

LA TEORIA CLASICA DE LA REGULACION DE MONOPOLIOS NATURALES

*Luis Rodríguez Romero**

El trabajo presenta una revisión de las condiciones de existencia de una situación de monopolio natural, el tipo de óptimos de primer y segundo orden alcanzables en dicha situación y las características de los mecanismos regulatorios básicos susceptibles de ser empleados en su consecución: regulación al coste del servicio y precios no lineales. La exposición se restringe a los que se pueden considerar como mecanismos tradicionales de regulación, sin considerar explícitamente la presencia de asimetrías de información.

Palabras clave: *regulación económica, monopolio natural, fallos del mercado, eficiencia técnica, análisis teórico.*

I. Introducción

En un principio podría considerarse un contrasentido el fuerte impulso que ha experimentado la teoría de la regulación durante los últimos años, coincidiendo con una clara revitalización del papel del mercado como mecanismo básico de asignación de recursos.

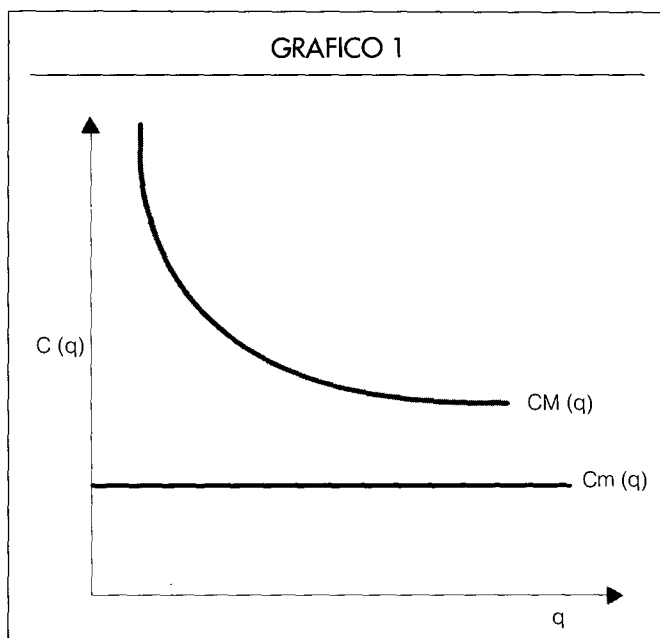
Dicha aparente paradoja encubre, sin embargo, una profunda lógica interna en la medida en que de la revitalización del papel del mercado se deriva inmediatamente una conocida y doble inferencia vinculada, por una parte, con la acotación de sus propios límites así como, por otra parte, con los condicionantes que impone su funcionamiento a cualquier política pública orientada a influir en el comportamiento de los agentes económicos. Ambos aspectos han constituido tradicionalmente las referencias fundamentales de la teoría de la regulación, dirigida a dos objetivos básicos:

1) Delimitar las condiciones que aconsejan la necesidad de una acción regulatoria, por parte del sector público; y 2) anali-

zar los problemas implícitos en dicha acción, vinculados con las características de la situación de partida, el tipo de interacción entre regulador y regulado, y el esquema general de incentivos existente en dicha relación.

Al igual que en otros muchos aspectos, la transmisión del interés sobre la teoría de la regulación a nuestro país llega con un cierto desfase y, lo que es más importante, se encuentra con un profundo vacío fruto de la falta de tradición de este tipo de análisis. En general se podría establecer que en el sector público español, como en mayor o menor medida ha sucedido en otros de tradición europea continental, ha primado un comportamiento intervencionista más que regulador. Esto ha supuesto en muchas ocasiones una recurrencia a la presencia pública directa en los mercados, con lo que esto implica respecto a la ambivalencia e indefinición de los papeles respectivos de regulador y regulado, muchas veces coincidentes en su titularidad pública. No es extraño, por lo tanto, que gran parte de los desarrollos sobre la actuación de los mecanismos reguladores haya tenido su origen en EEUU, país en el que tradicionalmente ha existido una clara delimitación de los papeles respectivos de regulador y regulado, mientras que en el caso europeo se haya concedido

* Universidad Carlos III de Madrid.



una mayor prioridad al análisis de la actuación de las empresas públicas, atendiendo en especial a los sistemas de fijación de precios para servicios públicos.

En general, la justificación de cualquier medida de carácter regulador se relaciona explícita o implícitamente con la existencia de algún tipo de fallo de mercado motivado por alguna de las tras razones fundamentales normalmente esgrimidas a este respecto: la existencia de efectos externos, la presencia de algún tipo de problema de información imperfecta o el carácter de los condicionamientos tecnológicos que configuran una situación de monopolio natural.

El primer caso normalmente se relaciona con la existencia de una situación de bien público y constituye la razón básica de la acción reguladora en actividades como la tecnología, el medio ambiente, la sanidad, etcétera, aunque también afecta a aspectos parciales de la acción en actividades como telecomunicaciones, transporte, etcétera. El segundo caso supone la base de las iniciativas dirigidas a paliar la existencia de asimetrías de información entre productores y usuarios a través del empleo de procesos de certificación, control o imposición de estándares, afectando especialmente a actividades de servicios pero también a otras

de carácter industrial como es la industria farmacéutica. Por último, la situación clásica de monopolio natural surge cuando por razones de eficiencia técnica es aconsejable limitar el número de productores en un mercado, articulando medidas adecuadas para hacer frente al efecto negativo sobre la asignación de recursos de una restricción de la competencia en el mismo.

El objetivo del presente trabajo se restringe al último de los casos mencionados, monopolio natural, planteando la plasmación a este caso concreto de las dos cuestiones básicas de cualquier situación reguladora anteriormente mencionadas: las razones que justifican una situación de monopolio natural, por una parte, y los problemas generales de carácter regulador que de dicha situación se derivan.

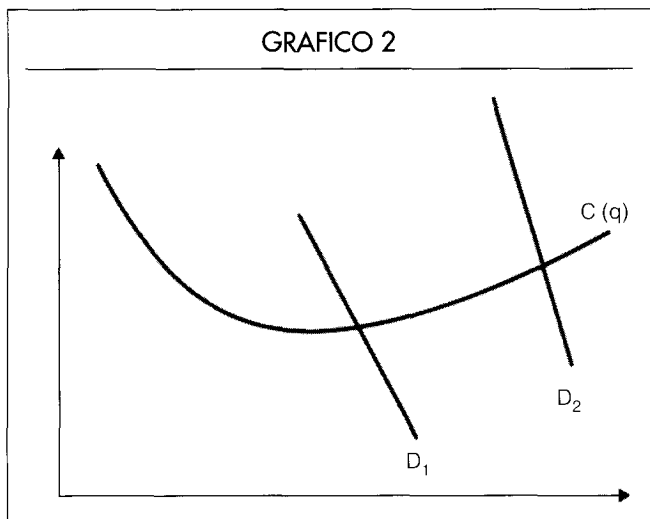
Así, en el siguiente apartado II se considerarán las características de costes y de demanda que definen una situación de monopolio natural en un marco tanto uniproducto como multiproducto. En el apartado III se definirán las condiciones generales de óptimo de primer orden y de segundo orden derivadas de dicha situación, así como los problemas planteados para su consecución. En los apartados IV y V se revisarán los dos tipos de mecanismos reguladores básicos: regulación al coste del servicio y precios no lineales; analizando a través de su actuación el tipo de restricciones y condicionamientos implícitos en la relación reguladora. Por último, el apartado VI intentará resumir algunas de las conclusiones derivadas de la revisión realizada.

