

Universidad de Oviedo



**MEDICIÓN DEL OUTPUT EN LOS SERVICIOS
PÚBLICOS**

Marián García Valiñas
e-mail: marian@econo.uniovi.es

MEDICIÓN DEL OUTPUT EN LOS SERVICIOS PÚBLICOS

Marián García Valiñas
Universidad de Oviedo
e-mail: marian@econo.uniovi.es

RESUMEN

En un contexto de restricción de recursos, los gobiernos se ven en la necesidad de introducir nuevas formulas de gestión que precisan de una clara definición de objetivos en términos de producción y resultados (OCDE, 1990). Este trabajo refleja diversas dificultades ligadas a la definición y cuantificación del output representativo de la prestación de servicios en el ámbito público. Objetivos ambiguos, la diferenciación entre producto intermedio y final, output bruto y valor añadido o problemas de agregación en contextos multiproducto son algunos aspectos metodológicos abordados. Asimismo, se proponen medidas en servicios como la sanidad, la educación y la seguridad pública, que constituirán aproximaciones al producto, tanto en términos cuantitativos como cualitativos.

Palabras clave: gestión pública, servicios públicos, cuantificación, valor añadido, output intermedio, output final.

1. INTRODUCCIÓN

“Los outputs de los servicios públicos son, conceptualmente, aquellas cantidades, expresadas básicamente en unidades físicas, que se obtienen del proceso productivo”. De esta manera definía Hirsch (1973, pág. 440) el producto derivado de las actuaciones públicas. La literatura económica señala la necesidad de una detallada especificación de las características de los bienes y servicios, con objeto de discernir los meros aumentos de los costes de aquellos asociados a la modificación de las condiciones de la prestación (Levitt y Joice, 1987). Ya por los años cincuenta, la Oficina Estadística Central del Reino Unido sostenía que, incluso una medida rudimentaria del output sería preferible a la utilización de los costes como índice de producto.

Es posible vaticinar, por tanto, que cada vez se hace más necesaria la determinación de medidas de producto adecuadas, y ello por motivos diversos. En primer lugar, son imprescindibles para facilitar y favorecer una buena gestión en el ámbito gubernamental, cuestión que parece cobrar gran importancia en las últimas reformas organizativas practicadas en el seno de la administración pública, encaminadas hacia el control del gasto (OCDE, 1990). Así, las variables de producto y resultados serán imprescindibles para la aplicación de nuevas formas de presupuestación basadas en objetivos, conformando igualmente, una buena base para la toma de decisiones a todos los niveles y el diseño de sistemas de incentivos encaminados hacia la motivación y la coordinación de los agentes económicos. Asimismo, los indicadores de producto tendrán mucho que aportar en los procesos de evaluación de las actuaciones públicas, sobre todo desde el momento en el que se empiezan a establecer como exigencia normativa. La medición del output se presenta como fundamental en procesos de control contractual, interno y externo (Smith, 1996). En relación a esta última tipología de control, los contribuyentes estarán en condiciones de apreciar los resultados que obtienen a cambio del dinero que aportan al sector público. La expresión inglesa *“value for money”* hace referencia a esa idea, de manera que los ciudadanos tendrán argumentos más sólidos para valorar las actuaciones gubernamentales.

El protagonismo creciente de estudios que analizan aspectos como la eficiencia o la productividad en el sector gubernamental constituye otro argumento adicional a considerar. Así, algunas teorías en relación al crecimiento del sector público realizan explícitamente la asunción de que la productividad en dicho contexto no experimentará incremento alguno [Baumol (1967); Skolka, (1977)]. La medición de variaciones en la productividad de los servicios públicos permitirá constatar dichos supuestos. En este sentido, la precisión en la determinación y diseño de las variables mejorará, asimismo, la validez de los resultados de ese tipo de estudios. Si se definen incorrectamente, es probable que las estimaciones de eficiencia y productividad en la prestación de servicios públicos presenten algún tipo de sesgo.

El trabajo se centrará en las actuaciones de suministro de servicios públicos, analizando la actuación pública en algunas prestaciones relevantes. En este sentido, es

destacable la contribución de dichas actuaciones, tanto con relación a la renta como al empleo. En 1997 el volumen de servicios públicos supuso una media del 11,5 % del Producto Nacional Bruto de diversos países de la OCDE, registrando la mayor parte de ellos porcentajes superiores al 10%. En cuanto al empleo, la contribución media se situaba en torno al 17% (OCDE, 1999). A la luz de estas cifras, puede advertirse que los servicios públicos absorben un elevado volumen de recursos, generando, al mismo tiempo, una actividad económica considerable. De esta forma, el artículo se estructurará como sigue: en primer lugar, serán abordadas algunas cuestiones de metodología, para pasar, seguidamente, a analizar índices representativos de la producción generada en la prestación de algunos servicios públicos representativos. Para finalizar, se expondrán los aspectos más destacados a modo de recapitulación.

2. ASPECTOS GENERALES Y METODOLÓGICOS

La larga lista de obstáculos en relación a la determinación del output en el sector público podría reducirse a dos grupos, tal como señalaban Fletcher y Snee (1985) en su análisis de la medida del output en el sector servicios, y cuyas conclusiones son perfectamente extrapolables al ámbito de los servicios públicos:

- a) Problemas de *definición*, ligados a la determinación de los objetivos.
- b) Problemas de *cuantificación*, vinculados a la aproximación del output mediante índices.

En este apartado se comentarán algunos problemas genéricos relacionados con la medición del producto en el ámbito de los servicios públicos, presentando bastantes afinidades con las prestaciones de carácter privado. No obstante, la propia naturaleza de la actividad pública complicará en gran medida los procesos de medida del output, tal como se podrá apreciar de inmediato.

2.1 Definición de objetivos y actividades

En opinión de Pliatzky (1986, pág. 192), *“la administración de un país es una labor bastante más compleja que la dirección de una empresa privada porque, mientras que una compañía que actúa en el mercado tiene, por lo general, el propósito de conseguir un beneficio, los gobiernos, por su naturaleza, aspiran a ciertos objetivos diferentes y, con frecuencia, en conflicto, de ahí que el sector público, a sus más altos niveles, trata de resolver dilemas, lo cual es aún más difícil que la administración de los recursos (...)”*. Un primer paso, previo a la cuantificación, supondrá definir los objetivos que la organización pública aspira a alcanzar, así como las actividades a realizar para su consecución. En este sentido, la ambigüedad es una característica inherente al proceso de definición del output público. Se habla con frecuencia de agencias públicas con múltiples objetivos, en ocasiones contrapuestos, reflejo de la existencia de numerosos “usuarios” que compiten por unos recursos dados. Al final, la determinación de objetivos se reduce a un proceso de elección de algunos de ellos, elección en la que influirán aspectos muy diversos. La existencia de conflictos en los procesos de decisión y elección es una nota presente en todo tipo de organización. Sin embargo, se

puede afirmar que en organizaciones sin ánimo de lucro, el peso de las presiones e intereses externos es mucho mayor que en el resto (Schwenk, 1990).

2.2 Cuantificación del output

Una vez definidos los objetivos a que aspiran las actuaciones públicas, el paso siguiente consistirá en definir variables que aproximen el producto ofertado por las mismas. Dentro del campo del suministro público, el mayor peso es acaparado por las actividades de servicios, entendidos estos como prestaciones de utilidad que pretenden satisfacer necesidades de los consumidores o usuarios, en los cuales existe un claro predominio del factor trabajo. Se estarían considerando aquellas transacciones entre varios agentes económicos que conducen a un cambio en las condiciones de la persona o del bien (Hill, 1977). De forma genérica, cualquier tipo de servicio presenta un conjunto de características que hacen que su medida de su output sea aún más complicada.

En primer lugar los servicios son intangibles. Esto impide que sean aplicables los métodos tradicionales de medición basados en el cómputo de unidades físicas. En ocasiones, determinados bienes tangibles pueden ser representativos del servicio –sería el caso de los formularios del impuesto sobre la renta en el servicio de gestión tributaria-, pero no constituyen el servicio en sí mismos. Por otro lado, presentan cierta heterogeneidad, debido al componente personal tan intenso que poseen. Su output es, por tanto, difícil de generalizar. En el extremo, cada unidad producida es diferente a las demás. La calidad es variable, y no se puede determinar a priori, solo se apreciará una vez suministrado el servicio. El carácter heterogéneo de los servicios no será, sin embargo, la cualidad verdaderamente distintiva de los servicios¹. Por lo general, un mayor grado de heterogeneidad complicará los procesos de cuantificación del output. Adicionalmente, los servicios, se consumen mientras se realizan, y, generalmente, el cliente se haya implicado en el proceso. Es más, en la mayoría de los servicios sociales, el proceso productivo se desarrolla sobre el propio cliente. Este constituye a un tiempo un input y un output, de tal forma que ambas dimensiones se solapan. Así, la materia prima sobre la que se trabaja es el cliente, intentando mejorar su estado inicial. Finalmente, es necesario remarcar que la mayor parte de los servicios no son almacenables, luego, si un servicio no se utiliza cuando está disponible, la capacidad del mismo se pierde. El carácter efímero de estos², junto al rasgo anterior, determina la verdadera diferencia con los bienes. Todos estos rasgos propios de las prestaciones de servicios complican en gran medida el análisis. A continuación comentaremos algunos aspectos adicionales a tener en cuenta en la elección y diseño de dichas variables.

