Universidad de la República
Facultad de Ciencias Económicas y Administración
Microeconomía Avanzada
Notas Docentes

Monopolio Natural y Regulación Económica

1. INTRODUCCION

1.1 Teoría de la Regulación.

Regulación: intervención del Estado que limita o reglamenta la libre acción de las empresas.

Empíricamente, se constata usualmente intervención en varios sectores, en los que la acción del Estados afecta los equilibrios a los que se hubiese llegado libremente:

- Servicios públicos
- Salud
- Transporte
- Sector financiero
- Calidad en sectores varios

La cuestión es analizar en que circunstancias esta intervención está justificada desde el punto de vista teórico.

1.2 Mercados

El curso venía analizando una serie de estructuras de mercado, o formas en que oferentes y demandantes interactúan en los mercados. Así se analizó la estructura de mercado competitiva, después la estructura de mercado monopólica. Un caso particular de estructura monopólica lo constituye el monopolio natural. La existencia de monopolio natural es una de las justificaciones de la existencia de regulación en la economía.

Aunque la teoría de la regulación abarca otros aspectos que la regulación de los monopolios naturales, una parte importantante está dedicada a este tópico.

1.3 Servicios Publicos (utilities)

Los sectores de servicios públicos se caracterizan por:

- Son de consumo masivo por parte de la población
- Requerir importantes costos de inversión (costos fijos)
- La inversión es específica

La primer característica los hace ser políticamente sensibles.
Las otras características los hace sensibles a comportamientos oportunistas.

Ejemplos: telecomunicaciones, electricidad, agua,

1.4 Criterios

Para analizar en que situaciones está justificada la intervención pública, recurrimos a las recomendaciones normativas de la economía. Desde el punto de vista normativo, el criterio que se aplica para determinar la conveniencia de una situación es el de la eficiencia económica. Eficiencia se mide con el criterio de Pareto.

En un merco de equilibrio parcial, ya se analizó como un mercado competitivo lleva a una asignación de recuros eficiente en el sentido de Pareto mientras que un monopolio no es eficiente según el mismo criterio. El análisis de equilibrio parcial no considera sin embargo los efectos que cambios en un mercado producen en los demás, ya sea por la vía de la complementariedad o sustituibilidad entre bienes como por el efecto sobre la renta de los consumidores. En un marco de equilibrio general se cosideran estas interrelaciones y por lo tanto las conclusiones son más generales.

La piedra angular de las recomendaciones de la economía en cuanto a la intervención en el funcionamiento del mercado se encuentra en los teoremas del bienestar. Los mismos se deducen en un marco de equilibrio general.

Primer teorema: La competencia (solución descentralizada) me lleva a la eficiencia (con muy pocos requisitos de información).

Segundo teorema: Si también me preocupa la distribución, no necesito renunciar al mercado como mecanismos para llegar a la eficiencia, siempre que pueda tener otros mecanismos como transferencias de ingreso de suma fija.

La intervención pública se recomienda cuando:

- Hay fallas de mercado
- Hay poder de mercado

La intervención cuando hay poder de mercado tiene dos aspectos

- Políticas de competencia
- Regulación en sectores específicos de servicios públicos, y en especial monopólios naturales

En las clases que dedicaremos al tema nos centraremos en la conceptualización del monopolio natural y en los aspectos claves de la regulación de los mismos. Sobre el final se discutirán algunos aspectos de la reforma regulatória que ha producido en los servicios públicos en los últimos 25 años en el mundo occidental, así como los aspectos centrales de la dicha reforma en Uruguay.

2. CARACTERIZACIÓN DEL MONOPOLIO NATURAL

2.1 Monopolio Natural

Existe un monopolio natural cuando los costos de producción son tales que para los demandantes del mercado es es más barato obtener la producción de una empresas que de muchas. En esta situación es óptimo desde el punto de vista de costos, de que exista una empresa que muchas.

Definición (MN = Subaditividad)

Se dice que en la producción de un bien existe monopolio natural cuando la función de costos exhibe subaditividad para las cantidades demandadas, es decir cuando una sola firma es capaz de producir la cantidad que se demanda del bien en cuestión, a un costo menor o igual al que tendrían a dos o más firmas.[1]

[1] Esta definición vale para el caso que la firma venda un único producto o varios productos.

Definición subaditividad de costos
Si llamamos C(q) a la función de costos con la tecnología disponible, se dice que C(q) es subaditiva en $Q \subset R$, si para todo $q \in Q$ y todo N, si tomamos q_1, q_2, ...q_N cualesquiera tales que

$$\sum_i q_i = q \;,$$

entonces: $$\sum_i C(q_i) > C(\sum_i q_i)$$

Un monopolio natural surge de dos fuentes: economías de escala y economías de alcance. Separamos la discusión para el caso de que la firma produzca un solo producto o varios.

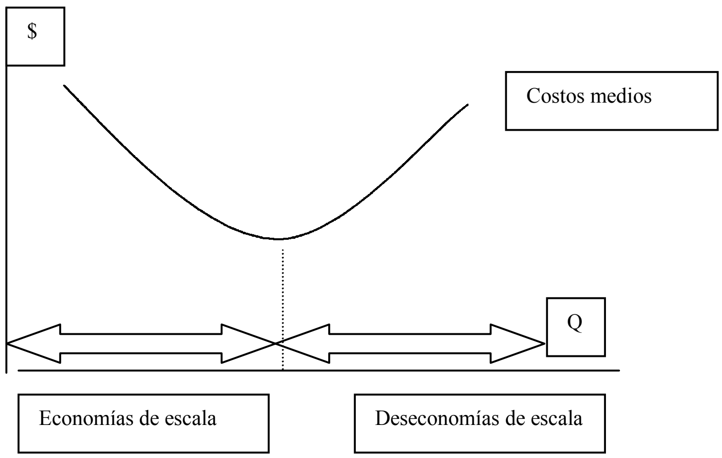
2.2 Caso de empresas monoproducto

En el caso de que la empresa produzca un solo producto el monopolio natural proviene básicamente de la existencia de economías de escala. Economías de escala existen cuando los cosos medios de producción decrece a medida que el producto aumenta (son una condición suficiente para la existencia de monopolio natural).

Definición (Economías de escala = costo medio decreciente)
La función de costos C(q) presenta economías de escala en un conjunto $Q \subset R$ si y sólo si, el costo medio es decreciente en todo punto de Q , es decir si para todo q_1, $q_2 \in$ Q, tal que $q_2 > q_1$, entonces se cumple que:

$$C(q_1)/q_1 > C(q_2)/q_2$$

Economías de escala existen cuando los cosos medios de producción decrecen a medida que el producto aumenta. Una función de costos puede (suele) tener tramos en dónde existen economías de escala y tramos dónde existen deseconomías de escala.



La fuente principal de economías de escala son los costos fijos, esto es, los costos en que debe incurrirse, no importa la cantidad de producto producida. Ejemplo típico son las plantas de generación eléctrica. A medida que aumenta la producción, los costos fijos por unidad de producto decrecen y los costos medios totales pueden decrecer.