Como se desprende del anterior esquema, la exposición se restringirá básicamente a los mecanismos tradicionales de regulación, sin entrar en los desarrollos más recientes sobre diseño de mecanismos óptimos que incorporen explícitamente la existencia de asimetrías de información entre regulador y regulado¹.

II. Condiciones de existencia de un monopolio natural

Tradicionalmente, la existencia de una situación de monopolio natural se ha relacionado con la presencia de economías de escala

¹ Véase el reciente libro de LAFFONT, J. J., y TIROLE, J. (1993) que recoge y ordena los últimos desarrollos en este campo.



en la tecnología de referencia, lo que implica un menor nivel de costes si una única empresa atiende a la demanda total del mercado.

El Gráfico 1 muestra el caso arquetípico de una situación semejante en la que la empresa se enfrenta a un elevado nivel de costes fijos iniciales junto a un coste marginal constante respecto al volumen de producción, correspondiendo, por lo tanto a una función de costes:

$$C(q) = C_F + c q$$

siendo $C(q)$: coste total; C_F : coste fijo; c : coste marginal; q : producción.

Como casos ejemplificadores de esta situación normalmente se consideran los servicios públicos (*public utilities*) en los que la duplicación del capital fijo que se derivaría de la existencia de más de una empresa en el mercado supondría, por lo tanto, un desperdicio de recursos, en el sentido de que cualquier situación de equilibrio obtenida con dos o más empresas podría lograrse con un menor coste con la existencia de una única empresa.

Ahora bien, como claramente se desprende de los trabajos de autores como Baumol, Bailey, Panzar o Willig a finales de los años setenta², la existencia de economías de escala (elasticidad de coste menor que la unidad) constituye una condición suficiente aunque no necesaria para la existencia de una situación de monopolio natural en un marco uniproducto. La condición necesaria y suficiente se derivaría de la existencia de subaditividad de costes definida como:

$$C(q) < C(q_1) + \dots + C(q_n); \sum_{i=1}^n q_i = q_n$$

Como resulta evidente, dicha condición es local a un punto determinado, dependiendo su rango de aplicación de la función de producción de referencia.

La aplicación de la condición de subaditividad en un marco uniproducto tiene el interés de subrayar la doble vinculación de la existencia de una situación de monopolio natural a condiciones tecnológicas y de demanda. Tal y como se refleja en el Gráfico 2, una situación semejante puede darse en un tramo de la función de producción con rendimientos decrecientes de escala y, lo que es más importante, puede dejar de producirse sin ningún tipo de transformación de la tecnología de partida debido a variaciones en el nivel de producción motivadas por desplazamientos de la demanda. Este tipo de matizaciones ha tenido una gran importancia en términos prácticos ya que ha supuesto la introducción de una nueva dimensión en las discusiones de casos regulatorios concretos, habiendo sido empleada, por ejemplo, referida al sector de las telecomunicaciones en el proceso que culminó con la desmembración de la A.T&T a principios de la década de los años ochenta³.

De cualquier forma, la condición de subaditividad adquiere especial relevancia referida a una situación multiproducto. En dicha situación, la existencia de costes medios vectoriales decrecientes (*declining ray average cost*), concepto que correspondería a la transcripción a dicho marco de la existencia de economías de coste en un plano uniproducto, y que se define como:

$$\frac{c(\lambda q)}{\lambda q} < \frac{c(q)}{q}; q = (q_1, \dots, q_n)$$

² Véase BAUMOL, W. J. (1977); PANZAR, J. C., y WILLIG, R. D. (1981); BAUMOL, W. J.; BAILEY, E. E., y WILLIG, R. D. (1977).

³ Véase EVANS, D. S., y HECKMAN, J. J. (1983).

no constituye una condición ni necesaria ni suficiente para la configuración de una situación de monopolio natural:

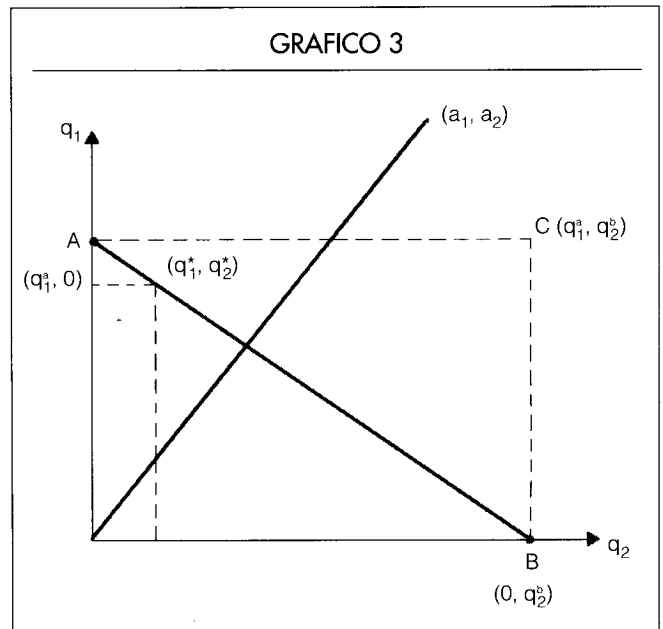
Ello es debido básicamente a la existencia de una nueva dimensión en la función de costes derivada del efecto sobre el nivel de los mismos de la producción conjunta de distintos tipos de productos. El concepto que recoge este tipo de efectos es el de economías de alcance, definido como:

$$C(q_1, q_2, \dots, q_n) < C(q_1, 0, \dots, 0) + C(0, q_2, \dots, 0) + \dots + C(0, 0, \dots, q_n)$$

es decir, la existencia de subaditividad restringida a las combinaciones resultantes de la producción de cada uno de los tipos de productos en solitario.

De cualquier forma, y en contra de lo que se podía suponer en un principio, la simple unión del supuesto de economías de alcance al de economías de escala no supone una condición suficiente para la existencia de monopolio natural en el caso multiproducto como cabría de esperar de la simple transcripción de las condiciones existentes en un marco uniproducto. Esto es debido al componente de economía de escala recogido en la propia definición de economías de alcance, lo que puede dar lugar a un descenso de costes en la producción conjunta derivado del efecto de escala resultante de unir las producciones individualizadas de cada uno de los productos sin existir una real complementariedad de costes entre los mismos y, por lo tanto, sin que dicho efecto reductor de costes se dé para variaciones marginales.

