

ESPECIFICANDO MATRICES DE PONDERACIONES SOCIALES PARA REDES DE INVESTIGADORES: EL CASO DE LOS ECONOMISTAS ACADÉMICOS

Katarina ZIGOVA (*)

Universidad de Constanza

Resumen

Este artículo muestra cómo se pueden utilizar redes de colaboración y redes de citas para especificar las matrices de ponderaciones sociales en relación con una comunidad de investigadores. Para ello, se utilizan dos conceptos competidores –*comunicación* y *comparación*–, que funcionan a modo de vehículos por los que el medio social ejerce influencia en el comportamiento individual. La autora argumenta que, en el caso de las redes de investigación, la colaboración captura el canal *comunicación*, mientras que la cita captura el canal *comparación*. Además, opina que este último concepto podría ser el principal determinante social de la productividad investigadora de los individuos, y sugiere una matriz de ponderaciones sociales de referencia basada en dicho principio. A continuación, procede a testar dicha matriz de referencia con varias especificaciones alternativas utilizando un modelo de comparación bayesiana, concluyendo que la matriz de referencia ofrece mejores resultados que las especificaciones alternativas. Esta conclusión respaldaría las teorías socioeconómicas que subyacen a la especificación de la matriz de ponderaciones sociales de referencia.

Palabras clave: redes de colaboración y citación, matriz de ponderaciones sociales, autocorrelación por la red, productividad investigadora.

Abstract

This paper shows how collaboration and citation networks can be used to specify social weight matrices for a community of researchers. I use two competing theories of social influence on individual behavior, namely communication and comparison. I argue that in research networks collaboration captures communication, while citation captures comparison. I further argue that the comparison principle is likely to be the main social driver of individual research productivity and suggest a benchmark social weight matrix based on this principle. I test the benchmark matrix against several alternatives using a Bayesian model comparison procedure and conclude that the benchmark matrix outperforms alternative specifications. This result lends support to socio-economic theories underling the benchmark specification of the social weight matrix.

Key words: collaboration and citation networks, social weight matrix, network autocorrelation, research productivity.

JEL classification: C21, D83, I23.

I. INTRODUCCIÓN

EL establecimiento de redes relacionales o de contactos (*networking*) es un fenómeno inherente a toda comunidad social. El análisis de redes sociales reconoce que el comportamiento y las creencias de un individuo se ven influidos por el comportamiento y las creencias de los demás individuos, especialmente por aquellos que están directamente relacionados con él. Esta influencia o efecto de la red recibe diversas denominaciones, tales como *contagio social* (p. ej., Leenders, 2002), *efecto social* (p. ej., Manski, 1993), *efectos inter pares* o *vecindad* (p. ej., Moffit, 2001), así como muchos otros términos alternativos que invariablemente hacen alusión al efecto que las interrelaciones sociales tienen sobre el comportamiento de los individuos. En el presente trabajo, nos centramos en las representaciones de las redes de investigación en forma de matriz.

Dentro de la comunidad investigadora, las redes surgen de forma espontánea debido a la

existencia de numerosos temas de interés común en los que los investigadores colaboran. De hecho, la coinvestigación se ha vuelto más frecuente en las últimas décadas (p. ej., Laband y Tollison, 2000; Bosquet y Combes, 2013). La investigación conjunta puede comenzar como un intercambio informal de ideas, adoptar más tarde la forma de debates estructurados sobre áreas de mutuo interés a investigar, y traducirse finalmente en una colaboración propiamente dicha. Idealmente, esta clase de interacción desemboca en una publicación conjunta. El conjunto de todas las actividades colaborativas en una comunidad investigadora articula una red en la que los nodos están constituidos por los investigadores, y las coautorías representan los vínculos entre ellos. Las publicaciones conjuntas proporcionan una definición precisa de los lazos que se desarrollan en el seno de la red (Newman, 2001a), los cuales contrastan claramente con el concepto más etéreo de «contacto» o «amigo» que caracteriza a las redes sociales virtuales.

En el mundo de la investigación, además de las interrelaciones bilaterales o multilaterales, también tienen importancia las interacciones unilaterales, es decir, la asimilación y adopción de la literatura ya existente. Esta clase de interacción puede (y, con frecuencia, es así) excluir el contacto personal y está, asimismo, exactamente definida a través de las referencias. Las citas son vínculos explícitos que se establecen entre dos estudios que comparten algún contenido importante (Hummon y Doreian, 1989). Frente al carácter bidireccional de los vínculos entre los coautores, las citas son «unidireccionales», esto es, al citar publicaciones de otro investigador, no tiene por qué haber reciprocidad por parte del citado.

Las publicaciones son un indicador de la productividad individual. Ello facilita que la productividad pueda ser objeto de medición. Los modelos econométricos sobre los efectos de la red (p. ej., Marsden y Friedkin, 1993) ofrecen un medio para analizar el efecto de las interacciones de red sobre las variables de resultado observables, como la productividad investigadora. La clave en estos modelos econométricos radica en cómo definir la denominada matriz de ponderaciones **W**, que operacionaliza las relaciones de influencia generadas a lo largo y ancho de la red (Leenders, 2002; Corrado y Fingleton, 2012).

En este estudio, demostramos cómo es posible derivar la especificación de una matriz de ponderaciones sociales a partir de una estructura de redes sociales dada, aplicando las teorías de *comunicación* y de *comparación* procedentes de la sociología. Para ilustrar este enfoque, nos basamos en el comportamiento documentado de colaboración y citación de todos los economistas adscritos a una universidad o un instituto de investigación en Austria, Alemania y la zona germanófono de Suiza. Para ello, hemos revisado los trabajos publicados en revistas por dichos economistas durante el período 2000-2010.

La estructura de las redes de colaboración de los investigadores ha sido estudiada en muchas disciplinas, empezando por Newman (2001a, 2001b, 2004), que analizó patrones de colaboración en las áreas de las matemáticas, la física y la biología. Otros ejemplos incluyen la filosofía y la psicología (Cronin, Shaw, y La Barre, 2003), la sociología (Moody, 2004), la economía (Goyal, van der Leij, y Moraga-González, 2006), y las ciencias de la empresa (Acedo *et al.*, 2006). Por oposición a los estudios sobre redes de colaboración, los estudios sobre redes de citación no se enfocan en aspectos estructurales. Se originaron a partir de las

ciencias de la información, y su objetivo consiste en identificar los estudios más influyentes. De ahí que la mayoría de los estudios que analizan las redes de citación enlacen artículos (o revistas), en lugar de individuos. Un ejemplo clásico y bien conocido es el estudio de Derek J. de Solla Price (1965) que analiza la red mundial de citas de artículos científicos a través de la base de datos Science Index Citation de Garfield de 1961 (Garfield, 1963). Hummon y Doreian (1989) utilizan la red de citas de publicaciones para investigar la secuencia cronológica de los estudios que condujeron al descubrimiento de la estructura del ADN. La investigación más reciente realizada por White, Wellman y Nazer (2004) se acerca más al enfoque seguido en el presente trabajo. En ella analizan la evolución del patrón de citación entre 16 investigadores internacionales (procedentes de siete disciplinas), todos los cuales trabajan en un proyecto común.

En el presente estudio, la red es vista como un *input* intermedio para operacionalizar la matriz de ponderaciones sociales. Nos apoyamos para ello en conceptos de influencia social, los cuales distinguen entre contagio vía *comunicación* y contagio vía *comparación* (Leenders, 2002). En la comunidad investigadora, el contagio vía *comunicación* tiene lugar cuando dos o más investigadores colaboran entre sí. El concepto de comunicación asume un contacto directo entre varios individuos en aras de un objetivo común. Además, dicho contacto es siempre mutuo. En cambio, la idea del contagio social vía *comparación* no implica contacto directo entre los individuos, sino que presupone que estos imitan el comportamiento de otros individuos con una imbricación similar a la suya en la red. Empleamos la red de citación como representación del contagio vía *comparación*, y argumentamos que los investigadores utilizan las citas como medio para alinear su trabajo investigador al de los autores citados por ellos. Burt (1987, 2010) demuestra que, en entornos competitivos, la *comparación* es el determinante principal del contagio en las redes. Este hallazgo es directamente aplicable al entorno universitario, dada la evidencia abrumadora a favor de una elevada competitividad entre los investigadores (Stephan, 2010). Por tanto, exploramos las redes de colaboración y citación para una misma comunidad investigadora a fin de poder comparar la importancia relativa de los dos tipos de efectos de red. Además, estas dos redes, representativas de los conductos *comunicación* y *comparación* que inciden en la productividad investigadora, nos permiten construir la representación de la matriz subyacente.

A las matrices basadas en estas redes se las denominará en el presente trabajo «matrices de ponderaciones sociales» pues, técnicamente, se corresponden con la matriz de ponderaciones espaciales (1). Siguiendo la tradición, usamos la letra W (por *weight*, «ponderación» en inglés) para designar la matriz de ponderaciones. W es una matriz cuadrada y, en el caso que nos ocupa, de tamaño igual al de la red. El elemento i, j -ésimo de esta matriz mide la proximidad entre el nodo i -ésimo y el j -ésimo. El análisis econométrico de los efectos de red utiliza el mismo aparato matemático que la econometría espacial, con la diferencia de que en lugar de unidades geográficas tenemos individuos, y la matriz W replica la estructura de red que subyace a estos. Una de las cuestiones que se afrontan al resolver un problema de autocorrelación por la red (o, para el caso, un modelo econométrico espacial) estriba en seleccionar, de entre un universo de matrices de ponderaciones compatibles con las conceptualizaciones teóricas, aquella matriz que de hecho determina el proceso de autocorrelación social observado.

El enfoque econométrico espacial estándar que se sigue para seleccionar la matriz de ponderaciones es testar varias matrices razonables mediante pruebas o contrastes estadísticos (p. ej., LeSage y Pace, 2009: cap. 6). En respuesta a los argumentos críticos que reclaman un enfoque más sustantivo en la especificación de W (Corrado y Fingleton, 2012), hemos seguido un enfoque inverso: especificamos la matriz de ponderaciones sociales W de referencia que es compatible con el concepto *comparación*, es decir, con la naturaleza competitiva de la producción de artículos de investigación. Al operacionalizar la matriz W de referencia, incorporamos además la idea sociológica relativa a la distancia sociométrica y tenemos en cuenta la intensidad de las conexiones bilaterales entre los individuos. Para determinar la bondad estadística de la W de referencia, a continuación se especifican cinco matrices alternativas, cada una de las cuales proviene de relajar un criterio impuesto sobre la matriz W de referencia. Por último, aplicamos un enfoque de modelo de comparación bayesiana (LeSage y Pace, 2009) al objeto de testar la eficacia de la matriz W de referencia y compararla con las alternativas. La matriz de referencia demuestra ser superior a todas las especificaciones alternativas en cuatro tipos de modelos econométricos espaciales. Este resultado apoya la tesis de que la *comparación*, y de un modo más general la competencia, es el motor social determinante de la productividad investigadora.