Las economías de escala pueden existir para niveles bajos de producto y no existir a niveles mayores de producción; esto da como resultado la forma de U de curva de costos medios que se muestra en el gráfico anterior.

Teorema (Economías de escala => subaditividad de costos)
Si C(q) tiene costos medios decrecientes (tiene economías de escala) para $q \leq q_M$, entonces es subaditiva para todo $q \leq q_M$.

Sea la cantidad a producir q, tal que $q \leq q_M$ y un conjunto $\{q_i\}$,i=1,..n, tales que

$$\sum_i q_i = q$$

Por ser la función de costo medio decreciente para cantidades menores que q_M:

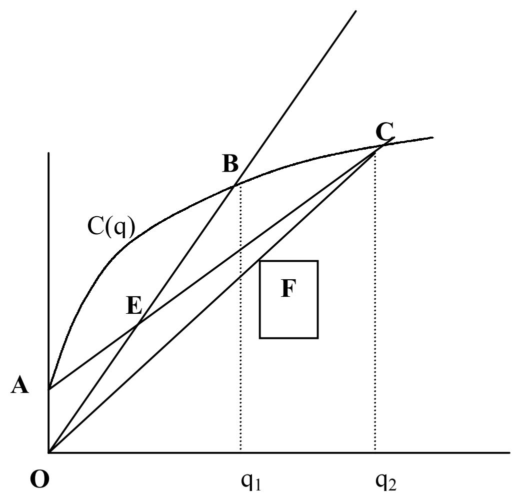
$$C(q_i)/q_i > C(q)/q \quad \text{para todo } i$$

$$C(q_i) > C(q).\, q_i /q$$

$$\sum_I C(q_i) > \sum_i (C(q).\, q_i /q) = (C(q).\, /q\,).\sum_i q_i = C(q)$$

En resumen $\Sigma\, C(q_i) > C(q)$, para todo $q \leq q_M$ y todo conjunto $\{q_i\}$, tal que $\Sigma q_i = q$, es decir hay subaditividad.

Función de costos cóncava (técnico)
Teorema: Si la función de costos de producción de un solo bien es cóncava, entonces es subaditiva para cualquier cantidad a producir (Se muestra gráficamente).



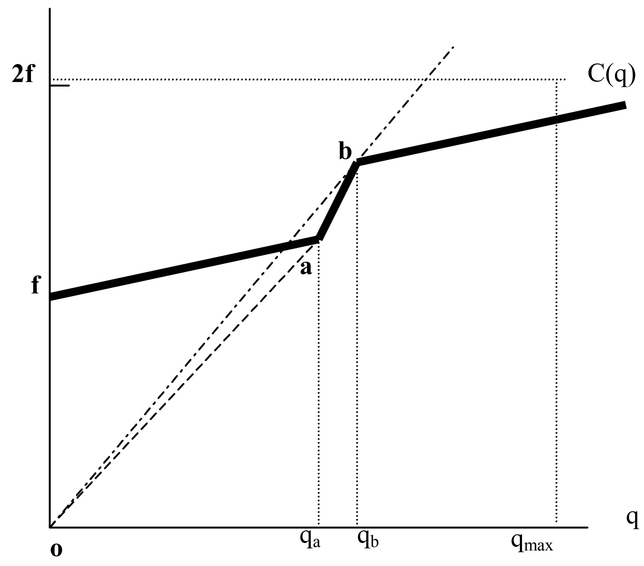
A son los costos fijos.
q_1 y q_2 son dos cantidades cualquiera.

Si se produce q_1 los costos totales están dados por B. Si se produce q_2 los costos totales están dados por C. Los costos medios por producirse q_1 son la pendiente de la recta que parte del origen y pasa por B (la altura son los costos totales y la base la cantidad producida). Análogamente para la cantidad q_2. Esta última pendiente la puedo calcular también como la altura del punto F dividido q_1. Esto muestra que para la cantidad mayor q_2 los costos medios son menores. (De forma más simple, si el rayo que parte del orígen y pasa por la curva de costo total tiene mayor pendiente, entonces los costos medios son mayores). Esto muestra que si la curva de costos es cóncava, entonces hay economías de escala.

Entonces una función de costos cóncava es de costos medios decrecientes y por lo tanto subaditiva.

Se muestra con un ejemplo que economías de escala no son necesarias para que exista subaditividad de costos
La función de costos puede ser subaditiva en un conjunto $Q \subset R$, sin que los costos medios sean decrecientes en dicho conjunto.

Veamos el ejemplo gráfico siguiente:



En la curva de costos totales mostrada, que se caracteriza por un fuerte costo fijo, para producir cualquier rango entre q_a y q_b, es intuitivo que una sola empresa lo hará a costo menor que dos o más. Si dos empresas debiesen producir la misma cantidad, deberían incurrir entre ambas en el costo fijo **2f.** Con el dibujo que hicimos el costo de producción de las dos empresas será superior al de una sola, para cualquier producto entre q_a y q_b. Es decir que entre 0 y qmáx existe monopolio natural.

Sin embargo, el costo medio, dado por la pendiente de los rayos oa y ob , es creciente en el tramo de producción q_a, q_b.

La existencia de monopolio natural depende del rango de economías de escala en relación al tamaño de la demanda del mercado. En particular, un monopolio natural existe en la producción de un bien solo si las economías de escala existen en un rango de producto suficientemente grande en relación a la demanda.

Figura I.3. Train.

(a) Monopolio Natural. Los costos medios son decreciente para todo nivel de producto que puedan ser demandados a cualquier precio, esto es, parta todo el tramo a la derecha de la curva de demanda.
(b) Monopolio Natural incluso cuando las economías de escala existen en un rango menor respecto a la demanda. Si las empresas se dividen el mercado de forma igual, producirán a costos mayores que una sola firma. Lo mismo es imaginable ocurra para cualquier otra división del mercado. Dicho de otra manera, la demanda corta en un tramo dónde no hay economías de escala, pero sigue habiendo subaditividad de costos.
(c) Duopolio natural. Las economías de escala aparecen en un rango menor respecto de la demanda. Dos empresas produciendo Q_2 cada una tienen un costos AC_2, mientras que una empresa sola produce $2Q_2$ a un costo AC_1 mayor que que AC_2, por lo que no hay monopolio natural.
(d) Las economías de escala son muy chicas en relación a la demanda, por lo que el mínimo costo se alcanza con múltiples empresas.

Atención: para la producción conjunta de más de dos bienes, la concavidad no es condición suficiente de subaditividad. Imaginemos una función de costos cóncava para producir las cantidades (x,y) que valga cero si x=0 o y=0 y sea positiva en todo otro punto del primer cuadrante. $C(x,0) = C(0,y) = 0 < C(x,y)$, no es subaditiva.

