A15. La empresa: el tamaño eficiente y sus límites. Mención especial de la economía de los costes de transacción. La Teoría de la Organización Industrial: barreras a la entrada y mercados impugnables

1. Teoría de la empresa

La **teoría de la empresa** estudia cuatro ámbitos: su existencia y sus objetivos, por qué nacen las empresas y qué persiguen; sus límites, qué determina su tamaño; su organización interna, y su comportamiento en los mercados.

1.1. Existencia y objetivos

1.1.1. Enfoques tradicionales

Maximización de beneficios: la empresa existe como unidad de decisión que combina factores para maximizar la diferencia entre ingresos y costes. Esta visión es operativa y potente para la teoría de mercados, pero supone una unidad monolítica, sin conflictos internos, y con toda la información necesaria para optimizar.

Escuela austriaca: para Menger, Hayek y Kirzner, el orden económico es un descubrimiento incesante y no un equilibrio en reposo. El empresario no optimiza mecánicamente, sino que está *alerta* a las ocasiones de beneficio, y un beneficio extraordinario indica una oportunidad aún sin aprovechar, no un monopolio. Hayek (1945) subrayó, además, que el sistema de precios es un mecanismo descentralizado para transmitir información dispersa. Competir es, para ellos, un proceso que transcurre en el tiempo, no un estado.

1.1.2. Enfoque directivista o gerencial

Origen: desde Berle y Means (1932), la propiedad, de los accionistas, está separada de la gestión, de los directivos. Si estos tienen discrecionalidad, la empresa puede perseguir objetivos distintos del beneficio.

Baumol (1959): los directivos maximizan los ingresos por ventas, de los que dependen su remuneración y prestigio, sujetos a un beneficio mínimo que satisfaga a los accionistas:

$$\max_x \ \mathrm{IT}(x) = p(x)\,x \quad \text{s.a.:} \quad \pi(x) \geq \pi_{\min}$$

El lagrangiano es $\mathcal{L} = p(x)\,x + \lambda\,[\pi(x) - \pi_{\min}]$, y las condiciones de Kuhn-Tucker, necesarias y suficientes con ingreso total y beneficio cóncavos, son:

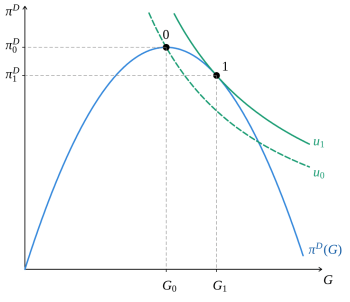
$$\mathrm{Img}(x) + \lambda\,\pi'(x) = 0$$
$$\lambda \geq 0, \qquad \lambda\,[\pi(x) - \pi_{\min}] = 0$$

Con la restricción activa, $\mathrm{Img}(x_B) = -\lambda\,\pi'(x_B) > 0$, de modo que la empresa opera a la derecha del máximo de beneficios, con $x_\pi < x_B < x_{\mathrm{IT}}$ y $p_B < p_\pi$. Produce más y vende más barato que el maximizador de beneficios, con un beneficio igual a $\pi_{\min}$. Si la restricción no es activa, $\lambda = 0$, produce donde $\mathrm{Img}(x_{\mathrm{IT}}) = 0$, con elasticidad unitaria, lo que solo es factible si ese beneficio cubre el mínimo. Cuando $\pi_{\min}$ tiende al beneficio máximo, el modelo converge al tradicional.

Estática comparativa: un impuesto de suma fija, o un aumento de los costes fijos, no altera la producción ni el precio del maximizador de beneficios, pero reduce la producción y sube el precio del maximizador de ventas. Que las empresas trasladen a precios los costes fijos fue el principal apoyo empírico de Baumol. La razón es que el impuesto desplaza hacia abajo la curva de beneficios, cuyo corte derecho con $\pi_{\min}$ se mueve a la izquierda.

Otras hipótesis: sustituyen las ventas por otros objetivos de los directivos.

- *Utilidad gerencial (Williamson, 1964)*: los directivos maximizan una utilidad que depende del gasto en plantilla, G, y de los beneficios discrecionales, π^D, sujetos a un beneficio mínimo. Eligen la tangencia de la frontera de beneficios con su curva de indiferencia más alta, con un gasto G_1 mayor que el que maximiza el beneficio, G_0 (*preferencia por el gasto*). La frontera crece mientras el gasto es productivo y decrece desde su máximo, el punto 0 que preferirían los propietarios, y en el punto 1 el directivo acepta un beneficio π_1^D menor.



Williamson (1964): el directivo elige un gasto en plantilla G_1 superior al que maximiza el beneficio, G_0.