El artículo se estructura conforme se detalla a continuación. En la siguiente sección, se presentan y visualizan las propiedades de las redes de colaboración y citación del conjunto de investigadores considerado. En la tercera sección, se justifica la elección de la matriz de ponderaciones W de referencia y se establecen las alternativas. A continuación, en la cuarta sección se compara la eficacia de la matriz de ponderaciones de referencia con las alternativas. La última sección se ocupa de las conclusiones.

II. REDES DE INVESTIGADORES: ANÁLISIS DESCRIPTIVO

En el entorno de la investigación académica tiene lugar una variedad de interacciones sociales tanto formales como informales. Ya a principios de los años setenta (Crane, 1972: cap. 3) se identificaron cuatro grandes tipos de interacciones sociales en el seno de la comunidad investigadora: debates informales de investigación, publicaciones conjuntas, relaciones docente-alumno, e influencia ejercida por las publicaciones de otros científicos en la elección de proyectos y técnicas de investigación. Todas estas interrelaciones, en caso de documentarse debidamente, pueden verse como articuladoras de una red (social) en la que los nodos representan a los investigadores, mientras que los vínculos entre ellos se ilustran en forma de relaciones. Las relaciones informales son más difíciles de registrar que las formales porque los investigadores, aun cuando respondan al cuestionario formulado, podrían no recordar a todas las personas con las que debatieron sobre su investigación. Crane (1972), por ejemplo, observó en su muestra de matemáticos y sociólogos que, al preguntárseles sobre las personas que habían influido en su trabajo, el 70 por 100 de los nombres aportados se reflejaban en sus citas. De igual modo, Melin y Persson (1996) estudiaron la colaboración entre investigadores en la Universidad de Umea (Suecia), y descubrieron que solo un 5 por 100 de la colaboración investigadora no desembocaba en una publicación. Estos hechos nos permiten concentrarnos en las interacciones sociales engendradas por dos vías: la de la colaboración, que cristaliza en coautorías, y la de las citas, pues dichas interacciones capturan la práctica totalidad de los lazos sociales que se generan en el mundo de la investigación.

A la red basada en las coautorías nos referiremos como la red de colaboración. Los nodos están formados por los investigadores, y los vínculos entre ellos expresan las coautorías publicadas. Probablemente, el autor que con más intensidad se

ha dedicado al estudio de las redes de colaboración científica es Newman (p. ej., 2001a), constatando que los vínculos colaborativos son consistentes, en el sentido de que cada vínculo está bien documentado y todos son del mismo tipo (2). Las redes de citación también son consistentes en este mismo sentido: las referencias son fácilmente observables y del mismo tipo. Las redes de colaboración y citación son, por tanto, fáciles de manejar, ya que pueden representarse mediante lo que se conoce como la matriz de contigüidades $A = a_{ij}, \forall ij$. Si la matriz A representa una red de colaboración, el elemento $a_{ij} = a_{ji} = 1$ si existe un vínculo colaborativo entre i y j , y 0 en caso contrario. La matriz de contigüidades A de una red de citación es, en general, no simétrica, puesto que en ella el elemento $a_{ij} = 1$ si i cita a j , pero $a_{ji} = 1$ implicaría que j también cita i (lo que no tiene por qué ocurrir).

El presente estudio se basa en los registros de publicaciones de todos los economistas académicos adscritos a universidades o institutos de investigación en Austria, Alemania y las regiones germanófonas de Suiza. La fuente de la que se han obtenido los datos es la base de datos donde se registran todas las investigaciones (3). Esta base de datos comprende, además de los economistas del ámbito académico, los investigadores del campo del *management* (gestión de empresas) y las finanzas. Por un lado, registra las características individuales de los científicos relevantes, como edad, sexo, título de doctorado o no, cohorte, universidad donde realizó el doctorado, y campo de investigación. Por otro lado, proporciona información relativa a los artículos que cada científico ha publicado en revistas, como el nombre de la revista, el año de publicación, el título y los (co)autores. La base de datos está auspiciada por la Asociación de Economía alemana (4) y se utiliza para elaborar los *ranking* que publica el periódico económico alemán *Handelsblatt*. No obstante, su principal finalidad es apoyar la investigación bibliométrica (5).

Para los estudios sobre redes, es importante que todos los sujetos pertenecientes a la red objeto de análisis se incluyan en la representación empírica. Al incluirse todas las unidades, ya no se trata de una muestra aleatoria, por lo que nos referiremos a ella con el término «conjunto», en lugar de muestra. El conjunto de datos analizado se compone de los 1.572 economistas académicos con grado de doctor o superior que trabajaban en los países germanófonos a finales de 2010. Se compilaron las redes sociales de dichos economistas a partir de 13.242 artículos publicados por ellos en revistas especializadas durante el período 2000-2010. En adelante,

para identificar en el presente trabajo a estos individuos o a sus publicaciones, utilizaremos el adjetivo «relevante».

1. Red de colaboración

Para la red de colaboración, se recopilaron las colaboraciones entre los economistas relevantes tal y como se documenta en sus publicaciones. El promedio de publicaciones por individuo es de alrededor de 10,3 con una desviación típica de 12,9. Cerca del 35 por 100 de estas publicaciones tienen un único autor, un 45 por 100 son coautorías con un economista externo al conjunto relevante y el 20 por 100 restante contienen vínculos de coautoría con economistas relevantes. Solo se utilizan estas coautorías relevantes para construir la red de colaboración. Pese a que el porcentaje del 20 por 100 pueda parecer bajo, aun así incluye a 1.036 (66 por 100) individuos en la red con casi 3.500 vínculos de colaboración entre ellos. Cerca de 2.000 de dichos vínculos conectan a los mismos dos investigadores, de modo que, en total, la red de colaboración consiste en 1.544 vínculos únicos no direccionados. La red se subdivide en varios componentes: el de mayor tamaño, también denominado el componente gigante, comprende al 54 por 100 de los economistas relevantes; además, hay 75 componentes más pequeños. Junto a ellos, existen 536 vértices aislados, que representan a investigadores que solo han publicado con coautores externos al conjunto de economistas relevantes (54 por 100), investigadores que solo han publicado estudios bajo su propio nombre (18 por 100), e investigadores que o bien no han publicado ningún artículo en revistas especializadas (11 por 100) o bien no lo hicieron durante el período considerado (17 por 100) (6). Los nodos de un grafo difieren en el número de vínculos colaborativos de cada investigador, es decir, en el número de aristas de un nodo, lo que también se conoce como «grado» o «valencia» en análisis de redes. El 6 por 100 superior de individuos con alto grado (ocho o más), se encuentran localizados en puntos más centrales. Dicho rol puede demostrarse excluyendo del grafo a estos individuos altamente conectados. En tal caso, el componente gigante se desagrega en más de 90 componentes más pequeños.

2. Red de citación

Para construir la red de citación, se recopilaron los vínculos por citas entre las publicaciones de los

economistas relevantes durante el período 2000-2010, utilizando datos de citación procedentes de la base de datos RePec (7). Alrededor de 4.500 de las publicaciones relevantes contienen citas a otro artículo publicado por un miembro de la red. Alrededor de 1.200 citas conectan al mismo par de autores, de modo que, en total, la red de citación se compone de unos 3.300 vínculos únicos direccionados. Comparada con la red de colaboración, la red de citación es, por tanto, mucho más densa. Esto es lo que cabría esperarse, ya que la incidencia de la citación es mucho mayor que la incidencia de la colaboración. Como resultado, solo hay nueve componentes más pequeños además del componente gigante. En comparación con la red de colaboración, hay más vértices aislados (668 que contienen economistas que no publicaron (22 por 100), economistas cuyas publicaciones no están indexadas en la base de datos CITec (30 por 100), economistas que no citan o no son citados por ningún miembro de la red (38 por 100), y economistas que solo tienen autocitas (10 por 100) (8). En la red de citación, los vínculos son direccionados debido a la naturaleza de las citas. Cada vínculo direccionado (también llamado «arco» en teoría de grafos) desde el vértice i hasta el vértice j surge si el economista i cita en una de sus publicaciones alguna de las publicaciones del economista j . Para revelar la existencia de posibles conglomerados (*clusters*) de citación, es interesante observar la red de citación desde la perspectiva de los campos de investigación. El coeficiente de *clustering* puede reflejar esto. Atendiendo a dicho parámetro, solo el coeficiente de *clustering* del campo de la microeconomía resulta ser significativamente mayor que el *clustering* total en la red de citación. La exclusión del 6 por 100 superior de los economistas más prestigiosos, es decir, la red de citas más frecuentemente citadas, no produce una desagregación del componente gigante, como sí sucedía en el caso de la colaboración. Además, incluso si se mantuviera únicamente a los individuos de elevado prestigio seguiríamos obteniendo una red bien interconectada.

3. Propiedades de mundo pequeño de las redes

En el análisis de redes y en los modelos que estudian los efectos de red, es importante cubrir la totalidad del universo del tópico objeto de investigación. En el presente estudio, dicho universo –formado por todos los economistas académicos de Austria, Alemania y la zona germanófono de Suiza– comparte una cultura académica homo-

gánea. Aun cuando la profesión en los países de habla alemana forma parte de la red económica mundial, estos economistas siguen estando «más cerca» unos de otros que del resto, debido a que comparten instituciones académicas similares. Esta coherencia regional tiene su reflejo en el hecho de que, durante el período 2000-2010, casi un tercio de las publicaciones conjuntas tuvo como coautor a otro miembro de la red, y la mitad tuvo como coautor a alguien con fuertes vínculos con la región germanófono (9). Además, ambas redes constituyen lo que se denominan «mundos pequeños» (Watts y Strogatz, 1998). Nos basamos en la definición de «mundo pequeño» acuñada por Goyal, van der Leij y Moraga-González (2006) durante el estudio de la evolución de las redes de colaboración entre economistas, para lo cual se requieren las siguientes propiedades:

- i) N es muy grande comparado con el grado promedio.
- ii) El componente gigante o principal existe y abarca una amplia sección de N .
- iii) La distancia promedio en el seno del componente gigante es pequeña, es decir, de orden $\ln(N)$.
- iv) Se aprecia un elevado agrupamiento o *clustering* en comparación con los grafos aleatorios con el mismo N y el mismo grado promedio.

El cuadro n.º 1 reporta los estadísticos descriptivos que confirman que las redes de colaboración y citación del presente artículo reúnen los atributos propios de un «mundo pequeño»: $N=1.572$ es muy grande comparado con el grado promedio de 2, el componente gigante existe y en ambas redes agrupa a más de la mitad de los nodos.

