ESTUDIOS REGIONALES Nº 40 (1994), PP 167-188

Una metodología general para la elaboración de índices complejos de dotación de infraestructuras.

José Ramón Cancelo de la Torre
Universidad de La Coruña
Pilar Uriz Tomé
Universidad de Castilla-La Mancha

1. INTRODUCCION

El estudio del papel de la infraestructura en el desarrollo regional ha cobrado un protagonismo creciente desde comienzos de los ochenta. La consideración explícita del stock de infraestructura como input del proceso productivo ha puesto de relieve que una trayectoria estable y equilibrada de desarrollo no se puede mantener sin una adecuada provisión de capital público. De faltar esta última se puede hablar todavía de crecimiento, pero en todo caso de crecimiento desequilibrado y con una utilización ineficiente de los demás recursos productivos.

Este planteamiento del proceso productivo a escala regional ha tenido su máximo exponente en el llamado enfoque del Potencial de Desarrollo Regional (Biehl, 1980), según el cual cada región dispone de una dotación de factores productivos propios o específicos que la hacen distinta a las demás, y que son los que determinan que a largo plazo los agentes económicos se encuentren en puntos concretos del territorio nacional.

Entre estos factores figura la infraestructura, que ocupa además un papel destacado ya que es el único de estos recursos específicos cuya dotación puede modificar a medio plazo la administración pública, y en ese sentido se configura como el instrumento más potente a la hora de diseñar la política de desarrollo regional.

Las implicaciones de esta tesis han sido analizadas desde diversos puntos de vista: encontramos así estudios teóricos[1], aplicaciones descriptivas[2] o estimacio-

1. Véase por ejemplo Biehl (1980), Yu (1981) o Blum (1982).
2. Entre otros podemos citar Carbonell y otros (1990), Márquez (1991), Pérez Touriño (1992), Bandrés (1993) y Cancelo (1993).

nes empíricas de la relación entre producción e infraestructura utilizando variables directamente observables como aproximaciones de esta última[3].

Sin embargo, poco a poco el interés se ha ido desplazando hacia procedimientos de cuantificación del stock de grandes categorías de infraestructura a través de índices complejos, cuyos valores en ocasiones se incorporan a cuasifunciones de producción con el fin de lograr estimaciones más acusadas de la contribución de la infraestructura a la producción potencial[4].

Este artículo se centra en el análisis de algunas de las propiedades de los índices empleados para cuantificar la dotación de infraestructuras en una región concreta. La discusión no es trivial ni de segundo orden, ya que todas las estimaciones empíricas de la contribución de la infraestructura al desarrollo están condicionadas a la forma en que aquélla se ha cuantificado. En consecuencia, si distintas medidas conducen a diferentes conclusiones en el estudio de la relación infraestructura-desarrollo, es evidente que la discusión de las propiedades de medidas alternativas ocupa un lugar central en el debate.

El trabajo se estructura de la siguiente manera: el siguiente epígrafe desarrolla algunas consideraciones generales sobre la elaboración de índices de infraestructura. A continuación se revisan las principales características de la metodología de cuantificación actualmente más utilizada, la propuesta por Biehl. El epígrafe cuarto se centra en procedimientos lineales de agregación, proponiendo el uso de técnicas estadísticas de análisis de datos para ponderar las variables que forman el índice. El quinto epígrafe generaliza la forma funcional de la relación de agregación, que a su vez está relacionada con el grado de sustituibilidad admisible entre los elementos que forman el índice. El trabajo termina con una aplicación a la elaboración de un índice general de infraestructura de tipo económico.

2. CONSIDERACIONES GENERALES SOBRE LA ELABORACION DE INDICES DE INFRAESTRUCTURA

El mayor problema asociado a la utilización del stock de infraestructura en trabajos aplicados, tanto si nos referimos a la infraestructura en conjunto como a grandes categorías de la misma, es que por su propia definición se trata de una variable no directamente observable.

3. Looney y Frederiksen (1981), Costa (1988).
4. Véase Navarre y Prud'homme (1984), Biehl (1986, 1988), Castillo y otros (1987), Bosca y otros (1990) o Cutanda y Paricio (1991, 1992).

El cuadro 1 muestra de forma sintética y jerarquizada las distintas categorías de infraestructura; en un primer nivel distinguimos –Hansen (1965)– dos grandes categorías, infraestructura de tipo económico e infraestructura de tipo social. A continuación, en un segundo nivel, cada una de éstas se descompone en subcategorías: siguiendo por ejemplo a Biehl (1988), tenemos tres subcategorías para las infraestructuras de tipo económico y cuatro para las de tipo social. A su vez, en muchas de las subcategorías de nivel 2 se pueden definir nuevas divisiones, como las que aparecen ocupando el nivel 2.1.

Todas las categorías de las que hemos estado hablando hasta ahora son conceptos teóricos con pleno sentido económico, pero no directamente medibles. Las variables que realmente podemos observar son distintas manifestaciones parciales de las subcategorías de los niveles 2 y 2.1.

Supondremos en adelante que en total tenemos p variables observables o manifestaciones parciales de infraestructura, cuyo valor para la región k-ésima se denotará por W_{ik} (i=1, 2, ..., p; k=1, 2, ..., n). Además, en el resto del trabajo el subíndice i siempre denotará una variable, en tanto que el subíndice k se referirá a regiones.

Así planteado, el problema es similar al que se trata en la literatura de números índices. Las categorías de infraestructura de niveles 0, 1, 2 y 2.1 son magnitudes complejas no directamente observables, y lo que el analista observa son magnitudes simples que entran en la formación de la correspondiente magnitud compleja. En consecuencia, se trata de construir índices complejos que permitan comparar la situación de una región con el elemento de referencia e, indirectamente, con las demás regiones. Como ocurre con todos los índices, la cifra concreta que toma el índice para una determinada región es irrelevante, y lo realmente informativo es cómo se compara esa cifra con los valores correspondientes a las demás regiones.

La construcción de índices de infraestructura plantea, además de los problemas de carácter general comunes a todos los números índices, una serie de cuestiones específicas a las que hay que dar una solución. Concretamente, es preciso establecer los criterios a emplear en: la depuración del efecto tamaño; el tratamiento de las unidades de medida; la ponderación asignada a cada variable observable en el índice; y la forma funcional de la relación de agregación[5].

