( \lambda + \mu = 1 \),

\[
C(\lambda x + \mu y) = C(\mu y) > \mu C(y).
\]

Thus, for any \( \alpha \in (0, 1) \), any \( y \in (0, \infty) \),

\[
C(\alpha y) > \alpha C(y).
\]

If \( y = y_1 + y_2 \), \( y_1 = \alpha y \), \( y_2 = (1 - \alpha) y \), for \( \alpha \in (0, 1) \), then \( C(y_2) = C((1 - \alpha) y) > (1 - \alpha) C(y) \). This implies that \( C(y) - C(y_2) < \alpha C(y) \).

Thus \( C(y) < C(y_1) + C(y_2) \) because \( C(y_1) > C(\alpha y) > C(y) - C(y_2) \).

This proves the strict subadditivity of \( C \).

Subadditivity can then be taken as the obvious criterion of natural monopoly. That is,

**Definition 2 : Natural Monopoly** An industry is said to be a natural monopoly if over the entire relevant range of outputs, the firms’ cost func-

1) Let \( C \) be homogenous function of degree \( k \), then for \( y = (y_1, \ldots, y_n) \)

\[
k C(y) = \sum_{j=1}^{n} y_j \frac{\partial C}{\partial y_j}.
\]

2) A concave function \( C \) on \( \mathbb{R}^n \) is said to be strictly concave, if \( \lambda C(x) + \mu C(y) < C(\lambda x + \mu y) \) for \( x, y \in \mathbb{R}^n \), \( x \neq y \), \( \lambda > 0 \), \( \mu > 0 \), \( \lambda + \mu = 1 \).
tion is subadditive.

III. Sustainable Industry Configuration

Next, we investigate a notion of sustainable industry configuration proposed by Boumol et al.

**Definition 3: Feasible Industry Configuration** A feasible industry configuration is composed of $m$ firms respectively producing the non-negative output quantities $y^1, \ldots, y^m$ for sale at a price $p$ such that

$$\sum_{i=1}^{m} y^i = Q(p) \text{ and } py^i - C(y^i) \geq 0 \text{ for } i = 1, \ldots, m.$$

That is to say, a configuration is feasible if production is sufficient to meet demand and no firm losing money.

**Definition 4: Sustainable Industry Configuration (Sustainability)** A feasible industry configuration with price $p$ and output $y^1, \ldots, y^m$ is sustainable if $p^e y^e \leq C(y^e)$ for all $p^e \leq p$ and $y^e \leq Q(p^e)$, where $p$ denotes incumbent's price, $p^e$ denotes entrant's price.

This means that no outside potential competitor can enter by cutting prices and make money supplying quantities that do not exceed total market demands.

We can find a price satisfying sustainability under particular restrictions on cost function. For example, cost function is as follow.

$$C(y) = \begin{cases} 
\frac{1}{2}y^2 + \frac{1}{2} & (0 \leq y < 2) \\
\frac{5}{8}y + \frac{5}{4} & (2 \leq y < 3) \\
\frac{1}{8}y^2 + 2 & (3 \leq y)
\end{cases}$$
If there are four firms, we have an only sustainable price, \( p = 1 \). For output 1 of each firm leads sustainable industry configuration (see Figur 1).

Sustainability plays the crucial role in defining equilibrium in contestable markets. Notwithstanding this fact, it is noted that there are not necessarily industry sustainable configuration. We can illustrate this as follows. Let reverse demand function be \( p = - (1/\alpha) y + b \) where \( \alpha \) and \( b \) both are positive and constant \((p < b)\), and let cost function be \( C(y) = (1/\alpha) y + c \) where \( c \) is non-negative and constant. Take a \( \epsilon > 0 \) satisfying \( p^* = p - \epsilon > 0 \). Then

\[
\begin{align*}
    y^* &= -\alpha p^* + ab = -\alpha (-p + \epsilon + b) > 0 \\
    p^* y^* &= a (p - \epsilon) (-p + \epsilon + b) \\
    C(y^*) &= (-p + \epsilon + b).
\end{align*}
\]