Además, el segundo componente más grande tras el componente gigante es, en ambos casos, un «enano», con tamaños de 7 y 3, respectivamente. La distancia promedio de 6,9 se aproxima a los «seis grados de separación» (Milgram, 1967) y también está muy cerca de $\ln(1.572)=7,4$. El diámetro, es decir, la distancia más larga entre dos individuos, es también pequeña para ambas redes. Finalmente, el *clustering* en la red de colaboración es cerca de 80 veces más elevado que lo que predicen los grafos aleatorios con los mismos parámetros, y más de 40 veces más elevado en el caso de la red de citación (10). Las redes de colaboración y citación del presente artículo cumplen las propiedades que

CUADRO N.º 1

PROPIEDADES DE MUNDO-PEQUEÑO DE LAS REDES DE COLABORACIÓN Y CITACIÓN

RED COMPLETA	COLABORACIÓN	CITACIÓN	
Número de vértices (N).....	1.572	1.572	
Grado promedio	1,96	2 (in/out)	
Vértices aislados.....	536 (34%)	668 (42%)	
Número de componentes (w/o) aislados.....	76	10	
Componente gigante	845 (54%)	885 (56%)	solapamiento: 649 conjunto: 1.081
Segundo componente más grande.....	7	3	
COMPONENTE GIGANTE SOLO	COLABORACIÓN	CITACIÓN	
Grado promedio	3,36	3,54 (dentro/fuera)	
Diámetro	24	13*	
Distancia promedio	6,9	4,92*	
Agrupamiento (1-vecino).....	0,32	0,17	

Notas: La autora usa el software Pajek (de Nooy, Mrvar, and Batagelj 2005) para calcular la medidas relativas a las propiedades de mundo-pequeño. * Medidas calculadas solo para vértices alcanzable en la red de citación directa.

esperaríamos encontrar en redes sociales de gran tamaño. Esta característica tiene una implicación importante: aun cuando los datos analizados cubren únicamente un segmento o región de toda la investigación económica mundial, dicho segmento refleja las propiedades del todo, siendo apto, por tanto, para llevar a cabo un análisis de los efectos de red. Newman (2001a) propone que las propiedades de «mundo pequeño» representan una característica imprescindible en cualquier comunidad científica funcional.

III. ESPECIFICANDO LAS MATRICES DE PONDERACIONES SOCIALES EN LAS REDES DE INVESTIGACIÓN

El análisis de los modelos estadísticos que explican el efecto de las redes sociales en el comportamiento individual se ha convertido hoy en día en el enfoque estándar en las ciencias sociales (p. ej., Friedkin, 1990; Leenders 2002; Burt 2010). En paralelo a este fenómeno, cabe reseñar el auge de popularidad de los modelos econométricos espaciales, comenzando por el seminal manual sobre econometría espacial publicado por Anselin en 1988 y la refinación de dicho concepto por el mismo autor treinta años más tarde (Anselin, 2010). Los modelos de redes y de econometría espacial comparten la misma estructura y utilizan las mismas técnicas de estimación, aun cuando difieren en sus motivaciones. Por consiguiente, en el presente trabajo se realiza una exposición sucinta de la estructura general del modelo econométrico espa-

cial basado en datos de corte transversal y posteriormente se explican las analogías con los modelos estructurales dedicados al estudio de los efectos de red.

Adoptando la notación de LeSage y Pace (2009), el modelo econométrico espacial general puede expresarse como sigue:

$$\begin{aligned} y &= \rho W y + \alpha + X \beta + W X \theta + u \\ u &= \lambda W u + \varepsilon, \quad \varepsilon \sim iid(0, \sigma^2 I), \end{aligned} \quad [1]$$

donde y es un vector de dimensión $N \times 1$ de la variable endógena, la matriz X de orden $N \times k$ es la matriz de variables exógenas, mientras que β es el vector de parámetros subyacente y α es un parámetro constante optativo. Los términos multiplicados por la matriz cuadrada W son los componentes de las interacciones espaciales. La matriz W de orden $N \times N$ es la matriz de ponderaciones espaciales y refleja las hipótesis explícitas del analista sobre la intensidad y la reciprocidad de las relaciones entre las unidades bajo análisis, que suelen consistir en países, regiones o ciudades (Corrado y Fingleton, 2012). Los parámetros escalares ρ , λ y el vector θ de orden $k \times 1$ miden los efectos promedio de las unidades interrelacionadas en la variable que se desea explicar y . Si $\rho \neq 0$, el modelo contiene efectos de interacción endógenos, y si $\theta \neq 0$, el modelo contiene efectos exógenos. Un λ distinto de cero implica que el modelo incluye efectos correlacionados. El modelo [1] podría incorporar uno cualquiera de los parámetros expresivos de interacciones, dos cualesquiera o bien los tres. Para una taxonomía

completa de las variantes del modelo [1] véase, p. ej., Elhorst (2010).

Los modelos que tratan sobre la influencia de las redes (p. ej., Friedkin, 1990; Marsden y Friedkin, 1993) difieren de los modelos espaciales en dos aspectos sustanciales: primero, la matriz de ponderaciones espaciales es sustituida por la denominada matriz de ponderaciones sociales y, segundo, las unidades de análisis son individuos en lugar de entidades geográficas. Los individuos retratados están imbricados en alguna red social explícitamente conocida, y el analista está interesado en determinar el efecto de esta red sobre una variable conductual y. La matriz de ponderaciones sociales cuantifica la estructura de la red social conocida. Claramente, la matriz de ponderaciones espaciales o, cuando corresponda, de ponderaciones sociales juega un papel crucial en la estimación del modelo [1] porque el tamaño de los efectos interacción ρ , λ y θ depende de la elección de \mathbf{W} (Leenders, 2002). Debido a ello, y también como respuesta a ciertas críticas recientes, se ha instado a que los estudios de econometría espacial fundamenten la elección de la matriz \mathbf{W} en argumentos sustantivos (Corrado y Fingleton, 2012).

En el presente artículo, se asume dicho reto especificando la matriz de ponderaciones sociales \mathbf{W} en base a teorías de influencia social (Burt, 1987; Leenders, 2002), justificando la distancia sociométrica mínima para considerar que existe influencia (Burt, 2010), y dando cuenta de la intensidad de las conexiones entre los investigadores siguiendo las directrices sugeridas por Newman (2001b). Dichas teorías nos permiten decidir quién está vinculado con quién y, en segundo lugar, la intensidad del vínculo. En las siguientes subsecciones desarrollamos con más detalle el enfoque aplicado en nuestro trabajo.

1. Teorías sobre influencia social: comunicación y comparación

Las teorías acerca de la influencia social exploran los procesos sociales que operan en la formación de las actitudes de un individuo o la adaptación de su comportamiento. El origen de estas teorías puede buscarse en el trabajo de Homans (p. ej., 1961) sobre formas elementales de comportamiento social dentro de los grupos. Partiendo de la base del análisis de Homans, los sociólogos empíricos estudiaron, por ejemplo, el papel que la influencia social desempeña en las decisiones sobre adopción de la innovación

médica (Burt, 1987) y en las decisiones sobre donaciones a asociaciones benéficas sin ánimo de lucro (Galaskiewicz y Burt, 1991). Para evaluar cómo influyen los aspectos sociales en los mecanismos de adopción, debemos conocer el entorno social en el que está insertado el individuo. Todos los contactos sociales entre todos los individuos de la comunidad bajo estudio constituyen la red social de dichos individuos. La pregunta pertinente desde el punto de vista de cada «ego» individual de la red representativa es cuáles del resto de individuos que conforman la red social (también llamados «alteri», o en plural «alteri»), influyen en su comportamiento. La teoría de redes sociales distingue dos grupos de procesos sociales entre cada «ego» y sus «alteri»: *comunicación* y *comparación* (p. ej., Leenders 2002). Estudiamos los procesos de comunicación y comparación en una comunidad de investigadores mediante el escrutinio de sus redes de colaboración y citación. El objetivo es especificar correctamente la matriz de ponderaciones, que mide la influencia que ejerce la comunidad de investigadores, interrelacionados a través de lazos de colaboración y/o citación, en la productividad investigadora de cada individuo.

La influencia vía *comunicación* engloba los contactos sociales, directos e indirectos, entre el «ego» y sus «alteri». Esta es una forma natural de retratar la influencia social, y asume que los contactos sociales entre el ego y los alteri modificarán probablemente el comportamiento del «ego». La mayoría de los estudios de influencia social asumen que la *comunicación* es el proceso dominante (Leenders, 2002). El proceso de *comparación* asume que el ego adapta su comportamiento comparándose a sí mismo con los individuos que ocupan una posición similar en la estructura social (Burt, 1987). La *comparación* subraya la existencia de competencia entre el «ego» y los «alteri» porque aquel adapta su comportamiento para mantener o mejorar su estatus (Burt, 2010). Los individuos que son objeto de comparación por el ego no coinciden necesariamente con aquellos con los que éste se comunica.

En general, la influencia social queda formalizada a efectos operativos en la matriz \mathbf{W} de orden $N \times N$, cuyos elementos w_{ij} igualan las medidas de proximidad estandarizadas (Burt, 1987).

$$w_{ij} = \frac{\text{proximity } i \text{ to } j}{\sum^k (\text{proximity } i \text{ to } k)}, \quad \forall k \neq i. \quad [2]$$

La medida de proximidad utilizada, que habitualmente se denota por d_{ij} para representar la

distancia entre los individuos i y j , depende de la clase de influencia social que se desea mostrar con el estudio.

Para operacionalizar la influencia social que opera vía *comunicación*, los investigadores aplicaron una medida de proximidad basada en un parámetro de vinculación (cohesión) que incluye información sobre el número, la longitud y la intensidad de los vínculos existentes entre los miembros de la red (Leenders, 2002). Como medida de la proximidad cohesiva entre i y j , Burt (1987) utiliza inversas de distancias de la ruta que conecta al individuo j con el individuo i (11). La proximidad es uno si i y j son vecinos directos, es inferior a uno si se encuentran separados por una distancia mayor, y es cero si no existe una ruta directa que conecte a i y j . Esta medida de proximidad es distinta de cero para cualquier par de miembros de la red con independencia de cuántos intermediarios existan entre ellos, y no depende de la intensidad de los vínculos entre ambos individuos. La medida de cohesión propugnada por Leenders (2002) limita los valores en proximidad cohesiva distintos de cero solo a los pares de individuos físicamente adyacentes, siendo cero para todos los demás pares, si bien su medida sí incluye la intensidad de las conexiones a efectos de cálculo. En el presente estudio, se utiliza una medida de proximidad que está a mitad de camino entre los dos extremos. Se sigue a Leenders (2002) y se utiliza la intensidad de los vínculos, pero no solo se incluye al primer vecino del ego, sino también al de segundo orden, y se normaliza la medida tal y como sugiere Burt (1987).

La idea original del contagio social vía *comparación* presume que un individuo i se comporta igual que otros individuos que tienen una imbricación similar en la red. Utilizando la misma red que también se utilizó para operacionalizar la *comunicación*, la *comparación* se operacionaliza definiendo la proximidad de [2] como el grado en el que i y j son estructuralmente equivalentes. Por ejemplo, Burt (2010) utiliza la medida de la distancia euclídea

$$d_{ij} = [\sum_k (z_{ik} - z_{jk})^2 + \sum_k (z_{ki} - z_{kj})^2]^{1/2}, \quad [3]$$

donde z_{ij} es la distancia de ruta estandarizada entre los individuos i y j , y las sumas se computan para todo k distinto de i y j . Nótese que d_{ij} es cero para los individuos perfectamente equivalentes, es decir, para aquellos cuyas distancias hasta y desde todos los demás miembros de la red son iguales. Cuanto mayor sea la proximidad d_{ij} , menos equivalente es la

posición de i y j en la red. Para obtener una medida que aumente conforme aumenta la proximidad, Burt (2010) representa el elemento i, j -ésimo de la matriz de ponderaciones sociales mediante el término

$$w_{ij} = \frac{d_{\max} - d_{ij}}{\sum_k (d_{\max} - d_{kj})}, \quad k \neq i. \quad [4]$$

Leenders (2002) sugirió una medida alternativa de equivalencia estructural que también está basada en la distancia euclídea:

$$d_{ij} = \frac{((\hat{i} - \hat{j}) \cdot (\hat{i} - \hat{j}))^{1/2}}{\sqrt{2n}}, \quad [5]$$

donde \hat{i} es el vector acumulado de la fila i -ésima y la columna i -ésima de la matriz de contigüidades A , $\hat{j} = 1 - \hat{i}$, y n es el número de miembros de la red. d_{ij} es igual a 0 en individuos completamente no equivalentes, mientras que d_{ij} es 1 en el caso de perfecta equivalencia estructural.