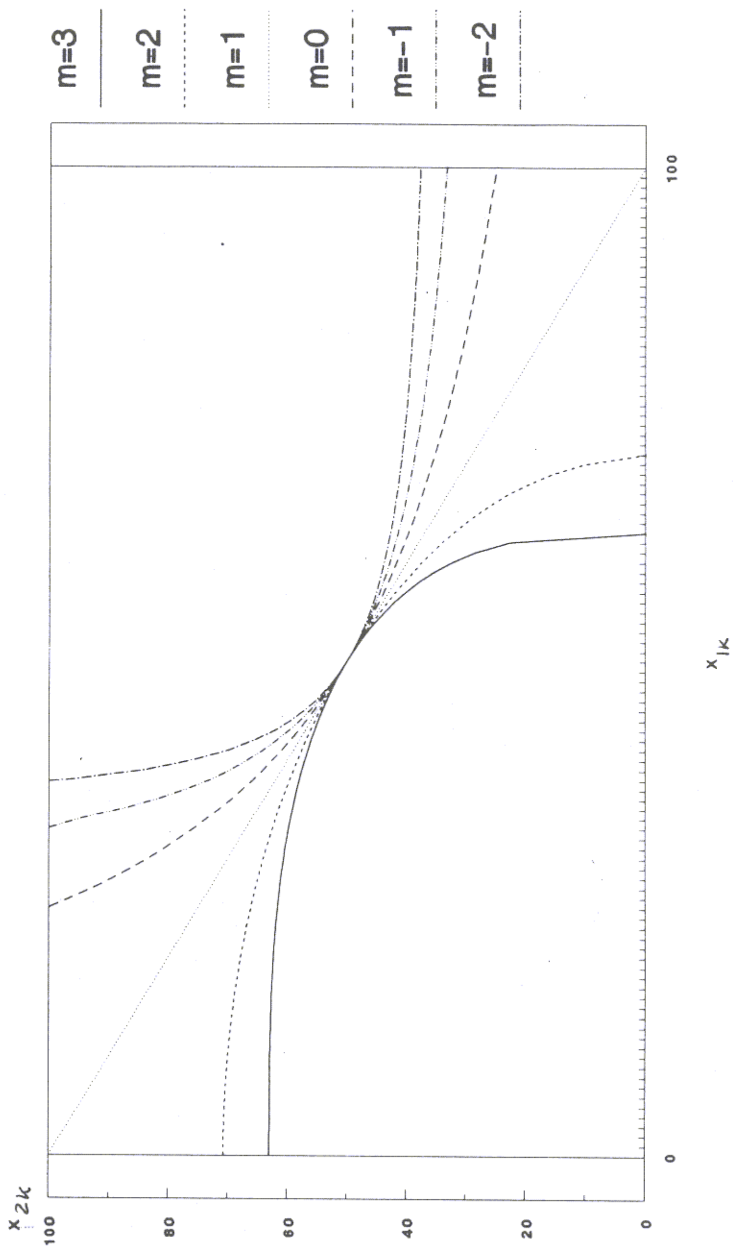
5. Una característica adicional de los índices de infraestructura es que la solución tradicional, basada en monetizar los valores de las magnitudes simples, no es aplicable aquí, ya que no existe una relación directa entre el coste de una determinada manifestación de infraestructura y su capacidad de servicio; el ejemplo de Biehl (1988, epígrafe tercero) es esclarecedor en este punto.

CUADRO 1

INFRAESTRUCTURAS: CATEGORIAS Y MANIFESTACIONES

NIVEL O	INFRAESTRUCTURA GENERAL						
NIVEL 1	INFRAESTRUCTURA DE TIPO ECONOMICO			INFRAESTRUCTURA DE TIPO SOCIAL			
NIVEL 2	TRANSPORTES	COMUNICACIONES	ENERGIA	SANITARIA	EDUCACION	SOCIAL	CULTURAL
NIVEL 2.1.	* CARRETERA * FERROCARRIL * MARITIMO		* ELECTRICA * GAS * OLEODUCTOS		* BASICA Y MEDIA * UNIVERSITARIA * FORMAC. PROF.		
NIVEL 3	- kms carret. - kms ferroc. - long. muelles - etc.	- lín. teléf. - estac. radio - líneas telex - etc.	- pot. cent. hidr. - kms oleodutos - kms gaseoductos - etc.	- num. camas hosp. - num. ambulatorios - num. farmacéuticos - etc.	- puestos EGB - puestos talleres FP - num. profes. univ. - etc.	- num. res. 3ª edad - num. plazas guard. - cent. planif. fam. - etc.	- num. teatros - num. bibliot. - fond. bibliogr. - etc.

GRAFICO 1
CURVAS ISOMEDIA PARA DOS VARIABLES



A continuación, discutimos cada uno de estos puntos por separado.

1. Depuración del efecto tamaño

En general cuanto mayor sea la región mayor será la dotación de infraestructura en términos absolutos, y en consecuencia mayores serán los valores de las variables observables. Ahora bien, a efectos de analizar su incidencia en la actividad económica lo relevante es la *capacidad potencial de servicio* o, como lo llama Biehl (1988), su *propiedad de estrangulamiento,* para cuya medición es necesario depurar el efecto tamaño de los datos originales y plantear el análisis en términos de regiones homogéneas en cuanto al tamaño.

Este tema ha sido extensamente discutido en la literatura, y el acuerdo es general: la solución consiste en relativizar los valores observados utilizando como indicador del tamaño de la región su superficie o su población, en función de si la variable observable tiende a tomar un valor mayor a medida que crece el espacio o a medida que aumenta la población.

Así, a partir de las magnitudes originales W_{ik} definimos la variable i-ésima depurada del efecto tamaño, Z_i, como

$$Z_{ik}=W_{ik}/S_k \quad \text{o bien} \quad Z_{ik}=W_{ik}/P_k$$

donde S_k (P_k) es la superficie (población) de la región k, y se utiliza una u otra según sea la variable W_i. En la práctica en algún caso la decisión puede no ser inmediata (como en el número de ambulatorios, por ejemplo), pero en general la depuración del efecto tamaño no suele plantear grandes controversias.