- *Crecimiento (Marris, 1964)*: maximizan la tasa de crecimiento de la empresa sujetos a una restricción de valoración bursátil. Si la acción cae demasiado, la amenaza de una adquisición hostil disciplina a los directivos, y el mercado de control corporativo actúa así como mecanismo disciplinante externo.
- *Ineficiencia-X (Leibenstein, 1966)*: sin presión competitiva y con información asimétrica, nadie se esfuerza al máximo y la empresa opera por debajo de su frontera técnica, aunque se critica que ello puede ser el óptimo, dado el coste del control.

Conclusiones: Baumol deriva por primera vez la conducta de la empresa de un objetivo distinto del beneficio, con una predicción contrastable, y el enfoque anticipa la teoría de la agencia. Pero los objetivos se postulan en lugar de derivarse de los incentivos, el beneficio mínimo es exógeno y la competencia tiende a expulsar a quien no maximiza el beneficio. Por eso la discrecionalidad gerencial solo se sostiene con poder de mercado y barreras a la entrada.

En suma, estos enfoques son complementarios y explican qué persigue la empresa, pero no por qué ciertas transacciones se organizan dentro de ella y no en el mercado.

1.2. Límites

1.2.1. Enfoque neoclásico: la caja negra

El enfoque neoclásico atiende a la tecnología, los costes y el tamaño del mercado. El tamaño de la empresa lo fija el mínimo del coste medio a largo plazo, en forma de U. La dimensión óptima es única, y la empresa es, en lo demás, una caja negra tecnológica.

1.2.2. Enfoque tecnológico de la TOI

Horquilla de capacidad (Stigler, 1939): las empresas mantienen una reserva de capacidad que les da flexibilidad ante la incertidumbre. La evidencia apunta a un coste medio a largo plazo en forma de L, es decir, alcanzada la escala mínima eficiente (EME), es casi constante. No hay dimensión óptima única, lo que explica la coexistencia de tamaños muy distintos.

Subaditividad (Baumol, Panzar y Willig, 1982): una sola empresa produce el *output* total a menor coste que cualquier reparto entre varias, definición moderna de **monopolio natural**. En el caso uniproducto, las economías de escala son condición suficiente pero no necesaria. El límite del tamaño eficiente lo fija la subaditividad, no la pendiente del coste medio. En el caso multiproducto, con vectores de producción $\mathbf{x}_i$, la subaditividad exige:

$$C\left(\sum_i \mathbf{x}_i\right) < \sum_i C(\mathbf{x}_i)$$

Economías de alcance (Panzar y Willig, 1981): producir varios bienes en una misma empresa cuesta menos que hacerlo por separado, por los *inputs* públicos internos, como la marca o el conocimiento, las indivisibilidades y las estructuras comunes. Explican la empresa multiproducto y fijan el límite horizontal del tamaño.

1.2.3. La empresa como estructura de contratos

Pregunta: Coase (*The Nature of the Firm*, 1937; Nobel en 1991) planteó por qué existen las empresas si el mecanismo de precios asigna eficientemente los recursos y, a la inversa, por qué una sola no organiza toda la producción.

Idea: usar el mercado tiene **costes de transacción**: búsqueda de contrapartes, negociación, supervisión y previsión de contingencias. La empresa existe porque, para ciertas actividades, organizar internamente la producción cuesta menos que recurrir al mercado.

Razonamiento: Williamson (1975; Nobel en 2009) distingue costes *ex ante*, de previsión, redacción y negociación de los contratos, y *ex post*, de regateo, ejecución y litigios. El coste de acudir al mercado decrece con la dimensión y el de organización interna crece por las deseconomías gerenciales. La dimensión óptima minimiza su suma, y en ella ambos costes marginales se igualan:

$$\min_x \ \mathrm{CT}(x) = \mathrm{CE}(x) + \mathrm{CI}(x) \quad \Rightarrow \quad \mathrm{CI}'(x^*) = -\mathrm{CE}'(x^*)$$

Desarrollo gráfico: el coste total es la suma vertical de ambas curvas y tiene forma de U, y la dimensión óptima, x^*, corresponde a su mínimo, C^*. Con curvas simétricas, ese punto está sobre el corte de $\mathrm{CE}(x)$ y $\mathrm{CI}(x)$, donde $\mathrm{CE}(x^*) = \mathrm{CI}(x^*)$.

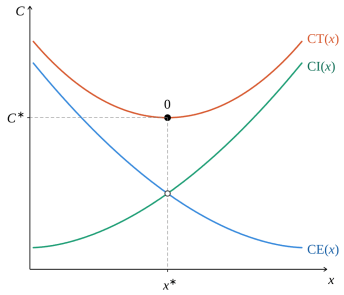
Conclusiones: sitúa los costes de usar el mercado en el centro del análisis. Pero el concepto es amplio y poco preciso, lo que permite explicar *a posteriori* cualquier frontera, y no aclara qué permite a la organización interna evitar esos costes.