El concepto *comparación* muestra cierto grado de solapamiento con el concepto *comunicación*. Esta falta de discriminación ha llevado a varios investigadores a descartar arbitrariamente o bien la *comunicación* o bien la *comparación* como mecanismo de influencia. Leenders (2002) ilustra este dilema con cuatro estudios consecutivos (Coleman, Katz y Menzel, 1966; Burt, 1987; Marsden y Podolny, 1990; Strang y Tuma, 1993) que analizan la misma red de médicos y su influencia en la adopción de un nuevo fármaco. Los mencionados trabajos llegan a cuatro respuestas diferentes, a saber, que la adopción vino impulsada por la *comunicación*, por la *comparación*, por ninguna de las dos, y por ambas.

La confusión entre los conceptos *comparación* y *comunicación* proviene del hecho de que habitualmente ambos han sido operacionalizados con la ayuda de la misma red social, pero utilizando diferentes medidas de distancia entre los pares de individuos. En el presente estudio, adoptamos un enfoque diferente y utilizamos la disponibilidad de dos redes distintas integradas por la misma comunidad de investigadores. Argumentamos que una de ellas captura el proceso *comunicación*, mientras que la otra captura el proceso *comparación*. Por tanto, a diferencia de la mayoría de los estudios previos sobre matrices de ponderaciones sociales, aplicamos medidas de cohesión y de equivalencia estructural a ambas redes (es decir, a la red de citación y a la red de colaboración), y demostramos que, en lugar de depender de la medida elegida, lo que tipifica el

principio comparación, o el principio comunicación, es la propia red por sí misma.

Además, se justifica la preferencia *ex ante* de la influencia social vía *comparación* como motor predominante de la productividad investigadora. La *comunicación* implica que dos individuos hablan directamente entre sí. Los científicos necesitan comunicarse directamente con sus colaboradores. De ahí que sea natural asumir que la red de colaboración refleja patrones de comunicación. Además, la colaboración en el marco de la investigación requiere tiempo. Por eso no es raro que los lazos de comunicación deriven en relaciones a largo plazo, a veces incluso de amistad. Con independencia de la profundidad de la relación, la comunicación implica reciprocidad.

La *comparación*, en cambio, asume simplemente que los individuos se observan unos a otros, sin necesidad de conversación, conocimiento mutuo, y, lo que es más importante, reciprocidad entre ellos. La red de comparación puede ser, por tanto, altamente impersonal. En los estudios sobre innovación médica descritos por Burt (1987, 2010), la red de profesionales médicos se basaba en lazos profesionales. Algunos de dichos lazos eran formales, otros eran estratégicos, y algunos otros pueden haber estado basados también en la amistad. Esta red servía, por tanto, para analizar tanto el efecto comunicación como el efecto comparación. Sin embargo, una red de colaboración pura no ofrece realmente una síntesis del concepto comparación, pues captura muchos lazos que exceden claramente de la simple comparación. La red de citación, por el contrario, captura información de tipo no personal relativa a la literatura que sirve como *input* para el trabajo propio del autor. Las citas constituyen vínculos explícitos entre dos estudios con algún contenido importante en común (Hummon y Doreian, 1989). Si bien no es irracional pensar que

los investigadores se comuniquen con los autores a los que citan, esto es más bien la excepción a la norma general. Es precisamente el carácter en su mayor parte anónimo de las redes de citación lo que las convierte en idóneas a los fines de modelizar el efecto comparación sobre la productividad investigadora.

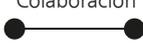
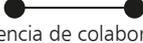
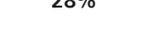
Una evidencia importante para este argumento es la práctica ausencia de solapamiento entre las redes de colaboración y de citación. El cuadro n.º 2 enumera posibles tipos de patrones de colaboración y citación. La mayoría de los economistas o bien colaboran pero no se citan unos a otros, o bien no colaboran pero se citan, o bien son citados por otro economista relevante (12). Estas combinaciones «limpias» conforman más del 90 por 100 de todos los vínculos existentes en las dos redes (cf. cuadro n.º 2), razón por la cual estas dos redes representan lazos en gran parte diferentes.

La cuestión central de nuestro análisis radica en saber si la productividad de los investigadores está determinada por la red de colaboradores o por la red de autores a los que citan. Nuestra hipótesis es que el efecto *comparación* es más importante que el efecto *comunicación*. La razón estriba en la propia naturaleza de la investigación. El sistema investigador es competitivo (Carayol, 2008; Stephan, 2010). En la competencia entre los investigadores, los aspectos pecuniarios juegan un papel, pero este es secundario respecto al prestigio, la estima profesional y el reconocimiento del propio trabajo por la comunidad académica (véase, p. ej., Merton, 1973). Burt (1987) también se muestra convencido de la importancia del efecto comparación para que exista competencia.

En una *survey* de experimentos donde se proponía a los participantes la resolución de un rompecabezas, Homans (1961) descubrió que, cuando estos

CUADRO N.º 2

DISTRIBUCIÓN DE PATRONES DE INTERACCIÓN EN COLABORACIÓN Y CITAS

FRECUENCIA DE PARES	CITAS	AUSENCIA DE CITAS
 Colaboración	 2%	 8%
 Ausencia de colaboración	 3%	 59%
		 28%

Nota: Las combinaciones «limpias» están destacadas en negrita.

operaban en un entorno cooperativo, en lugar de intentar resolver todo el rompecabezas ellos solos se especializaban en tareas concretas. La creciente especialización y la naturaleza cada vez más interdisciplinar de la investigación son las principales razones para la colaboración científica y su crecimiento constante en el tiempo (Laband y Tollison, 2000; Goyal, van der Leij y Moraga-González, 2006).

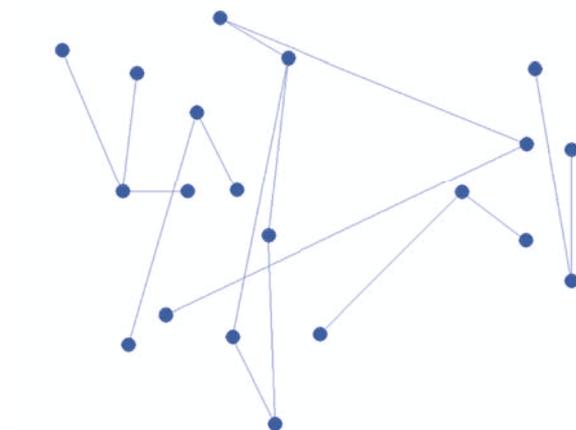
Por otro lado, Homans (1961) también argumenta que quienes compiten tienen más probabilidades de volverse más similares en términos de sus conocimientos que los que cooperan entre sí. En teoría económica existe un concepto análogo llamado «diferenciación mínima» (Hotelling, 1929), el cual vendría a decir que la competencia entre dos vendedores hace que el grado de diferenciación de su oferta a los clientes se reduzca. El principio de la «diferenciación mínima» también es aplicable en el caso de n vendedores si los gustos de los consumidores son suficientemente heterogéneos (Palma De *et al.*, 1985). En el caso de investigadores competitivos, esto implica que con el tiempo su productividad investigadora se vuelve similar, respaldando adicionalmente nuestra hipótesis de que la *comparación* juega un papel esencial en la investigación.

El efecto *comparación* será probablemente tanto más fuerte cuanto más productivos sean los individuos, debido a que su recompensa es superior. Elhorst y Zigova (2014) muestran que la competencia entre los principales institutos económicos es el doble de intensa que entre los institutos en general. El gráfico 1 muestra la red académica de los 50 economistas más prominentes en los países germanófonos según el *ranking Handelsblatt* en 2011 (13). Se observa que las dos redes difieren completamente entre sí: la colaboración entre estos individuos es meramente anecdótica, mientras que la red de citación muestra una red de interconexiones densa (14), una evidencia que vuelve a sugerir que la *comparación* es el principal motivador social de la productividad científica. En consecuencia, argumentamos que la matriz de ponderaciones sociales de referencia debería basarse en la red de citación.

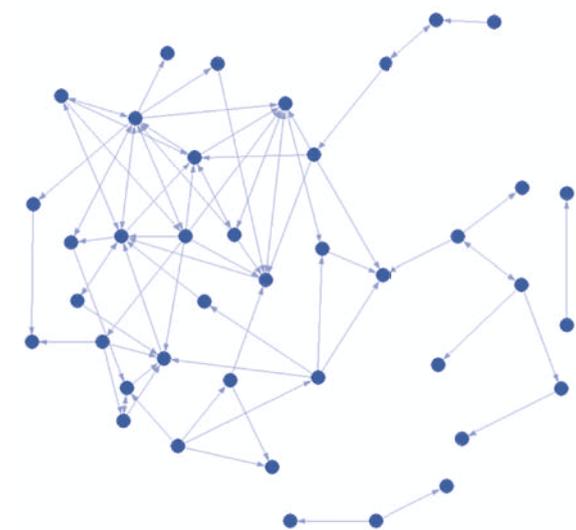
2. Horizonte de red

En la sección anterior, hemos argumentado que la influencia vía *comparación*, representada por la red de citación, es el principal proceso social que determina el comportamiento investigador. Sin embargo, continúan pendientes de abordarse otras dos

GRÁFICO 1
REDES DE COLABORACIÓN (A) Y DE CITAS (B) DE
LOS ECONOMISTAS DEL RANKING TOP-50



(a)



(b)

Notas: Los círculos rellenos representan economistas que están conectados, como mínimo, a otro economista Top a través de una publicación común (A) o bien a una cita interna o externa (B). La selección de los economistas Top se basa en el ranking de 2011 de productividad investigadora:

[HTTP://TOOL.HANDELSBLATT.COM/TABLE/INDEX.PHP?ID=79&PC=250](http://tool.handelsblatt.com/table/index.php?id=79&pc=250)

cuestiones relativas a la especificación de la matriz de ponderaciones sociales óptima: ¿cuán lejos llega el horizonte de red relevante?, y ¿cuál es la magnitud de influencia de los homólogos incluidos dentro de los límites de ese horizonte? El horizonte de red es la distancia sociométrica máxima entre dos individuos para que uno pueda ejercer influencia sobre el otro. La magnitud de la influencia comprende factores de red que refuerzan o debilitan el vínculo entre dos individuos. En el caso de la matriz \mathbf{W} , el horizonte de influencia determina qué elementos w_{ij} son iguales a cero y cuáles no, mientras que la magnitud de influencia determina el valor de los elementos w_{ij} distintos de cero.