Hasta este punto no hemos especificado la naturaleza de la función de costos C empleada para la definición de monopolio natural. Si se desea aplicar esta definición de monopolio natural para decidir la pertinencia de regular un sector, debe considerarse la **función de costos de largo plazo**, ya que es la que tiene en cuenta la posibilidad de cambios en la cantidad de capital fijo y la que refleja la posibilidad de entrada y salida de empresas al sector.

2.3 Caso de empresa multiporducto

Hasta ahora hemos considerado solamente el caso de un monopolio natural que produce un único producto.

No obstante, la mayor parte de los monopolios regulados son empresas multiproducto Ejemplos:

- A veces producen bienes o servicios que en su naturaleza son diferentes: telefonía básica urbana, llamadas interurbanas, teléfonos celulares, enlaces digitales para trasmisión de datos, etc.
- A veces ofrecen un bien o servicio homogéneo pero suministrado en distintos puntos geográficos y en distintos momentos: llamadas telefónicas entre el punto a y el b, a la hora tal del día; carga por tren del punto a al b en tal horario.

Muchas veces entre los productos vendidos por una empresa regulada, existen algunos vendidos en situación de monopolio y otros en los que existe competencia. Este es uno de los problemas más interesantes de la regulación, que no trataremos en detalle.

Igual que en el caso monoporducto, decimos que existe monopolio natural en un sector cuando existe subaditividad de las función de costos dada la tecnología de ese sector.

La definición de subaditividad para un sector que produce n bienes, es análoga a la que vimos para el caso de función de producción de un solo producto.

Definición: Subaditividad de la función de costos de producción de n bienes.

Sea $C:R^n \rightarrow R$ la función de costos en un sector, que obviamente depende de las cantidades de los n bienes que se produzcan.

Vamos a llamar con Q mayúscula a vectores de cantidades producidas de los n bienes.

Se dice que la función de costos es **subaditiva** si para cualquier entero m > 0, y para cualquier conjunto $Q^1, Q^2 ......Q^m$, se cumple:

$$C(Q^1) + C(Q^2)......+C(Q^m) > C(Q^1 + Q^2 .....+ Q^m)$$

Es decir que puede siempre realizarse la producción del vector de n bienes
$Q = Q^1 + Q^2 .....+ Q^m$, a un costo menor en una sola empresa que en más de una.
Atención: denotamos con superíndices y Q mayúsculas a los distintos vectores de producto y con subíndice y q minúscula a las distintas componentes escalares o cantidades de cada uno de los bienes, así será $Q^i=(q^i_1,{}^iq_2,......q^i_n)$.

Para la producción de un solo bien, habíamos encontrado una relación inmediata:
Costos medios decrecientes (economías de escala) $\Rightarrow$ subaditividad , si bien el recíproco no era cierto.

En el caso de que la empresa produzca varios productos, el monopolio natural proviene de la existencia de economías de escala o de alcance.

Definición:
Decimos que el vector de insumos y productos $(X,Q)= (x_1, x_2.....x_r ; q_1,q_2,......q_n)$ es factible, si con la tecnología del sector es posible obtener el vector de productos q, empleando los insumos x.

Definición: Economías de escala estrictas
Decimos que existen economías de escala en la producción de n bienes empleando r insumos, si dado cualquier vector de insumos y productos factible

$(x_1, x_2.....x_r ; q_1,q_2,......q_n)$

y dado cualquier w > 1, es también factible el vector de insumos y productos

$(w.x_1, w.x_2.....,w.x_r ; v_1q_1, v_2q_2,......, v_nq_n)$

donde $v_1,v_2,.....v_n$ son todos mayores a w.

Expresado en palabras, **existen economías de escala estrictas** si, al incrementar en una proporción w todos los insumos, **cada uno de los productos puede incrementarse en una proporción estrictamente mayor.**

Cuando se producen n bienes, no es inmediato definir un costo medio, ya que no podemos sumar cantidades de productos diferentes para obtener un producto total. Los costos medios decrecientes pueden definirse, pero si se mantiene una proporción dada en la producción de cada uno de los n bienes.

Definición: Costo medio decreciente por semirrectas (declining ray average cost)
Se dice que una función de costos en la producción de n bienes $C(q_1,q_2,......q_n)$ presenta costos medios decrecientes en una semirrecta, si para cualquier $q_1,q_2,......q_n$ no idénticamente nulo, y para cualquier w > 1, se cumple:

$$C(w.q_1,w.q_2,.....w..q_n) < w.C(q_1,q_2,.......,q_n)$$

Es decir que al aumentar en igual proporción todos lo productos, el costo conjunto crece menos que proporcionalmente.

Ejemplo: En la producción de autos y camionetas de una marca dada, y en un amplio rango de cantidades, si se producen 10000 autos y 2000 camionetas, el costo es menos que mil veces superior que si se producen 10 autos y 2 camionetas.

Teorema:
Si la función de costos C(q) es creciente estricta en cada una de sus variables, entonces la existencia de economías de escala implica que existe costo medio decreciente por semirrectas.

Demostración:

Sea $q=(q_1,q_2,......q_n)$ no idénticamente nulo, y un escalar $w > 1$:

Sea x el vector de insumos necesarios para producir q.

Por existir economías de escala estricta, existen $v_1, v_2,...,v_n$ todos ellos mayores que w tales que con el vector de insumos w.x pueden producirse cantidades de los n bienes $v_1q_1, v_2q_2,......, v_nq_n$

➔ $C(v_1q_1, v_2q_2,......, v_nq_n) = w.C(q)$ (1)

porque el vector de insumos w.x cuesta w veces lo que cuesta el vector x.

Pero por ser C(.) estrictamente creciente:

$w.q_i < v_i . q_i$ ➔ $C(w.q_1, w.q_2,......, w.q_n) < C(v_1q_1, v_2q_2,......, v_nq_n)$ (2)

Combinando (1) y (2)

$C(w.q_1,w.q_2,.....w..q_n) < w.C(q_1,q_2,.......,q_n)$ es decir hay costos medios decrecientes por semirrecta.

Sea *f(x,y)* la función de costos de producir dos bienes en cantidades *x* e *y* respectivamente y f(x,0) el costo total de producir uno de los bienes por una firma en una cantidad *x*, y *f(0,y)* el costo de producir el otro bien por otra firma en una cantidad *y*. Hay economías de alcance si *f(x,y)>f(x,0)+f(0,y)*.

Es posible igual que en el caso de economías de escala, de que existan economías de alcance para algunos niveles de producción y no para otros. Es posible que sea más barato para un firma producir varios bienes cuando las cantidades son pequeñas pero no para grandes cantidades. Cuando es conveniente producir con una empresa o con varias depende de la relación de las zonas de economías de alcance con la demanda por los bienes.