En suma, el enfoque tecnológico y el contractual convergen en que la empresa minimiza el coste total de producir e intercambiar, pero dejan abierta su organización interna.

1.3. Organización

1.3.1. Enfoque de derechos de propiedad

Idea: la propiedad no es el bien, sino un conjunto de derechos sobre él, de uso, de apropiación de sus rendimientos y de transferencia, y quién



Coase (1937): la dimensión óptima, x^*, minimiza la suma del coste de mercado y del de organización interna.

los tiene determina los incentivos.

Primera generación (Alchian, 1965; Demsetz, 1967; Alchian y Demsetz, 1972): los derechos aparecen cuando la ganancia de internalizar una externalidad supera el coste de establecerlos. En la producción en equipo la contribución individual no se observa por separado y surge el escaqueo (*shirking*). La solución es un supervisor que sea **perceptor residual** del beneficio. El empresario capitalista emerge así como respuesta eficiente a un problema de medición. Para Alchian, la conducta depende de la estructura de los derechos, y en la empresa pública, cuya propiedad no puede venderse, quien decide no carga con las ganancias ni con las pérdidas de sus decisiones.

Nuevo enfoque (Grossman y Hart, 1986; Hart y Moore, 1990): con contratos incompletos, la propiedad asigna los **derechos residuales de control**, sobre todo lo que el contrato no especifica. Las inversiones específicas se hacen antes de negociar, y cada parte solo se apropia de una parte de su rendimiento, de modo que invierte menos de lo eficiente (*hold-up*). La propiedad óptima va a la parte cuya inversión específica es más importante y reúne los activos complementarios. Los costes de integrar nacen del mismo mecanismo que sus beneficios. La propiedad de un activo refuerza la posición negociadora de su titular y sus incentivos a invertir, pero debilita los de la otra parte, que por eso invierte menos.

Conclusiones: da una razón de eficiencia a la propiedad privada y precisa los costes de la integración vertical. Pero la primera generación es verbal, y los fundamentos de los contratos incompletos los discuten Maskin y Tirole (1999).

1.3.2. Enfoque transaccional

Idea: en el marco de Williamson, la estructura de gobernanza óptima de cada transacción, mercado, formas híbridas o jerarquía, depende de tres atributos.

1. *Frecuencia*: las transacciones recurrentes justifican amortizar estructuras especializadas, y las ocasionales se resuelven mejor en el mercado.
2. *Incertidumbre*: deja más contingencias sin prever y encarece las renegociaciones, lo que, con activos específicos, favorece la jerarquía.
3. *Especificidad de los activos*: las inversiones no reutilizables fuera de la relación generan cuasirrentas apropiables y exponen al *hold-up*, lo que favorece la integración vertical. La cuasirrenta apropiable es el exceso del valor del activo en su uso actual sobre el de su mejor uso alternativo (Klein et al., 1978).

Razonamiento: la elección entre mercado y jerarquía se convierte en un problema de minimización de costes de transacción, que explica la integración vertical, las relaciones a largo plazo y las formas híbridas, como las franquicias o las alianzas.

Conclusiones: hace operativa la intuición de Coase y da predicciones contrastables sobre la integración vertical. Pero es en gran medida verbal, y la especificidad de los activos es difícil de medir. La racionalidad limitada y el oportunismo se introducen como supuestos sin formalizar.

1.3.3. Enfoque de contratos completos

Idea: aplica la teoría de la agencia (Ross, 1973; Jensen y Meckling, 1976) a la relación entre un principal, el propietario, y un agente, el directivo o trabajador, con contratos que contemplan todas las contingencias verificables.

Planteamiento: el principal, neutral al riesgo, elige el esfuerzo e y el contrato salarial $w(x)$ que maximizan su beneficio esperado, sujeto a la restricción de participación del agente, con unas probabilidades $p_i(e)$ que el esfuerzo desplaza hacia los resultados altos:

$$\max_{e,\, w(\cdot)} \ \mathbb{E}[\pi] = \sum_i p_i(e)\ [x_i - w(x_i)]$$

CPO: el producto marginal esperado del esfuerzo iguala la compensación que el agente exige por su desutilidad marginal.

CSO: con resultado esperado cóncavo en el esfuerzo y utilidad del agente cóncava, el programa es cóncavo y la CPO es necesaria y suficiente.

Solución: si el esfuerzo es observable, se alcanza el *first best*, en el que el principal exige el esfuerzo eficiente y paga al agente averso al riesgo un salario fijo.

Conclusiones: si el esfuerzo no es observable, el pago debe depender del resultado y se sacrifica aseguramiento, lo que da un *second best*. Los **costes de agencia** suman la supervisión, la garantía y la pérdida residual, que el gobierno corporativo reduce sin eliminar.