Existen distintos supuestos alternativos para lograr una traslación de la condición de subaditividad al caso multiproducto, expresada en términos de costes incrementales⁴. La posibilidad más sencilla, o, quizá, sería mejor decir menos complicada, es sustituir el supuesto de economías de alcance por el de convexidad transvectorial (*transray convexity*) que supone la



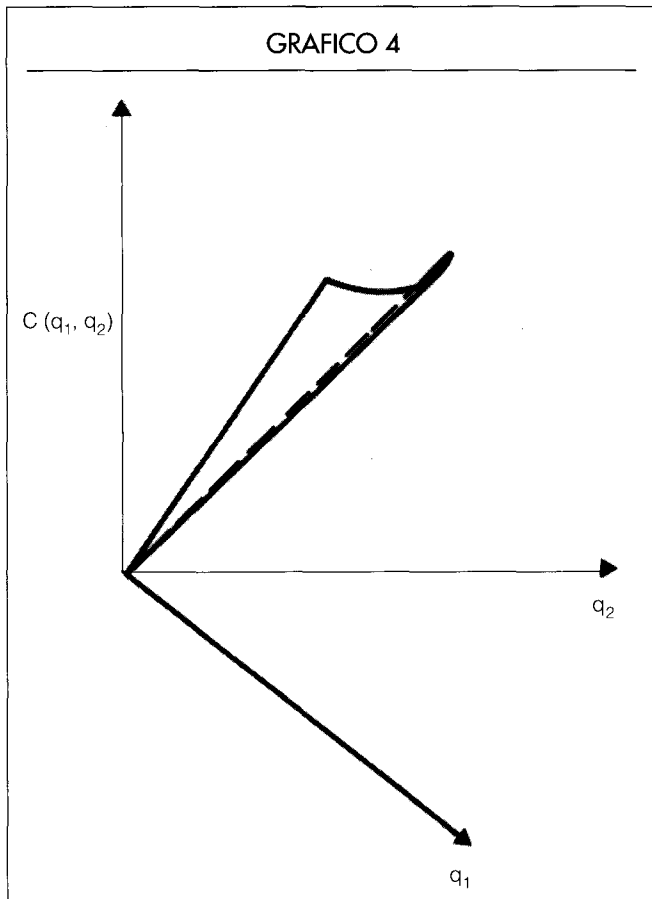
comparación del coste de combinaciones equivalentes de productos.

Así, se define que una función de costes $C(q_1, q_2, \dots, q_n)$ presenta convexidad transvectorial para la combinación (q_1^*, \dots, q_n^*) , cuando existe al menos un vector determinado (a_1, \dots, a_2) , $a_i > 0$, respecto al que la función de costes de las combinaciones correspondientes al conjunto de productos perpendiculares al mismo y que contiene a (q_1^*, \dots, q_n^*) es convexa. Es decir, existe al menos un conjunto H en la que la función de costes es convexa, siendo H :

$$H(a, q^*) = \left\{ q_i \geq 0; \sum_{i=1}^n a_i q_i = \sum_{i=1}^n a_i q_i^* \right\}$$

En el Gráfico 3 el segmento AB mostraría las distintas combinaciones de producto para las que la función de costes debería de ser convexa en caso de existir convexidad transvectorial. Como puede observarse, de esta forma se evita el problema de cambio de escala que se derivaría de la definición habitual de economías de alcance, según el cual se aplicaría la condición de subaditividad a las combinaciones A y B en comparación con la

⁴ Véase SHARKEY, W. W. (1982), capítulo 4.



combinación C. El Gráfico 4 refleja para el caso de dos inputs la conformación de una función de costes con convexidad transvectorial.

En cualquier caso, es conveniente recordar que la unión de economías de escala vectoriales junto a convexidad transvectorial representa únicamente una de las condiciones suficientes, aunque no necesarias, para la existencia de una situación de monopolio natural en un marco multiproducto, pudiéndose dar muchas otras combinaciones específicas, tecnológicas y de demanda, de las que se derive una situación local de subaditividad. Esto resalta la dificultad de «contrastar» econométricamente la existencia de una situación de monopolio natural a través de la estimación de funciones de costes y de demanda, así como la provisionalidad con la que debe tomarse cualquier resultado a dicho respecto.

III. Monopolio natural y bienestar: óptimos de primer y segundo orden

Una vez planteada una situación de monopolio natural, se deduce de inmediato una clara discrepancia entre las posibilidades de obtención de una situación de eficiencia tecnológica en el proceso productivo y los incentivos necesarios para la existencia de una situación de eficiencia en la asignación de recursos, de la que se deriva, en último término, el problema regulador. En efecto, como es evidente, de la propia definición de monopolio natural se infiere una mayor eficiencia tecnológica de la existencia de una única empresa que, sin embargo, puede aprovechar el poder de mercado que tal posición le confiere para fijar unos precios que vulneren la eficiencia asignativa. Tal contradicción sustancia la intervención reguladora que debería conjugar la eficiencia tecnológica que posibilita la existencia de una única empresa, junto con una restricción adecuada de su poder de mercado que permita una eficiencia asignativa.

La regla general de fijación de precios que se deriva de tal situación en un marco de información completa por parte del regulador, óptimo de primer orden, es inmediata, surgiendo de la simple maximización del bienestar total, considerado como la suma del excedente de consumidores y productor, es decir:

$$\begin{aligned} \text{Max } E_t &= \int_0^q p(q) dq - p(q)q + p(q)q - C(q) & [1] \\ \frac{\delta E_t}{\delta q} &= p(q) - C_m(q) = 0; p(q) = C_m(q) \end{aligned}$$

siendo E_t : excedente total; $p(q)$: función de demanda del mercado; $C(q)$: función de costes de la empresa; $C_m(q)$: coste marginal.

Como es evidente, la regla resultante a imponer por el regulador con objeto de maximizar el bienestar total, o primer óptimo, es la igualación entre precio y coste marginal. Ahora bien, en una situación de rendimientos crecientes de escala en el punto considerado, o, en términos más generales, siempre que se

cumplan algunas de las condiciones suficientes para la existencia de un monopolio natural mencionadas en el apartado anterior, la aplicación de dicha regla implicaría de forma inmediata la existencia de pérdidas por parte de la empresa. Por otra parte, en otro tipo de situaciones la misma regla puede dar lugar a la existencia de beneficios extraordinarios, restringiendo la producción respecto a la que podría surgir de una situación de igualdad entre el precio y el coste medio.

Ante una situación semejante caben dos posibles opciones: 1) La articulación por parte del regulador de algún método de subsidiación que permita cubrir las posibles pérdidas experimentadas manteniéndose un precio igual al coste marginal; y 2) la imposición de una restricción de igualdad entre rentas y costes de la empresa, asumiendo las posibles distorsiones asignativas a las que esto puede dar lugar.

La primera de las soluciones apuntadas implica la existencia de un trasvase de rentas del resto de la economía a los consumidores del mercado considerado, lo que permite el mantenimiento de un óptimo de primer orden desde un punto de vista asignativo; mientras que la segunda plantea una situación factible de equilibrio, en la que la empresa cubre sus costes a cambio de una cierta distorsión en la asignación de recursos.

De cualquier forma, en la primera de las soluciones señaladas se está empleando el supuesto implícito de que el trasvase de recursos desde el resto de la economía al sector implicado no genera ningún tipo de coste. En caso de computar un precio o coste sombra derivado de las distorsiones producidas por el procedimiento específico para articular tal trasvase, la solución se ve alterada, introduciéndose una desviación respecto a una situación de óptimo de primer orden cuya entidad dependerá de la amplitud del precio sombra supuesto. Así, si replanteamos la solución general [1] suponiendo un precio sombra de las subvenciones igual a λ , tendremos:

$$\text{Max } E_t \int_p^{\infty} q(p) dp + q(p)p - C(q) - \lambda [C(q) - q(p)p]$$

cuya condición de primer orden es igual a:

$$\frac{\delta E_t}{\delta p} = -q(p) + (1+\lambda) \left(q(p) + p \frac{\delta q}{\delta p} - \frac{\delta c}{\delta q} \frac{\delta q}{\delta p} \right)$$

$$\frac{p-C_m}{p} = \frac{\lambda}{1+\lambda} \frac{1}{\epsilon} \quad [2]$$

siendo ϵ : elasticidad de la demanda de una referencia.