Definición: Economías de alcance (scope).
Decimos que en la producción de n bienes existe economías de alcance, si con la tecnología empleada, para todo $q_1, q_2, ...q_n$, la función de costos de producción $C(Q):R^n \to R$ cumple:

$$C(q_1, q_2, ...q_n) < C(q_1,0,...0) + C(0,q_2, ..,0) +.....+ C(0,0,...,q_n)$$

Es decir que existen economías de alcance si el costo de producir cantidades cualesquiera de los n bienes en forma conjunta en una misma empresa, es menor que la suma de los costos de producir los n bienes en forma independiente.

Ejemplos: Existen economías de alcance en la producción de servicios financieros diversos: servicios para las empresas, préstamos personales, préstamos hipotecarios, etc. Una gran parte de la infraestructura de locales, sistemas de información, etc. son compartidas por la provisión de cada uno de los productos financieros mencionados.

Teorema: Si existe subaditividad, entonces existen econonías de alcance.
Lo demostraremos para n=2 para simplificar la notación.

Sean las cantidades q_1 y q_2 de dos bienes a producir. Sean $Q^1=(q^1,0)$, $Q^2=(0,q^2)$ los dos vectores de producción si se producen los dos bienes en empresas separadas.

Los costos de producción respectivos en empresas separadas son:
$C(Q^1) = C(q_1,0)$
$C(Q^2) = C(0,q_2)$

Por la subatidividad:

$C(Q^1)+C(Q^2) > C(Q^1+Q^2)$

$Q^1+Q^2 = (q_1,0) + (0,q_2) = (q_1, q_2)$

Pero el primer miembro es $C(q_1,0) + C(0,q_2)$

Y el segundo miembro de la desigualdad es $C(q_1, q_2)$, entonces la subatidividad termina implicando:

$$C(q_1,0) + C(0,q_2) > C(q_1, q_2)$$

es decir las economías de alcance.

Las economías de alcance requieren que $C(Q) < C(Q^1) + C(Q^2)$

donde $Q^1 = (q_1,0)$ y $Q^2=(0,q_2)$ y por lo tanto $Q^1 + Q^2 = Q$

La subaditividad requiere que C(Q) < C(OP) + C(PQ) , donde los vectores OP y PQ son cualquier forma de descomponer la producción del vector Q. Como se ve, un caso particular de esa "descomposición" es lo requerido por las economías de alcance.

Ejemplo: El recíproco no es cierto, pueden existir economías de alcance pero no subaditividad.

Consideremos por ejemplo la función de costos $C(q_1,q_2) = q_1^2 + q_2^2 - q_1q_2$
Existen economías de alcance:
$C(q_1,0) + C(0,q_2) = q_1^2 + q_2^2 > q_1^2 + q_2^2 - q_1q_2 = C(q_1,q_2)$
Sin embargo no existe subaditividad. Basta considerar repartir la producción de las cantidades (q_1,q_2), en n empresas de igual función de costos.
$n. (q_1/n , q_2/n) = n.(q_1^2 + q_2^2 - q_1q_2)/n^2 = C(q_1,q_2)/n < C(q_1,q_2)$

La economía de escala (aumento más que proporcional del producto al aumentar proporcionalmente todos los insumos) **no es ni condición necesaria ni suficiente de subaditividad.** Para el caso monoproducto, la economía de escala era condición suficiente de subaditividad.

Veamos un contraejemplo:

Consideremos la producción conjunta de q_1 y q_2 empleando un solo insumo en cantidad L. La función de producción es:

$$q_1^{1/2} + q_1^{1/4}\, q_2^{1/4} + q_2^{1/2} = L$$

Esta función de producción presenta economías de escala, porque si tomamos una cantidad de insumos $\square L$, con $\square>1$, la producción de cada uno de los dos productos se puede incrementar en $\square^2 > 1$. Verifiquémoslo:

$$(\square^2 q_1)1^{1/2} + (\square^2 q_1)^{1/4}\, (\square^2 q_2)^{1/4} + (\square_2 q_2)^{1/2} = \square\ (q_1^{1/2} + q_1^{1/4}\, q_2^{1/4} + q_2^{1/2}) = \square\ .L$$

Sin embargo la función de producción no es subaditiva. Supongamos precio unitarios del insumo único.

$C(q_1,0) = q_1^{1/2}$

$C(0,q_2) = q_2^{1/2}$

$C(q_1,0) + C(0,q_2) = q_1^{1/2} + q_2^{1/2}$

$C(q_1,q_2) = q_1^{1/2} + q_1^{1/4}\, q_2^{1/4} + q_2^{1/2}$

$C(q_1,0) + C(0,q_2) < C(q_1,q_2)$ porque $q_2^{1/2} >0$ para cualquier $q_2 > 0$.

La relevancia de las economías de alcance depende muchas veces de como se definen los bienes. Si tomamos por ejemplo el mercado de las telecomunicaciones, los abonados locales se comunican con una unidad de intercambio local que dirige la llamada a su destino. Las llamadas se dirigen a esa unidad, no importando que se dirijan a la misma ciudad (llamada local) o al exterior (llamada internacional), por lo que el cable que une al usuario con la central es único no importando para que se use.

Si defino a la telefonía internacional como un servicio distinto que la telefonía local, existen economías de alcance considerables, ya que una sola empresa tendrá costos mucho menores para proveer el servicio que dos empresas (dos cableados son muy costosos e implican equipamiento redundante). Esta es la filosofía adoptada en Estados Unidos antes de la reforma que desintegró AT&T, que habilitaba a la misma a brindar ambos servicios de forma monopólica en la mayor parte de las regiones de Estados Unidos.

Si defino que los servicios que se brindan son el acceso local en cada ciudad (de origen y destino de la llamada internacional) y el transporte de las comunicaciones entre ambas ciudades, entonces las economías de alcance no existen, ya que nada indica que una sola empresa tenga costos menores que tres empresas que se encarguen cada cada parte de la comunicacion (originación de la llamada hasta la unidad de intercambio, otra del transporte y otra de terminacion desde la unidad de intercambio de destino y el usuario destinatario de la comunicación). En este caso no existen redundancias de infraestructura, ya que solo un cable va de usuario a usuario. Este tipo de argumento está atrás de la desintegración de AT&T. Existen hoy compañías telefónicas locales y *carriers* proveedores de servicios de larga distancia.

¿Pero que pasa a nivel de cada uno de los mercados? En el mercado de la larga distancia las economías de escala se agotan a un nivel menor respecto a la demanda, la competencia es beneficiosa y es alentada en general. En el mercado local (conexión de los usuarios con la unidad de intercambio) es previsible que las economías de escala existen y por lo tanto hay un monopolio natural de este tramo y una empresa única es razonable que exista. (Debe atenderse en este caso a otros argumentos como los mercados contestables y la convergencia de tecnologías que relativizan este agumento). ESTE ARGUMENTO ES VALIDO PARA LIBERALIZAR LA INTERNACIONAL PERO NO PARA PROTEGER LA LOCAL.