Thus

\[
    p^* y^* - C(y^*) = (-p + \epsilon + b) \{a (p - \epsilon) - 1\}
\]

If \( \alpha \) is a large number such that \( a (p - \epsilon) - 1 \geq 0 \), then \( p^* y^* \geq C(y^*) \). Hence there never exist \( p \) satisfying sustainability. Therefore, in general, if the elasticity of reverse demand function is extremly large and cost function is
slightly increasing, definition 3 is no significant.

**Proposition 3.** Let demand function $Q(p)$ and cost function $C(y)$ be continuous\(^3\). Then in a sustainable industry configuration involving two or more producing firms, all firms must produce output at which $py_i = C(y_i)$, $i = 1, \ldots, m \geq 2$.

**Proof.** Let $p$ be a sustainable price. If $py_i > C(y_i)$ for $i \in \{1, \ldots, m\}$, then by feasibility $py > C(y)$, namely $pQ(p) - C(Q(p)) > 0$. Since $Q$ and $C$ are both continuous, $F(p) = pQ(p) - C(Q(p))$ is continuous. Hence for enough small $\varepsilon > 0$, $F(p - \varepsilon) > 0$. This contradicts sustainability. \(\blacksquare\)

---

**N. Multiproduct Cost Structure**

In this section we introduce and discuss the cost concepts in the multiproduct case.

**Definition 5:** Economy of Scope

Let $P = \{T_1, \ldots, T_k\}$ denote a nontrivial partition of $S \subset N$. $\cup_i T_i = S$, $T_i \cap T_j = \emptyset$ for $i \neq j$, $T_i \neq \emptyset$ and $k > 1$. There are economies of scope at $y_S$ with respect to the partition $P$ if

$$\sum_{i=1}^{k} C(y_{T_i}) > C(y_S),$$

where $(y_{T_1}^1, \ldots, y_{T_k}^n)$ is a $n-$vector for which

$$y_{T_i}^j = \begin{cases} y_{T_i}^j > 0 & j \in T_i \\ 0 & j \in S - T_i \end{cases}$$

The term on the left side of (1) is the total cost of producing the bundle of goods $y_S$ in $k$ distinct firms, while the term on the right is the cost of producing $y_S$ with a single firm. By definition subadditivity implies economies of

---

\(^3\) Baumol. et al.'s proposition has not this restriction. They give a descriptive proof to it but not mathematical proof. It is obvious that their proposition do not necessarily hold without conditions on cost function.
Contestable Market And Cost Structure

scope, so that formally economies of scope can be interpreted as a restricted form of subadditivity.

**DEFINITION 6:** *Ray Average Cost* Let \( y^0 = (y^0_1, \ldots, y^0_n) \) be the unit bundle for a particular mixture of output and let \( t \) be the number of unit in the bundle \( y = ty^0 \), such that \( \sum_{i=1}^{n} y^0_i = 1 \). Then the ray average cost (RAC) is defined as

\[
RAC(y) = \frac{C(ty^0)}{t}.
\]

One can regard RAC as the average of the composite good and measure the its absolute quantity. Figur 2 describes RAC.

**DEFINITION 7:** *Degree of Scale Economies.* The degree of scale economies defined over the entire product set \( N = \{1, \ldots, n\} \), at \( y = (y_1, \ldots, y_n) \), is given by

\[
S_N(y) = \frac{C(y)}{y} \cdot \nabla C(y) = \frac{C(y)}{\sum_{i=1}^{n} y_i C_i(y)}
\]

where \( \nabla C(y) = (\frac{\partial C}{\partial y_1}, \ldots, \frac{\partial C}{\partial y_n}) \), \( C_i(y) = \frac{\partial C(y)}{\partial y_i} \).