¹ De este modo, el aspecto de heterogeneidad, y en cierta medida, la intangibilidad podrían ser extendidos a todos los bienes, en el sentido de que incorporan características y propiedades que los hacen deseables. Esta idea es recogida por la *Teoría de las Características* (Lancaster, 1966), según la cual, las características físicas de todos los outputs son extraordinariamente complejas, incluso en los casos aparentemente de mayor sencillez. Así, aunque generalmente se considere que el output de la industria del automóvil es el output físico de un automóvil, existen numerosos tipos, marcas y modelos, de ahí que el output se reduzca a un vector de características determinadas, más o menos complejas según los casos.

² Existen excepciones, como es el caso de algunos servicios relacionados con la información.

Output final versus output intermedio

Cabría establecer una diferenciación importante en relación a la etapa de la cadena productiva que se esté considerando. Inicialmente, los entes públicos necesitan contar con recursos de diversa índole -inputs-, que emplearían en la realización de una serie de procesos o actividades intermedias -throughputs-, las cuales, a su vez permitirían el suministro del producto/s o servicio/s -output/s-. Adicionalmente, en el sector público sería imprescindible avanzar un paso más, considerando el impacto o resultado de sus actuaciones -outcome/s-, lo que constituiría el output final de las mismas. Este deberá incorporar aspectos cualitativos y relacionados con la equidad -distribución espacial y personal de bienes y servicios-. Es necesario señalar, que la relación existente entre el output intermedio y los resultados de las actividades públicas está, en muchas ocasiones, fuertemente condicionado por el entorno, luego en dicho esquema desempeñarían un papel fundamental los factores ambientales. A todo ello habría que añadir la existencia fuertes interrelaciones entre las diferentes agencias públicas. Así, será muy frecuente que en la prestación de un determinado servicio intervengan varios departamentos gubernamentales. Esquemas input-output de estas características han sido planteados desde un punto de vista teórico por autores como Bradford, Malt y Oates (1969), Butt y Palmer (1985), Arvidsson (1986), Levitt y Joyce (1987) y Murray (1992), entre otros.

El empleo de los inputs como variable proxy del output, constituye una práctica tan frecuente como incorrecta. Por lo general, estamos en condiciones de afirmar que, cuanto mayor sea la proximidad al final de la cadena de valor, más adecuada será la unidad de medida, pero mayor dificultad presentará su cuantificación. Bradford, *et al.* (1969) expusieron una primera versión bastante simple del enfoque señalado anteriormente, como introducción teórica al ejercicio de análisis de los costes asociados a la prestación de servicios públicos a nivel local. Distinguieron dos etapas, definiendo en la primera de ellas el output público (Y) como función de un vector de inputs (X):

$$Y = f(X) \quad (2.1)$$

A continuación, especificaron una segunda etapa, basada en el supuesto de que los ciudadanos-votantes demandan servicios públicos en función de resultados - C - , de tal forma que la función de utilidad representativa de las preferencias individuales poseería la estructura que sigue:

$$U = U(C, Z) \quad (2.2)$$

denotando con Z al resto de bienes y servicios. Así pues, podríamos representar la segunda fase del proceso de producción pública como

$$C_k^i = c_k^i(Y, E) \quad (2.3)$$

donde C representa el resultado o impacto que sobre los ciudadanos tiene la prestación del servicio. E denotaría un vector de factores ambientales que aproximan el entorno en el que se efectúa la actividad. El subíndice k hace referencia al grupo social al que

pertenece el individuo, pudiendo incorporarse un superíndice i que refleje las particularidades de cada individuo.

No sería posible determinar a ciencia cierta el signo de la variación que, sobre el resultado final provocaría un cambio en el entorno. Es preciso recordar que E es un vector que incluye diversos componentes ambientales, cuyos efectos podrían compensarse. Por su lado, la relación entre C e Y parece ser no negativa, aunque de cuantía indeterminada. La cuantificación del producto final presenta complicaciones adicionales en cuanto a dinámica se refiere, puesto que, generalmente, los resultados no son observables en el corto plazo. En este sentido, es posible que se produzcan desajustes temporales. Este aspecto podrá ser apreciado en algunos índices de producto propuestos en el apartado siguiente del artículo.

Producto bruto versus valor añadido

No será indiferente la elección entre índices de producción bruta o de valor añadido excluyendo, con la última opción, los consumos intermedios. Parece adecuado en un principio, pongamos por caso, en los servicios educativos, diseñar medidas de valor añadido para aproximar el producto atribuible a los centros que proporcionan este tipo de servicios, determinando las mejoras formativas de sus alumnos. A este respecto, un índice construido por diferencia entre los resultados académicos al final y al principio de curso podría ser más representativo de la labor ejercida por los centros educativos.

No obstante, cuando se opta por este tipo de procedimientos, se están imponiendo, implícitamente, fuertes restricciones sobre la tecnología. Así, Arrow (1974) llegaba a demostrar que la sustitución de la producción bruta por valor añadido equivale a suponer que la función de producción es débilmente separable en los consumos intermedios, lo cual implica que la relación marginal de sustitución entre capital y trabajo es independiente de la cantidad de consumos intermedios empleados en el proceso productivo.

De forma similar, Bruno (1978) estudiaba las consecuencias derivadas de la utilización de medidas de valor añadido en la estimación del cambio en la productividad total de los factores, estableciendo una serie de condiciones para poder efectuar la sustitución de unas medidas por otras. Así, para que se verifique la equivalencia, en términos de productividades marginales relativas, de producción bruta y valor añadido, es necesario que se cumplan alguna de las condiciones siguientes: *a)* que las materias primas se empleen en proporciones fijas respecto al output bruto, *b)* que los precios relativos de dichos consumos intermedios permanezcan constantes o *c)* que la función de producción definida a partir del output bruto sea funcionalmente separable en grupos de inputs³. En el caso de que ninguno de dichos supuestos se satisfaga, las estimaciones de las variaciones en la productividad estarían claramente sesgadas⁴.

³ Nótese que la primera condición estaría incluida en la tercera. Si los inputs intermedios se emplean en proporciones fijas, la función de producción adoptaría la forma siguiente:

$$Y = \min \{z(L), a_1 M_1, a_2 M_2, \dots, a_m M_m\}$$

Tecnología multiproducto: agregación

Frecuentemente, el servicio principal prestado por los centros públicos de suministro suele presentar multidimensionalidad, a lo que habría que añadir, en otras ocasiones, servicios de tipo secundario⁵. En este contexto, podemos estar interesados, con fines de condensación de la información, en sintetizar varios outputs o dimensiones del mismo en un indicador único. Lo adecuado sería fijar ponderaciones considerando el porcentaje que suponen los ingresos procedentes de cada uno de los outputs en relación a los ingresos totales (Norsworthy y Jang, 1992).

El sector público suministra un gran número de servicios no destinados a la venta en el mercado, siendo imposible, por tanto, disponer de los precios como ponderaciones para construir índices agregados en contextos multiproducto. Y cuando se establecen precios –tasas, precios públicos-, es, cuanto menos, dudoso, emplearlos como ponderaciones adecuadas, pues son precios regulados no competitivos. La solución más recurrida pasa por estimar funciones de costes. La obtención de los costes marginales por derivación de la función de costes parece ser la opción más aconsejable (Fisher y Shell, 1972). Otras metodologías alternativas pueden ser utilizadas para agregar o reducir variables. Así, la técnica multivariante de *Análisis Factorial de Componentes Principales*, inicialmente propuesta por Hotelling (1933), persigue agrupar variables basándose en criterios de varianza y covarianza. El sistema se basa en la existencia de relaciones entre variables, como fundamento para la reducción de las mismas a unas pocas factores latentes no observables, con la mínima pérdida de información posible. Los factores finalmente resultantes de la aplicación de la técnica se caracterizan por ser linealmente independientes, lo cual elimina problemas de multicolinealidad en posteriores regresiones. Esta técnica es habitualmente aplicada en procesos de evaluación de la *calidad*, buscando la reducción de un elevado número de variables a unas pocas dimensiones cualitativas representativas. El principal inconveniente de esta técnica radica en la no observabilidad de los factores, aspecto que complica su identificación e interpretación.

Una forma alternativa de agregación es la proporcionada por la metodología de *Análisis Envolvente de Datos*, comúnmente conocido por las siglas DEA -*Data Envelopment Analysis*-. El DEA es un método no paramétrico, aplicado básicamente en la evaluación de la eficiencia, desarrollado inicialmente por Charnes, Cooper y Rhodes (1978). A través de programación lineal restringida, y partiendo de observaciones

Con esta modelización de la tecnología, se está asumiendo implícitamente que la función de producción es débilmente separable en L y M .

⁴ Bruno (1978) demuestra, que, bajo ciertos supuestos, la medida de los cambios en la productividad derivada de una función de valor añadido a precios constantes constituiría una infraestimación de la productividad real. Este argumento tiene importantes implicaciones para las aplicaciones que estiman las variaciones en la productividad del sector público. Si la elección de variables no es correcta, es posible que las conclusiones sean erróneas, estimando cambios en la productividad inferiores a los que realmente se estarían produciendo.