Esta solución supone la obtención de un óptimo de segundo orden y es, por lo tanto, similar en su estructura lógica a la derivada de la segunda de las opciones anteriormente indicadas, es decir la introducción en [1] de una restricción adicional que asegure que los ingresos obtenidos cubran los costes de producción. Las implicaciones de tal proceder en el caso uniproducto son evidentes, dando lugar a la igualdad entre el precio y el coste medio de producción, lo que constituye la referencia básica de los procedimientos regulatorios clásicos que siguen el principio general de fijación de precios al coste del servicio (*cost of service regulation*).

En el caso multiproducto, la imposición de la restricción de igualdad de rentas da lugar a la regla de fijación de precios conocida como Ramsey-Boiteaux, que surge de la aplicación de los principios generales de teoría impositiva derivados por el primer autor al marco regulatorio. En dicho caso, la empresa tiene un mayor margen de maniobra para la igualdad entre ingresos y rentas ya que cada uno de los precios de los productos que ofrece constituye una potencial variable de decisión⁵. Por lo tanto, la decisión de la empresa que se enfrenta a n demandas diferenciadas surgirá de:

$$\text{Max } E_t = \sum_{i=1}^n \int_{p_i}^{\infty} q_i(p_i) dp_i +$$

$$+ \sum_{i=1}^n p_i q_i(p_i) - C(q_1, \dots, q_n)$$

⁵ Este esquema resultaría igualmente aplicable al caso de un producto homogéneo y diversos mercados diferenciados en los que se pueda efectuar una discriminación de precios.

$$s.t. \sum_{i=1}^n p_i q_i(p_i) = C(q_1, \dots, q_n)$$

cuya condición de primer orden será:

$$-q_i(p_i) + (1+\lambda) \left(q_i(p_i) + p_i \frac{\delta q_i}{\delta p_i} - \frac{\delta C}{\delta q_i} \frac{\delta q_i}{\delta p_i} \right) = 0$$

que, una vez reordenada, da lugar a la conocida expresión:

$$\frac{p_i - Cm_i}{p_i} = \frac{\lambda}{1+\lambda} \frac{1}{\epsilon_i} \quad [3]$$

resultando la conocida regla que establece una divergencia entre el precio y el coste marginal inversamente proporcional a la elasticidad de la demanda del mercado del producto del que se trate.

En el supuesto de que las funciones de demanda involucradas fueran lineales la solución derivada implica una misma reducción relativa de la producción de cada mercado respecto a una situación de primer óptimo, en la que el precio fuera igual al coste marginal. En otras situaciones, la reducción únicamente será aproximadamente igual para cada uno de los bienes.

Si, por otra parte, las funciones de demanda de los distintos productos no son independientes, la divergencia entre precio y coste marginal debe tener en cuenta la elasticidad cruzada entre los mismos, de tal forma que la solución [3] se transformaría, convirtiéndose la relación entre el precio de los distintos productos en:

$$\frac{p_i - Cm_i}{p_i} (\epsilon_i - \epsilon_{ij}) = \frac{p_j - Cm_j}{p_j} (\epsilon_j - \epsilon_{ij})$$

siendo ϵ_{ij} : elasticidad de demanda cruzada entre los productos «i» y «j».

En conclusión, de lo anterior se deduce que la aplicación de la solución general o de primer óptimo a una situación de monopolio natural con información completa y sin problemas de incentivos puede verse restringida por la conformación específica de

las condiciones de costes y de demanda del mismo, lo que puede dar lugar a una situación de pérdidas o beneficios extraordinarios. Si el regulador considera que tal situación no es posible ni deseable, o si se computa un coste de los fondos públicos transmitidos a través de cualquier tipo de subvención, el resultado final se separará de una situación de óptimo de primer orden, dando lugar a las diferentes situaciones específicas analizadas como óptimos de segundo orden.

Los distintos mecanismos reguladores propuestos en la literatura intentan ofrecer pautas de comportamiento para el regulador que permitan aproximarse a algunas de las soluciones consideradas, difiriendo tanto en el tipo de pautas propuestas para poner en práctica la regulación como en el carácter de la solución alcanzable a través de ellas. En los siguientes apartados consideraremos algunos de estos mecanismos, restringiéndonos a lo que se ha dado en llamar como regulación clásica, es decir, aquella que no se deduce de un marco general en el que se considera explícitamente la existencia de asimetrías de información entre regulador y regulado.

IV. Regulación al coste del servicio

En general, la denominada regulación al coste del servicio engloba al tipo de mecanismos más empleados en la práctica reguladora hasta muy recientemente. El objetivo propuesto explícita o implícitamente en dicho tipo de regulación consiste en asegurar la viabilidad y estabilidad de la actividad de que se trate, proporcionándole un rendimiento adecuado que le permita cubrir sus costes y financiar las inversiones necesarias para mantener una situación de eficiencia tecnológica dinámica, y, todo ello, limitando los posibles beneficios extraordinarios derivados de su poder de mercado.

El ejemplo arquetípico de este tipo de regulación es el control de la tasa de rendimiento (*rate of return regulation*), cuyo comportamiento como mecanismo regulador fue analizado en el clásico artículo de Averch, H. y Johnson, L. (1962). La importancia de dicho trabajo es a veces subestimada en la medida en que, a pesar de todas las limitaciones que posteriores autores han

subrayado convenientemente, constituye un primer intento de considerar las consecuencias de un proceso regulador específico, derivando un marco formal y explícito de análisis aplicable a otras situaciones y del que se derivan hipótesis susceptibles de contrastación empírica.

La situación de partida se define como la de un productor que atiende a un mercado en una situación de monopolio natural, empleando únicamente dos tipos de inputs: variables (x_1) y fijos (x_2); y enfrentándose a una regulación que limita el rendimiento máximo que puede obtener en términos de su inversión en inputs fijos (r^*). La situación planteada es, por lo tanto, similar al problema de maximización condicionada considerando en el apartado anterior, variando únicamente la restricción específica con la que se enfrenta el productor. De hecho, la característica distintiva de este tipo de regulación, así como sus efectos distorsionados, se derivan del tratamiento desigual que dicha restricción otorga a los dos tipos de inputs empleados en la empresa. Esto resulta evidente si se procede a una simple reformulación de la restricción impuesta a la empresa. Según dicha condición:

$$r^* \geq \frac{p q(x_1, x_2) - w_1 x_1}{x_2} = \frac{\pi - w_2 x_2}{x_2} \quad [4]$$

$$\pi \leq (r^* - w_2) x_2$$

siendo π : beneficio de la empresa; w_1 : retribución del input variable; w_2 : retribución del input fijo.