La existencia de un monopolio natural debe ser analizada viendo la situación total de los costos y no la existencia de economías de escala o alcance por separado. Las economías de escala pueden existir con o sin economías de alcance y viceversa. El concepto que debe asociarse a la existencia de un monopolio natural es el de subaditividad de costos (existe un monopolio natural cuando la función de costos es subaditiva para cierto nivel de producto, por lo que subaditividad es sinónimo de monopolio natural).

La función de costos es subaditiva para un cierto nivel de producción, cuando el costo de producir ese nivel de producto es menor con una firma que con más de una firma, para todas las posibles formas de dividir la producción entre las fimas (división por productos o por cantidades o una combinación). Por ejemplo, si existen dos bienes A y B, la función de costos es subaditiva si para las demanda de los bienes, la producción de ambos bienes por una firma es mas barata que toda otra combinación, las que pueden inclir: producción de A por una firma y B por otra firma (razonable si existen economías de escala), que dos firmas produzcan A y B en cantidades iguales, una empresa produce 20% de A y 80% de B y la otra 80% de A y 20% de B, o cualquier otra combinación. En resumen, el concepto de subaditividad de costos incorpora consideraciones de escala y alcance e identifica, dadas todas las consideraciones, cuando una empresa en más barata.

Igualmente que las economías de alcance y escala, la subaditividad puede existir para algunos niveles de producto y no para otros.

3. MONOPOLIO NATURAL Y REGULACION

3.1 Resultado financiero del monopolista

El análisis de equilibrio parcial anterior nos indica que la maximización de la suma de excedentes de los consumidores y del monopolista se logra si se fija administrativamente el precio igual al costo marginal.

Recordemos que si el monopolista fija su producción de modo que p = CMg, está reproduciendo el comportamiento de firma tomadora de precios que conducía a un óptimo de Pareto en el equilibrio competitivo.

Sea q^* tal que $CMg(q^*) = P(q^*) =: p^*$, es decir la cantidad a vender según la regla marginalista.

El problema que se presenta inmediatamente es el del equilibrio financiero del monopolista sujeto a la regla marginalista, con lo que naturalmente surgen tres casos:

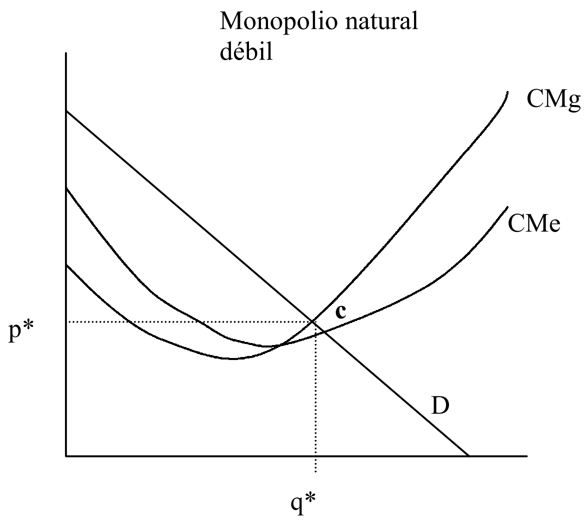
- $p^* = CMg(q^*) > CMe(q^*)$, es decir el monopolista tiene un beneficio positivo con la regla marginalista. A esta situación se la denomina **monopolio natural débil.**

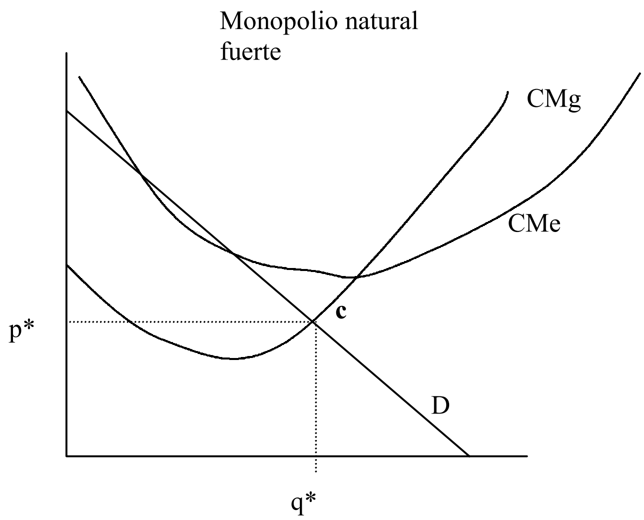
- $p^* = CMg(q^*) < CMe(q^*)$, es decir el monopolista tiene un beneficio negativo con la regla marginalista. A esta situación se la denomina **monopolio natural fuerte.**

- $p^* = CMe(q^*)$, esta situación es un evento que si bien no es imposible tiene probabilidad probabilidad nula, de modo que basta considerar las dos anteriores.

Supongamos que la curva de costos medios tiene a lo sumo un mínimo relativo y tiene un tramo inicial decreciente. Es decir para simplificar nuestro análisis restringimos la forma de la curva de costos medios.

Consideremos primero el caso en que la curva de costos medios tiene un mínimo local, lo que ocurre en el punto en que se cortan las curvas de costos medios y marginales.





Los dos casos posibles son:

- **Monopolio natural débil** $p^* = CMg(q^*) > CMe(q^*)$

  En este caso el costo medio es creciente en el rango de producción relevante en las proximidades de q*. Es decir que la curva CMg corta a CMe antes de q*

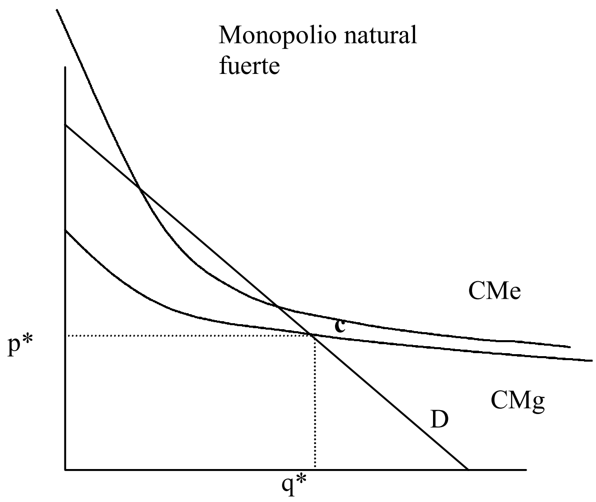
  El monopolista regulado con la regla marginalista obtiene beneficios extraordinarios

- **Monopolio natural fuerte** $p^* = CMg(q^*) < CMe(q^*)$

  En este caso el costo medio es decreciente en el tramo de producción relevante en las proximidades de q*. La curva CMg corta a CMe después de q*.