La distancia sociométrica entre dos individuos i y j en una red se define como el número mínimo de vínculos que son necesarios para llegar a j desde i . Los vecinos directos tienen una distancia sociométrica de uno. Los pares de individuos que no son vecinos directos y que comparten al menos un vecino común tienen una distancia sociométrica de dos. La distancia sociométrica media en las redes sociales es sorprendentemente pequeña dado el tamaño de toda la red. Milgram (1967) alude a esto con la expresión de «mundo pequeño». En nuestro estudio la distancia media es de siete en la red de colaboración, y de cinco en la red de citación (véase cuadro n.º 1).

El círculo de vecinos de un individuo es claramente más pequeño que la distancia media de una red, ya que de otro modo casi todo el mundo sería relevante para los demás. En cambio, como observa Milgram (1967), «cinco intermediarios representan una distancia psicológica enorme entre dos sujetos. Si la red juega algún papel para las relaciones sociales, el círculo de influencia debe incluir al menos a los vecinos inmediatos del individuo y ser claramente más pequeño que el círculo definido por la distancia sociométrica media; la cuestión es cuánto más pequeño. Leenders (2002) propone limitar la distancia de influencia a dos o tres como máximo. Burt (2010) dedica todo un libro al análisis de las denominadas redes de vecinos y discute numerosos ejemplos de redes profesionales en el mundo de los negocios, la política y la medicina, centrándose en la importancia de los lazos de vecindad para las decisiones que el individuo toma en su profesión, y, por tanto, aboga por que la distancia sociométrica relevante se establezca en dos.

Crane (1972: cap.3) reporta evidencias procedentes de las matemáticas y la sociología. En dichas disciplinas, más de dos terceras partes de los inves-

tigadores están vinculados a un académico prolífico con una distancia no superior a un intermediario. En la red de citación analizada en nuestro estudio, el número medio de investigadores a los que se puede llegar a través de dos vínculos es de 25, mientras que los situados a tres vínculos de distancia ascienden hasta los 92 (15). Parece razonable limitar en esta red el círculo de influencia a una distancia de dos, ya que, dado que el determinante social de la productividad investigadora es la comparación, resulta plausible asumir que existen en torno a 20 homólogos con los que un investigador puede desear compararse. En la matriz \mathbf{W} , fijamos, por tanto, en dos la distancia sociométrica que importa para un investigador.

3. Matriz de ponderaciones sociales de referencia

En las anteriores secciones hemos justificado la elección de la red de citación y una distancia sociométrica de dos como los pilares para la construcción de la matriz de ponderaciones de referencia \mathbf{W} . Queda pendiente de decidir la medida de proximidad, es decir, si emplear la cohesión o la equivalencia estructural. Con base en la literatura pasada sugerimos aquí una medida de cohesión y otra de equivalencia estructural para construir la red de citación. Más tarde mostramos que la red de citación es siempre estadísticamente preferida a la red de colaboración, con independencia de la medida aplicada.

Cualquiera que sea la medida de proximidad elegida, esta se cuantifica para todos los pares de miembros de la red con la ayuda de la matriz de contigüidades A . En el caso de la red de citación, el elemento i, j -ésimo de la matriz A es igual a 1 si i cita a j al menos en una ocasión, y 0 en caso contrario. La matriz A es asimétrica, porque la fila i -ésima de A indica a qué investigadores cita i y la columna i -ésima indica qué investigadores están citando a i .

Hasta ahora, hemos asumido que todos los vínculos de citación son igual de fuertes. Sin embargo, un autor puede citar a otro autor más de una vez, y el artículo donde se realiza la cita puede estar firmado por varios autores. Newman (2001b), abordando este tema en las redes de colaboración, definió el vínculo mutuo entre dos colaboradores i y j mediante el siguiente término

$$c_{ij} = \sum^k \frac{\delta_i^k \delta_j^k}{n_k - 1}, \quad [6]$$

donde se calcula el sumatorio de todas las publicaciones registradas en coautoría entre i o j , y el indicador δ_i^k es igual a uno si el investigador i ha colaborado en la publicación k ésima escrita por n_k investigadores. La matriz $\{c_{ij}, \forall i, j\}$ se denomina matriz de contigüidades ponderada. El razonamiento para la especificación [6] es aumentar la intensidad de la relación en base a la multiplicidad de la colaboración y reducirla mediante su multiplicación por el término $1/n_k-1$, implicando esto último que i divide el tiempo dedicado al estudio de k interactuando con n_k-1 otros investigadores. La especificación [6] de Newman (2001b) puede ser adaptada a las redes de citación del siguiente modo:

$$c_{ij} = \sum^k \frac{\delta_{ij}^k}{n_k}, \quad [7]$$

donde el indicador δ_{ij}^k es igual a uno si i citó a j en la publicación k ésima escrita por n_k autores. Al igual que en [6], el vínculo direccionado entre i y j es nuevamente aumentado en base a la multiplicidad de la citación y reducido mediante su multiplicación por el término $1/n_k$, que mide la probabilidad de que i fue responsable de citar a j en la publicación k .

Para hacer efectiva la multiplicidad en la matriz de referencia, en [7] multiplicamos la respectiva matriz de contigüidades A por el término c_{ij} . Utilizando la matriz de contigüidades ponderada de la red de citación, los elementos de \mathbf{W} se definen como sigue:

$$w_{ij} = \begin{cases} c_{ij} & \text{si } i \text{ cita a } j \\ \frac{1}{2} \max_{\forall k} \left\{ c_{ik} + \frac{c_{kj}}{2} \right\} & \text{si } i \text{ cita a } k \text{ y } k \text{ cita a } j, \end{cases} \quad [8]$$

donde la $k=(1,2,\dots,K)$ representa el conjunto de investigadores a los que cita i . La medida en [8] descuenta la citación indirecta recalculando el valor por la mitad. Esto está en línea con las distancias de ruta estandarizadas utilizadas por Burt (2010). La idea detrás de la suma máxima de rutas ponderadas entre i y j es que i está familiarizado con las publicaciones de j a través de k «fuentes» diferentes. Dado que la información sobre j no se acumula sino que más bien se solapa, i no puede saber más sobre j de lo que conoce sobre él a través de la suma máxima de sus conexiones con k y la conexión de k con j . La matriz W definida en (8) es la matriz de ponderaciones de referencia \mathbf{W} antes de su estandarización.

El paso final en el proceso de operacionalización de la matriz de ponderaciones sociales consiste en

una estandarización. Aplicamos una estandarización por fila $w_{ij} = w_{ij}/w_i$, donde w_i es igual a la suma de todos los elementos de la fila i -ésima de \mathbf{W} . La estandarización por fila redimensiona la influencia recibida, mientras que la estandarización por columna redimensiona la influencia ejercida (Leenders, 2002). En nuestro contexto, la estandarización por fila significa que la suma de las influencias recibidas de todos los investigadores con influencia sobre i es igual a la unidad, pero la intensidad relativa de cada influencia individual se mantiene inalterada.

4. Matrices de ponderaciones sociales alternativas

En las secciones anteriores, se ha presentado la secuencia de argumentos que nos llevan a definir la operacionalización preferida de la matriz de ponderaciones sociales [8]. El procedimiento estándar para seleccionar la matriz de ponderaciones en econometría espacial consiste en testar el comportamiento de varias matrices con la ayuda de test estadísticos de comparación de modelos (p. ej., LeSage y Pace, 2009: cap. 6).

Como respuesta a los argumentos que reclaman aplicar enfoques más sustantivos (Corrado y Fingleton, 2012), en primer lugar especificamos la matriz de ponderaciones sociales \mathbf{W} de referencia basándonos en conceptos socioeconómicos y en esta sección procedemos a testar su comportamiento comparándolo con varias matrices alternativas. Cada una de estas matrices alternativas se obtiene relajando uno de los criterios impuestos sobre la matriz \mathbf{W} de referencia.

La primera matriz alternativa, W_1 , operacionaliza la matriz de ponderaciones mediante una medida de cohesión en la red de colaboración. Utilizando la matriz de contigüidades ponderada [6] de la red de colaboración, los elementos de la W_1 se definen como:

$$w_{ij} = \begin{cases} c_{ij} & \text{si } i, j \text{ son vecinos directos} \\ \frac{1}{2} \max_{\forall k} \left\{ c_{ik} + \frac{c_{kj}}{2} \right\} & \text{si } i, j \text{ distancia soc. de 2 vía } k, \end{cases} \quad [9]$$

De modo similar a [8], la $k=(1,2,\dots,K)$ representa el conjunto de investigadores con los que colaboraron tanto i como j . La medida [9] descuenta las colaboraciones indirectas recalculando el valor por la mitad. A resultados del test de \mathbf{W} frente a W_1 se obtiene la preferencia estadística del proceso social,

esto es, *comparación* o *comunicación*, que influye en la comunidad de investigadores.

La segunda matriz alternativa, W_2 , se refiere a la red de citación, pero ignora en la medida de cohesión [8] la multiplicidad de citas y el número de coautores [7]. De ahí que la matriz W_2 resulte de aplicar [8] a la matriz de contigüidades plana A basada en la red de citación. El rechazo de la especificación W_2 frente a \mathbf{W} equivale a la preferencia estadística de tener en cuenta la «intensidad» de la conexión de i a j .

Las dos siguientes matrices alternativas, W_3 y W_4 , modifican el horizonte de influencia. W_3 expande la distancia sociométrica máxima a tres. Modificamos la medida de cohesión [8] para reflejar distancias de hasta tres, tal como se muestra en [10], y la aplicamos a la red de citación.

$$w_{ij} = \begin{cases} c_{ij} & \text{si } i, j \text{ son vecinos directos} \\ \frac{1}{2} \max_{\forall k} \left\{ c_{ik} + \frac{c_{kj}}{2} \right\} & \text{si } i, j \text{ distancia soc. de 2 vía } k, \\ \frac{1}{3} \max_{\forall k, l} \left\{ c_{ik} + \frac{c_{kl}}{2} + \frac{c_{lj}}{3} \right\} & \text{si } i, j \text{ distancia soc. de 3 vía } k \text{ y } l. \end{cases} \quad [10]$$

Por su parte, W_4 limita el horizonte de influencia únicamente a las citas directas. El rechazo de W_3 o W_4 frente a \mathbf{W} determina la preferencia estadística por una distancia socioeconómica entre los investigadores de 2.