⁵ En este sentido, es necesario destacar el papel jugado por los denominados servicios periféricos. Se alude con ellos a aquellas actividades, que no constituyen el verdadero objetivo del centro productivo, cuya prestación se realiza de forma simultánea a este. Contribuyen a definir el ambiente en el que se desarrolla la prestación, a pesar de estar hablando de tareas de tipo marginal. Un buen ejemplo serían los servicios de cafetería en los hospitales, facultades, los comedores de los colegios o los aparcamientos en museos u otros centros. A pesar de no constituir el foco de atención del análisis, son actividades que contribuyen a determinar la calidad de la prestación global del servicio.

muestrales homogéneas, construye una frontera de posibilidades de producción, formando una envolvente con las unidades eficientes. En esencia, el sistema consiste en optimizar –considerando ciertas restricciones- un ratio de productividad total. Dicho ratio se construye dividiendo la suma ponderada de los outputs entre la suma ponderada de los inputs, siendo generadas, por el propio programa de optimización, las ponderaciones a emplear en dicho sumatorio. Estas son fijadas de la forma más favorable a la eficiencia de cada unidad, de tal manera que las ponderaciones variarán entre unidades en función de cuales sean sus niveles de outputs e inputs. No obstante, la lógica económica llevará a que en numerosas ocasiones sea necesario restringir dichas ponderaciones⁶. Este procedimiento de restricción es ilustrativo de una de las prácticas habituales en el contexto público, consistente en la introducción de *juicios de valor* en la fijación de ponderaciones con vistas a la agregación, tan vez en función de criterios alejados de cuestiones de eficiencia y/o justicia distributiva.

3. SERVICIOS PÚBLICOS

En el contexto público, el abanico de servicios es muy amplio: *servicios “preferentes”* -sanidad, educación, etc.-, servicios de *seguridad pública* -bomberos, policía, etc.-, servicios de *utilidad pública* -suministro de luz, agua, etc.-, servicios de *apoyo* -recogida de basuras, limpieza, información, etc.-, *comunicaciones* -transporte público, construcción y mantenimiento de carreteras- *servicios culturales y recreativos* -museos, bibliotecas, etc.- pueden ser buenos ejemplos de ello. En este capítulo se analizarán varios servicios cuya importancia en términos de gasto puede apreciarse en el *cuadro 2.1*.

⁶ Si las ponderaciones varían sin ningún tipo de límite, se estará introduciendo, indirectamente, cierto grado de heterogeneidad. Piénsese, por ejemplo, en una unidad productiva que, en teoría, y al igual que el resto de unidades contenidas en una muestra determinada, debería producir dos outputs, puesto que ambos se consideran relevantes; sin embargo, supóngase que dicho centro decide especializarse en uno de ellos: en el extremo, decide destinar todos sus inputs a producir uno de los outputs, descuidando casi por completo el otro. En esta situación, sería deseable “penalizar” en términos de eficiencia al centro productivo, ya que el producto descuidado podría tener la misma importancia –o incluso superior- que el producido masivamente. Tal como señalaban Charnes et. al. (1990) la construcción de una aproximación a la función de producción puede verse derrumbada por la ausencia de variables realmente relevantes. De ahí que, entre las numerosas extensiones experimentadas por el modelo DEA, haya surgido la idea de establecer límites a las ponderaciones asignadas por el sistema, evitando excluir del análisis variables importantes.

CUADRO 2.1

Gasto público como porcentaje del Producto Nacional Bruto en algunos países de la OCDE (1997): sanidad, educación y seguridad

	<i>SANIDAD</i>	<i>EDUCACIÓN</i>	<i>SEGURIDAD Y ORDEN PÚBLICO</i> ⁷
Australia	3.1	3.6	1.3
Austria	5.2	4.3	0.9
Bélgica	0.5	6.1	1.6
Dinamarca	5.0	5.6	1.0
Francia	3.4	5.2	0.9
Alemania	6.3	3.7	1.6
Islandia	6.6	4.0	1.3
Italia	3.4	4.2	1.7
Nueva Zelanda	2.9	3.8	1.5
Noruega	4.9	5.5	0.9
Portugal	3.4	5.4	2.0
España	3.9	3.1	1.2
Suecia	4.8	5.3	1.5
Reino Unido	5.7	4.5	2.0

Fuente: Adaptado de OCDE (1999), *OCDE in Figures*. París.

A continuación se presentarán medidas de output en los servicios sanitarios, educativos y de protección. En todos ellos se discutirá, en primer lugar, el carácter heterogéneo del servicio, para pasar seguidamente al diseño de índices de producto final e intermedio.

3.1 Servicios sanitarios

Los servicios sanitarios constituyen un sector de actividad de características únicas, tanto por sus dimensiones en relación con las economías de cada país como por la propia naturaleza de su actividad de restauración de la salud. El objetivo de los mismos consistirá en el mantenimiento o la obtención de mejoras en la *salud* de los individuos, definida esta por la Organización Mundial de la Salud como un “*estado de completo bienestar físico, mental y social*”.

Constituyen servicios caracterizados por un alto grado de heterogeneidad, tanto en relación a la tipología de los centros de suministro, como a los casos que estos tratan. De ahí que, de forma previa a la elección y diseño de variables de producto sea preciso clasificar dichos casos para reflejar la casuística⁸ -*case-mix*- del centro sanitario. Los

⁷ El cómputo excluye gastos en defensa.

⁸ Feldstein (1967) mostró que las diferencias en la casuística explicaban un 27,5 % de las variaciones del coste por caso para una muestra de hospitales británicos. No obstante, es posible que dicho valor fuese subestimado al establecer un número pequeño de categorías (Wagstaff, 1989b).

criterios de clasificación son diversos: el departamento de admisión del paciente [Feldstein, (1967), Wagstaff (1989a)], la patología -*Clasificación Internacional de Enfermedades*-, a lo que podría añadirse el consumo de recursos -*Grupos Homogéneos de Diagnóstico*⁹, en función de la dolencia y el coste del tratamiento-. Una vez efectuada la clasificación, es posible definir variables de output referidas a cada una de las categorías de servicios o tratamientos, o elaborar algún índice que refleje la complejidad del hospital. En esta última dirección se encaminan los trabajos de Feldstein (1967) y Evans y Walker (1972), en los que se diseñan índices ilustrativos de la heterogeneidad del centro sanitario. En ambos casos, se trata de medidas sintéticas obtenidas a partir de las proporciones de pacientes atendidos en los diferentes servicios o categorías de diagnóstico suministradas por el hospital (González y Villalobos, 1993).

Output final

Una vez realizada la clasificación de los casos, será necesario determinar la variable más adecuada para aproximar el producto sanitario. La literatura recoge, por lo general, medidas de una de las dimensiones de las actividades sanitarias. Nos referimos a las mejoras del estado de salud de los pacientes. Puesto que constituye la actividad más importante, será el centro de atención en la construcción de índices de producto. De todas formas, determinados indicadores de producto final, al considerar la variable salud, estarán reflejando a la vez actividades curativas, preventivas e incluso de investigación.

En los años 70 comienza a desarrollarse una corriente teórica enfocada a la construcción de índices de salud, que podrían ser interpretados como medidas de producto sanitario final, no sin ciertas precauciones¹⁰. Su construcción está basada en la agregación de los años de vida ponderados por su calidad. Para Torrance (1986), esta aproximación representa un importante instrumento de medición de los programas sanitarios, en especial cuando se diseña en términos de valor añadido (Williams, 1985). Con los métodos basados en un enfoque de demanda y funciones de utilidad, se proporciona un método para incorporar formalmente consideraciones relativas a la calidad de vida –a parte de la cantidad-. Así, no parece adecuado omitir aspectos cualitativos, empleando otros índices bastante más incompletos¹¹, argumentando dificultades de medida y de construcción. Se trata de medidas subjetivas, que ponderan en mayor medida la perspectiva del paciente que los datos clínicos objetivos. La nota

⁹ Los *Grupos Homogéneos de Diagnóstico* (Fetter *et al.*, 1980) fueron el fruto de una investigación iniciada a mediados de los años 60 en la Universidad de Yale y concluida a principios de los 80. Similar orientación tienen las *Patient Management Categories* (Young *et al.*, 1982).

¹⁰ La mayor parte de la literatura les otorga la categoría de medidas de utilidad [Torrance, (1976); Pliskin *et al.* (1980)]. Si bien el hecho de considerarlo como una medida de producto o de utilidad no tiene ningún tipo de repercusiones a nivel individual, puede tener cierta enjundia a nivel agregado o social. Así, si se desean agregar los índices individuales, y se interpretan como medidas de utilidad, surgirá el problema de las comparaciones interpersonales de utilidad, cuestión que no se plantearía si fuesen considerados como índices de producto. En este sentido, Johannesson M. (1999), propone un método de agregación de índices de salud construyendo ponderaciones a partir de funciones de utilidad en las que se incluye como argumento los estados de salud del resto de individuos de la sociedad. Considera, por tanto, las variaciones en la salud propia y de los demás. Dolan (1998) había realizado un ejercicio similar, teniendo en cuenta en el cálculo de las ponderaciones únicamente el último tipo de variaciones.