  El monopolista sujeto a la regla marginalista incurre en pérdidas

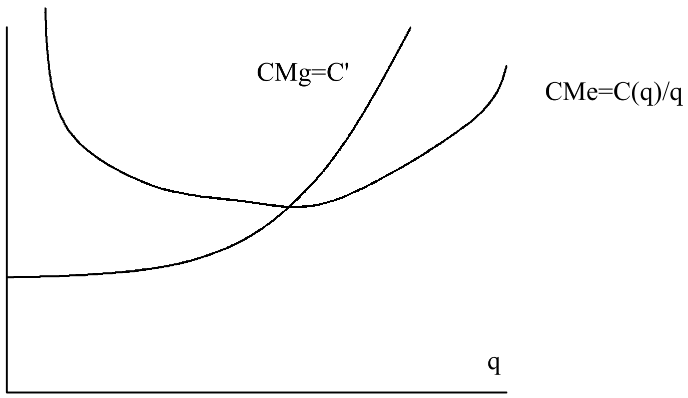
Si la curva de costos medios no tiene ningún mínimo local, como inicialmente es decreciente, lo será en todo punto y tenemos otro caso de monopolio natural fuerte



En resumen, con las simplificaciones realizadas respecto a la curva de costos medios, tenemos que los tres gráficos anteriores resumen las posibilidades de evolución de los costos medios y marginales.

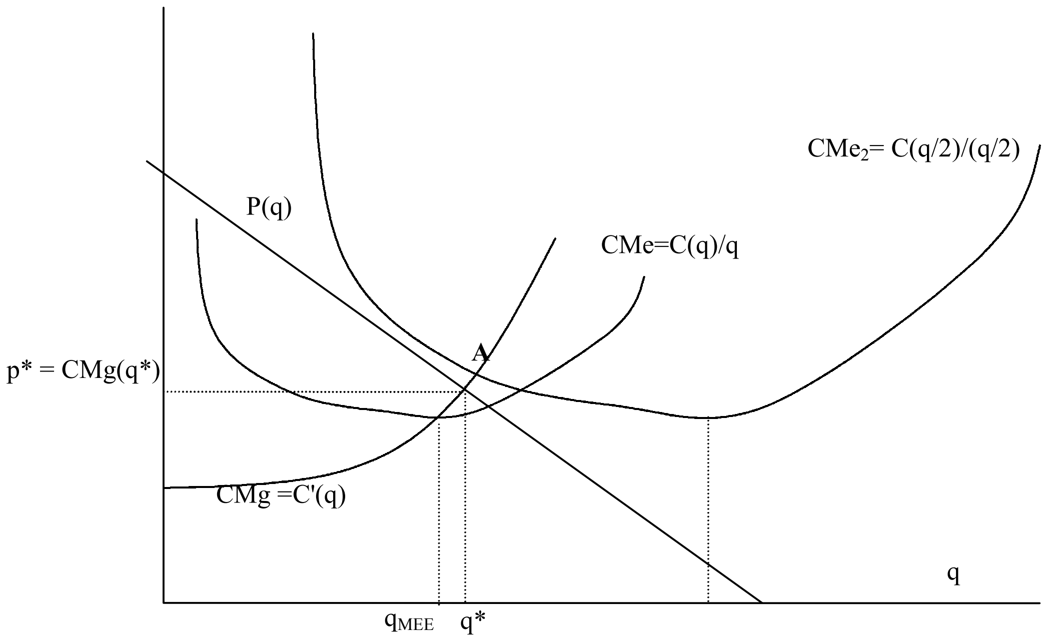
Un ejemplo de monopolio natural débil

El monopolio natural débil es una situación no totalmente intuitiva. La empresa está operando en el tramo de costos medios crecientes, pero aún así la función de costos es subaditiva. Supongamos que la función de costos en un sector es la que se presenta en el dibujo (que representa un caso de monopolio natural débil):



Como en este caso los costos marginales son siempre crecientes, **si hay dos o más empresas** en el mercado, la forma de repartir la producción entre ellas, de modo de minimizar los costos es en cantidades iguales.

Llamemos q_{MEE} a la mínima escala eficiente, es decir al punto en el que con una sóla empresa se obtiene el costos medio mínimo.



Para verificar la subaditividad para el rango relevante de la demanda, debemos verificar que es más barato producir una cantidad q dada en una sola empresa y no en dos. Si el costo marginal es creciente la forma mejor de producir en dos empresas es repartir la producción por mitades.

Llamemos CMe(q) a la curva de costo medio produciendo en una empresa y $CMe_2(q)$ a la curva de costo medio produciendo óptimamente en dos empresas.

$CMe_2(q) = 2C(q/2)/q = C(q/2) / (q /2) = CMe(q/2)$

Existirá subaditividad en todos los q que verifiquen:

$C(q)/q < 2.C(q/2) / q$

Es decir que el costo medio produciendo en una sola empresa es menor que el costo medio produciendo en dos empresas de la manera óptima, es decir repartiendo la producción por mitades.

En el dibujo que mostramos existe subaditividad y por lo tanto monopolio natural, para cantidades a la izquierda del punto **A** de intersección de las curvas de costos medios de una y dos empresas.

Consideremos ahora un ejemplo numérico.

Sea $C(q) = F + q^2/2$

$CMg = q$, el costo marginal es creciente

$CMe = F/q + q/2$

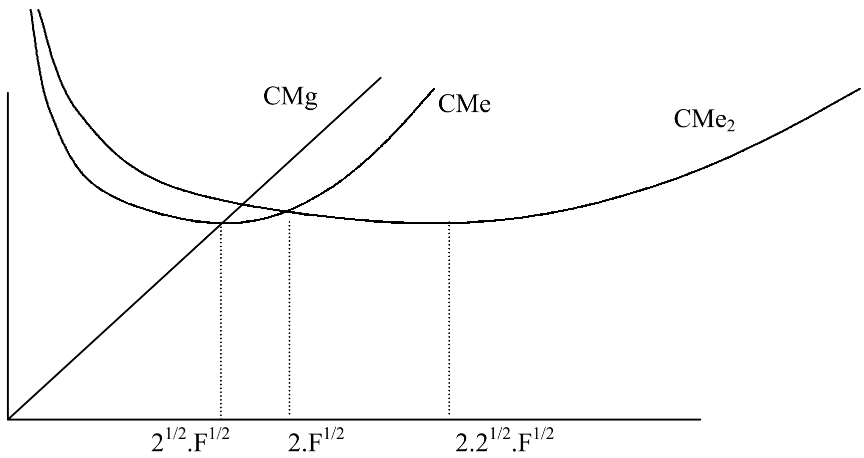
El mínimo costo medio se da para $q = F/q + q/2 \Rightarrow q/2 = F/q \Rightarrow q = (2.F)^{1/2}$

El costo medio de producir en dos empresas es:

$CMe_2 (q) = CMe(q/2) = F/(q/2) + q/4 = 2.F/q + q/4$

Existe monopolio natural si:

$CMe_2 (q) > CMe(q) \Rightarrow 2.F/q + q/4 > F/q + q/2 \Rightarrow F/q > q/4 \Rightarrow q^2 < 4F \Rightarrow q < 2.F^{1/2}$



A su vez, si la demanda corta a la curva CMg en $[(2F)^{1/2}, 2.F^{1/2}]$, se trata de un monopolio natural débil.