1.4. Comportamiento

1.4.1. Modelo de variaciones conjeturales

Idea: los modelos de oligopolio difieren en lo que cada empresa espera que haga la rival cuando altera su producción. Recogida esa expectativa en la **variación conjetural** (Bowley, 1924; Frisch, 1933), $v_i \equiv \partial x_j/\partial x_i$, todos son casos de una misma CPO.

Planteamiento: cada empresa elige su cantidad tomando la conjetura como dada:

$$\max_{x_i} \ \pi_i = p(x_i + x_j)\cdot x_i - \mathrm{CT}_i(x_i)$$

CPO: como $dX/dx_i = 1 + v_i$, la empresa iguala su coste marginal a un ingreso marginal percibido o, dividiendo por el precio, con la cuota s_i y la elasticidad de la demanda, fija un margen de Lerner que depende de la conjetura:

$$p(X) + p'(X)\,(1+v_i)\,x_i = \mathrm{CT}_i'(x_i)$$
$$\iff \frac{p - \mathrm{CT}_i'}{p} = \frac{(1+v_i)\,s_i}{|\varepsilon_{Xp}|}$$

La empresa anticipa que el precio caerá por su propia expansión y por la reacción que conjetura en la rival, de modo que, cuanto mayor es la conjetura, más restringe la producción y mayor es el margen. El valor de v_i selecciona el modelo:

- Con $v_i = -1$, la rival compensa la variación (competencia perfecta o Bertrand, con margen nulo).
- Con $v_i = 0$, la rival no reacciona (Cournot).
- Con $v_i = \varphi_j'(x_i)$, la rival sigue su función de reacción (líder de Stackelberg).
- Con $v_i = 1$, la rival replica la variación (cártel).

Con n empresas simétricas, el margen es $[1+(n-1)v]/(n\,|\varepsilon_{Xp}|)$: crece con la conjetura y decrece con la elasticidad y, salvo en el cártel, con el número de empresas. El **parámetro de conducta**, $\theta \equiv 1+(n-1)v \in [0,n]$, es el que estima la NEIO.

Conclusiones: ofrece una lectura común de los modelos estáticos de oligopolio y un índice continuo del poder de mercado. Pero carece de fundamento estratégico, porque es pseudodinámico y sus conjeturas son *ad hoc*. Con conjeturas consistentes (Bresnahan, 1981), demanda lineal y coste marginal constante, solo cabe $v = -1$, y la teoría de juegos lo sustituyó. Con $p = a - X$ y coste marginal c, la rival que conjetura v responde con $x_j = (a-c-x_i)/(2+v)$, de pendiente $-1/(2+v)$, que solo coincide con la conjetura si $v = -1$. Con costes marginales crecientes, la conjetura consistente queda entre -1 y 0.

1.4.2. Diferenciación horizontal (Hotelling, 1929)

Los consumidores se distribuyen uniformemente en un segmento, con costes de transporte, y las empresas eligen localización y precio. Con precios fijos, se aglomeran en el centro (el principio de **mínima diferenciación**). Con precios endógenos y costes cuadráticos, se separan al máximo para relajar la competencia en precios (d'Aspremont et al., 1979). El resultado es muy sensible a los costes de transporte, y con costes lineales puede no existir equilibrio. La localización es así una decisión estratégica, porque acercarse al rival capta demanda pero intensifica la competencia en precios.

1.4.3. Innovación (Schumpeter, 1942; Arrow, 1962)

Schumpeter: la innovación crece con el tamaño y el poder de mercado, porque solo la gran empresa financia la I+D y diversifica su riesgo, y solo el beneficio extraordinario permite recuperar la inversión.

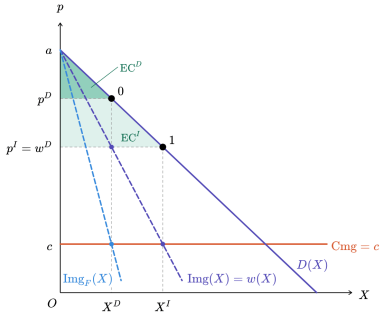
Arrow: el monopolista innova menos por el **efecto reemplazo**, porque la innovación canibaliza el beneficio que ya obtiene. Con una innovación de proceso no drástica, el monopolista gana la diferencia entre dos beneficios positivos y la empresa competitiva, la diferencia entre uno positivo y cero. La competitiva, que parte de un beneficio nulo, cobra algo menos que el coste anterior, se queda con todo el mercado y gana el ahorro de coste sobre toda la demanda.

Como el excedente social supera a ambos, el orden de incentivos es planificador, competencia y monopolio. Hay dos matices: el *efecto eficiencia* (Gilbert y Newbery, 1982), porque el monopolista amenazado puja más por la patente, y la financiación de la I+D, núcleo del argumento schumpeteriano.