2. Tratamiento de las unidades de medida

Las unidades en que están expresadas las variables observables no son, por la propia naturaleza de los hechos que se trata de medir, comparables. Sin embargo, y al igual que el anterior, éste es un punto no conflictivo en la medida en que existe un acuerdo generalizado sobre la forma de proceder. El criterio habitual consiste con convertir las variables en magnitudes adimensionales mediante la normalización

$$X_{ik}= \frac{Z_{ik}}{\max\limits_{t} Z_{it}} \cdot 100$$

De esta manera, X_{ik} mide la dotación real de la variable i-ésima para la región k, expresada en tanto por ciento del valor de la dotación real de la región de referencia. Como región de referencia se toma la que tiene mayor dotación de la magnitud que mide la variable Z_i; o, equivalentemente, es la región con mayor dotación de la variable directamente observable W_i una vez que se ha depurado de los valores de ésta el efecto tamaño. Tal y como se ha definido, X_{ik} puede tomar valores entre 0 y 100.

Es posible utilizar normalizaciones alternativas, entre las cuales podríamos destacar la tipificación:

$$Y_{ik}= \frac{Z_{ik} - \overline{Z_i}}{DT_i}$$

donde

$$\overline{Z_i} =\sum_{k=1}^{n} \frac{Z_{ik}}{n} \quad , , \quad DT_i = \sqrt{\frac{\sum_{k=1}^{n} (Z_{ik} - \overline{Z_i})^2}{n}}$$

en cuyo caso la región de referencia es aquélla para la cual Y_i es igual a cero, o lo que es lo mismo, para la que el valor de la magnitud observable coincide con el valor medio una vez eliminado el efecto tamaño.

En este trabajo desarrollamos únicamente la normalización más habitual, basada en expresar Z_{ik} como porcentaje respecto al máximo valor de esa variable. En Cancelo y Uriz (1993) se exponen algunos resultados referidos a la transformación Y_{ik}[6].

3. La ponderación asignada a cada variable en el índice

Uno de los aspectos más críticos en la construcción de todo índice se refiere a la ponderación a asignar a cada variable que entra en su composición. En el caso concreto de los índices de infraestructura, la asignación se ve complicada porque se plantea en dos contextos diferentes:

a) Las variables que forman el índice son completamente comparables desde un punto de vista cualitativo, y la necesidad de ponderar surge porque su capacidad de servicio, desde un punto de vista puramente cuantitativo, es diferente. Esto es lo que ocurre, por ejemplo, al considerar distintos tipos de carreteras.

6. En este trabajo, página 533, se ha deslizado una errata en la definición de la media; la expresión correcta es

$$\overline{Z_i} = \sum_{k=1}^{n} \frac{Z_{ik}}{n}$$

En este caso es posible llegar a introducir ponderaciones objetivas, en el sentido de que la mayor parte de los analistas pueden estar de acuerdo con ellas. En el caso de las carreteras los técnicos en tráfico pueden proporcionar las capacidades de las vías de distinto tipo, y en base a ellas se pueden ponderar los kilómetros de las diversas clases de carreteras.

b) Las variables observables no son idénticas desde un punto de vista cualitativo: en este caso tenemos una serie de variables que constituyen aproximaciones parciales a la dotación total de una determinada categoría de infraestructura, y lo que se discute con la ponderación es hasta qué punto cada magnitud observable aporta información relevante sobre lo que se pretende medir.

Compárese por ejemplo el número de camas de hospital con el número de ambulatorios: ¿deberían entrar con el mismo peso en un índice general de infraestructura sanitaria? Evidentemente cualquier decisión que se tome aquí, incluyendo la más habitual en el análisis aplicado consistente en no ponderar y tratar por igual a todas las magnitudes observables, contiene un componente subjetivo bastante elevado.

4. La forma funcional de la relación de agregación

Este punto está relacionado con la sustituibilidad entre distintas manifestaciones o, en un plano más general, entre distintas categorías de infraestructura[7].

A lo largo de este trabajo emplearemos la expresión *grado de sustituibilidad* para referirnos al estudio de hasta qué punto los elementos que se agregan pueden variar sin que lo haga el índice. Esta es una propiedad básica de cualquier procedimiento de agregación que se proponga para obtener medidas indirectas de la dotación real de infraestructura, y la decisión en este punto condiciona toda cuantificación que se realice, tanto más cuanto más dispares sean los elementos que se agregan.

De las cuatro cuestiones específicas que hemos tratado existe, como hemos señalado, un acuerdo más o menos generalizado en las dos primeras, y las decisiones que implican mayor grado de subjetividad se centran en la ponderación de las distintas variables observables y en la forma funcional de la relación de agregación. Por esta razón el resto del artículo se centrará en la discusión de estos dos últimos puntos.

7. Las relaciones de sustituibilidad también se pueden introducir a través de las ponderaciones de una combinación lineal, pero esto no deja de ser una aproximación local al empleo de formas funcionales más generales.

3. LA METODOLOGIA PROPUESTA POR BIEHL

La mayor parte de los trabajos aplicados han seguido la metodología de elaboración de índices de infraestructura propuesta por Biehl, cuyas principales características se exponen en Biehl (1988), páginas 300-302. Dada su relevancia en el análisis aplicado, en este epígrafe revisamos brevemente los supuestos implícitos que dicha metodología incorpora.

La propuesta de este autor se puede sintetizar en el siguiente procedimiento secuencial en cuatro etapas:

ETAPA 1) Cuando existen coeficientes técnicos en base a los cuales definir ponderaciones (por ejemplo, capacidad de distintos tipos de carreteras), se obtiene una media aritmética ponderada de las magnitudes observables. Esta media sustituye en el resto del trabajo a las magnitudes que han entrado en su cómputo.

Si no existen coeficientes técnicos con los que ponderar las magnitudes observadas, éstas pasan directamente a la segunda etapa.

ETAPA 2) Se depura el efecto tamaño, expresando las variables observables o las medias resultantes de la etapa primera, según los casos, en porcentaje respecto al máximo valor de cada una de ellas. Con ello se obtienen las variables que en el epígrafe anterior denotábamos por X_i.