3.2 Segundo óptimo y regulación del monopolio monoproducto

Si nos encontramos frente a una situación de monopolio natural fuerte, la aplicación de la regla de óptimo primero, es decir fijar un precio p*, que conduce a una demanda $q^*=D(p)$, tales que $p^* = CMg(q^*)$, conduciría a un beneficio negativo para el monopolista, ya que:

$$p^* = CMg(q^*) < CMe(q^*)$$

Si el monopolista no puede recurrir a subsidios para permitir la sostenibilidad del óptimo primero (p*,q*), su problema es encontrar un precio p, o una cantidad q, que maximice la suma de excedentes US, sujeto a que el beneficio del monopolista sea no negativo.

$$US(q) = \int_0^q P(x).dx - C(q)$$

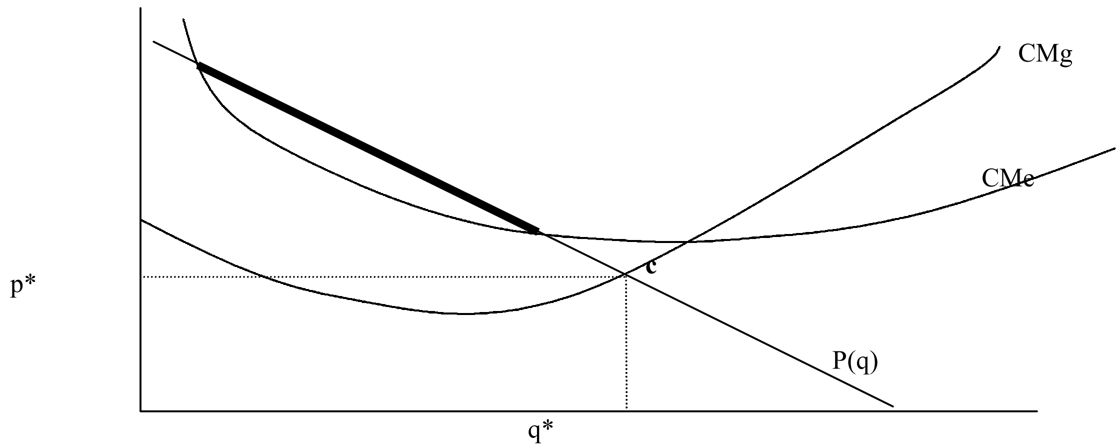
$$dUS/dq = P(q) - C'(q) = P(q) – CMg(q)$$

Supongamos que para $q < q^*$ $P(q) > CMg(q)$.

Entonces $dUS/dq = P(q) – CMg(q) > 0$ para $q < q^*$.

El problema de maximizar el excedente sujeto a la restricción presupuestal es entonces

Máx q
s.a: $P(q).q \geq C(q)$ o lo que es equivalente $P(q) \geq CMe(q)$



El problema será semejante al planteado en el gráfico, donde los puntos de la curva de demanda en los que el monopolista tiene beneficios no negativos son los indicados con el trazo grueso

La solución será un punto en el que la restricción está operativa, es decir en donde P(q) = CMe(q) y dentro del conjunto de puntos de la demanda que cumplen esta propiedad, el que tenga una cantidad máxima entre ellos.
Existen tres aspectos a considerar por un regulador al analizar la regulación a aplicar en un monopolio monoproducto:

- **La magnitud de la ineficiencia asignativa creada por el monopolio.** Podría ocurrir que el peso muerto del monopolio fuese muy reducido y no justificase la intervención regulatoria. Esta es una consideración más bien teórica, ya que lo probablemente incida sobre la decisión de regular la magnitud del beneficio del monopolista y las quejas de los consumidores respecto al precio.

- **La situación financiera que crea la regla tarifaria del costo marginal, que conduce al óptimo primero.** Como vimos, la aplicación de esa regla puede dar lugar a pérdidas para el monopolista, si se trata de un monopolio natural fuerte.

- **La existencia o no de barreras a la entrada.** La necesidad de regular dependerá de la amenaza de entrada de competidores. Si existen barreras a la entrada, y la entrada está bloqueada, al no existir amenaza de entrada potencial la única restricción a la conducta monopólica en la fijación de precios procederá de la regulación. Si por el contrario, y situándonos en el caso totalmente opuesto, existiese la posibilidad de entrada y salida sin costos hundidos, entonces la competencia potencial podría tal vez contribuir a disciplinar al monopolista.

En la práctica el regulador de un monopolio natural fuerte tiene varias alternativas:

- **Aplicar el precio tal que p=CMg y subsidiar al monopolista con fondos públicos**

  Esta alternativa conduce a maximizar el excedente social si los fondos públicos para subsidio pueden obtenerse mediante el sistema impositivo, sin ninguna ineficiencia. Es decir si por cada peso de subsidio alcanza con limitar en un peso el excedente de los contribuyentes.

  Esta alternativa es en general políticamente inviable, especialmente si el monopolista es una empresa privada, a menos que el subsidio tenga por objeto permitir prestar un servicio público en una región o a un grupo de usuarios de bajos ingresos.

  Por otra parte aunque fuese posible implementar el subsidio, desde el punto de vista microeconómico, la obtención de fondos públicos mediante impuestos genera ineficiencias asignativas ya que difícilmente es posible obtener dichos impuestos sin alterar las decisiones de los agentes económicos.

  Hay una discusión clásica en la década del 30 a partir de los trabajos de Hottelling, que aparece resumida en el punto 2.5 del libro de Berg y Tschirhart. En la perspectiva institucional actual, la posibilidad de subsidio parece mucho menos aceptable que cuando la discusión tuvo lugar.

- **Fijar un precio que se iguale al costo medio**, apartándose de la tarifa marginalista y sin emplear en absoluto los subsidios.

  Esto puede hacerse siempre y cuando la curva de costos medios corte a la curva de demanda en algún punto. En el caso de monopolio natural fuerte, podría no existir punto de corte.

  Si existe más de un punto de corte, será preferible aquél en el que la producción sea mayor, porque la producción está más próxima a q*.

  Este criterio p=CMe conducirá a un precio inferior al que fijaría el monopolista por su cuenta, y a una cantidad mayor y por lo tanto sería preferible a no regular

- **Recurrir a un precio intermedio entre los dos anteriores**, subsidiando parcialmente al monopolista y permitiendo un precio superior al costo marginal, dada la cantidad producida.

- **Recurrir a alguna otra forma de tarifa no lineal** (esto es, que no tenga un único precio variable para todas las unidades)

  Si fuese posible una tarifa con un cargo por unidad consumida igual al costo marginal, con lo que los consumidores compran las cantidades del óptimo p=CMg, y con un cargo fijo que permitiese cubrir el déficit generado por la tarifa marginalista, se resolvería el problema del financiamiento manteniendo el máximo excedente colectivo posible. Este tipo de soluciones, de **tarifas no lineales,** será objeto de un análisis más adelante y como veremos la solución no es tan sencilla.