Es decir, la restricción impuesta es equivalente a imponer una cota superior a los beneficios de la empresa que implica un trato asimétrico de los inputs empleados, al formularse únicamente en función de uno de ellos: el fijo.

Tomando la igualdad en términos estrictos, el problema de maximización que se le plantea a la empresa a la hora de determinar sus precios será:

$$\begin{aligned} & \text{Max } p q(p) - w_1 x_1 - w_2 x_2 \\ \text{s.t. } & p q(p) - w_1 x_1 - w_2 x_2 - (r^* - w_2) x_2 = 0 \end{aligned}$$

Obteniéndose la condición de primer orden:

$$(1-\lambda) \left(q(p) + p \frac{\delta q}{\delta p} - w_1 \frac{\delta x_1}{\delta q} \frac{\delta q}{\delta p} - w_2 \frac{\delta x_2}{\delta q} \frac{\delta q}{\delta p} \right) =$$

$$- \lambda (r^* - w_2) \frac{\delta x_2}{\delta q} \frac{\delta q}{\delta p}$$

siendo λ : multiplicador de Lagrange.

Reordenando la anterior expresión se puede llegar a un resultado similar en sus líneas generales al deducido en las expresiones [2] y [3] del anterior apartado, donde se indica la divergencia óptima entre precio y coste marginal derivada del cumplimiento de la restricción impuesta:

$$p - \left[w_1 \frac{\delta x_1}{\delta q} + \left(w_2 - \frac{\lambda}{1-\lambda} (r^* - w_2) \right) \frac{\delta x_2}{\delta q} \right] \frac{\delta q}{\delta p} = -q(p) \quad [5]$$

El resultado en este caso tiene una interpretación muy interesante, ya que plantea una situación similar al equilibrio de un monopolista que se enfrenta a un coste marginal en el que el precio del capital aparece distorsionado debido a las condiciones de regulación⁶. En el caso en que la tasa máxima de rendimiento (r^*) sea superior al coste del factor fijo (w_2), el coste implícito del uso de dicho factor para la empresa disminuirá en la medida en que un mayor empleo del mismo permite la obtención de mayores niveles de beneficios. El resultado final será una mayor producción de la correspondiente a una situación no restringida de monopolio junto a una distorsión en el uso relativo de factores, empleando en mayor medida de lo que sería eficiente el input fijo sobre el que se establece la restricción de beneficios.

De cualquier forma, el efecto de la restricción impuesta sobre la utilización relativa de factores resulta más evidente en la formulación tradicional del modelo referida a la demanda derivada de inputs. Así, si replanteamos el problema de maximización planteado en [4] como:

⁶ Véase: BÖS, D. (1986), capítulo 9.

$$\begin{aligned} \text{Max } & p q(x_1, x_2) - x_1 w_1 - x_2 w_2 \\ \text{s.t } & q(x_1, x_2) = x_1 w_1 - r^* w_2 \end{aligned}$$

en el que a partir de las condiciones de primer orden referidas a los inputs x_1 y x_2 :

$$\begin{aligned} (1-\lambda) \text{Im}(q) \frac{\delta q}{\delta x_1} - (1-\lambda) w_1 &= 0 \\ (1-\lambda) \text{Im}(q) \frac{\delta q}{\delta x_2} - w_2 + \lambda r^* &= 0 \end{aligned} \quad [6]$$

podemos obtener la restricción final:

$$\frac{Pm_1}{Pm_2} = \frac{w_2}{w_1} - \frac{\lambda}{(1-\lambda)} \frac{(r^* - w_2)}{w_1} < \frac{w_2}{w_1} \quad [7]$$

siendo $\text{Im}(q)$: función de ingreso marginal; Pm ; productividad marginal.

El resultado indica, por tanto, que si la tasa establecida de rendimiento máximo del input fijo es superior al precio de dicho input, la distorsión de costes introducida por el mecanismo regulador llevará a una mayor utilización relativa de dicho input respecto a lo que correspondería en una situación de eficiencia. Esto es debido a la distorsión resultante de relacionar, tal y como se expresó en la restricción [4], el nivel de beneficio obtenible con la entidad absoluta empleada del factor fijo, de forma que la empresa toma en consideración dicho efecto positivo a la hora de determinar su utilización relativa de factores y emplea una mayor cantidad relativa de dicho input de la que correspondería en una situación de eficiencia.

Ahora bien, lo anteriormente establecido no supone que, en un contexto de información perfecta, la empresa tenga incentivos para una utilización ineficiente, desde un punto de vista tecnológico, del factor fijo sobre el que se basa la regulación, recurriendo, por ejemplo, a inversiones superfluas o innecesarias. Como hemos visto en [5], de las condiciones de maximización se deriva que la empresa actuará como un monopolista restringido igualando su ingreso marginal con el coste marginal distor-

sionado por el procedimiento regulatorio. De cualquier forma, la empresa seguirá teniendo incentivos para actuar en un tramo elástico de la función de demanda donde tenga ingresos marginales positivos. En dichas condiciones, cualquier situación de ineficiencia tecnológica, es decir, la obtención de un menor volumen de producción del máximo factible con los inputs empleados, daría lugar a un descenso de sus beneficios. Por lo tanto, la ineficiencia creada por el procedimiento regulador se refiere al empleo de una utilización relativa de factores inadecuada dado el precio de los mismos, pero no supone una ineficiencia tecnológica derivada de la obtención de un nivel de producto inferior al técnicamente posible.

Del resultado obtenido en [7], puede parecer derivarse que, dado que la distorsión en el empleo relativo de factores surge de la discrepancia entre la tasa de rendimiento permitida respecto al empleo del factor fijo y el precio de dicho factor, una mayor aproximación de dicha tasa al precio del factor fijo debería implicar una mejora en la situación de referencia. Esta idea intuitiva que dio lugar a algún tipo de sugerencias prácticas para el ajuste dinámico de la tasa de rendimiento establecida en el procedimiento regulador es claramente errónea, ya que al mantenerse la discrepancia entre dicha tasa y el precio del factor fijo, disminuyendo, sin embargo, su entidad, la empresa tiene incentivos a distorsionar aún más su utilización relativa de factores con objeto de compensar con una mayor base de referencia (factor fijo) la disminución experimentada. Algebraicamente, dicho resultado resulta inmediato si se diferencia la restricción regulatoria con la que se enfrenta la empresa:

$$p q(x_1, x_2) - w_1 x_1 - r^* x_2 = 0$$

respecto al empleo de los factores (x_1 y x_2) y la entidad de la tasa máxima de rendimiento (r^*), obteniéndose:

$$\left(\text{Im} \frac{\delta q}{\delta x_1} - w_1 \right) dx_1 + \left(\text{Im} \frac{\delta q}{\delta x_2} - r^* \right) dx_2 = x_2 dr^*$$

Teniendo en cuenta las condiciones de optimalidad en la utilización del input variable que se deducen de [6], se obtiene:

$$\frac{dx_2}{dr^*} = \frac{x_2(1-\lambda)}{w_2-r^*} < 0$$

Es decir, en caso de reducir la tasa de rendimiento máximo establecida por el sistema regulador, la entidad de capital empleado tendería a aumentar en vez de reducirse.