¹¹ En relación a la medición del producto final sanitario, se ha planteado la utilización de las tasas de mortalidad. Así, Auster *et al.* (1969) emplearon dicho indicador, ajustado por edades. El problema principal de este índice es que no considera la severidad ni la complicación de la enfermedad.

común a todos los métodos englobados en este enfoque radica en la inclusión de las valoraciones de los diferentes estados de salud.

Culyer *et al.* (1971) resaltaban la necesidad de articular el diseño de estos índices en varias etapas, que comprenderían, fundamentalmente la definición de salud y sus dimensiones, así como el cálculo de valoraciones correspondientes a cada estado de salud. En este sentido, Torrance (1986) señalaba tres formas posibles de cálculo de dichas valoraciones. Una primera modalidad se centra en la construcción de escalas graduales que recogen las diferencias relativas entre los diversos estados de salud. El procedimiento a seguir para su diseño precisa, como primer requisito, de la definición de funciones y capacidades asociadas a la buena salud, y al resto de situaciones, introduciendo, asimismo, valoraciones individuales de cada uno de ellas, comprendidas, en principio, en una escala entre 0 –muerte- y 100 –perfecto estado de salud-, pudiendo ser reescalado a ponderaciones comprendidas entre 0 y 1. Este procedimiento es adoptado por Rosser y Kind (1978), cuya matriz de valoraciones se recoge en el *cuadro 2.2*. Estos autores describieron los estados de salud en función de dos parámetros: discapacidad y dolor¹². Dichas descripciones determinan un total de 29 estados diferentes, cada cual recibe una valoración¹³: algunos son valorados incluso con ponderaciones inferiores al asignado a la muerte. Este trabajo no es el único que describe y valora los estados de salud. Existen otros estudios mucho más detallados y complejos, en relación al número de estados descritos y al tamaño de la muestra¹⁴.

¹² La definición anteriormente señalada, atribuida a la Organización mundial de la Salud, permite apreciar las diferentes dimensiones que conviven en el concepto de salud. La salud no solo consiste en la ausencia de enfermedad, sino que existen aspectos adicionales a tener en cuenta. Así, Culyer *et al.* (1971) aluden a dimensiones como la intensidad del dolor y duración de la enfermedad. No obstante, es necesario considerar adicionalmente el estado emocional del individuo, así como los obstáculos que la enfermedad genera a la actividad social, a la movilidad y a la capacidad física (Rosser y Watts, 1978).

¹³ La matriz recoge valores medios de las diferentes partes implicadas en las encuestas (pacientes, doctores, etc.), un total de 70 encuestados. La pregunta a responder era la siguiente: ¿cuántas veces es peor el estado X en relación al estado de referencia Y?. Para mayor detalle, véase Rosser y Kind (1978).

¹⁴ A este respecto, véase Culyer, A.J. y Maynard, A. (1997).

CUADRO 2.2
Valoraciones de diferentes estados de salud

	DOLOR / ANGUSTIA			
DISCAPACIDAD	Inexistente	Leve	Moderado	Fuerte
Ninguna discapacidad	1,0	0,995	0,99	0,97
Ligera discapacidad social	0,99	0,986	0,973	0,932
Severa discapacidad social	0,98	0,972	0,956	0,912
Elección laboral seriamente limitada	0,964	0,956	0,942	0,87
Incapacidad de obtención de un empleo	0,946	0,935	0,9	0,7
Confinado a permanecer sentado	0,875	0,845	0,68	0,0
Confinado a permanecer en la cama	0,677	0,564	0,0	-1,486
Inconsciencia	-1,028	-	-	-

1 = perfecta salud 0 = muerte

Fuente: adaptado de Rosser, R. y Kind, P., (1978) "A Scale of Valuations of States of Illness: Is There a Social Consensus?", *International Journal of Epidemiology*, 7(4), págs. 347-358.

Existirían formas alternativas de valoración de las patologías. Una posibilidad consistiría en considerar el riesgo como medida de valor. La versión básica de este procedimiento establece alternativas con probabilidades asociadas *-standard gamble-*. Para ilustrar el método, supongamos que un individuo tiene dos posibilidades de actuación: puede, por un lado –primera alternativa-, someterse a un tratamiento médico para alcanzar un perfecto estado de salud, con cierto riesgo –probabilidad $(1-p)$ - de perecer durante la aplicación del mismo o, por otro lado –segunda alternativa -, permanecer en su actual estado de salud –estado x - con total seguridad. En estas circunstancias la probabilidad p reflejaría la valoración del estado x por parte de dicho individuo.

Otra metodología de valoración, el sistema de trade-off temporal, al igual que el anterior, se basa en la comparación de pares de condiciones, aunque su aplicación es más adecuada en contextos en los que los resultados de los tratamientos médicos presentan alto grado de certidumbre. Si el individuo se declara indiferente entre vivir durante n_p años en perfecto estado de salud o vivir durante un periodo temporal superior, n_x , en un estado de salud peor, x , la valoración de dicho estado se obtendrá como el cociente $\frac{n_p}{n_x}$.

No existe un consenso respecto al método a utilizar en la valoración de los estados de salud. Hay quien opina que el método más apropiado es el juego standard, puesto que es el que más se acerca a los axiomas de la teoría de la utilidad esperada (Torrance y Feeny, 1989). Otros, por el contrario, se decantan por el trade-off temporal (Richardson, 1994) o por la escala (Broome, 1993). Generalmente, el juego standard suele obtener ponderaciones superiores al resto de metodologías, al introducir aversión al riesgo en relación a los diferentes estados de salud (Culyer y Wagstaff, 1993). La cuantía inferior para las ponderaciones suele ser objeto de las escalas graduales. Las diversas modalidades de estos índices –*Quality Adjusted Life Years* (QALYs) y *Healthy Years Equivalents* (HYEs)- emplean las tres variantes¹⁵.

Output intermedio

Constituyen índices más manejables, pero con ellos se pierde precisión. Su relación con el producto final –el impacto en la salud- se percibe como incierta. Así, el *número de casos* es una medida empleada frecuentemente en la literatura. Feldstein (1967) señaló su mayor adecuación como proxy del output hospitalario, puesto que el centro sanitario trata pacientes, no maximiza estancias. Así planteada, esta medida no considera diferencias entre casos atribuidas a la duración de la estancia. En ocasiones, se incluye el *número de consultas* como un índice de presión asistencial. Incluso podría ayudar a aproximar aspectos cualitativos, mediante el cálculo del porcentaje de consultas programadas sobre el total, indicativo de posibles retrasos en la prestación del servicio.

Otra opción consistiría en tratar de reflejar los casos resueltos. La proxy comúnmente empleada a estos efectos es el número de *altas*. Numerosos trabajos aplicados consideran esta variable como un índice de output [Prior y Sola (1993); Ehreth (1994)]. No obstante, esta interpretación falla en diversas ocasiones, lo cual pone en duda su capacidad como indicador de producto. Así, el hecho de que aumenten las altas puede venir forzado por falta de capacidad, de tal forma que podrían darse de alta casos sin estar totalmente resueltos, por necesitar las camas para pacientes en situación de mayor gravedad. En este sentido, el porcentaje de reingresos por altas podría ser un indicador adecuado para aproximar la calidad, a emplear junto a otros índices, como el porcentaje de traslados a otros centros, o el porcentaje de revisiones.

¹⁵ Los QALYs suelen aplicar el método de trade-off temporal (Pliskin *et al.*, 1980) o de escalas (Williams, 1985), mientras que los HYE, recurren, por lo general, a probabilidades (Gafni y Mehrez, 1991).

La *duración de la estancia* es una medida menos empleada, al menos de forma aislada, puesto que tan solo recoge los servicios de alojamiento por parte del centro hospitalario, una de las múltiples dimensiones del producto sanitario. Por lo general suele emplearse junto con otras variables, como el número de casos (Steele y Gray, 1982) o las consultas externas (Prior y Solá, 1993). La relación entre mejoras en la salud y días de hospitalización no parece clara; es un índice que puede crear incentivos perversos, si se establece como un objetivo a maximizar. En este sentido, puede conducir a la generación de estancias excesivamente largas e innecesarias. Es más, las tendencias actuales parecen considerar más eficiente la reducción de la duración de las estancias, que su alargamiento, por considerar este último hecho como sintomático de la existencia de ineficiencias.

3.2 Servicios educativos

La educación ha sido definida como un proceso de transformación de cantidades fijas de inputs en individuos con cualidades distintas (Hanushek, 1971). Al igual que sucedía con la variable salud, el resultado final sobre la educación se verá fuertemente condicionado por factores exógenos diversos. Cada individuo es un mundo, en el que interaccionan una multiplicidad de fuentes educativas, y no solo los centros educativos oficiales¹⁶.