ETAPA 3) A continuación se calculan índices de las categorías de infraestructura de los niveles 2 y 2.1 del cuadro 1, a partir de una media aritmética sin ponderar de las pertinentes variables X_i.

Una vez obtenida la media, el correspondiente índice se obtiene normalizando el valor resultante de la forma habitual, es decir, expresándolo en porcentaje respecto al valor de la región que arrojó una media aritmética más alta; se tiene así[8]

$$I_k^{(1)} = \frac{x_{1k} + X_{2k} + \ldots + X_{pk}}{\max\limits_{t} (X_{1t} + X_{2t} + \ldots + X_{pt})} \cdot 100$$

8. Puesto que nos estamos refiriendo a una categoría concreta de infraestructura, en la construcción del índice sólo entran parte de las p variables observables X_i, las directamente relacionadas con esta categoría. No obstante, para no complicar en exceso la notación todas las fórmulas incluirán a la totalidad de las variables, aún cuando en una aplicación concreta sólo se utilizaría el subconjunto de variables relevantes para la categoría que se estudia.

donde la notación $I_k^{(1)}$ está relacionada con la formulación general del epígrafe cinco.

ETAPA 4) Por último, se obtienen índices de categorías de infraestructura de nivel 1 y un índice global de infraestructura aplicando medias geométricas sin ponderar a los índices de infraestructura de nivel 2, y normalizando los resultados de la forma habitual..

Las dos primeras etapas no plantean mayor discusión, ya que están relacionadas con los dos puntos del epígrafe anterior sobre los que decíamos que el acuerdo era general. La tercera incorpora tres supuestos implícitos, relacionados entre sí, que merecen mención especial:

1) Para cualquier índice de las categorías de infraestructura del nivel 2 que se construya directamente a partir de variables observables, a las manifestaciones parciales normalizadas X_i se les asigna el mismo peso en la elaboración del índice. Por lo tanto es preciso llevar a cabo previamente una selección rigurosa de estas manifestaciones, ya que no necesariamente un aumento de información se traduce en un mejor índice.

2) Para aquellas categorías de infraestructura de nivel 2 cuyo índice se construya a partir de los índices de infraestructura de las categorías pertenecientes al nivel 2.1., el índice resultante es el resultado de ponderar por igual a las distintas subcategorías. Así, y tomando por ejemplo la infraestructura de transporte, en el valor final del índice es tan importante la infraestructura de transporte por carretera como la de transporte aéreo.

3) El grado de sustituibilidad implícito en una media aritmética simple es bien conocido: caídas en uno de los elementos que la forman pueden ser compensadas por aumentos en la misma cuantía en cualquier otro sumando, de manera que el índice resultante no varíe.

En cuanto a la cuarta etapa, en la que se obtienen índices más generales de infraestructura mediante medias geométricas, también merece un comentario. Como señala Biehl, la media geométrica admite menos sustituibilidad entre los elementos que la forman que la media aritmética: "la media geométrica refleja esto en tanto en cuanto la media geométrica y la media aritmética son idénticas si dos indicadores son iguales, pero se desvían cada vez más a medida que esos valores difieren. Cuanto mayor sea la diferencia entre capacidades relativas, menor será la sustituibilidad" (Biehl, 1988, pág. 301).

Esta elección entre media aritmética y media geométrica, con el fin de permitir diferentes grados de sustituibilidad en los índices de infraestructura correspondientes a distintos niveles que aparecen en el cuadro 1, significa una mejora considerable respecto a un tratamiento uniforme basado en una única relación de agregación. No obstante aparenta ser excesivamente limitado en la medida en que sólo contempla dos posibles relaciones de agregación y, por otra parte, no se considera explícitamente cómo la reducción en el grado de sustituibilidad admisible repercute en los valores finalmente resultantes del índice.

4. GENERALIZANDO LAS RELACIONES DE AGREGACION LINEALES: UNA PROPUESTA PARA LA ASIGNACION DE PONDERACIONES OBJETIVAS

En el epígrafe anterior se ha señalado que en las etapas 3 y 4 del procedimiento habitualmente utilizado se introducen determinados supuestos que en ocasiones pueden ser difíciles de justificar. En consecuencia sería útil generalizar los procedimientos allí propuestos, diseñando esquemas de agregación más flexibles.

En este epígrafe mantendremos la linealidad de la relación de agregación y consideraremos combinaciones lineales generales en las que las ponderaciones de las variables están determinadas mediante técnicas estadísticas de análisis de datos[9].

Una revisión de literatura estadística sobre elaboración de índices e indicadores sintéticos lleva a agrupar las propuestas potencialmente utilizables en este contexto en tres grandes familias: 1) índices construidos a partir de una media aritmética de los componentes, en la línea de lo propuesto por Biehl y discutido en el epígrafe anterior; 2) índices basados en la primera componente principal, como solución clásica al problema estadístico de reducción de la dimensionalidad[10]; y 3) índices que se obtienen a partir de la distancia de Ivanovic modificada, que da origen a la familia de índices DP2[11].

Como comentario general, podemos señalar que también los nuevos índices propuestos presentan ciertas limitaciones. Su principal aportación reside en que proporcionan una visión de los datos originales desde una perspectiva diferente a la que se tiene a partir de la media aritmética simple; por tanto el cálculo y comparación de índices que se basan en criterios distintos puede aportar información de interés, que no se detectaría si utilizásemos un único criterio como base de la cuantificación de la dotación de infraestructura. En ese sentido, las tres familias de índices que consideramos en este epígrafe no se deben ver como mutuamente

9. Esta sección está basada en Cancelo y Uriz (1993), donde se extiende la discusión aquí resumida.
10. Véase Mardia y otros (1979).
11. Véase Ivanovic (1974) y Pena (1977).

excluyentes sino como complementarias, con independencia de que en última instancia el analista se vea obligado a decidir por la cifra que arroja un índice concreto.