En el caso de las matrices W_5 y W_6 , adaptamos la medida de equivalencia estructural [5] propugnada por Leenders (2002) a las redes de citación y colaboración. Con la red de citación en mente, el valor de d_{ij} en [5] mide el subgrupo de investigadores a los que citan tanto i como j (o a los que no cita ninguno de los dos), y de los investigadores que citan tanto a i como a j (o que no citan a ninguno de ellos). De ahí que d_{ij} sea una medida de similitud en el patrón de cita, o según el caso, de no cita. d_{ij} es exactamente igual a 1 si i y j citan exactamente al mismo grupo de investigadores y son citados por el mismo grupo de investigadores (no teniendo por qué coincidir necesariamente ambos grupos). d_{ij} es inferior a 1 si i y j citan y son citados por grupos de autores que se solapan solo en parte. d_{ij} es 0 si i cita precisamente a investigadores a los que j no cita y es citado por aquellos investigadores que no citan a j . El inconveniente de d_{ij} en [5] es que puede ser distinto de cero también para individuos bastante distantes entre

sí. Su equivalencia estructural puede ser grande simplemente porque ambos no citan a otros muchos investigadores. En consecuencia, Leenders (2002) aboga por calcular la medida de equivalencias estructurales [5] únicamente para los pares de individuos que son capaces de observarse el uno al otro. Por dicha razón, limitamos solo a los pares de individuos separados por una distancia sociométrica de uno o dos la posibilidad de que la equivalencia estructural tome valores distintos de cero, y, además, basamos la medida de equivalencia estructural de dichos pares únicamente en una subred compuesta por todos sus citantes y citados directos. Esta estrategia está en línea con los argumentos anteriormente esgrimidos a favor de la distancia sociométrica de dos. En resumen, modificamos la medida de equivalencia estructural de Leenders (2002), que se define en [5], como sigue

$$\tilde{d}_{ij} = \begin{cases} \frac{((\hat{s}_i - \tilde{s}_j) \cdot (\hat{s}_i - \tilde{s}_j))^{1/2}}{\sqrt{2n_{ij}}} & \text{para } i \neq j \text{ y } A_{ij}^2 > 0 \\ 0 & \text{de lo contrario,} \end{cases} \quad [11]$$

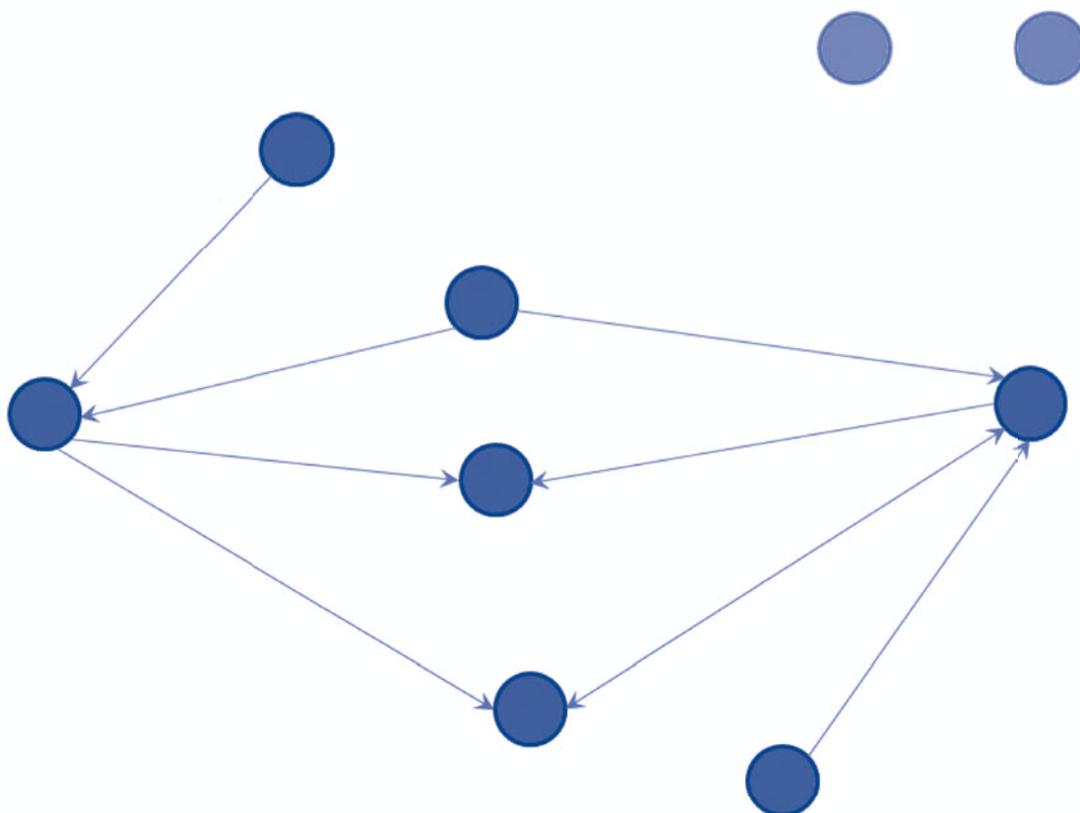
donde la matriz $A_2 = A \cdot A$, contiene valores distintos de cero para los pares de individuos ij separados por una distancia de ruta de uno o dos. n_{ij} es el tamaño de la subred en torno a i y j compuesta por citados y citantes de i y j . El vector \hat{s}_i está formado por la acumulación de los elementos seleccionados de la fila i -ésima y la columna i -ésima de la matriz de contigüidades A subyacente a la subred en torno a i y j . Como en [5], $\tilde{s}_j = I_{n_{ij}} - \hat{s}_j$. La matriz de equivalencia estructural $\tilde{D} = \{\tilde{d}_{ij}, \forall i, j\}$ es simétrica y mide, a través de la distancia euclídea, la ratio entre el número de citados y citantes comunes y el número de todos los citados y citantes de cualquier par. El gráfico 2 ilustra y discute la medida d_{ij} para un par hipotético ij .

Para ajustar consecuentemente la matriz de equivalencia estructural \tilde{D} , multiplicamos cada elemento de \tilde{D} por

$$C_{ij} = \sum_l c_{il}, \forall l: A_{ii} = 1 \text{ y } A_{jj} = 1, \quad [12]$$

donde se calcula el sumatorio para todas las citas l que i y j tienen en común. La intensidad de la equivalencia estructural definida de este modo puede interpretarse en el sentido de cuánto desea el investigador i alinearse con el investigador j . La matriz de ponderaciones sociales W_5 contiene entonces elementos w_{ij} distintos de cero con respecto a todos los pares de investigadores no separados por más de dos vínculos, tal que w_{ij} es el resultado

GRÁFICO 2
ILUSTRACIÓN DE LA MEDIDA DE EQUIVALENCIA ESTRUCTURAL PARA UN PAR HIPOTÉTICO DE UNA RED QUE CONSTA DE NUEVE NODOS



Notas: Los nodos 1 a 5 son autores citados o que citan y pertenecen a la subred de trabajo. Los nodos 6 y 7 no son parte de la red entera, debido a algunos vínculos con los nodos 1 a 5 (aquí no es relevante como son). Los nodos 6 y 7 afectan al [5], pero no a la medida [11].

de multiplicar la equivalencia estructural [11] por la intensidad de esta equivalencia [12], es decir,

$$w_{ij} = C_{ij} \tilde{d}_{ij} \quad [13]$$

La matriz W_6 aplica la medida de equivalencia estructural [13] en la red de colaboración, teniendo en cuenta la intensidad de los lazos de colaboración c_{ij} definidos en [6]. El test de \mathbf{W} frente W_5 puede resolver estadísticamente la cuestión, tanto si la

comparación está suficientemente representada, mediante el uso de la medida de cohesión simple en la red de citación, como si no. Además, rechazar W_6 frente a \mathbf{W} y W_5 significa una clara preferencia estadística por la red de citación con independencia de cuál sea la medida utilizada.

El cuadro n.º 3 resume la notación y presenta algunos estadísticos descriptivos relativos a la matriz \mathbf{W} de referencia y a las matrices alternativas W_1 a W_6 .

CUADRO N.º 3

ESTADÍSTICOS DESCRIPTIVOS DE LA MATRIZ SOCIAL DE PESOS DE REFERENCIA W Y LAS MATRICES ALTERNATIVAS W_1 A W_6 PREVIAS A LA NORMALIZACIÓN POR FILAS. MEDIAS Y RANGO BASADOS EN VALORES NO NULOS

MATRIZ	DESVIACIÓN DE W	CELDA NO NULAS	RANGO (MEDIA)
W	Cohesion en la red de.....	19,086	0,10 - 15,79 (0,67)
W_1	Cohesión y colaboración en red.....	12,670	0,03 - 26,50 (1,33)
W_2	como W , pero ignorando multiplicidad y coautorías.....	19,086	0,75 - 1,75 (0,83)
W_3	como W , distancia sociométrica hasta 3.....	69,561	0,10 - 15,79 (0,58)
W_4	como W distancia sociométrica hasta 1.....	3,297	0,10 - 11,00 (0,78)
W_5	Comparación en red de citación.....	17,517	0,07 - 16,58 (0,83)
W_6	Comparación en red de colaboración.....	11,254	0,01 - 22,05 (0,74)

IV. TESTANDO LA W EN EL MODELO DE AUTOCORRELACIÓN POR LA RED

En esta sección comparamos las matrices alternativas W_1 a W_6 con la matriz W de referencia con la ayuda de modelos que estudian los efectos de red. Téngase en cuenta que en el caso que estos test dieran como resultado el rechazo de la matriz preferida, no podría inferirse automáticamente que W no está correctamente especificada y que las teorías subyacentes son erróneas. Dicho rechazo también podría significar que los criterios exclusivamente estadísticos podrían ser insuficientes para seleccionar la matriz de ponderaciones sociales óptimas. Debido a que los test se basan en el tamaño y la significación de los parámetros del efecto espacial, se selecciona la matriz con mejor comportamiento en términos puramente estadísticos. Tal y como señalan Corrado y Fingleton (2012), el tamaño y la significación del parámetro del efecto espacial pueden traslucir un mensaje engañoso, ya que estos valores podrían simplemente estar recogiendo el efecto de la correlación espacial existente entre variables dependientes omitidas del modelo.

La razón última para operacionalizar la matriz de ponderaciones sociales con base en las teorías de la influencia social consiste en obtener estimaciones correctas del efecto de las redes sociales profesionales sobre la productividad de los investigadores. En el campo de las ciencias económicas, el valor (calidad) y el número (cantidad) de las publicaciones de un investigador suele considerarse el baremo más creíble de la productividad de dicho científico. En esta sección, comparamos el comportamiento de la matriz de ponderaciones sociales W de referencia con sus alternativas W_1 a W_6 valiéndonos de un modelo de autocorrelación por la red sobre la productividad, y aplicamos un procedimiento de contraste bayesiano.