Podrían diferenciarse dos enfoques respecto al resultado que pretenden alcanzar los procesos educativos. La elección condicionará la medida de output elegida. Una de las corrientes teóricas otorga a la educación el carácter de un mero *filtro* [Arrow, (1973); Spence (1973)], actuando como un instrumento de discriminación o criba. En este caso, el proceso educativo no contribuye a alterar las capacidades individuales, sino que constituye una señal de las mismas. No obstante, parece una opción bastante más aceptada en la literatura la interpretación de la educación como una *inversión en capital humano* [Schultz (1961); Becker (1964)]. De hecho, la educación posee características atribuibles a los bienes de consumo y de inversión¹⁷. Según esta teoría, la educación contribuiría a mejorar la capacidad productiva del educando, incrementando la probabilidad de obtener rentas –superiores– en periodos futuros. En estos casos, la medida del output deberá ir encaminada hacia los resultados del proceso educativo.

Los centros educativos producen, fundamentalmente, dos tipos de outputs: educación e investigación. Este último tipo de producto es atribuido a las unidades superiores de enseñanza. Así, en las enseñanzas básicas, el foco de atención radicará fundamentalmente en conocimientos genéricos y el desarrollo psicomotriz del individuo, mientras que en enseñanzas superiores, y, en cierta medida, secundarias, primará la difusión de conocimientos específicos y la investigación. Por otro lado, la

¹⁶ Así, una parte importante de la educación proviene de focos externos a las instituciones oficialmente dedicadas a su provisión, tales como la familia, las empresas, las relaciones sociales o los medios de comunicación. Es la llamada educación informal, que junto a la educación adquirida en los centros oficiales configuran el producto educativo. Bridge *et al.* (1979) establecieron la distinción entre producto escolar, en referencia al producto exclusivamente atribuible al centro, y producto educativo, mucho más amplio, haciendo alusión al output aportado por todo tipo de experiencias educativas.

¹⁷ Así, los rasgos de bien de consumo hacen referencia al disfrute derivado de la educación recibida, mientras que los de inversión conectarían con la teoría del capital humano, pues implicarían el incremento de la capacidad productiva del individuo derivada de dicha educación (Schultz, 1961). Esta diferenciación se halla estrechamente ligada a los productos a corto y largo plazo derivados de la educación.

enseñanza básica y la secundaria, pueden concebirse como un output intermedio a generar para acceder a estudios superiores. Así se pronunciaba Becker (1964), señalando que el capital humano acumulado se convierte en un input para producir unidades adicionales de capital. En todos los casos, se tratará de recoger medidas de producto intermedio y final que incluyan aspectos cuantitativos y cualitativos.

Output final

El enfoque de la educación como inversión continua en capital humano da pie a la construcción de medidas de output final, en función de los resultados deseables o el impacto de los procesos educativos. En este sentido, la educación es un medio de alcanzar fines más ambiciosos, ya sean pecuniarios o no pecuniarios¹⁸. Entra en juego, por tanto, el mercado de trabajo, es decir, hasta que punto la educación implica ventajas a la hora de encontrar un empleo y obtener rentas. Este enfoque considera el output como el rendimiento derivado de la inversión en capital humano. Desde este punto de vista, la educación sería una condición necesaria pero no suficiente de acceso al mercado de trabajo y a mejores niveles de ingresos. En este último caso estaríamos hablando de efectos no inmediatos derivados de la educación que habrán de ser tenidos en cuenta con precaución.

Las medidas del producto formativo final estarían estrechamente ligadas con las características de los puestos de trabajo a los que se podría acceder de contar con la formación aportada por los centros educativos. Se suele tomar como referencia el primer destino laboral de los graduados, valorando y ponderando sus principales aspectos -remuneración, estabilidad, peligrosidad, etc-, en términos absolutos o relativos. Así, Johnes y Taylor (1990) definen varios indicadores, reflejando diferentes niveles de éxito de los graduados. En su aplicación a la Educación Superior en el Reino Unido, emplean la información ligada al primer destino, clasificando a los graduados en función de su situación laboral¹⁹.

Este tipo de medidas ha recibido diversas críticas, tanto teóricas como en el campo aplicado. El principal problema de este enfoque radica en que el diferencial de ingresos o ventajas derivada del acceso a puestos de trabajo superiores podría ser atribuida a otros factores añadidos al margen de la educación: la existencia de habilidades innatas, la suerte, raza, sexo, o características del mercado de trabajo serían algunas de ellas. Taubman (1976) concluye que la educación no constituye el principal determinante de las desigualdades económicas. Además, a la hora de abordar el análisis se suele tomar

¹⁸ Habría que destacar beneficios no monetarios tales como el prestigio social asociado a la educación, el afán de superación, etc. (Cohn y Millman, 1975).

¹⁹ Las categorías, recogidas por el *Careers Advisory Service*, y empleadas en la construcción de los indicadores serían las siguientes: A. Graduados con trabajo fijo; B. Graduados que continúan sus estudios; C. Graduados en paro; D. Graduados con trabajo temporal. Partiendo de estos datos, construyen los siguientes indicadores:

$$\begin{aligned}
 & \text{- Porcentaje de graduados con trabajo} = \frac{A + D}{A + B + C + D} \times 100 \\
 & \text{- Porcentaje de graduados con trabajo fijo} = \frac{A}{A + B + C + D} \times 100 \\
 & \text{- Porcentaje de graduados que continúan sus estudios} = \frac{B}{A + B + C + D} \times 100
 \end{aligned}$$

como referencia el primer empleo, procedimiento que no muestra la posibilidad de la consecución de un trabajo mejor considerando un horizonte temporal algo más amplio.

En relación a la actividad de investigación, es desarrollada fundamentalmente por centros educativos superiores, los cuales compatibilizan dicha actividad con labores docentes, estando poco clara, en ocasiones, la línea divisoria entre las dos funciones²⁰. Harris (1989), a parte de incluir aspectos de cantidad y calidad de la investigación, hacía referencia al impacto de la misma como la influencia real de los resultados científicos en actividades de investigación próximas o relacionadas. Luego esos tres aspectos habrán de ser considerados a la hora de diseñar variables de producto en este campo. Una suma de los trabajos de investigación realizados por el centro, ponderada en función de los méritos relativos asociados a las diferentes formas de publicación (Smith and Fielder, 1971) podría constituir una buena medida del output investigación. Asimismo se plantean medidas alternativas que aproximarían la calidad y el impacto del trabajo de investigación, tales como la frecuencia en la que se citan las publicaciones de miembros del centro o departamento, o el conjunto de subvenciones captadas por el mismo.

Output intermedio

Partiendo de la consideración del producto educativo como output intermedio, o requisito necesario para el acceso a etapas posteriores, podrían definirse una serie de índices de producto. Estos son frecuentemente empleados en educación básica y media, aunque también se recurre a ellas en niveles superiores. Se caracterizan por su parcialidad, pues suelen centrarse lo más posible en los aspectos atribuibles propiamente a los diferentes centros educativos.

Así, el *número de alumnos* representa la dimensión cuantitativa del producto educativo. Esta variable constituye una expresión física del producto de los centros educativos, que ofrece como ventaja principal la simplicidad de su diseño. Aunque algunas aplicaciones empíricas en enseñanza superior se decantan por medidas absolutas, parece más adecuada su utilización en términos relativos, como el porcentaje de alumnos que hayan superado el curso en relación a los matriculados. En ocasiones, se emplea la *tasa de abandono* como medida de output, reflejando el enfoque del filtro en relación a la educación²¹. Entre los principales inconvenientes de este tipo de medidas figuran la no consideración de aspectos cualitativos -tiempo empleado en la obtención de la graduación, calificaciones obtenidas, etc-. De ahí que esta variable, no se emplee de forma aislada, sino acompañada de medidas de calidad en la graduación, tal como sucede en los trabajos de Gray *et al.* (1984) y Walberg y Fowler (1987), entre otros.

Por su lado, la calidad del output educativo puede ser aproximada a través de las *calificaciones* del alumnado²². No obstante, no faltan opiniones críticas en relación a la

²⁰ En este sentido, la línea divisoria entre investigación y educación no es nítida en estudios de postgrado.

²¹ Así, esta orientación es recogida nítidamente en la aplicación empírica de Strauss y Sawyer (1986).

²² Indicadores alternativos de la calidad de los procesos y productos educativos pueden construirse a partir de encuestas del alumnado sobre el profesorado. En el caso de las enseñanzas básicas, sería necesario recurrir a las opiniones de los padres.

misma, advirtiendo que un especial énfasis en los resultados de los exámenes puede desviar la atención de aspectos no menos importantes, de carácter no cognitivo, como son las facultades personales y sociales²³. Esta claro que son medidas imperfectas, pero son aquellas que aproximan de la mejor manera posible la producción educativa (Hanushek, 1986). En este sentido, lo más adecuado sería incorporar las puntuaciones de aquellas asignaturas específicas impartidas por los centros, puesto que, en opinión de Madaus *et al.* (1979), los test de desarrollo intelectual general no miden la producción directamente atribuible a los centros educativos, sino sus efectos secundarios. El objetivo concreto de los mismos se plasma habitualmente en la transmisión de los conocimientos específicos de las asignaturas del plan de estudios y que es por esa función por la que se les va pedir responsabilidades. De todas formas, no faltan los partidarios de la inclusión de materias básicas como reflejo de conocimientos necesarios para el desenvolvimiento de tareas individuales y colectivas en una sociedad moderna.