Este resultado viene a subrayar la contradicción básica con la que se enfrenta el establecimiento de una tasa de rendimiento máximo como mecanismo regulador. Como se ha comprobado, en el caso de establecer dicha tasa a un nivel superior al del precio de mercado del factor fijo, la empresa tiene incentivos a utilizar dicho factor en mayor medida que la indicada por los precios relativos vigentes. Una eventual disminución de la tasa establecida de rendimiento no mejora dicha situación, tendiendo, por el contrario, a empeorarla. En el caso extremo, si la tasa establecida iguala al precio del factor, la empresa no tendría ningún incentivo especial a desarrollar la actividad de que se trate y, en caso de desarrollarla, sería completamente indiferente ante cualquier combinación de factores y productos, eficiente o ineficiente, que diera lugar a un beneficio nulo. Por último, si la tasa se establece a un nivel inferior al precio de mercado, la empresa no llevará a cabo la actividad.

En último término, el interés del trabajo inicial de Averch, H. y Johnson, L. (1962) radica principalmente en la propuesta metodológica planteada respecto al análisis formal de las consecuencias de un procedimiento concreto de regulación, en su caso la imposición de una tasa de rendimiento máxima sobre la utilización del factor fijo, a partir de sus efectos en las condiciones maximizadoras de la empresa. Dicho esquema básico ha sido, de hecho, el aplicado posteriormente en la consideración de cualquier modelo regulador, complejizándose las condiciones supuestas respecto al entorno y las características de la interacción entre regulador y regulado.

El tipo de resultados obtenidos, por otra parte, tuvieron la virtualidad de generar análisis aplicados dirigidos a contrastar el

cumplimiento de las hipótesis deducidas del modelo. En dicha contrastación se han seguido estrategias muy distintas que van de la simple estimación de productividades marginales relativas y su comparación con el precio de mercado de los factores (Courville, L., 1974), hasta la estimación de precios sombra de los factores a y su comparación con los precios vigentes en el mercado (Atkinson, S. E., y Halvorsen, R., 1986), pasando por la simple comparación de la eficiencia relativa de empresas reguladas y no reguladas (Pescatrice, D. R., y Trapani, J. M., 1980). Gran parte de estos análisis se han referido al sector eléctrico en EEUU.

Una extensión inmediata del tipo de regulación sobre la tasa máxima de rendimiento que elimina sus efectos distorsionadores sobre la utilización relativa de factores, es el emplear como término de referencia para establecer dicha tasa cualquier otra variable operativa de la empresa que no suponga un tratamiento desigual de los factores productivos. Dicha tasa límite se puede establecer en relación a la producción, a las ventas o, incluso, a los costes de la empresa (*cost plus*), dando lugar a diferentes esquemas regulatorios que comparten una misma estructura y resultados semejantes⁷.

Así, por ejemplo, en el caso más general de una tasa de rendimiento máximo sobre el nivel de producción, el problema de maximización al que se enfrenta la empresa será:

$$\begin{aligned} \text{Max} \quad & I(q) - C(q) \\ \text{s.t.} \quad & r^* \geq \frac{I(q) - C(q)}{q} \end{aligned}$$

siendo $I(q)$: función de ingresos; $C(q)$: función de costes

La solución de primer orden a dicho problema es:

$$[Im(q) - Cm(q)] = -\frac{1}{1-\lambda} r^* < 0$$

siendo $Im(q)$: función de ingreso marginal; $Cm(q)$: función de costes marginales.

⁷ Véase: TRAIN, K. T. (1991), capítulo 1.

El mecanismo regulador lleva, por tanto, a que la empresa mantenga una discrepancia entre sus ingresos y costes marginales, lo que supone producir una cantidad superior a la derivada de una situación no restringida de monopolio, y cuya entidad depende de la tasa máxima de rendimiento sobre producto establecida. La distorsión introducida por el procedimiento regulador es independiente, en este caso, de las cantidades empleadas de cada uno de los factores, lo que sugiere una correcta utilización relativa de los mismos, tal y como se puede fácilmente demostrar a través de la obtención de las correspondientes condiciones de primer orden. Finalmente, diferenciando la restricción respecto al nivel de producto y la tasa de regulación se obtiene:

$$(Im(q) - Cm(q) - r^*) dq = q dr^*$$

$$\frac{dq}{dr^*} = \frac{q}{Im(q) - Cm(q) - r^*} < 0$$

Es decir, al igual que en el caso anterior, un descenso de la tasa reguladora lleva a un incremento de la base sobre la que dicha tasa actúa, que, en esta ocasión, corresponde al nivel de producto. De hecho, si el regulador establece una tasa de rendimiento lo suficientemente pequeña, la empresa se aproximará infinitesimalmente a una situación de óptimo de segundo orden en el que el precio fuera igual al coste medio. En cualquier caso, dicha situación ($r^*=0$) no llegará a alcanzarse, ya que supondría, una vez más, que la empresa sería indiferente ante cualquier combinación de factores y producto que proporcionara un beneficio nulo.

V. Precios no lineales

La tarificación a través de un sistema de precios no lineales constituye la segunda de las grandes orientaciones posibles dentro de un marco regulador para posibilitar la obtención de un óptimo de primer o segundo orden en condiciones de monopolio natural. Los recursos necesarios para garantizar la viabilidad de la actividad de que se trate provienen, en dicho caso, de la posibili-

dad que tal mecanismo confiere de transferir al productor parte del excedente de los consumidores, aumentando en dicha forma sus ingresos y permitiéndole una aproximación del precio al coste marginal. Este sistema es muy utilizado de hecho en actividades que se consideran próximas a una situación de monopolio natural, como son los denominados servicios públicos: agua, teléfono, gas, electricidad, etcétera; remontándose su uso al siglo pasado en el que la denominada tarifa Hopkinson, aplicada a las primeras redes de distribución eléctrica, ya distinguía entre energía total suministrada (Kw/h) y demanda máxima instantánea o potencia (Kw), considerándolas como dos productos diferentes.

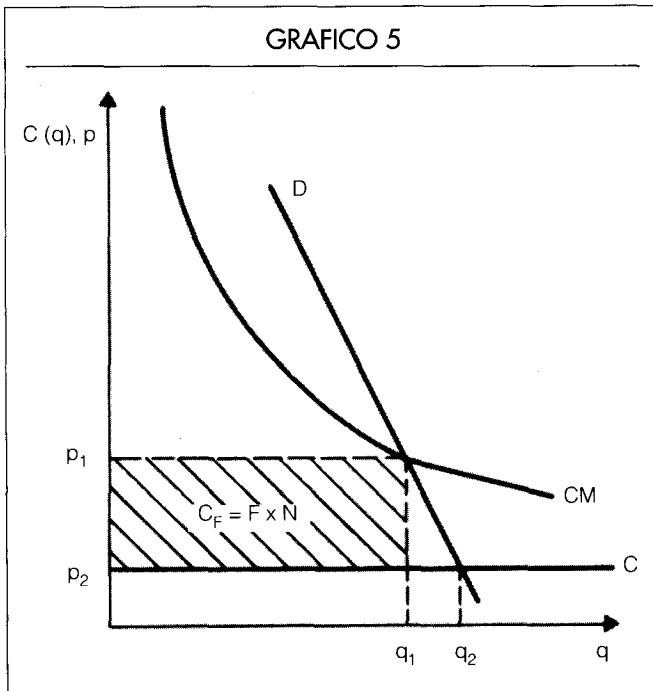
Como es bien sabido, un sistema de precios no lineales constituye en último término un sistema de discriminación de segundo orden, según el cual el precio de un bien varía en función de la cantidad consumida. Si se toma como referencia más general a este respecto un esquema de tarifa en dos partes en el que el coste total de la compra de una determinada cantidad (q) del bien que se trate consta de dos elementos, uno fijo (F), que no depende de la cantidad consumida, y otro variable, proporcional a dicha cantidad, es decir:

$$T(q) = F + p q$$

se pueden diferenciar dos casos característicos según que el número de consumidores esté dado —siendo, por lo tanto, nula su elasticidad respecto tanto a la cuota fija (F) como al precio unitario (p)—o bien dependa de la tarifa establecida.