Entrando ya en la definición de los índices propuestos, y comenzando con los índices basados en la primera componente principal, se trata de identificar el eje principal de la nube de puntos que forman las observaciones originales, calcular la proyección del punto correspondiente a cada región sobre dicho eje, y finalmente normalizar dichas proyecciones de manera que el valor máximo del índice sea igual a 100. Se tiene así para cada región k:

$$U_k^{(CP)} = b_1X_{1k} + b_2X_{2k} + \dots + b_pX_{pk}$$

donde $b = (b_1\ b_2\ \dots\ b_p)'$ es el autovector asociado al mayor autovalor de la matriz de dispersión de las variantes X_i, sujeto a la restricción de que su módulo sea la unidad. Puesto que esta combinación lineal no cumple la condición de normalización, el índice final está definido como

$$I_k^{(CP)} = \frac{U_k^{(CP)}}{\max\limits_t U_t^{(CP)}} \cdot 100$$

En cuanto a los índices basados en la distancia de Ivanovic modificada, se define la combinación lineal

$$U_k^{(DIM)} = X_{1k} + X_{2k}\,(1-R^2_{2;1}) + \dots + X_{pk}\,(1-R^2_{p;1,2,\dots,\,p-1})$$

donde DIM es el acrónimo de Distancia de Ivanovic Modificada, y sin pérdida de generalidad hemos puesto que las variables X_i están ordenadas de tal manera que la ponderación de la variable i-ésima es $1-R^2_{i;1,2,\dots,\,i-1}$. En esta expresión $R^2_{i;1,2,\dots,\,i-1}$ representa el coeficiente de determinación lineal de la regresión de la variable i-ésima sobre las variables 1, 2, ..., i-1; por lo tanto $1-R^2_{i;1,2,\dots,\,i-1}$ es el porcentaje de varianza de X_i no explicada por $X_1, X_2, \dots, X_{i-1}$.

El índice final se obtiene normalizando $U^{(DIM)}$ de manera que el valor del índice para la región con mayor dotación sea igual a 100:

$$I_k^{(DIM)} = \frac{U_k^{(DIM)}}{\max\limits_t U_t^{(DIM)}} \cdot 100$$

En Cancelo y Uriz (1993) se desarrollan las propiedades, los problemas asociados a su cálculo, y un ejemplo de aplicación a datos reales de estos tres tipos de índices.

5. GENERALIZANDO LAS RELACIONES DE AGREGACION NO LINEALES: SU INCIDENCIA EN EL GRADO DE SUSTITUIBILIDAD IMPLICITO EN EL INDICE

Hemos señalado anteriormente que la distinción entre media aritmética y media geométrica permite incorporar diferentes supuestos sobre el grado de sustituibilidad de las magnitudes a combinar en el índice. Ahora bien, no hay razones de peso, ni teóricas ni empíricas, por las que limitarse a estos dos tipos de medias.

Nuestra propuesta consiste en extender las dos posibilidades habitualmente contempladas al conjunto de medias definidas por la fórmula de Foster[12]: se llama media de orden m a la expresión[13]

$$X_k^{(m)} = \left(\sum_{i=1}^{p} X_{ik}^m \ f_i \right)^{1/m}$$

donde f_i representa la ponderación asignada a la variable X_i. Como casos particulares, para m=1 tenemos la media aritmética X_k, para m=0 la media geométrica (G_k), para m=-1 la media armónica y para m=2 la media cuadrática.

Si todas las variables reciben el mismo peso, $f_i = 1/p$, la expresión anterior toma la forma

$$X_k^{(m)} = \left(\sum_{i=1}^{p} \frac{X_{ik}^m}{p} \right)^{1/m}$$

Puesto que en las aplicaciones se utilizan medias simples, en lo que sigue nos centraremos en el caso particular de ponderaciones iguales, aunque los resultados que se deriven pueden extenderse sin dificultad a medias de orden m ponderadas.

La fórmula de Foster define una familia de relaciones de agregación, cuyos miembros se diferencian en que permiten diferentes grados de sustituibilidad entre los elementos que forman la media. Tomando como referencia m=1, la media aritmética, se tiene que:

12. Véase por ejemplo Calot (1974) páginas 73-78, aunque él no emplea la expresión "fórmula de Foster".
13. Estas medias no se suelen calcular a partir de las variables directamente observables, sino sobre los índices de infraestructura de nivel 2. Sin embargo, y con el fin de no complicar en demasía la notación, seguimos denotando por X_i a los elementos a partir de los cuales se obtiene el índice.

1) Para valores dados de las variables X_i correspondientes a una región concreta, cuanto mayor es el valor de m mayor es la media resultante. En términos del problema que nos ocupa, para valores dados de las manifestaciones parciales a medida que m aumenta mayor es la dotación implícita de infraestructura.

2) Supongamos que p-1 valores de las variables de una región están dados (sin pérdida de generalidad los correspondientes a las p-1 primeras variables), y que se trata de determinar el p-ésimo valor, X_{pk}, que hace que la media de orden m sea igual a un valor predeterminado c. En ese caso, se demuestra que a medida que m disminuye X_{pk} tiene que estar cada vez más alejado de los valores extremos o, equivalentemente, estar más cerca de las medidas de posición central de la distribución formada por los valores que para esa región toman las p-1 variables restantes.

Esta segunda propiedad establece que a medida que m disminuye el grado de sustituibilidad decrece, en el sentido de que para mantener el valor de la media constante los valores de las p variables tienen que ser cada vez más parecidos entre sí.

Estas ideas se pueden formalizar definiendo las llamadas *curvas isomedia,* es decir, el conjunto de valores del vector $(X_{1k}\ X_{2k}\ \ldots\ X_{pk})'$ que dan lugar al mismo valor de $X_k^{(m)}$. Para el caso general

$$X_k^{(m)} = \left(\sum_{i=1}^{p} \frac{X_{ik}^m}{p} \right)^{1/m} = c$$

implica que se ha de cumplir la restricción

$$X_{pk} = \left(pc^m - \sum_{i=1}^{p-1} X_{ik}^m \right)^{1/m}$$

En el caso particular de la media aritmética obtenemos la conocida relación lineal,

$$X_{pk} = pc - \sum_{i=1}^{p-1} X_{ik}$$

y en el caso de la media geométrica, tomando logaritmos, tenemos la misma relación en logaritmos

$$1nX_{pk} = p\ 1nc - \sum_{i=1}^{p-1} 1nX_{ik}$$

Como se desprende de su expresión general, para m=1 la curva isomedia es un hiperplano, para m>1 una hipersuperficie cóncava y para m<1 una hipersuperficie convexa, siendo la concavidad (convexidad) tanto más acusada cuanto mayor (menor) es m.