En la mayor parte de los trabajos empíricos se emplean las puntuaciones agregadas, ya sea referidas a una materia o a varias. Al considerar diversas asignaturas, la construcción de un índice que incluyera todas ellas haría ganar operatividad. En este sentido, se podría contemplar una suma ponderada, fijando las ponderaciones teniendo en cuenta la futura orientación académica de los alumnos. En otras ocasiones se definen medidas de valor añadido, como diferencia entre las calificaciones de un examen realizado a principios de curso o etapa y las obtenidas en un examen final [Cohn, (1968); Summers y Wolfe, (1977); San Segundo, (1985); Ribera *et al.* (1998)]. El empleo de puntuaciones medias en las diferentes asignaturas del curriculum aparece con menor frecuencia: considerada de forma aislada, puede contener desigualdades, especialmente problemáticas en las enseñanzas básicas. De ahí que sea necesario incorporar medidas de dispersión como la desviación típica (Brown y Saks, 1975).

3.3 Servicios de protección

Generalmente, cuando se habla de servicios de seguridad, el producto final se identifica con una reducción de la aparición de sucesos desfavorables. En este sentido, el output tendrá carácter negativo (Shoup, 1969), puesto que el objetivo final de dichos servicios consistirá en evitar ese tipo de circunstancias. En este apartado se analizarán, concretamente, dos servicios públicos que responden a ese perfil: las actividades policiales y los servicios de protección y prevención contra incendios. En ambos casos, se pretende incrementar los niveles de seguridad de los ciudadanos en diversos ámbitos.

²³ Mientras que en la introducción de conocimientos parece existir cierto grado de consenso, no sucede lo mismo en relación a aspectos no cognitivos –actitudes, motivación, autoestima, etc-. Por un lado, algunos autores [Averch *et al.*, (1974); Johnson, (1974)] les otorgan una importancia significativa, incluso superior al peso de factores cognitivos (Gintis, 1971). Argumentan que estas cuestiones constituyen un producto adicional ofrecido por los centros educativos, y que será necesaria su consideración en las medidas de output (Levin, 1974). Las aplicaciones empíricas de Cohn y Millham (1975), o Bordman *et al.* (1977), incorporan como variables de producto algunas dimensiones no cognitivas, como la motivación, las expectativas, la autoestima o la creatividad. Todos coinciden en que estos factores influirán en gran medida sobre el futuro socioeconómico del alumno. Otros, sin embargo, afirman que los componentes afectivos no atraen la atención de los diferentes agentes implicados en el proceso educativo –padres, alumnos, etc.-, los cuales parecen más preocupados por los resultados académicos (Madaus *et al.* 1979).

Output final

Shoup (1969) afirmaba que la producción policial, "en sí misma no es agradable de consumir o recibir, pero no obstante se valora porque impide o debe impedir algo aún más desagradable". Generalizando, las medidas de output final diseñadas en el campo de la prevención y protección consisten en la reducción de siniestros –delitos, incendios-. Suelen construirse, por tanto, índices que consideran dichos sucesos en relación a la población de riesgo. En el caso de los servicios policiales, se intenta reflejar la reducción en el nivel de delitos cometidos, así como la disminución de la probabilidad de ser víctima. Su construcción respondería a la siguiente expresión:

$$\begin{array}{ll} \textit{Tasa de delitos registrados} & \textit{Tasa de victimización} \\ TD_R = \frac{D_R}{P} & TV_R = \frac{V_R}{P} \end{array} \quad (3.1)$$

siendo D_R el número de delitos registrados por las instituciones policiales, V_R el número de víctimas registradas y P el tamaño de la población en ambos casos. El empleo de variables registradas constituye una aproximación a siniestros realmente acontecidos, procedimiento no exento de problemas²⁴. Similar orientación podría ser aplicada en el caso de los servicios de protección contra incendios²⁵. Suele considerarse a esta como la principal actividad del cuerpo de bomberos, luego una medida válida consistirá en calcular el *cociente del número de incendios entre la población*.

No obstante, las medidas de output final hasta ahora planteadas presentan el inconveniente de no considerar la *gravedad* de los siniestros. Con la finalidad de reflejar la severidad de los casos, se propone la *valoración monetaria de las pérdidas* asociadas a los mismos. De esa forma proceden Duncombe y Yinger (1993), en su aplicación a los servicios de protección y prevención de incendios, considerando el peso del valor de las pérdidas por siniestros (L) en relación al valor total de las propiedades (V)²⁶. Esta

²⁴ Dicha aproximación ha hecho que numerosos autores cuestionen la validez de estas medidas, debido a la aparición de "efectos desplazamiento" (Mehay, 1977) y "efectos denuncia" (Carr-Hill y Stern, 1979). Con los primeros se hace referencia a la modificación de la conducta del delincuente con la finalidad de reducir la probabilidades de ser detenido. Esto llevaría a que las tasas de delitos podrían reducirse sin que ello significara una menor incidencia de la criminalidad. Respecto al segundo tipo de efecto, implica cambios de actitud de los ciudadanos para informar de los acontecimientos delictivos; su aparición determinará la calidad de la aproximación efectuada, influyendo directamente en la brecha existente entre los delitos denunciados y los delitos cometidos. Por tanto, un incremento en el número de delitos registrados puede significar bien que la gente comete más crímenes o bien que se denuncian más hechos delictivos que antes. Es decir, que un aumento de los siniestros registrados puede ser consecuencia de un mayor número de denuncias por parte de los ciudadanos, si que ello suponga un crecimiento de los delitos cometidos.

²⁵ En la mayor parte de las aplicaciones de los dos tipos de servicios, se opta por invertir las tasas comentadas anteriormente, con el propósito de expresar la medida del output en términos positivos.

²⁶ Duncombe y Yinger (1993) descomponen el índice anteriormente señalado, en el producto de dos subíndices, recogiendo a un tiempo actividades de supresión y prevención:

$$\frac{L}{V} = \frac{L}{F} \cdot \frac{F}{V}$$

donde F es el número de incendios. Así, una reducción de las pérdidas podría deberse a una disminución en las pérdidas por incendio, resultado de los servicios de supresión –primero de los subíndices- o a una disminución del número de incendios consecuencia de los servicios preventivos –segundo de los subíndices-. No obstante, este indicador es parcial, al introducir solamente pérdidas materiales y no personales.

alternativa de valoración también presenta obstáculos. Consideremos, en el caso de los servicios policiales, las pérdidas asociadas al robo. Así, Becker (1968) afirma que la ganancia obtenida por el ladrón es equivalente a la pérdida de la víctima. Sin embargo, Tullock (1967) calcula la pérdida social como dos o tres veces superior a la cantidad robada, puesto que tanto la víctima como el delincuente destinan recursos para evitar el éxito de la otra parte. Lind y Lipsky (1971) piensan que los bienes robados no pueden distribuirse por canales legales y experimentan reducciones de valor. El problema se complica en el caso de los delitos contra las personas, a pesar de la ayuda proporcionada por disciplinas tales como la economía de la salud: así, Shoup y Mehay (1971) recomiendan tener en cuenta tanto los costes de tratamiento médico como las ganancias que se dejan de obtener. No obstante, Williams y Anderson (1975) señalaban que dicha valoración deja fuera costes relevantes para las víctimas como el dolor y el sufrimiento y los sentimientos de indignación ante la violación de las normas. Existiría una vía alternativa para tener en cuenta la gravedad de los siniestros. Consistiría en ponderarlos en función de su importancia, ponderaciones basadas en juicios de valor²⁷. El resultado ofrece una escala de siniestros ordenados según su severidad. Este método, introduce alto grado de subjetividad, aspecto que ha acarreado las críticas más importantes.

Output intermedio

Las dificultades asociadas a la medición del output final llevan a la elección de medidas de output intermedio. Una de las deficiencias comunes a todas ellas consiste en que centran su atención en las actividades de represión y protección, dejando un tanto marginadas las actividades preventivas. En el caso de los servicios policiales, el número de arrestos practicados es una aproximación frecuentemente empleada. En este sentido, la *tasa de arrestos* muestra las detenciones practicadas por la policía en relación al número de delitos registrados. Este índice, tal como está construido, -el numerador refleja individuos y el denominador incidentes- puede llevar a una sobrevaloración del output, ya que en el caso de arrestos múltiples por un mismo incidente aumenta la tasa de arrestos, pero no los delitos solucionados.

Un indicador alternativo para reflejar el producto derivado de la actuación policial es el *número de casos resueltos* o esclarecidos en relación a los delitos registrados [Carr-Hill y Stern (1979), Levitt y Joice (1987)]. En ambos casos, la principal debilidad radica en la existencia de un desfase temporal entre el numerador y el denominador, en la medida en que los arrestos o los casos resueltos durante un periodo no tienen por qué coincidir con los delitos registrados en ese periodo. Este rasgo puede dar lugar a que en ocasiones la tasa se sitúe en torno al 100% o en un nivel superior, sin que ello signifique una buena actuación. Adicionalmente, ni a través de la tasa de arrestos ni de la tasa de delitos esclarecidos se informa sobre la calidad, y por ello, algunos autores se inclinan por considerar como medida de output intermedio la tasa de condenas²⁸, ya que, aunque supone una combinación del output policial y del output judicial, permite conocer si la actuación policial ha dado como resultado una condena. De todas formas, sería posible incluir algún indicador de calidad compatible con los índices anteriormente señalados,

²⁷ Para el caso de los servicios policiales, pueden considerarse las opiniones de las víctimas potenciales (Sellin y Wolfgang, 1964) o los juicios valorativos de los miembros de la policía (Bryant *et al.*, 1968).