**PROPOSITION 4.** Returns to scale at the output point \( y \) are increasing, decreasing, or locally constant if and only if \( S_N > 1 \), \( S_N < 1 \), or \( S_N = 1 \) respectively.
This is a general version of proposition 1. We can easily prove this proposition by using Euler’s theorem.

**Proof.** If $C$ is a homogeneous function of degree $k$, then $C(ty) = t^k C(y)$ where $t$ is positive and constant. This yields $\sum_{j=1}^{n} y_j \frac{\partial C}{\partial y_j} = kC(y)$. Thus, since $\frac{\partial C}{\partial y_j} > 0$ for $j = 1, \ldots, n$, we have

\[
S_N = 1 \iff C(y) > \sum_{j=1}^{n} y_j \frac{\partial C}{\partial y_j} = 1 \iff C(y) > \sum_{j=1}^{n} y_j \frac{\partial C}{\partial y_j} \iff k = 1.
\]

\[
S_N > 1 \iff C(y) > \sum_{j=1}^{n} y_j \frac{\partial C}{\partial y_j} > 1 \iff C(y) > \sum_{j=1}^{n} y_j \frac{\partial C}{\partial y_j} \iff k < 1.
\]

\[
S_N < 1 \iff C(y) < \sum_{j=1}^{n} y_j \frac{\partial C}{\partial y_j} < 1 \iff C(y) < \sum_{j=1}^{n} y_j \frac{\partial C}{\partial y_j} \iff k > 1.
\]

**Definition 9:** The degree of scale economies specific to the product set $T \subset N$ at $y$ is given by $S_T(y) = I C_T(y) = \sum_{j=1}^{n} y_j C_j(y)$, where $I C_T(y) = C(y) - C(y_{N-T})$.

**Proposition 5.** $C$ is increasing return to scale with respect to $T$ at $y$ if and only if $S_T > 1$, decreasing if and only if $S_T < 1$, and constant if and only if $S_T = 1$.

**Proof.** Since $C$ is a homogeneous function of degree $k$, for $y_T = (y_1, \ldots, y_t, 0, \ldots, 0)$

\[
C(t y_T) = t^k C(y_T).
\]

We have, by Euler’s theorem,

\[
\sum_{j=1}^{n} y_j \frac{\partial C}{\partial y_j} = \sum_{j=1}^{t} y_j \frac{\partial C}{\partial y_j} = kC(y_T),
\]

If $S_T(y) > 1$, then

\[
C(y) - C(y_{N-T}) > \sum_{j=1}^{t} y_j \frac{\partial C}{\partial y_j}.
\]

Thus we have, by (2)

\[
C(y) - C(y_{N-T}) = C((y_1, \ldots, y_t, y_{t+1}, \ldots, y_n)) - C((0, \ldots, 0, y_{t+1}, \ldots, y_n)) > \sum_{j=1}^{t} y_j \frac{\partial C}{\partial y_j} = kC(y_T) = kC((y_1, \ldots, y_t, 0, \ldots, 0)).
\]
Hence

\[ C(\langle y_1, \ldots, y_t, y_{t+1}, \ldots, y_n \rangle) > C(\langle 0, \ldots, 0, y_{t+1}, \ldots, y_n \rangle) \]

\[ + k C(\langle y_1, \ldots, y_t, 0, \ldots, 0 \rangle). \]  

(5)

It is necessary for (4) that \( k < 1 \), so that \( C \) is increasing return to scale by (2).

The proof of cases where \( S < 1 \) and \( S = 1 \) are perfectly analogous.