Consideremos el caso particular en que se combinan dos variables, en el cual es posible representar gráficamente las curvas isomedia. Para c=50 las curvas correspondientes a distintos valores de m figuran en el gráfico 1. En dicho gráfico es inmediato comprobar que se verifican los siguientes resultados, que se generalizan para cualquier número de variables:

1) La curvatura es tanto más acusada cuanto más se aleje de uno el valor de m.

2) Cuanto mayor sea m, mayor es el grado de sustituibilidad que se admite entre las dos variables que componen el índice.

3) Valores de m mayores (menores) que uno implican un grado de sustituibilidad entre los componentes mayor (menor) que el implícito en la media aritmética, es decir, en una combinación lineal de las magnitudes que forman el índice.

4) Las curvas sólo coinciden en el punto en que todas las variables que forman el índice toman el mismo valor.

5) Si consideramos todas las combinaciones de X_{1k} y X_{2k} que están sobre la curva isomedia para valores dados de m y c, se demuestra que la función $X_{1k} + X_{2k}$ tiene un máximo (mínimo) para $X_{1k} = X_{2k}$ si m es mayor (menor) que uno[14].

La interpretación de este último resultado en nuestro contexto es de gran interés: cuando se admite un grado de sustituibilidad superior al implícito en la media aritmética, salen perjudicadas las regiones en las que los valores que forman el índice son muy parecidos entre sí. Así por ejemplo si m=2 y c=50, de todas las combinaciones que dan lugar a una media cuadrática igual a 50, la expresión $X_{1k} + X_{2k}$ es máxima para $X_{1k}=X_{2k}=50$. En consecuencia una región equilibrada, en el sentido de no mostrar grandes diferencias entre los valores de los componentes del índice, se ve perjudicada respecto a regiones con grandes diferencias entre los valores de las variables X_i.

14. De forma esquemática, se trata de optimizar la expresión $X_{1k}+X_{2k}$ sujeto a la restricción que impone la curva isomedia y a que X_{1k}, X_{2k} tomen valores en el intervalo [0, 100]. Las condiciones de primer orden (véase por ejemplo Kamien y Schwartz (1981), apéndice A) apuntan a $X_{1k}=X_{2k}$ como posible extremo, en tanto que las condiciones de segundo orden indican que dicho extremo corresponde a un máximo si m>1 y a un mínimo si m<1.

En cambio si el grado de sustituibilidad admitido es inferior al implícito en la media aritmética, las regiones más equilibradas salen beneficiadas. Así, si m=0 y c=50, la combinación (50,50) es ahora un mínimo de la función $X_{1k} + X_{2k}$.

En este punto sería informativo poder cuantificar el distinto grado de sustituibilidad implícito en los diferentes valores de m. Ahora bien, cualquier medida que se pueda definir depende necesariamente de los valores concretos de las variables a agregar, y en consecuencia el grado de sustituibilidad implícito en distintas medias sólo se puede medir a posteriori y para datos referentes a un problema concreto: no resulta posible establecer de antemano un valor deseable del grado de sustituibilidad, y a partir de él obtener el tipo de media a emplear.

Teniendo en cuenta esta limitación, el procedimiento propuesto en este trabajo consiste en:

1) Calcular las medias de orden m para distintos valores de m, incluyendo m=1; para la región k hallaríamos $X_k^{(m)}$ para valores seleccionados de m.

2) Calcular los correspondientes índices de infraestructura, normalizando con el criterio habitual las medias resultantes:

$$I_k^{(m)} = \frac{X_k^{(m)}}{\max_t X_t^{(m)}} \cdot 100$$

3) Definir el coeficiente de sustituibilidad de la media de orden m para la región k como

$$CS_k^{(m)} = \frac{I_k^{(m)}}{I_k^{(1)}} \cdot 100$$

Este coeficiente compara, para cada región, el valor del índice al que conduce la media de orden m con el que se tendría si se hubiese utilizado la media aritmética, que actúa como relación de agregación de referencia. Valores de CS_k (m) mayores (menores) que 100 indican que el grado de sustituibilidad admitido entre componentes es superior (inferior) al implícito en la media aritmética.

Obsérvese que esto tiene consecuencias claras: si para un determinado valor de m CS_k (m) es mayor que 100, se está afirmando que los mismos valores de los componentes llevarían a una estimación más pequeña del stock de infraestructura si los componentes se hubiesen agregado linealmente; si CS_k (m) es menor que 100, la conclusión es la contraria. Así, la aportación principal de este coeficiente es que permite cuantificar la sobrevaloración / infravaloración respecto a la referencia lineal.

Como hemos señalado anteriormente, sería interesante poder fijar un valor deseado de CS_k (m), por ejemplo 75 para todas las regiones, y determinar qué valor de m deberíamos utilizar para alcanzar dicho grado de sustituibilidad. Puesto que ello no es posible, en la práctica el analista debe calcular distintos índices para diversos valores de m y, a partir de los correspondientes valores de los coeficientes CS, escoger la media que considere que se aproxima más al grado de sustituibilidad que está dispuesto a admitir.

Nótese también que cada región tiene su propio grado de sustituibilidad: un mismo valor de m es compatible con que distintas regiones muestren grados de sustituibilidad diferentes, en función de los valores concretos que para dichas regiones tomen los componentes sobre los que se calcula la media. En ese sentido también es interesante conocer qué regiones están viendo primadas o reducidas sus estimaciones del stock de infraestructura por la relación de agregación seleccionada.

6. APLICACION A LA ELABORACION DE UN INDICE GENERAL DE INFRAESTRUCTURA ECONOMICA

Presentamos a continuación un ejemplo para ilustrar el uso en el análisis aplicado de los resultados teóricos que acabamos de derivar.