A tal efecto, utilizamos versiones del modelo espacial general [1], donde $\rho \neq 0$ o $\lambda \neq 0$. De acuerdo con la taxonomía de los modelos espaciales (p. ej., Elhorst, 2010), existen cinco modelos posibles:

$$\begin{aligned}
 \mathbf{y} &= \rho \mathbf{W} \mathbf{y} + \alpha \mathbf{1} + \mathbf{X} \beta + \varepsilon && \text{(SAR)} \\
 \mathbf{y} &= \alpha \mathbf{1} + \mathbf{X} \beta + \mathbf{u}; \mathbf{u} = \lambda \mathbf{W} \mathbf{u} + \varepsilon && \text{(SEM)} \\
 \mathbf{y} &= \rho \mathbf{W} \mathbf{y} + \alpha \mathbf{1} + \mathbf{X} \beta + \mathbf{u}; \mathbf{u} = \lambda \mathbf{W} \mathbf{u} + \varepsilon && \text{(SAC)} \\
 \mathbf{y} &= \rho \mathbf{W} \mathbf{y} + \alpha \mathbf{1} + \mathbf{X} \beta + \mathbf{W} \mathbf{X} \theta + \varepsilon && \text{(SDM)} \\
 \mathbf{y} &= \alpha \mathbf{1} + \mathbf{X} \beta + \mathbf{W} \mathbf{X} \theta + \mathbf{u}; \mathbf{u} = \lambda \mathbf{W} \mathbf{u} + \varepsilon && \text{(SDEM)}
 \end{aligned}$$

donde $\varepsilon \sim iid(0, \sigma^2 \mathbf{I})$ (16). Estos modelos asumen que la red influye directamente en la productividad de los investigadores por la vía de la producción de los demás investigadores que conforman la comunidad analizada (si $\rho \neq 0$) o bien a través de factores no observados (si $\lambda \neq 0$). En los modelos arriba citados, imponemos una matriz de ponderaciones sociales W de orden $N \times N$ estandarizada por fila. Esta puede ser o bien la matriz de ponderaciones W de referencia o una de las seis matrices alternativas. $\mathbf{1}$ es un vector $N \times 1$ de unos asociado al término constante α y \mathbf{X} es una matriz de covariables de productividad investigadora de orden $N \times k$ que mide las características individuales e institucionales (17). Por otro lado, \mathbf{y} es un vector de productividades investigadoras $N \times 1$. La productividad investigadora suele medirse por el número de artículos publicados en revistas ajustados por la calidad. En el presente trabajo, primero calculamos, para cada individuo i , su productividad investigadora en cada año t del período 2000-2010 como

$$y_{it} = \sum_{k=1}^{N_{it}} \frac{q_k}{a_k} \tag{14}$$

La expresión [14] consiste en la suma de los ratios entre las ponderaciones de calidad de cada revista q_j y el número de autores a_j de la k ésima publicación

del científico i . Dicho sumatorio se refiere a todas las publicaciones que el economista i ha publicado en el año t . Para $q \in [0,1]$, utilizamos las ponderaciones de calidad de cada revista propuestas por Combes y Linnemer (2010). Estos autores indexan unas 1.200 revistas económicas. Por último, calculamos, para cada individuo i , su productividad investigadora anual media a fin de suavizar las productividades volátiles anuales y de tener en cuenta el diferente número de años en activo del investigador. Tal productividad media se convierte en los modelos (SAR - SDEM) en la variable dependiente que se desea explicar a partir del conjunto de variables de control centradas en el efecto de red capturado por ρ o por λ .

El cuadro n.º 4 reporta las estimas de las ρ y λ para todas las matrices consideradas en los cinco modelos espaciales. Las ρ estimadas son altamente significativas para todas las especificaciones de matrices de ponderaciones y el efecto de red que señalizan es positivo, oscilando entre 0,03 y 0,44, pero siendo en la mayoría de los casos superior a 0,2. El efecto de ρ es inferior solo en los modelos con variables explicativas sujetas a correlación por la red (SDM), debido a que parte del efecto de red actúa vía covariables. Un patrón similar puede observarse con las λ estimadas. En la mayoría de los casos muestran un valor superior a 0,2 en los modelos SEM «limpios». Solo en los modelos SDEM, donde las perturbaciones correlacionadas por la red se estiman junto a las variables explicativas correlacionadas por la red, son las λ más pequeñas, aunque en la mayoría de los casos siguen siendo significativas. Una perspectiva interesante la ofrecen los modelos SAC, donde la autocorrelación por la red se combina con las perturbaciones correlacionadas por la red. Aquí las ρ apenas difieren de las estimadas en el modelo SAR, salvo por los dos casos en los que compensan las λ significativas.

Si bien el foco de este estudio se centra en la correcta especificación de la matriz de ponderaciones sociales, existe una interpretación del parámetro de red estimado. Este puede describirse, en el caso de la matriz \mathbf{W} de referencia y el modelo SAR, como sigue. Si un investigador perteneciente a la comunidad publica un artículo adicional en una revista de calidad q , la productividad de los demás investigadores que están influidos por dicho investigador aumenta de media en $0,16 \times q$ (cf. [res_prod]) artículos. El coeficiente 0,16 es la suma media por filas (o columnas) de los elementos de fuera de diagonal de la matriz $(I - \rho W)^{-1}$, con $\rho = 0,264$ y $W = \mathbf{W}$ (véase detalles en LeSage y Pace 2009).

Para comparar la matriz de ponderaciones sociales \mathbf{W} de referencia con las alternativas, utilizamos probabilidades de un modelo bayesiano posterior. El enfoque bayesiano no requiere formular ninguna hipótesis nula y permite contrastar simultáneamente cualquier número de matrices de ponderaciones a fin de encontrar la estadísticamente preferida. Este método, que parte de la teoría bayesiana de comparación de modelos, es propugnado por LeSage y Pace (2009: cap. 6) cuando existe incertidumbre sobre la especificación exacta de la matriz de ponderaciones espaciales.

Supónganse M matrices de ponderaciones alternativas, $W_i, i = 1, 2, \dots, M$. Cada una de ellas produce en un modelo con autocorrelación por la red (de SAR a SDEM) un valor de la función de probabilidad (18). Asumamos ahora una distribución apriorística del espacio de parámetros $\theta = (\rho, \beta, \lambda, \sigma^2)$. Las probabilidades posteriores para cada W_i -modelo son

$$p(W_i|y) = \frac{p(y|W_i)p(W_i)}{p(y)}. \quad [15]$$

CUADRO N.º 4

ESTIMACIONES DEL EFECTO DE LOS PARÁMETROS DE RED EN CUATRO TIPOS DE MODELOS ESPACIALES PARA LA MATRIZ DE REFERENCIA \mathbf{W} Y DE SEIS MATRICES ALTERNATIVAS \mathbf{W}_i

	SAR MODEL		SEM MODEL		SAC MODEL				SDM MODEL		SDEM MODEL	
	ρ	std. dev.	λ	std. dev.	ρ	std. dev.	λ	std. dev.	ρ	std. dev.	λ	std. dev.
\mathbf{W}	0,264	0,027	0,251	0,038	0,238	0,024	0,050	0,001	0,156	0,043	0,088	0,017
W_1	0,226	0,009	0,199	0,036	0,444	0,024	-0,343	0,002	0,163	0,032	0,146	0,036
W_2	0,262	0,028	0,25	0,040	0,248	0,047	0,026	0,091	0,145	0,011	0,082	0,017
W_3	0,290	0,033	0,294	0,051	0,260	0,056	0,061	0,115	0,183	0,065	0,085	0,021
W_4	0,149	0,007	0,127	0,011	0,176	0,001	-0,034	0,030	0,078	0,025	0,051	0,029
W_5	0,241	0,017	0,226	0,024	0,245	0,042	-0,005	0,078	0,077	0,043	0,039	0,012
W_6	0,116	0,007	0,07	0,011	0,180	0,002	-0,108	0,062	0,029	0,003	0,023	0,063

La $p(y|W_i)$ en (15) es la probabilidad marginal que se obtiene según (16) utilizando el algoritmo MCMC (19).

$$p(y|W_i) = \int p(y|\theta^i, W_i) p(\theta^i|W_i) d\theta^i \quad [16]$$

Las probabilidades $p(y|W_i)$ se utilizan para testar si la matriz **W** de referencia es estadísticamente superior para explicar los modelos centrados en la productividad investigadora. El cuadro n.º 5 reporta las probabilidades en un modelo bayesiano posterior de los test entre la matriz **W** de referencia y las alternativas. La matriz de referencia supera estadísticamente a todas las matrices alternativas en cada uno de los cuatro modelos. Así pues, con base en estos test bayesianos, podemos excluir estadísticamente la W_1 alternativa y, por consiguiente, la hipótesis de la influencia social vía *comunicación*, es decir, vía la red de colaboraciones. Tener en cuenta la intensidad de los vínculos al medir la equivalencia estructural es superior estadísticamente a ignorar dicha intensidad, ya que **W** también supera a W_2 . Las matrices W_3 y W_4 son igualmente rechazadas. Obsérvese que esto es así a pesar de los mayores efectos de red de la W_3 en los cuatro modelos. En consecuencia, concluimos que la distancia sociométrica de dos es necesaria y suficiente para explicar los efectos de red sobre la productividad investigadora. La matriz **W** de referencia es estadísticamente preferida a la matriz W_5 lo que corrobora nuestra proposición inicial de que el efecto comparación está suficientemente expresado por la matriz de citación específica. Además, un test por pares entre W_5 y W_6 (no mostrado aquí) rechaza la W_6 , lo que significa que la red de citación es estadísticamente superior a la red de colaboración también cuando se utiliza una medida de equivalencia estructural en ambas redes.

La argumentación estadística mediante probabilidades de un modelo bayesiano posterior nos ha ayudado a discriminar entre las especificaciones alternativas. Además, el análisis demuestra que, sea cual sea la medida utilizada, las matrices de ponderaciones basadas en la red de citación superan a las basadas en la red de colaboración, un indicativo de que el conducto dominante del efecto ejercido por la red social sobre la productividad investigadora se manifiesta a través de los incentivos competitivos.

V. CONCLUSIONES

En el presente trabajo proponemos un enfoque estructurado para la especificación de matrices de ponderaciones sociales que reflejen la influencia de las redes en el comportamiento de los individuos que forman parte de las mismas. Dicho enfoque se ilustra aplicándolo al estudio del impacto que las redes de colaboración y de citación tienen sobre la productividad investigadora de los economistas académicos.

Nuestro enfoque se compone de tres pasos. En primer lugar, reportamos por medio de estadísticos descriptivos las propiedades importantes de las redes analizadas para documentar que satisfacen los requisitos de «mundo pequeño», lo que implica que funcionan bien como redes sociales. En segundo lugar, nos apoyamos en dos teorías socioeconómicas para identificar el círculo de individuos influyentes. Estas dos teorías competidoras acerca de la influencia social en el comportamiento de los individuos describen el rol de la *comunicación* y la *comparación*. En nuestra ilustración, aplicamos estas teorías a la comunidad de investigadores, argumentando que debido sobre todo a las fuerzas competitivas que caracterizan el área de la investigación, el

CUADRO N.º 5

MODELO BAYESIANO POSTERIOR DE PROBABILIDADES ASUMIENDO VARIOS CONJUNTOS DE MATRICES DE PONDERACIÓN COMPETITIVAS

	SAR MODEL	SEM MODEL	SDM MODEL	SDEM MODEL
W	0,966	0,983	0,723	0,793
W_1	0,000	0,000	0,000	0,000
W_2	0,033	0,016	0,030	0,028
W_3	0,001	0,001	0,000	0,000
W_4	0,000	0,000	0,000	0,000
W_5	0,001	0,000	0,247	0,179
W_6	0,000	0,000	0,000	0,000

Notas: El modelo quinto fue el SAC, pero el posterior bayesiano aún no está, actualmente, finalizado. La suma de las probabilidades por columna es uno. La probabilidad más alta en cada columna indica la matriz de ponderaciones preferida para cada modelo espacial.

proceso de comparación resulta ser la palanca social preferida de la productividad investigadora. Estos dos primeros pasos nos conducen a una especificación de la matriz de ponderaciones sociales de referencia que se basa en la red de citación. Para cuantificar la proximidad entre los pares de individuos, utilizamos una medida de cohesión y una medida de equivalencia estructural.