**DEFINITION 10: Trans-Ray Convexity** A cost function \( C(y) \) is trans-ray convex through some point \( y^* = (y_1^*, \ldots, y_n^*) \) if there exists any vector of positive constraint \( w_1, \ldots, w_n \) such that for every two output vectors \( y^a = (y_1^a, \ldots, y_n^a) \) and \( y^b = (y_1^b, \ldots, y_n^b) \) that lie on the hyperplane \( \sum_{i=1}^{n} w_i y_i = w_0 \) through point \( y^* \) \(^4\) (so that they satisfy \( \sum_{i=1}^{n} w_i y_i^a = \sum_{i=1}^{n} w_i y_i^b = \sum_{i=1}^{n} w_i y_i^* \) for \( \alpha \) such that \( 0 < \alpha < 1 \) we have

\[ C[\alpha y^a + (1 - \alpha) y^b] \leq C(y^a) + (1 - \alpha) C(y^b). \]

---

4) Let \( P \) be hyperplane passing through \( y^* \) vertical to \( w \), then

\[ P = \{ X \mid (X - y^*) \cdot w = 0 \} \]

Moreover if \( P' \) is hyperplane passing through \( y^a \), we have

\[ P' = \{ X \mid (X - y^a) \cdot w = 0 \} \]

Thus \( Xw = y^a w = \sum_{i=1}^{n} y_i^a w_i \). Hence \( P = P' \) because \( y^a \) passes through \( P' \).
Trans-ray convexity requires that the production cost of a weighted average of a pair of output bundles, $y^a$ and $y^b$, is not greater than the weighted average of the cost of producing each of them in isolation. A graph of this concept can be drawn as Figure 3.

Baumol et al. suggest that trans-ray convexity and economies of scale at every point implies that there will be a monopoly. To provide an intuitive explanation, assume that the cost function has the property ray economies of scale at every point in conjunction with trans-ray convexity. In Figure 4, curve $OBy$ characterizes the ray economies of scale and curve $ABC$ representing trans-ray convexity has an only bottom point $B$. In this case there is a strong possibility of existence of monopolist. Let $p = (p_1, p_2)$ be fixed prices, and let $R$ be total revenue, then $R(y) = p_1 y_1 + p_2 y_2$. If there exists a monopolist, he can exclude entrant by setting prices to $p$ such that revenue hyperplane passes through $O$ and is tangent to curve $ABC$ at $B$ in Figure 4. For this prices satisfies sustainability.
V. Economic Implications

We have described the essentials of the theory of the perfect contestable market and given some propositions with reference to it. In this section, we examine economic implications of the theory.

First, we can use a new analytic methods of the theory as a standard public policy. There are various artificial barriers to entry, for instance, cartel formation, legal institution, or government intervention, in today's market. Almost all of companies defended by those barriers may obtain profits more than normal profits. They may charge high prices so much to consumers. The theory shows that if those barriers is removed, then prices fall and their excessive profits come to equal to normal profits when fixed capital is not sunk. This means that all consumers may get a benefit by competition.

Second, the analytic methods of the theory is useful for understanding of the meaning of local production, or decentralization of industries into the consuming place. Recently in our country, National Railway, the telecommunication company have divided and privatized, It seems that the divisions of those big enterprise is a result reverse to what the theory intends. But we can construe the implication of division of company with the theory. We show this by Figure 4. Suppose that firm A produces commodity $Q$ in place $X$, and that its cost curve is presented by $A_{\text{ya}}O$ in Figure 4. Then assume that its cost curve of commodity $Q$ in place $Y$ is given by $C_{\text{yb}}O$ in Figure 4. We replace $Q$ in place $X$ with $Q_x$ and $Q$ in place $Y$ with $Q_y$, interpreting that commodity $Q_x$ is different from commodity $Q_y$. Here we investigate what set of $X_A$ and $X_B$ are most economical. If $y_1^q = y_2^b$, for $y^q = (y_1^q, 0)$, $y^b = (0, y_2^b)$, we see that $\bar{y} = (y_1^q, y_2^b)$ is a set of $X_A$ and $X_B$ that may the make production cost of quantify $y_1^q$ smallest. Thus it is most costless for firm A to produce $y_1^q$ in place $X$ and $y_2^b$ in place $Y$. This may yield a benefit of division for firm A.
References