²⁸ Consúltense las aplicaciones de Carr-Hill y Stern (1979) o Willis (1983), entre otros.

como el tiempo de respuesta²⁹, la capacidad para disponer de suficiente evidencia probatoria para condenar al presunto delincuente³⁰ o alguna medida construida a partir de valoraciones de los ciudadanos.

El producto suministrado por los servicios de protección contra incendios y salvamento suele ser aproximado mediante el *número de intervenciones*, en sus diversas variantes, puesto que constituyen la principal actividad de estos servicios. De ahí que su número constituya una medida proxy del output utilizada con frecuencia. De todas formas, este indicador es criticado por no ofrecer una relación nítida con los niveles de seguridad. Así, Cuenca (1994), señalaba que un mayor número de siniestros declarados, aumenta el riesgo de que un ciudadano se vea afectado, y, por consiguiente, menor es la protección. Pero por otro lado, los siniestros de poca significación no serán declarados, luego se registrará un menor número de intervenciones. En este último supuesto, menos intervenciones no implican un aumento de la protección.

4. CONCLUSIONES

La medición del output en el sector público ha sido, sin duda, una actividad en expansión, intensificada en la década de los 80. Es un aspecto que está irremediamente ligado a las nuevas formas de funcionamiento organizativo en el sector público. Sistemas retributivos en función de la productividad, fórmulas presupuestarias basadas en objetivos y resultados, así como diversos mecanismos de control y evaluación de las actividades públicas precisan de instrumentos válidos para su buen funcionamiento, y la definición de medidas precisas de producto constituyen una condición necesaria, aunque no suficiente para lograrlo.

Tras el análisis de la medición del output en servicios públicos como la sanidad, la educación o la seguridad, parece claro que dicha tarea no se haya exenta de dificultades, comenzando por el establecimiento de los fines que persiguen las actuaciones públicas y continuando por la construcción de indicadores. Tal como señalaba Hirsch (1973), *“frecuentemente, es extremadamente complejo definir y medir unidades básicas de output, así como sus características cualitativas (...)”*. El sector público ha de aspirar a satisfacer numerosos intereses, frecuentemente contradictorios. Se ha hecho alusión a una función objetivo cambiante temporal y espacialmente. De forma adicional, las organizaciones públicas no han de limitarse a suministrar un output intermedio, sino que han de perseguir la consecución de resultados, procurando que el impacto de dichas prestaciones sea lo más beneficioso posible. Esta diferenciación existente entre output final e intermedio se ha puesto de manifiesto a lo largo de todo el artículo, siguiendo el esquema propuesto, entre otros, por Bradford *et al.* (1969) a este respecto. Asimismo, se han mostrado las mayores deficiencias de operatividad que presentan los índices de

²⁹ Véase, por ejemplo, en Ballart (1990).

³⁰ Greenwood, Chaiken y Petersilia (1977) consideran que la calidad del arresto vendrá dada por la existencia de evidencia suficiente para posteriormente acusar al delincuente. De esta forma, construyen un "índice de calidad" calculado en función de la información proporcionada a la acusación procedente de los informes del delito y del arresto, y de los interrogatorios a las víctimas, testigos y sospechosos en el momento del incidente, del arresto y de la investigación posterior.

output final. En la práctica, se opta por proxies del mismo, de ahí que sean utilizadas con mayor frecuencia medidas de output intermedio. No obstante, deben ser empleadas conociendo sus limitaciones, que se reducen, principalmente, a la parcialidad y la posibilidad de manipulación. En cualquier caso, se rechaza la alternativa de aproximar el producto a través de parte o la totalidad del gasto en que incurren los centros productivos.

Una nota común a todas las actividades radica en la inclusión de aspectos cualitativos. Se ha definido, por tanto un proceso de suministro público como un conjunto de etapas, que, partiendo de la utilización de factores, se prolonga hasta el nivel de bienestar de los individuos. Las opiniones de los usuarios en relación a la prestación del servicio cobran importancia en el contexto público. En este sentido, el lado de la demanda resulta claramente fortalecido, pues hay un cambio en la orientación de las políticas públicas de suministro. Intentando reflejar este aspecto se ha hecho alusión a índices basados, en mayor o menor medida, en las valoraciones de los usuarios.

REFERENCIAS BIBLIOGRAFICAS

Arrow, K.J. (1973) "Higher education as a filter", *Journal of Public Economics*, 2(3), págs. 193-216.

_____, (1974) "The measurement of real value added", en David, P. y Reder, M. (eds.), *Nations and Households in Economic Growth: Essays in honor of Moses Abramovitz*. Academic Press. New York.

Arvidsson, G. (1986) "Performance Evaluation", en Kaufmann, F., Majone, G. y Ostrom, V. (eds.), *Guidance, Control and Evaluation in the Public Sector*. De Gruyter. New York.

Auster, R., Levenson, I. y Sarachek, D. (1969) "The production of health, an exploratory study", *Journal of Human Resources*, 4 (4), págs. 411-436.

Ballart, X. (1990). "La evaluación de la actividad y rendimiento de la Policía", *Documentación Administrativa*, 224-225, págs. 361-405 .

Baumol, W. J. (1967) "Macroeconomics of unbalanced growth: the anatomy of urban crisis", *The American Economic Review*, 3, págs. 415-426.

Becker, G. S. (1964) *Human capital. A theoretical and empirical analysis with special reference to education*. Princeton University Press.

_____. (1968). "Crime and Punishment: An Economic Approach". *The Journal of Political Economy*, 76 (2), págs. 169-217.

Boardman, A.E., Davis, O.A. y Sanday, P.R. (1977) "A simultaneous equations model of the educational process", *Journal of Public Economics*, 7, págs. 22-49.

- Bradford, R.A., Malt, R.A. y Oates, W.E. (1969). "The rising cost of local public services: some evidence and reflections". *National Tax Journal*, 22 (2), págs. 185-202.
- Broome, J. (1993) "QALYs", *Journal of Public Economics*, 50, págs. 149-167.
- Brown, B.W. y Saks, D.H. (1975) "The production and distribution of cognitive skills within schools", *Journal of Political Economy*, 83(3), págs. 571-593.
- Bruno, M. (1978) "Duality, intermediate inputs and value-added", en Fuss, M. y McFadden, D. (eds.) *Production Economics: A Dual Approach to Theory and Applications*. North-Holland. Amsterdam.
- Bryant, J.W., Chambers, M.L. y Falcon, D. (1968). *Patrol effectiveness and patrol development*. University of Lancaster. Operations Research Department.
- Butt, H. y Palmer, B. (1985) *Value for Money in the Public Sector: the Decision Maker's Guide*. Basil Blackwell. Oxford.
- Carr-Hill, R.A. y Stern, N.H. (1973). "An econometric model of supply and control of recorded offences in England and Wales", *Journal of Public Economics*, 2, págs. 289-318 .
- ____ (1979) *Crime, the Police and Criminal Statistics*. Academic Press. London.
- Cave, M., Kogan, M. y Smith, R. (1990) *Output and Performance Measurement in Government. The State of the Art*. Jessica Kingsley Publishers. London.
- Cohn, E. (1968) "Economies of scale in Iowa high school operations", *The Journal of Human Resources*, 3(4), págs. 423-434.
- ____ y Millman, S.D. (1975) *Input-output Analysis in Public Education*. Ballinger Publishing Company. Cambridge.
- Cuenca, A. (1994) "Eficiencia técnica en los servicios de protección contra incendios". *Revista de Economía Aplicada*, 2 (5), págs. 87-109 .
- Culyer, A.J. (1980) *The Political Economy of Social Services*. Martin Robertson. Oxford.
- ____, Lavers, R.J. y Williams, A. (1971) "Social indicators: Health", *Social Trends*, 2, págs. 21-41.
- ____, y Maynard, A. (1997) *Being Reasonable about the Economics of Health. Selected Essays by Alan Williams*. Edward Elgar. Cheltenham, UK.
- ____ y Wagstaff, A. (1993) "QALYs versus HYE's", *Journal of Health Economics*, 12, págs. 311-323.