En el primer caso, el resultado es inmediato, tal y como fue establecido en Coase, R. (1948). La insensibilidad del número de consumidores respecto a la tarifa permite la obtención de una situación de óptimo de primer orden, similar a la derivada de una subsidiación pública sin coste que cubriese las pérdidas derivadas del establecimiento de un precio igual al coste marginal. Así, por ejemplo, si se parte de una función de costes similar a la establecida en el apartado II, en la que el coste total está compuesto de un coste fijo más un coste marginal constante:

$$C(q) = C_f + c \cdot q$$



la tarifa en dos partes que permitiría la obtención de óptimo de primer orden sería:

$$T(q) = F + p \cdot q$$

tal que:

$$F = \frac{C_F}{n}$$

$$c = p$$

siendo n : número de consumidores.

Según puede observarse en el Gráfico 4, este tipo de tarifa sería totalmente equivalente para la empresa al establecimiento de un precio igual al coste medio (p_1, q_1), en la medida en que ambas le proporcionan unos ingresos que cubren sus costes. Sin embargo, desde el punto de vista de la asignación de recursos, la adopción de una tarifa en dos partes (p_2, q_2), hace que su producción aumente respecto a dicha situación, hasta el punto de lograr una situación de primer óptimo igualando el precio unitario al coste marginal.

Por lo que respecta al consumidor individual, ambas situaciones no tienen por qué resultar totalmente equivalentes, favoreciendo, en general, la adopción de una tarifa en dos partes a los consumidores con mayores niveles de consumo, dada la menor incidencia relativa en su gasto total de la introducción de una cuota fija.

De cualquier forma, el supuesto de que el número de consumidores está dado, no viéndose afectado por la entidad de la tarifa, no parece en principio muy realista, en especial cuando se refiere a actividades, como suelen ser los servicios públicos susceptibles de ser consideradas como monopolio natural y que se caracterizan por un elevado nivel de costes fijos en comparación con su coste marginal. En el caso de que la demanda de acceso sea sensible al precio, nos encontramos con la segunda de las situaciones anteriormente mencionadas, siendo necesario incorporar de una forma explícita el efecto de la tarifa sobre el número de consumidores a la hora de establecer un diseño óptimo de la misma.

El problema planteado en dicha situación es muy similar a la derivación de una tarifa óptima en dos partes para un productor individual, sustituyendo la función a maximizar que, en este caso, sería la función de bienestar de los consumidores, restringida a la igualación de rentas por parte del productor que permita el mantenimiento de la actividad. La base de resolución de este problema parte de la consideración explícita de la distribución de preferencias de los consumidores respecto al bien de que se trate, planteando la sustituibilidad entre el componente fijo y variable de la tarifa en los ingresos del productor junto a su efecto diferencial en el número de demandantes que permanecen en el mercado.

Partiendo de una distribución de preferencias de los consumidores [8], normalizada en el entorno ($0 < \theta^* < 1$), desde un mayor a un menor nivel de preferencias y siendo:

$$n = \int_0^1 f(\theta) d\theta$$

el número total de consumidores del mercado, la empresa se enfrentará con una función de demanda correspondiente a cada tipo de consumidores:

$$q^* = q [p, F, y(\theta^*), \theta^*]$$

derivada de una función indirecta de utilidad:

$$v^* = v [p, F, y (\theta^*), \theta^*]$$

donde p : componente variable de la tarifa en dos partes; F : componente fijo; $y (\theta^*)$: renta de los consumidores tipo (θ^*) ; θ^* : preferencias de los consumidores tipo (θ^*) .

Dada esta situación, la tarifa en dos partes óptima desde el punto de vista del consumidor surgirá de maximizar una suma ponderada de la utilidad correspondiente a cada grupo de consumidores, sujeta a la restricción de igualdad de rentas por parte de la empresa suministradora, es decir:

$$\begin{aligned} & \text{Max} \int_0^{\hat{\theta}(p,F)} \omega (\theta) v [p, F, y (\theta), \theta] f(\theta) d(\theta) \\ & \text{s.t. } p \cdot Q + F \cdot n - C(q) = 0 \end{aligned}$$

siendo $\hat{\theta}$: último valor en la escala de preferencias que entra en el mercado dada la tarifa establecida, $\hat{\theta} = \hat{\theta} (p, F)$; Q : cantidad total consumida dado $\hat{\theta}$; $\omega (\theta)$: ponderación otorgada a las utilidades de los consumidores de preferencias θ .

En el caso más probable en el que, dadas las preferencias de los consumidores, la determinación del tipo marginal que entra en el mercado ($\hat{\theta}$), sea sensible respecto tanto de la cuota fija (F) como del precio marginal (p), se dará una divergencia entre dicho precio y el coste marginal ($Cm \neq p$), al contrario de lo que sucedía anteriormente. De cualquier forma, el signo de dicha divergencia no está determinado, dependiendo de la configuración de la función de distribución de las preferencias de los consumidores. En general, puede demostrarse⁸ que existen dos posibilidades dependiendo de la entidad relativa de la cantidad media consumida por el total de consumidores que entran en el mercado (\bar{q}) respecto a la cantidad demandada por el consumidor marginal (\hat{q}):

a) Si $\bar{q} > \hat{q} \rightarrow p > Cm$; $C_p / n > F > 0$

b) Si $\hat{q} > \bar{q} \rightarrow Cm > p$; $F > C_p / n$

Es decir, únicamente resultará conveniente incrementar el precio unitario respecto al coste marginal si la cantidad media demandada es superior a la del consumidor marginal, dado que de esta forma el incremento de renta derivado de los consumidores inframarginales compensará las pérdidas derivadas de disminuir la cuota fija de entrada al mercado. Por el contrario, en el caso en que la demanda del consumidor marginal sea superior a la demanda media, interesará establecer un precio unitario inferior al coste marginal, ya que las pérdidas derivadas de los consumidores inframarginales quedarán compensadas por el aumento de cuota fija.

En el caso de poder disponer de estimaciones independientes respecto a la demanda de acceso al mercado y la demanda de uso del bien, es posible plantear el anterior problema de una forma más inmediata considerando cada uno de estos aspectos como un bien independiente con determinadas elasticidades cruzadas. Así, si definimos (q_a) como el número de demandantes que acceden al mercado y (q_u) la cantidad total del bien consumida por todos los demandantes del mercado, la fijación de precios para ambos aspectos puede plantearse como un problema típico de precios Ramsey, resultante de la maximización del excedente total del consumidor teniendo en cuenta la igualdad de rentas y costes del productor, así como la relación existente entre ambos tipos de demanda.