Los datos empleados son los que proporcionan Cutanda y Paricio (1992) para distintas categorías de infraestructura de tipo económico. Concretamente en la página 85 presentan índices del stock de infraestructura de transportes, comunicaciones, oferta de energía y abastecimiento de agua para las 17 comunidades autónomas españolas. El objetivo es construir un índice del stock global de infraestructura de tipo económico combinando los cuatro valores asociados a cada CCAA.

El cuadro 2 muestra los valores originales de Cutanda y Paricio, las medias de orden m para m=3, 2, 1, 0, -1 y -2, y los índices finales una vez normalizadas estas medias. De estos últimos, el correspondiente a m=0 es el indicador económico propuesto en Cutanda y Paricio, véase su cuadro 2.

Se puede comprobar como se cumplen los resultados teóricos referidos a las medias de orden m: cuanto mayor es m, mayor es el valor de la media cualquiera que sea la región que consideremos. En cuanto a los índices finales, se comprueba que al aumentar el valor de m se favorece a las regiones peor dotadas, ya que se reduce la diferencia con respecto a la región con media más alta pues ésta, por definición, siempre toma el valor 100. Esta última característica también se puede comprobar a partir de la relación inversamente proporcional entre la dispersión de los índices, medida por el coeficiente de variación, y el orden m de la media empleada.

CUADRO 2

INDICADORES ALTERNATIVOS DE LA DOTACION DE INFRAESTRUCTURA DE TIPO ECONOMICO

	INDICES PARCIALES				MEDIA DE ORDEN m						INDICES FINALES DE INF. ECONOMICA					
CC.AA.	Transp	Comun	Energía	Agua	m=3	m=2	m=1	m=0	m=-1	m=-2	m=3	m=2	m=1	m=0	m=-1	m=-2
Andalucía	72.94	38.00	22.41	90.22	66.73	62.06	55.89	48.65	41.78	36.55	70.32	65.66	59.40	51.97	44.86	39.46
Aragón	24.77	57.93	22.29	96.76	65.55	58.80	50.44	41.94	35.45	31.44	69.07	62.21	53.61	44.80	38.06	33.94
Asturias	68.87	50.05	51.12	83.15	66.20	64.76	63.30	61.87	60.52	59.30	69.76	68.52	67.28	66.08	64.99	64.03
Baleares	92.57	76.83	28.26	72.39	74.42	71.61	67.51	61.76	54.78	48.10	78.42	75.77	71.76	65.96	58.82	51.93
Canarias	80.09	64.05	18.00	85.83	70.72	67.47	61.99	53.06	41.97	33.23	74.52	71.39	65.89	56.67	45.06	35.88
Cantabria	72.25	55.27	21.80	87.69	67.51	64.11	59.25	52.56	44.84	38.12	71.14	67.83	62.98	56.14	48.14	41.16
C. La Mancha	20.99	35.34	20.71	91.55	59.19	51.23	42.15	34.44	29.60	26.92	62.37	54.21	44.80	36.78	31.78	29.06
Cast. y León	34.17	42.70	18.84	87.84	58.53	52.59	45.89	39.42	34.14	30.32	61.68	55.64	48.77	42.10	36.66	32.74
Cataluña	74.48	73.55	70.58	94.92	79.61	78.98	78.38	77.83	77.33	76.87	83.88	83.56	83.31	83.13	83.03	83.01
C. Valenc.	63.40	60.52	21.32	92.57	68.39	64.63	59.45	52.46	44.44	37.54	72.06	68.38	63.19	56.03	47.72	40.53
Extremadura	22.26	28.61	14.64	91.99	58.87	49.98	39.38	30.43	25.15	22.33	62.03	52.88	41.85	32.50	27.00	24.11
Galicia	68.89	40.95	26.01	62.62	54.81	52.49	49.62	46.30	42.85	39.68	57.76	55.54	52.74	49.45	46.01	42.85
Madrid	78.69	100.00	100.00	97.66	94.90	94.51	94.09	93.63	93.14	92.61	100.00	100.00	100.00	100.00	100.00	100.00
Murcia	35.05	45.86	12.99	88.81	59.49	53.36	45.68	36.90	28.86	23.34	62.69	56.45	48.55	39.41	30.99	25.20
Navarra	31.83	55.11	29.75	100.00	67.42	61.11	54.17	47.80	42.93	39.64	71.05	64.65	57.58	51.05	46.09	42.80
País Vasco	100.00	71.39	74.86	93.63	86.67	85.83	84.97	84.11	83.25	82.43	91.33	90.81	90.31	89.83	89.39	89.00
Rioja	38.05	58.50	6.93	94.85	65.21	58.98	49.58	34.78	20.18	13.51	68.71	62.41	52.70	37.15	21.67	14.59
media											72.16	68.00	62.63	56.41	50.60	46.49
desv. típica											10.75	12.83	15.64	18.74	21.52	23.28
c. variación											0.15	0.19	0.25	0.33	0.43	0.50
max/min											1.73	1.89	2.39	3.08	4.62	6.86

FUENTE: Los índices parciales son los que proporcionan Cutanda y Paricio (1992), página 85; las medias de orden m y los índices finales de infraestructura de tipo económico se calcularon de acuerdo con lo expuesto en el epígrafe 5 de este trabajo.

CUADRO 3

COEFICIENTES DE SUSTITUIBILIDAD PARA DISTINTOS INDICES DEL STOCK DE INFRAESTRUCTURAS DE TIPO ECONOMICO

CC.AA.	m=3	m=2	m=0	m=-1	m=-2
Andalucía	118,4	110,5	87,5	75,7	66,4
Aragón	128,8	116,1	83,6	71,0	63,3
Asturias	103,7	101,8	98,2	96,6	95,2
Baleares	109,3	105,6	91,9	82,0	72,4
Canarias	113,1	108,3	86,0	68,4	54,5
Cantabria	113,0	107,7	89,2	76,4	65,5
C.-La Mancha	139,2	121,0	82,1	70,9	64,9
Cast. y León	126,5	114,1	86,3	75,2	67,1
Cataluña	100,7	100,3	99,8	99,7	99,6
C. Valenc.	114,0	108,2	88,7	75,5	64,2
Extremadura	148,2	126,4	77,7	64,5	57,6
Galicia	109,5	105,3	93,8	87,2	81,2
Madrid	100,0	100,0	100,0	100,0	100,0
Murcia	129,1	116,3	81,2	63,8	51,9
Navarra	123,4	112,3	88,7	80,1	74,3
País Vasco	101,1	100,6	99,5	99,0	98,6
Rioja	130,4	118,4	70,5	41,1	27,7
Media	118,2	110,2	88,5	78,1	70,8

Si calculamos los correspondientes coeficientes de sustituibilidad –cuadro 3–, comprobamos que son siempre mayores (menores) o iguales que 100 si m>1 (m<1); esto simplemente refleja que para la región k el valor del índice de orden m es mayor (menor) que el valor del índice basado en la media aritmética, salvo que la región sea precisamente la mejor dotada según ambos índices.