- Charnes, A, Cooper, W. W. y Rhodes, E, (1978). "Measuring the efficiency of decision making units", *European Journal of Operational Research* , 2, págs. 429-444.
- Diez-Ticio, A. (1997) *Análisis económico del delito y de la intervención policial. Una aplicación a España*. Tesis Doctoral. Universidad de Zaragoza.
- Dolan, P. (1998) "The measurement of individual utility and social welfare", *Journal of Health Economics*, 17, págs. 39-52.
- Duncombe, W. y Yinger, J. (1993) "An analysis of returns to scale in public production, with an application to fire protection", *Journal of Public Economics*, 52, págs.49-72.
- Ehreth, J.L. (1994) "The development and evaluation of hospital performance measures for policy analysis", *Medical Care*, 32(6), págs. 568-587.
- Evans, R.G. y Walker, M.D. (1972) "Information theory and the analysis of hospital cost structure", *Canadian Journal of Economics*, 5, págs. 398-418.
- Feldstein, M.S. (1967) *Economic Analysis for Health Service Efficiency: Econometrics Studies of the British National Health Service*. North-Holland. Amsterdam.
- Fetter, R.B., Shin, Y., Freeman, J.L. Averil, R.F. y Thompson, J.D. (1980), "Case mix definition by diagnosis-related groups", *Medical Care*, 18, Suplemento.
- Fisher, F.M. y Shell, K. (1972) *The Economic Theory of Price Indices. Two Essays on the Effects of Taste, Quality and Technological Change*. Academic Press. New York.
- Flecher, J. y Snee , H. (1985) "The Need for Output Measurement in the Service Industries: A Comment", *The Service Industrial Journal*, 5 (1), págs. 13-78.
- Gafni, A., Birch, S. y Mehrez, A. (1993) "Economics, health and health economics: HYEes versus QALYs", *Journal of Health Economics*, 12, págs. 325-339.
- Gafni, A. y Mehrez, A. (1989) "Quality-adjusted life years, utility theory and healthy years equivalents", *Medical Decision Making*, 9, págs.142-149.
- _____ (1991) "The healthy years equivalents: How to measure them using the standard gamble approach", *Medical Decision Making*, 11, págs.140-146.
- Gintis, H. (1971) "Education, technology and the characteristics of worker productivity", *The American Economic Review*, 61, págs. 266-279.
- González, B. y Villalobos, J. (1993) "Indicadores de actividad y costes en hospitales españoles", *Hacienda Pública Española*, monografías, 1, págs.127-141.
- Greenwood, P. W., Chaiken, J. M. y Petersilia, J. (1977) *The Criminal Investigation Process*. Lexington Books. Toronto.

- Hanushek, E.A. (1971) "Teacher characteristics and gains in student achievement: estimation using micro data", *American Economic Review*, 6 (mayo), págs. 280-288.
- ____ (1986) "The economics of schooling: production and efficiency in public schools", *Journal of Economic Literature*, 24 (septiembre), págs. 1141-1177.
- Harris, G. (1989) "Research output in Australian university research centres in economics", *Higher Education*, 18 (4), págs. 397-409.
- Hill, T.P. (1977) "On goods and services", *Review of Income and Wealth*, diciembre, págs. 315-338.
- Hirsch, W.Z (1973) *Urban Economic Analysis*. McGraw-Hill. New York. Traducción al castellano: *Análisis de Economía Urbana*. Instituto de Estudios de Administración Local, 1977.
- Hotteling, H. (1933) "Analysis of a Complex of Statistical Variables into Principal Components", *Journal of Education Psychology*, 24, págs. 417-441.
- Johnes, J. y Taylor, J. (1990) *Performance Indicators in Higher Education. UK Universities*. SRHE and Open University Press. Buckingham.
- Johannesson, M. (1999) "On aggregating QALYs: a comment on Dolan", *Journal of Health Economics*, 18, págs. 381-386.
- Johnson, D.W. (1974) "Affective outcomes", en Walberg, H. (ed), *Evaluating Educational Performance*. McCutchan Publishing Corp. Berkeley.
- Lancaster, K.J. (1966) "A new approach to consumer theory", *Journal of Political Economy*, 74, págs. 132-157.
- Levitt, M. S. y Joyce, M. A. S. (1987). *The growth and efficiency of public spending*. Cambridge University Press. Cambridge.
- Liebowitz, S.J. y Palmer, J.P. (1984) "Assesing the relative impact of economics journals", *Journal of Economic Literature*, 22, págs.77-88.
- Lind, R. C. y Lipsky, J.P. (1971). "The measurement of police output: conceptual issues and alternative approaches", *Law and Contemporary Problems*, 36 (4), págs. 566-88
- Madaus, G.F., Kellaghan, T., Rakow, E.A. y King, D.J. (1979) "The sensitivity of measures of school effectiveness", *Harvard Educational Review*, 49(2), págs. 207-230.
- Mancebón, M. J. (1996) *La eficiencia de los centros educativos públicos*. Tesis Doctoral. Universidad de Zaragoza.
- Mehay, S. L. (1977). "Interjurisdictional spillovers of urban police services". *Southern Economic Journal*, 43 , págs. 1352-59 .

- Murray, R. (1992) "Measuring Public-Sector Output: The Swedish Report", en Griliches, Z. (ed.), *Output Measurement in the Service Sector*, págs. 517-542. National Bureau of Economical Research. Conference on Research. The University of Chicago Press.
- Norsworthy, J.R. y Jang, S.L. (1992) *Empirical Measurement and Analysis of Productivity and Technological Change. Applications in High-Technology and Service Industries*. Elsevier Science Publishers. North-Holland. Amsterdam
- OCDE (1990) *Public Management Developments. Survey 1990*. Serie Public Management Studies, París.
- ____ (1999) *OCDE in Figures*. París.
- Pliatzky, L. (1986) "Can goverment be efficient?", *Lloyds Bank Review*, 159, págs. 22-32. Traducción al español en *Hacienda Pública Española*, 129, págs. 197-192.
- Pliskin, J., Shepard D.S. y Weinstein, M.C. (1980) "Utility functions for life years and health status", *Operations Research*, 28, págs.206-224.
- Prior, D. y Solá, M. (1993) *La eficiencia de los hospitales de Cataluña: una comparación entre los hospitales públicos y privados*. Departament de Sanitat i Seguritat Social, Generalitat de Cataluña.
- Ribera, M., Santamaría, L. y Vilardell, I. (1998) "Modelo de evaluación de la gestión en un centro educativo de secundaria", Documento presentado a las VII Jornadas de la Economía de la Educación. Santander. Septiembre.
- Richardson, J. (1994) "Cost utility analysis: what should be measured?", *Social Science Medical*, 39, págs. 7-20.
- Rosser, R. y Kind, P. (1978) "A Scale of Valuations of States of Illness: Is There a Social Consensus?", *International journal of Epidemiology*, 7(4), págs. 347-358.
- San Segundo, M.J. (1985) *Empirical studies of quality of schooling*. Tesis Doctoral. Universidad de Princeton.
- Schultz, T.W. (1961) "Investment in human capital", *American Economic Review*, 51, págs. 1-17.
- Schwenk, C.R. (1990) "Conflict in organizational decision making: an exploratory study of its effects in for-profit and not for-profit organizations", *Management Science*, 36(4), págs. 436-448.
- Sellin, T. y Wolfgang, M. E. (1964). *The Measurement of Delinquency*. John Wiley and Sons, Inc.
- Shoup, C. S. (1969) *Economía del Sector Público*. Madrid, Instituto de Estudios Fiscales, 1980.

- ____ y Mehay, S. L. (1971). *Program budgeting for urban police services*. Praeger. New York
- Skolka, J. (1977) "Unbalanced productivity growth and the growth of public services", *Journal of Public Economics*, 7(2), págs. 271-280.
- Smith, P (1996) *Measuring Outcome in the Public Sector*. Taylor & Francis. London.
- Smith, R. y Fielder, F. (1971) "The measurement of scholarly work: a critical review of the literature", *Educational Record*, 52, págs. 225-32.
- Spence, M. (1973) "Job market signaling", *Quarterly Journal of Economics*, 87(3), págs. 355-379.
- Steele, R. y Gray, A.M. (1982) "Statistical cost analysis: The hospital case" *Applied Economics*, 14, págs. 491-502.
- Taubman, P. (1976) "Personal characteristics and the distribution of earnings", en Atkinson A.B. (ed.), *The personal distribution of incomes*, págs. 193-226. Allen & Unwin. Londres.
- Torrance, G.W. (1976) "Social preferences for health states. An empirical evaluation of three measurement techniques", *Socioeconomic Planning Sciences*, 10, págs. 128-136.
- ____ (1986) "Measurement of health state utilities for economic appraisal: a review", *Journal of Health Economics*, 5, págs.1-30.
- ____ y Feeny, D. (1989) "Utilities and quality-adjusted life years", *International Journal of Technology Assessment in Health Care*, 5, págs. 559-575.
- Tuckman, H.P. (1971) "High school inputs and their contribution to school performance", *The Journal of Human Resources*, 6(4), págs. 490-509.
- Tullock, G. (1967)). "The welfare cost of tariffs, monopolies and theft". *Western Economic Journal*, Vol. 5 , págs. 224-32 .
- Wagstaff, A. (1989a) "Econometric studies in health economics. A survey of the British literature", *Journal of Health Economics*, 8, págs. 1-51.
- ____ (1989b) "Estimating efficiency in the hospital sector: a comparison of three statistical cost frontier models", *Applied Economics*, 21, págs. 659-672.
- Williams, A. (1985) "Economics of coronary artery bypass grafting", *British Medical Journal*, 291, págs. 326-329.
- ____ y Anderson, R. (1975). *Efficiency in the social services*. Basil Blackwell and Martin Robertson. London.

Willis, K.G. (1983) "Spatial variations in crime in England and Wales: testing an economic model", *Regional Studies*, 17 (4), págs. 261-72 .

Young, W., Skinkola, R. y Zorn, D. (1982) "The measurement of hospital case mix", *Medical Care*, 20, págs. 501-512.