El resultado será similar al indicado en el apartado III para más de un bien, indicando una divergencia entre el precio y el coste marginal que dependerá de la elasticidad propia de demanda y de la elasticidad cruzada, es decir:

$$\frac{p_a - Cm_a}{p_a} (\epsilon_a - \epsilon_{ua}) = \frac{p_u - Cm_u}{p_u} (\epsilon_u - \epsilon_{au})$$

siendo p_a = precio de acceso, es decir, cuota fija de la tarifa en dos partes; p_u = precio de uso, es decir, precio unitario del bien; ϵ_a : elasticidad respecto al propio precio; ϵ_{ua} : elasticidad cruzada.

En el caso en que la demanda de acceso estuviera dada siendo, por lo tanto, el número de consumidores y la demanda de

⁸ Véase BERG, S. B., y TSCHIRHART, J. (1988), capítulo 4.

uso insensibles a la cuota fija ($\varepsilon_a = 0$; $\varepsilon_{au} = 0$; $\varepsilon_{ua} = 0$), el resultado sería:

$$\frac{p_u - Cm_u}{p_u} \varepsilon_u = 0$$

lo que implica en último término la solución Coase anteriormente mencionada con un precio unitario (p_u) igual al coste marginal y una cuota fija (p_a) determinada por el resultado de repartir el coste fijo entre los consumidores existentes en el mercado.

Por otra parte, si la demanda de acceso no está dada, siendo, por lo tanto, sensible al esquema tarifario establecido, el resultado más habitual es el que corresponde con la solución *a*) anteriormente mencionada en este mismo apartado, caracterizado por un precio de acceso inferior al del caso Coase ($p_a < C_F/n$) financiado por un precio de uso superior a su coste marginal. De cualquier forma, también sería posible obtener una solución similar a la del caso *b*), resultante de una situación en la que la elasticidad cruzada entre acceso y uso supere a la elasticidad directa de acceso pero no a la de uso.

VI. Resumen y consideraciones finales

La revisión efectuada ha tratado de presentar un esquema resumido de lo que puede considerarse como la teoría clásica de la regulación de monopolios naturales.

Para ello se ha partido, en primer lugar, de las propias condiciones de existencia de una situación de monopolio natural, aspecto que tiende a obviarse o simplificarse en muchas de las presentaciones de este tema, pero que adquiere una singular importancia en un marco multiproducto, como es el que normalmente caracteriza a las actividades susceptibles de ser consideradas como monopolios naturales. De la resumida presentación efectuada se deduce la dificultad de delimitar el conjunto de condiciones necesarias de un monopolio natural, así como la complejidad que adquieren las condiciones suficientes de existencia del mismo con una tecnología multiproducto. Una posible conclusión sería la necesidad de realizar análisis detallados de

las bases tecnológicas existentes en cada caso particular, algo que la cantidad y calidad de los datos disponibles raramente permiten. De cualquier forma, de los intentos realizados en los últimos años a este respecto se deduce una necesaria revisión de algunos de los lugares comunes normalmente empleados en este tema, cuestionando la consideración total o parcial de monopolio natural de actividades normalmente consideradas como tales: telecomunicaciones, transporte aéreo, transporte ferroviario, sector eléctrico, etcétera.

Una vez planteada la existencia de una situación de monopolio natural, el resto del trabajo se ha limitado a revisar las dos soluciones básicas posibles en dicho marco: óptimos de primer y segundo orden; así como las características y actuación de los dos tipos básicos de mecanismos reguladores empleados para su obtención: regulación al coste del servicio y precios no lineales.

Como exponente de la actuación de ambos mecanismos reguladores se han empleado dos de sus formulaciones arquetípicas, como son el control de la tasa de rendimiento y el diseño de un sistema tarifario en dos partes. En ambos casos la exposición se ha tratado de centrar en aquellos aspectos que resultan reveladores del tipo de actuación de los mecanismos que representan, intentando ejemplificarlos a través de un caso específico. Para ello se ha insistido en sus principios básicos susceptibles de ser aplicados en diversas variantes.

Al restringirse la exposición a lo que normalmente se considera como teorías clásicas de regulación, el supuesto implícito empleado a lo largo de la misma es la ausencia de problemas de información para el regulador sobre las condiciones tecnológicas y de demanda con las que se enfrenta el regulado. El incumplimiento de dicho supuesto ha resultado evidente en la práctica reguladora, dando lugar a intervenciones «*ad hoc*» con objeto de paliar sus efectos. No ha sido, sin embargo, hasta muy recientemente cuando ha empezado a desarrollarse una teoría específica cuyo punto de partida ha sido la incorporación explícita de la existencia de asimetrías de información entre regulador y regulado a la hora de determinar el diseño óptimo de los mecanismos reguladores. Dicho diseño, en cual-

quier caso, parte de los principios generales establecidos en la teoría clásica, incorporando en ellos un esquema de incentivos que contribuya a una revelación por parte de la empresa de sus verdaderas condiciones de demanda y costes⁹. Desde este punto de vista pueden considerarse como un desarrollo lógico de los mismos, incorporando los nuevos elementos posibilitados por los recientes desarrollos en la teoría de la información.

Referencias

- ATKINSON, S. E., y HALVORSEN, R. (1980): «A Test of Relative and Absolute Price Efficiency in Regulated Industries», *Review of Economics and Statistics*, 62.
- AVERCH, H., y JOHNSON, L. (1962): «Behavior of the Firm Under Regulatory Constraint», *American Economic Review*, 52.
- BAUMOL, W. J. (1977): «On the Proper Cost Test for Natural Monopoly in a Multiproduct Industry», *American Economic Review*, 67.
- BAUMOL, W. J.; BAILEY, E., y WILLIG, R. (1977): «Weak Invisible Hand Theorems on the Sustainability of Multiproduct Natural Monopoly», *American Economic Review*, 67.
- BERG, S. V., y TSCHIRHART, J. (1988): *Natural Monopoly Regulation*, Cambridge University Press.
- BÖSS, D. (1986): *Public Enterprise Economics: Theory and Application*, North Holland.
- COURVILLE, L. (1974): «Regulation and Efficiency in the Electric Utility Industry», *Bell Journal*, 5.
- EVANS, D. S., y HECKMAN, J. J. (1983): «Multiproduct Cost Function Estimates and Natural Monopoly Test for the Bell System», en EVANS, D. S. (ed.), *Breaking Up Bell*. North Holland.
- LAFFONT, J. J., y TIROLE, J. (1993): *A Theory of Incentives in Procurement and Regulation*. The MIT Press.
- PANZAR, J., y WILLIG, R. (1981): «Economies of Scope», *American Economic Review*, 71.
- PESCATRICE, D. R., y TRAPANI, J. M. III (1980): «The Performance and Objectives of Public and Private Utilities Operating in United States», *Journal of Public Economics*, 13.
- SHARKEY, W. W. (1982): *The Theory of Natural Monopoly*. Cambridge University Press.
- TRAIN, K. T. (1991): *Optimal Regulation*, The MIT Press.

⁹ Véase LAFFONT, J. J., y TIROLE, J. (1993).