Evidencia: el gasto en I+D crece en proporción al tamaño, y Aghion, Bloom, Blundell, Griffith y Howitt (2005) hallan una **U invertida**: con poca competencia domina el efecto de escapar de la competencia, y con mucha, el schumpeteriano. Ni la competencia ni el monopolio alcanzan el incentivo óptimo.

1.4.4. Otros modelos de comportamiento

1. *Doble marginalización* (Spengler, 1950): si un fabricante monopolista vende con precio mayorista lineal a un minorista también monopolista, cada uno añade su margen sin internalizar el del otro. Con demanda lineal, la cadena vende la mitad que la empresa integrada, más cara, con un beneficio conjunto un 25 % inferior y la cuarta parte del excedente del consumidor. La integración, o una restricción vertical que elimine el segundo margen, mejora a empresas y consumidores (regla de la razón de Chicago), aunque puede cerrar el mercado a los rivales.



Doble marginalización con demanda lineal y coste marginal constante.

2. *Diferenciación y variedad* (Spence, 1976; Dixit y Stiglitz, 1977): margen constante, $p = \frac{\sigma}{\sigma-1}$ Cmg, y beneficio nulo con libre entrada (A18).
3. *Diferenciación vertical* (Shaked y Sutton, 1982): calidades distintas para suavizar la competencia en precios y número finito de empresas (A18).

En suma, la estructura de costes delimita cuántas empresas caben, los contratos dónde terminan y la agencia qué persiguen, y la conducta altera la propia estructura del mercado.

2. Teoría del mercado

El recorrido parte de Harvard y Chicago, sigue con los determinantes de la estructura de los mercados y termina con su regulación.

2.1. Antecedentes

2.1.1. Paradigma ECR: Harvard (años cuarenta)

Idea: para Mason y Bain, la estructura del mercado, número de empresas, concentración, barreras y diferenciación, determina la conducta, en precios, inversión o publicidad, y esta el resultado. La causalidad es unidireccional, de modo que toda industria con pocas empresas resulta sospechosa, y la política de competencia mide la concentración con el índice de Herfindahl-Hirschman y los ratios de concentración.

Conclusiones: inaugura el análisis empírico sistemático de los mercados y da instrumentos operativos. Pero es estático, sin conducta estratégica, y su causalidad es discutible, porque las empresas más rentables crecen y concentran el mercado.

2.1.2. Escuela de Chicago (años cincuenta y sesenta)

Idea: para Stigler, Demsetz, Posner y Bork, la concentración no indica necesariamente poder de mercado, por dos razones. Las grandes empresas pueden serlo por ser más eficientes, de modo que la causalidad va de la eficiencia a la concentración, y la forma en L de los costes medios implica que la gran dimensión no comporta desventaja. Además, la amenaza de entrada disciplina los precios incluso en mercados concentrados. Con libre entrada el mercado se autorregula, y la regulación debe limitarse a eliminar las barreras legales.

Conclusiones: subraya la eficiencia y la entrada potencial, que formalizarán los mercados impugnables. Pero ignora las barreras estratégicas y la predación, que puede ser racional. En suma, Harvard era demasiado estático y Chicago, demasiado optimista.

2.2. Estructura

2.2.1. Causas: las barreras a la entrada

Definición: las **barreras a la entrada** deciden si la concentración es inocua o lesiva, y su definición fija el perímetro de la intervención.

1. *Bain (1956)*: condiciones que permiten a las instaladas fijar persistentemente precios supracompetitivos sin inducir la entrada: economías de escala, diferenciación de producto y ventajas absolutas de costes. Con economías de escala, el entrante ha de elegir entre entrar pequeño con desventaja de costes o grande deprimiendo el precio, y la lealtad de marca le obliga a gastos de introducción ya amortizados por las instaladas.
2. *Stigler (1968)*: costes que soporta el entrante y no soportaron las instaladas. Las economías de escala ya no son barrera. Por eso Harvard era intervencionista y Chicago, abstencionista.
3. *Baumol y Willig (1981)*: los costes fijos no son barrera, pero sí los **costes hundidos**, que crean una asimetría de riesgo entre quien ya los comprometió y quien aún puede evitarlos.

Barreras estratégicas: las crean las instaladas, como el precio límite (Bain-Sylos-Modigliani), la inversión en capacidad como compromiso creíble (Spence, 1977; Dixit, 1980) o la proliferación de marcas (Schmalensee, 1978). Con las legales forman las barreras artificiales, frente a las naturales, que nacen de la tecnología y la demanda.

2.2.2. Mercados impugnables (Baumol, 1982)

Idea: Baumol retoma la distinción de Demsetz entre la competencia en el mercado y por el mercado y, con Panzar y Willig, define un mercado **impugnable** como aquel en que los entrantes pueden entrar y salir sin desventaja: igual acceso a la tecnología, costes fijos pero no hundidos, y entrada más rápida que la reacción de la instalada. La reacción se retrasa por causas como los costes de menú, mientras la demanda responde al instante al entrante, y sin costes hundidos la inversión es recuperable al salir.