Centremos el análisis en la propuesta habitual de agregación para obtener índices de infraestructura de nivel 1, el caso m=0. El grado de sustituibilidad medio, definido como la media de CS (0) para las 17 comunidades autónomas, es igual a 88.5%. Se comprueba que La Rioja se ve muy perjudicada respecto a las demás regiones por el hecho de que se use la media geométrica en vez de la media aritmética, ya que mientras la reducción media en el valor del índice de infraestructura de oferta de energía para esa región. Claramente por debajo de la sustituibilidad media también se encuentran, de más perjudicadas a menos, Extremadura, Murcia, Castilla-La Mancha y Aragón.

Se puede comprobar como al aumentar la convexidad de la curva isomedia, y por lo tanto reducir el grado de sustituibilidad admitido, se producen cambios cualitativos y cuantitativos. Si ahora nos centramos en el índice correspondiente a m=-1, el valor medio del coeficiente de sustituibilidad se reduce a 78.1%, lo que indica que en media los índices de infraestructura económica resultantes son un 22% más bajos que si se hubiese usado la media aritmética. La Rioja vuelve a ser la más perjudicada, seguida ahora por Murcia, Extremadura (que intercambian su posición) y Canarias.

BIBLIOGRAFIA

BANDRES, E. (1993): "Las infraestructuras: políticas y realizaciones", en J. L. GARCIA DELGADO (dir.), *España, economía,* Espasa Calpe, Madrid, 1051-1072.

BIEHL, D. (1980): "Determinats of regional disparities and the role of public finance", *Public Finance,* 35, 44-71.

BIEHL, D. (com., 1986): *The Contribution of Infrastructure to the Regional Development,* Final Report of the Infrastructure Study Group, Document, Commission of the European Communities, Parts 1 and 2, Office for Official Publications of the European Communities, Luxemburgo.

BIEHL, D. (1988): "Las infraestructuras y el desarrollo regional", *Papeles de Economía Española,* 35, 293-310.

BLUM, U. (1982): "Effects of transportation investments on regional growth: a theoretical and empirical investigation", *Papers of the Regional Science Association,* 49, 169-184.

BOSCA, J., J. PARICIO y F. PEREZ (1990): *Dotación de infraestructura y desarrollo espacial,* mimeo, Consellería de Obres Publiques y Urbanisme, Generalitat Valenciana.

CALOT, G. (1974): *Curso de Estadística Descriptiva,* Paraninfo, Madrid.

CANCELO, J. R. (1993): "La economía de Castilla-La Mancha y su potencial de desarrollo", *Ponencias y actas de la XVIII Reunión de Estudios Regionales,* Universidad de Castilla-La Mancha, Toledo, 71-90.

CANCELO, J. R. y P. URIZ (1993): "La cuantificación de la dotación de infraestructuras: análisis comparativo de distintas medidas", *Comunicaciones a la XIX Reunión de Estudios Regionales,* Asociación Castellano-Leonesa de Ciencia Regional, Salamanca, 528-543.

CARBONELL, A., V. SANCHEZ, J. PEÑA, L. MARQUINA, J. MARTIN, J. GIL y F. ALVAREZ-MIRANDA (1990): *Las infraestructuras en España: carencias y soluciones,* Instituto de Estudios Económicos, Madrid.

CASTILLO, F., M. LEZANA y M. ORTUZAR (1987): "El desarrollo económico de los municipios vascos", *Ekonomiaz,* 7-8, 223-258.

COSTA, J. (1988): "Le role des equipements collectifs dans le developpement regional", *Revue d'Economie Regionale et Urbaine,* 1, 143-157.

CUTANDA, A. y J. PARICIO (1991): *Las dotaciones de capital público: disparidades regionales y desarrollo espacial,* mimeo, Instituto Valenciano de Investigaciones Económicas.

CUTANDA, A. y J. PARICIO (1992): "Crecimiento económico y desigualdades regionales: el impacto de la infraestructura", *Papeles de Economía Española,* 51, 83-101.

HANSEN, N. (1965): "Unbalanced growth and regional development", *Western Economic Journal,* 4, 3-14.

IVANOVIC, B. (1974): "Comment établir une liste des indicateurs de developpement", *Revue de Statistique Appliquée,* vol. 22.

KAMIEN, M. I. y N. L. SCHWARTZ (1981): *Dynamic Optimization,* North Holland, New York.

LOONEY, R. y P. FREDERIKSEN (1981): "The regional impact of infrastructure equipment in Mexico", *Regional Studies,* 15, 285-296.

MARDIA, K. V., J. T. KENT y J. M. BIBBY (1979): *Multivariate Analysis,* Chapman and Hall, Londres.

MARQUEZ, C. (1991): "Política regional europea y desarrollo regional en Andalucía: el caso de las infraestructuras de transporte por carretera", *Revista de Estudios Regionales,* 29, 81-114.

NAVARRE, F. y R. PRUD'HOMME (1984): "Le role des infrastructures dans le dèveloppement regional", *Revue e'Economie Régionale et Urbaine,* 1, 5-22.

PENA, B. (1977): *Problemas de medición del bienestar y conceptos afines,* INE, Madrid.

PEREZ TOURIÑO, E. (1992): "Las infraestructuras como factor de despegue del desarrollo gallego", en F. GONZALEZ LAXE (coord.), *Estructura Económica de Galicia,* Espasa Calpe, Madrid, 429-488.

YU, U.S.H. (1981): "Regional factor specifity, wage diferential and resource allocation", *Regional Science and Urban Economics,* 11, 69-79.