3.3 Regulación del monopolio multiproducto

Primer Optimo

De manera totalmente análoga a lo que hicimos para el caso monoproducto, si el regulador quiere maximizar la suma de excedentes de los consumidores en los n mercados en los que vende el monopolista, más el beneficio del monopolista, la regla que conduciría al óptimo sería que los precios se igualen a los respectivos costos marginales.

Para que el excedente del consumidor en cada uno de los mercados quede inequívocamente definido dado el precio en el mismo, **supondremos demandas independientes para los n bienes que vende el monopolista.**

Sea:

$P_i(q_i)$ la función de demanda del bien i, nótese que no depende de las cantidades de los restantes bienes, ya que resulta de invertir $Q_i(p_i)$

C(q) la función de costos totales del monopolista, que depende del vector q con las n cantidades producidas y vendidas.

$EC_i(q) = \int_0^{q_i} P_i(x).dx - P_i(q_i).q_i$ es el excedente de los consumidores por el bien i

$B(q) = \sum_1^n P_i(q_i).q_i - C(q)$ es el beneficio del monopolista

$$US(q) := \sum EC_i(q_i)+ B(q) = \sum \int_0^{q_i} P_i(x).dx - C(q)$$

El regulador está maximizando la utilidad social total US(q), que resultó igual al área bajo la curva de demanda de los consumidores menos el costo de producción del bien.

Hallaremos ahora las cantidades q_i que maximizan el excedente total. Para ello buscaremos un extremo relativo, suponiendo que se da para valores estrictamente positivos de las ventas de cada uno de los n bienes.

$\partial US/\partial q_i = P_i(q_i) - \partial C/\partial q_i = 0$ ➔ $\mathbf{P_i(q_i) = CMg_i(q)}$

Como vemos vale también una regla marginalista.

Segundo Optimo

Tenemos ahora un monopolio multiproducto, es decir que el regulador debe fijar los precios de n bienes que vende el monopolista. Suponemos que la tarificación a precio marginal genera déficits para el monopolista. Es necesario recurrir a unas tarifas "second best", segundo mejor, es decir las óptimas entre las que permiten al monopolista cubrir sus costos.

Suponemos que la cantidad demanda de cada bien, es independiente de los precios de los restantes n -1 bienes que produce el monopolista.

El regulador está impedido de hacer transferencias al monopolista, y suponemos que sólo puede fijar tarifas lineales, esto es compuestas por distintos precios unitarios por el consumo de cada bien, independientes de la cantidad consumida.

La producción de cada uno de los n bienes tendrá un costo marginal diferente. El problema del regulador es **¿cuánto aparta cada uno de los n precios de sus respectivos costos marginales** de modo que el monopolista tenga beneficio nulo, en lugar de negativo?

Los precios Ramsay surgen de un análisis de equilibrio parcial en el que se trata de maximizar la suma de excedentes de los consumidores, sujeto al equilibrio presupuestal. Por definición el excedente del monopolista es idénticamente nulo y no entraría en la maximización, lo incluimos para darle una interpretación adecuada a la variable dual de la restricción del problema. Despreciamos las variaciones de excedente en los otros mercados.

El excedente total de los consumidores a maximizar será la suma de los n excedentes en los n mercados.

El problema del regulador es entonces:

$$\underset{\{q_i\}}{\text{Máx}} \quad \sum \left[\int_0^{q_i} P_i(x)dx - P(q_i).\, q_i \right] + \left[\sum P(q_i).\, q_i - C(q) \right] = 0$$

s.a.:

$$\sum P(q_i).\, q_i - C(q) \geq 0 \quad \text{o su equivalente} \quad C(q) - \sum P(q_i).\, q_i \leq 0$$

El objetivo es la suma del excedente de los consumidores en los n mercados, más el segundo sumando entre paréntesis recto que es el excedente del monopolista.

La restricción es que el excedente del monopolista sea no positivo nulo.

Las variables en las que se maximiza son cantidades por comodidad, el precio en cada mercado queda unívocamente determinado por la curva de demanda inversa en cada mercado.

Suponiendo que el máximo tiene lugar en un extremo relativo en el que los n bienes se venden en cantidades no nulas, se trata de hallar el lagrangeano del problema y anular sus derivadas respecto a las n cantidades. Sea λ el multiplicador de Lagrange de la restricción. Por tratarse de un problema de maximización sujeto a una restricción de menor o igual, en el óptimo el valor de la variable dual debe ser positivo.

El lagrangeano L es igual a:

$$L = \sum \left[\int_0^{q_i} P_i(x)dx - P_i(q_i).\, q_i \right] + (1+\lambda) \left[\sum P_i(q_i).\, q_i - C(q) \right] = 0$$

Para i = 1,...n, se tiene:

$$\underline{P_i(q_i)} - P_i'(q_i).\, q_i - \underline{P_i(q_i)} + (1+\lambda).\left[P_i'(q_i).\, q_i + P_i(q_i) - CMg_i(q) \right] = 0$$

$$(1+\lambda).\left[P_i(q_i) - CMg_i(q) \right] = -\lambda\, P_i'(q_i).\, q_i$$

$$\left[P_i(q_i) - CMg_i(q) \right] / P_i(q_i) = \left[\lambda/(1+\lambda) \right] \left[- P_i'(q_i).\, q_i / P_i(q_i) \right]$$

➔ **$[p_i - CMg_i]/ p_i = [\square\square\square\square|\square_i|$**

Esta fórmula, es conocida como la regla de Ramsay.

El aumento relativo del precio cada bien respecto a su costo marginal, es mayor cuanto menor sea su elasticidad en valor absoluto $|\square_i|$, es decir cuanto más inelástica sea su demanda.

A la expresión $[p_i - CMg_i]/ p_i$ se la suele llamar índice de Lerner.

El multiplicador $\square > 0$, es la ganancia de excedente total que se obtendría si se relaja en una unidad la restricción presupuestaria. Este

Otra forma de expresar la fórmula es:

$$\frac{[p_i - CMg_i]/ p_i}{[p_j - CMg_j]/ p_j} = \frac{|\square_j|}{|\square_i|} \qquad \text{para todo i, j}$$

Los apartamientos son inversamente proporcionales a las elasticidades de los bienes. Nuevamente vale aquí la intuición que mostramos en un punto anterior. Cuanto más inelástica la demanda, menor será el triángulo de pérdida de excedente colectivo por variar el precio en un porcentaje dado.

4. Bibliografía

Berg y Tschirhart. Natural Monopoly Regulation. Cambridge University Press 1998.

Bergara, Mario. La Regulación de Servicios Públicos. Departamento de Economía – Facultad de Ciencias Sociales – Universidad de la República

Special Report: Unscrambling Telecommunications Policy. Whashington Research Council, 1997

Train, K. Optimal Regulation: The Economic Theory of Natural Monopoly, MIT Press, 1991.