Razonamiento: incluso un monopolio natural fija un precio igual al coste medio, porque cualquier precio superior atraería una entrada relámpago (*hit-and-run*). La disciplina procede de la competencia potencial, y la estructura deja de ser el indicador relevante. Formalmente, solo es de equilibrio una configuración *sostenible*, que atiende la demanda sin pérdidas y no permite una entrada rentable a un precio igual o menor:

$$\sum_i x_i = X(p), \qquad p\,x_i - C(x_i) \geq 0$$
$$p^e x^e - C(x^e) \leq 0 \quad \forall\, p^e \leq p,\ x^e \leq X(p^e)$$

Ello implica beneficio nulo, eficiencia productiva y, con dos o más empresas, $p = \text{Cmg}$. El monopolio natural queda en $p = \text{Cme} > \text{Cmg}$, el *second best*.

Conclusiones: invierte la carga de la prueba del paradigma ECR y fue muy influyente en la desregulación del transporte aéreo y las telecomunicaciones. Pero la ausencia total de costes hundidos rara vez se da.

2.2.3. Mercados de varias caras (Rochet y Tirole, 2003)

Idea: una plataforma pone en contacto a dos grupos con **externalidades cruzadas de red**: el valor para un lado depende del número de participantes del otro, como en las tarjetas de pago, entre comercios y titulares, los sistemas operativos o los mercados en línea.

Razonamiento: la estructura de precios no es neutral, es decir, importa el reparto de los precios entre los dos lados y no solo su nivel agregado. Suele ser óptimo subvencionar al lado más sensible al precio, o al que genera más externalidad, incluso con precios nulos o negativos. Un precio inferior al coste en un lado no prueba por sí solo la predación, y el mercado relevante debe analizarse con ambos lados.

Conclusiones: explica los servicios gratuitos o subvencionados para uno de los lados. Pero los resultados dependen mucho de las externalidades y de si los usuarios operan en varias plataformas.

2.2.4. Predación (Areeda y Turner, 1975)

Idea: la **predación** sacrifica beneficios presentes con precios bajos para expulsar a un rival y recuperarlos después. La escuela de Chicago (McGee, 1958) la consideraba irracional.

Regla de Areeda-Turner: aproxima el coste marginal por el coste variable medio, y un precio inferior se presume predatorio. Es demasiado permisiva, porque la predación por reputación puede darse por encima de ese coste, y estática, porque ignora la recuperación. Bolton, Brodley y Riordan (2000) proponen dos fases: primero, la plausibilidad de la recuperación, como exige en EE. UU. *Brooke Group* (1993), y después, la prueba de costes. La segunda fase examina los precios por debajo del coste y la ausencia de justificación de eficiencia.

Doctrina europea: *AKZO* (1991) presume abusivos los precios inferiores al coste variable medio, y los situados entre este y el coste total medio, si forman parte de un plan de exclusión, sin exigir la recuperación. El contraste entre la recuperación exigida y la intención suficiente refleja pesos distintos de los errores de tipo I y II. Las reglas de costes dan seguridad jurídica, pero solo capturan la predación más burda, y las de dos fases incorporan la teoría estratégica a costa de mayor complejidad probatoria.

2.3. Regulación

Cuando una sola empresa es la estructura eficiente, el problema es regularla sin la información que ella tiene.

2.3.1. Competencia por el mercado (Demsetz, 1968)

Idea: subastar periódicamente el derecho exclusivo a servir el mercado (*franchise bidding*). Con licitación competitiva, el precio converge al coste medio y se replica una regulación de beneficio nulo sin regulador permanente ni conocer los costes de la empresa. Williamson (1976) objetó que los contratos de concesión son incompletos, que los activos específicos dificultan la reversión y que el incumbente acumula ventajas informativas en las relicitaciones.

2.3.2. Competencia por comparación (Shleifer, 1985)

Idea: el regulador fija el precio de cada empresa con los costes de empresas comparables de otros mercados, p. ej. distribuidoras regionales (*yardstick competition*). Fijarlo con los costes declarados por la propia empresa destruiría sus incentivos a reducirlos. Con la comparación, en

cambio, sus reducciones de costes no alteran su precio, y la empresa se beneficia íntegramente de ellas. Exige comparadores similares, o corregir econométricamente sus diferencias, y ausencia de colusión entre ellos.

2.3.3. Regulación por incentivos (Laffont y Tirole, 1986)

Idea: el regulador, el principal, no conoce la eficiencia de la empresa, el agente, y le ofrece un menú de contratos entre los que se autoselecciona, revelando su eficiencia a cambio de una renta informacional.

Supuestos: un monopolio natural tiene costes $\mathrm{CT}(x) = (\varphi - e)\,x + F$, donde φ es su eficiencia intrínseca y e su esfuerzo, ambos no observables, de modo que se combinan selección adversa y riesgo moral. El esfuerzo cuesta $\psi(e)$, convexa, y cada euro de fondos públicos cuesta $1+\lambda$.

Planteamiento: por el principio de revelación (Myerson, 1979), el regulador elige un mecanismo directo que maximiza el bienestar esperado sujeto a la participación y a la compatibilidad de incentivos de cada tipo. Con la producción normalizada a uno y sin costes fijos, el coste observado es $C = \varphi - e$, que el regulador reembolsa junto con una transferencia neta t, y la renta de la empresa es $U = t - \psi(e)$. Con W el bienestar que genera cada tipo, G y g la distribución y la densidad de los tipos, $\hat{\varphi}$ el tipo anunciado y $C(\hat{\varphi}) = \hat{\varphi} - e(\hat{\varphi})$ su coste asignado, el mecanismo resuelve:

$$\max_{\{t(\varphi),\,e(\varphi)\}} \int_{\underline{\varphi}}^{\bar{\varphi}} W(t(\varphi), e(\varphi); \varphi)\, g(\varphi)\, d\varphi$$

$$\text{s.a.:} \begin{cases} U(\varphi) = t(\varphi) - \psi(e(\varphi)) \geq 0 & \forall \varphi \\ U(\varphi) \geq t(\hat{\varphi}) - \psi(\varphi - C(\hat{\varphi})) & \forall \varphi, \hat{\varphi} \end{cases}$$

CPO: el ahorro marginal del esfuerzo iguala su desutilidad marginal más el coste de la renta informacional que hay que ceder a los tipos más eficientes:

$$1 = \psi'(e(\varphi)) + \frac{\lambda}{1+\lambda} \cdot \frac{G(\varphi)}{g(\varphi)} \cdot \psi''(e(\varphi))$$

Un poco más de esfuerzo ahorra una unidad de coste, pero si el regulador lo pide al tipo φ, su contrato atrae a los tipos más eficientes, de masa $G(\varphi)$, que podrían imitarlo con menos esfuerzo, y hay que subirles la renta. Cada euro de renta cuesta λ, frente al valor social $1+\lambda$ del ahorro, de ahí el factor $\lambda/(1+\lambda)$.

El esfuerzo se distorsiona a la baja respecto al *first best*, $\psi'(e) = 1$, tanto más cuanto mayor es λ, salvo para el tipo más eficiente (**no hay distorsión en la cima**).

Solución: el mecanismo se implementa con un **menú de contratos** lineales en el coste realizado, cuya potencia es la fracción del sobrecoste que recae sobre la empresa. Con un pago fijo a, un coste objetivo C^* y una potencia $b \in [0,1]$, que vale 1 en el precio fijo y 0 en el coste más margen, el menú es:

$$t(\hat{\varphi}, C) = a(\hat{\varphi}) - b(\hat{\varphi}) \cdot [C - C^*(\hat{\varphi})]$$

- La empresa eficiente elige un contrato de precio fijo, de modo que asume todo el riesgo de costes y retiene todo el ahorro de su esfuerzo, y percibe una renta informacional para que no le compense hacerse pasar por ineficiente.
- La ineficiente elige el contrato de menor potencia, el más próximo al de coste más margen: el regulador absorbe buena parte de sus sobrecostes, con un esfuerzo distorsionado a la baja, pero sin ceder renta.
- Los tipos intermedios combinan un componente fijo y un reparto parcial de los sobrecostes.

El menú resuelve el *trade-off* entre la eficiencia, que pide contratos de alta potencia, y la extracción de rentas. Como la empresa ajusta su esfuerzo hasta $\psi'(e) = b$, el menú induce el esfuerzo óptimo con $b(\hat{\varphi}) = \psi'(e(\hat{\varphi}))$, igual a 1 para el tipo más eficiente y decreciente con φ.

Conclusiones: fundamenta la regulación *price-cap* frente a la de tasa de retorno, y valió a Tirole el Nobel de 2014. Pero exige mucha información al regulador, es estático, expuesto al efecto trinquete, e ignora la captura del regulador.

En suma, la teoría del mercado ha pasado de tomar la estructura como dato a explicarla, y la regulación, de fijar precios a resolver el problema informativo del regulador.

3. Nueva teoría de la organización industrial empírica

Origen: Bain (1951) halla, para 42 industrias manufactureras estadounidenses de 1936-1940, beneficios mayores en las que las ocho mayores empresas superan el 70 % de las ventas. Demsetz (1973) objetó que, si fuera colusión, todas las empresas se beneficiarían, pero solo las grandes son más rentables, por lo que concentración y beneficio son endógenos. La lectura estructuralista veía en la correlación la colusión que la concentración facilita, y la de Demsetz, la eficiencia de Chicago, de modo que la disputa no podía zanjarse con regresiones entre industrias.

Idea: la **NEIO** (Bresnahan, 1989) sustituye esas regresiones por el análisis estructural de industrias concretas, especificando la demanda y la conducta para estimar directamente el poder de mercado.

Parámetro de conducta: la NEIO escribe el margen como $(p - \mathrm{Cmg})/p = \theta/(n\,|\varepsilon_{Xp}|)$, con $\theta = 0$ en competencia, 1 en Cournot y n en el cártel. Bresnahan (1982) y Lau (1982) muestran que θ está identificado si la demanda sufre rotaciones exógenas, de modo que el grado de competencia pasa a ser estimable. Así puede distinguirse la concentración que refleja eficiencia de la que refleja colusión.

3.1. Delimitación del mercado relevante

El análisis parte de delimitar el **mercado relevante**, de producto y geográfico, con el test del monopolista hipotético (SSNIP). Solo dentro de él tienen sentido los índices.

3.2. Medición del poder de mercado

3.2.1. Índices de poder de mercado

Índice de Lerner: el margen relativo entre precio y coste marginal, de 0 en competencia perfecta a 1. En Cournot, $I_{L,i} = s_i/|\varepsilon_{Xp}|$, y la media ponderada por las cuotas es la suma de sus cuadrados (el HHI en tanto por uno) dividida por la elasticidad (Cowling y Waterson, 1976):

$$I_{L,i} = \frac{p - \mathrm{Cmg}_i}{p} \in [0,1]$$

$$I_L = \sum_{i=1}^{n} s_i\, I_{L,i} = \frac{\sum_{i=1}^{n} s_i^2}{|\varepsilon_{Xp}|} \in [0,1]$$

En el monopolio, el índice iguala la inversa de la elasticidad de la demanda. Como el coste marginal no se observa, se aproxima con el margen contable, (ingresos − CV)/ingresos, o se estima con la NEIO.

***q* de Tobin**: el valor de mercado de la empresa entre el coste de reposición de sus activos. Con libre entrada tiende a 1, y un valor persistentemente superior indica rentas que la competencia no erosiona. No exige conocer el coste marginal, pero omite los intangibles, lo que lo sesga al alza.

3.2.2. Índices de posición

La **cuota de mercado**, $s_i = x_i/X$, va de 0 a 1. En la Unión Europea, una cuota del 50 % o superior permite presumir posición dominante (*AKZO*, 1991), aunque puede reflejar eficiencia (Demsetz, 1973).

3.2.3. Índices de concentración

Ratio de concentración: $\mathrm{CR}_k = \sum_{i=1}^{k} s_i$, la cuota de las k mayores empresas. Es sencillo, pero ignora el reparto de las cuotas. Va de k/n, con cuotas iguales, a 1, y Bain (1951) utilizó el CR_8.

Índice de Herfindahl-Hirschman: suma los cuadrados de las cuotas en porcentaje (de $1/n$ a 1 en tanto por uno, cuya inversa es el número equivalente de empresas). Una fusión de i y j lo eleva en $2\,(100\,s_i)(100\,s_j)$:

$$\mathrm{HHI} = \sum_{i=1}^{n} (100\,s_i)^2 \in \left[\frac{10\,000}{n},\ 10\,000\right]$$

- *Unión Europea (2004)*: los problemas son improbables con un HHI posterior inferior a 1000, entre 1000 y 2000 con un aumento inferior a 250, o superior a 2000 con un aumento inferior a 150.
- *Estados Unidos (2023)*: más de 1800 indica un mercado muy concentrado, y se presume lesiva la fusión que lo crea o refuerza con un aumento superior a 100.

3.3. Evidencia empírica

Lerner: Porter (1983) estima la demanda y la oferta del cártel ferroviario del Joint Executive Committee (1880-1886): en los periodos colusivos el margen se acerca al de Cournot, y las guerras de precios son castigos temporales que sostienen la colusión. Ellison (1994) halla un margen más próximo al de monopolio. Su modelo, más rico, estima una demanda más elástica e indicios de rebajas secretas de precios, que apoyan débilmente a Green y Porter (1984) frente a Rotemberg y Saloner (1986).

Cuota: Gal (2003) muestra que, en las economías pequeñas, la EME elevada hace inevitable e incluso eficiente una concentración alta, de modo que los umbrales no son trasladables entre jurisdicciones.

Otros: entre el 5,3 y el 7,4 % de las adquisiciones farmacéuticas buscan interrumpir un proyecto rival, las *killer acquisitions* (Cunningham et al., 2021), y Rysman (2004) estima externalidades cruzadas en las guías telefónicas y prefiere una estructura más competitiva al monopolio.