El tercer paso consiste en definir seis matrices de ponderaciones alternativas, las cuales se obtienen relajando en cada caso una hipótesis concreta subyacente a la matriz de ponderaciones sociales de referencia. A continuación, estimamos para cada una de las matrices alternativas cinco tipos de modelos con autocorrelación por la red enfocados en la productividad investigadora partiendo de la taxonomía de modelos econométricos espaciales. Luego, aplicamos un procedimiento bayesiano de contraste y selección de modelos para medir la superioridad de la matriz de referencia comparada con las matrices alternativas. En nuestro ejemplo ilustrativo, concluimos que la matriz de ponderaciones sociales de referencia supera a las alternativas en todos los casos. Además, el análisis bayesiano demuestra que las matrices de ponderaciones basadas en la red de citación superan a las basadas en la red de colaboración cualquiera que sea la medida utilizada.

Nuestros resultados sugieren que la red de citación, que captura el proceso sociológico de la *comparación*, es el retrato preferido de las interacciones entre los investigadores. Este resultado es acorde con la opinión de Burt (1987, 2010) en el sentido de que la comparación es el proceso social que actúa con mayor intensidad dentro de un grupo de individuos competidores.

Si bien se han utilizado datos bibliográficos relativos a economistas académicos de la región germanófona, el enfoque propuesto puede aplicarse directamente a investigadores de otras disciplinas, así como de otros contextos socioeconómicos, cuando el fenómeno a estudiar sea cómo influye el efecto de la red sobre el comportamiento de los individuos.

NOTAS

(*) La autora agradece a Toke Aidt, Heinrich Ursprung, Eylem Gevrek, Laurent Bach y Paul Elhorst sus comentarios a versiones anteriores de este artículo. También agradece a Simon Heß, Susann Adloff y Sebastian Kopf su apoyo en la investigación y a José Manuel Barrueco, responsable del mantenimiento de la base de datos CITec, por permitirle acceder a sus datos en formato simplificado.

(1) La matriz de ponderaciones espaciales representa la distribución espacial de las unidades analizadas y es un *input* crucial en los modelos econométricos espaciales (LESAGE y PACE, 2009).

(2) En otras redes sociales, los vínculos podrían no ser observados con una precisión del 100 por 100, porque las personas podrían no recordar a todos sus conocidos. Además, las relaciones de conocidos pueden ser de diferentes tipos: parientes, amigos, compañeros de trabajo, etc.

(3) www.forschungsmonitoring.org

(4) En alemán: *Verein für Socialpolitik*. www.socialpolitik.org

(5) P. ej., RAUBER y URSPRUNG (2008); FABEL, HEIN y HOFMEISTER (2008).

(6) Los economistas con «colaboradores solamente externos» (289) podrían ser parte del componente gigante, si incluyéramos a los colaboradores externos. Sin embargo, sucede que considerar a los colaboradores externos solo integra a menos de la mitad de ellos (125) en el componente gigante. El factor que motiva ser categorizado como economista aislado es un historial de publicaciones pobre: la mediana de publicaciones en revistas de los economistas aislados es 4 y su moda es 1, mientras que en el caso de los economistas no aislados es de 11 y 4, respectivamente.

(7) CITec = Citation in Economics: <http://citec.repec.org/>. La base de datos CITec contiene citas acerca de una tercera parte de los artículos RePec (July 2013). La consecuencia para los economistas relevantes es que solo en torno a 7.000 de los 13.242 artículos publicados en revistas forman parte de la base de datos CITec.

(8) Las autocitas harían que esos economistas estuvieran conectados a algunos miembros de la red de citación, pero solo por la vía de publicaciones de las que hubieran sido coautores. Por lo tanto, dichos vínculos se excluyen no solo para ellos sino también para todos los investigadores involucrados.

(9) Estos son colaboradores locales ajenos al campo de la economía o economistas académicos que trabajaron en algún país germanófono con anterioridad o después del año de corte 2010. Unos pocos son economistas no académicos, y algunos son estudiantes de doctorado, que no forman parte de los economistas relevantes de acuerdo con los criterios con los que se ha definido el conjunto.

(10) El coeficiente de *clustering* para grafos aleatorios de tamaño $N = 845$ (885), con grado promedio 3,36(3,54) es, según las predicciones del modelo, de $\sim 0,004$.

(11) La distancia de ruta (*path distance*) de dos miembros de una red viene dada por el número mínimo de vínculos existentes entre ellos. Para obtener datos estandarizados, se divide cada valor de distancia de ruta por alguna medida de distancia. Este es un enfoque equivalente al utilizado en econometría espacial.

(12) Recuérdese que las autocitas están excluidas de la red de citación.

(13) <http://tool.handelsblatt.com/tabelle/index.php?id=79&pc=250>

(14) La red de citación es unas tres veces más densa que la red de colaboración de los economistas más prestigiosos, incluso si se tiene en cuenta la mayor incidencia de los vínculos de citación en general.

(15) Estos cálculos se realizan respecto a los investigadores pertenecientes al componente gigante (cf. cuadro n.º 1).

(16) Las abreviaturas utilizadas para nombrar estos modelos son las habituales en la literatura sobre econometría espacial. SAR hace referencia al modelo espacial con autocorrelación; SEM significa el modelo espacial con errores correlacionados; SDM es el modelo espacial Durbin; SDEM es el modelo espacial Durbin con errores correlacionados; mientras que SAC no responde a unas siglas con un significado concreto.

(17) En referencia a nuestros anteriores trabajos centrados en datos de productividad investigadora (,), utilizamos el siguiente conjunto de variables de control: antigüedad académica, sexo, campo de investigación, cargo académico, país, instituto de investigación, tamaño del departamento, medida de colaboración externa, y porcentaje del clastro que no publica artículos.

(18) Como la probabilidad del modelo bayesiano posterior no está disponible para el modelo SAC, no consideramos esta especificación a los efectos del procedimiento de contraste bayesiano.

(19) Utilizamos aquí los códigos de Matlab `sar_g`, `sdm_g`, `sem_g`, desarrollados por LeSage. Se encuentran disponibles gratuitamente en <http://www.spatial-econometrics.com/>

BIBLIOGRAFÍA

ACEDO, F. J.; BARROSO, C.; CASANUEVA, C., y J. L. GALÁN (2006), «Co-authorship in management and organizational studies: An empirical and network analysis», *Journal of Management Studies*, 43: 957-983.

ANSELIN, L. (1988), *Spatial econometrics: Methods and models*, Kluwer, Dordrecht.

— (2010), «Thirty years of spatial econometrics», *Papers in Regional Science*, 89: 3-25.

BOSQUET, C., y P.-P. COMBES (2013), «Are academics who publish more also more cited? Individual determinants of publication and citation records», *Scientometrics*, 97: 831-857.

BURRIDGE, P.; ELHORST, J. P., y K. ZIGOVA (2016), «Group interaction in research: A general nesting spatial model with heteroskedastic shocks», en BALTAGI B., LESAGE J. P., y PACE R.K. (eds.): *Advances in Econometrics*, vol. 37, forthcoming.

BURT, R. S. (1987), «Social contagion and innovation: Cohesion versus structural equivalence», *American Journal of Sociology*, 92: 1287-1335.

— (2010), *Neighbor networks. Competitive advantage local and personal*, Oxford University Press.

CARAYOL, N. (2008), «An economic theory of academic competition: Dynamic incentives and endogenous cumulative advantages», en ALBERT M., S. VOIGT, y D. SCHMIDTCHEN (eds.): *Scientific competition. Conferences on new political economy*, Volume 25: 179-203. Mohr Siebeck, Tübingen.

COLEMAN, J.; KATZ, E., y H. MENZEL (1966), *Medical innovation*, Bobbs-Merrill, Nueva York.

COMBES, P.-P., y L. LINNEMER (2010), «Inferring missing citations: A quantitative multicriteria ranking of all journals in economics», GREQAM DT 2010-25.

CORRADO, L., y B. FINGLETON (2012), «Where is the economics in spatial econometrics?», *Journal of Regional Science*, 52: 210-239.

CRANE, D. (1972), *Invisible colleges: Diffusion of knowledge in scientific communities*, The University of Chicago Press.

CRONIN, B.; SHAW, D., y K. LA BARRE (2003), «A cast of thousands: Coauthorship and subauthorship collaboration in the 20th century as manifested in the scholarly journal literature of psychology and philosophy», *Journal of the American Society for Information Science and Technology*, 54: 855-871.

DE NOOY, W.; MRVAR, A., y V. BATAGELJ (2005), *Exploratory network analysis with Pajek*, Cambridge University Press.

ELHORST, J. P. (2010), «Applied spatial econometrics: Raising the bar», *Spatial Economic Analysis*, 5: 9-28.

ELHORST, J. P., y K. ZIGOVA (2014), «Competition in research activity among economic departments: Evidence by negative spatial autocorrelation», *Geographical Analysis*, 46: 104-125.

FABEL, O.; HEIN, M., y R. HOFMEISTER (2008), «Research productivity in business economics: An investigation of Austrian, German and Swiss universities», *German Economic Review*, 9: 506-531.

FRIEDKIN, N. E. (1990), «Social networks in structural equation models», *Social Psychology Quarterly*, 53: 316-328.

GALASKIEWICZ, J., y R. S. BURT (1991), «Interorganization contagion in corporate philanthropy», *Administrative Science Quarterly*, 36: 88-105.

GARFIELD, E. (1963), «Science citation index 1961», *Science Citation Index*, v-xvi.

GOYAL, S.; VAN DER LEIJ, M. J., y J. L. MORAGA-GONZÁLEZ (2006), «Economics: An emerging small world», *Journal of Political Economy*, 114: 403-412.

HOMANS, G. C. (1961), *Social behavior: Its elementary forms*, Harcourt: Brace & World, Inc.

HOTELLING, H. (1929), «Stability in competition», *The Economic Journal*, 39: 41-57.

HUMMON, N., y P. DOREIAN (1989), «Connectivity in a citation network: The development of DNA theory», *Social Networks*, 11: 39-63.

LABAND, D. N., y R. D. TOLLISON (2000), «Intellectual collaboration», *Journal of Political Economy* 108 (3), 632-662.

LEENDERS, R. T. A. J. (2002), «Modeling social influence through network autocorrelation: constructing the weight matrix», *Social Networks*, 24: 21-47.

LESAGE, J. P., y R. K. PACE (2009), *Introduction to spatial econometrics*, CRC Press\ Taylor & Francis Group, Boca Raton.

MANSKI, C. (1993), «Identification of endogenous social effects: The reflection problem», *Review of Economic Studies*, 60: 531-542.

MARSDEN, P. V., y N. E. FRIEDKIN (1993), «Network studies of social influence», *Sociological Methods and Research*, 22: 127-151.

MARSDEN, P. V., y J. PODOLNY (1990), «Dynamic analysis of network diffusion processes», en WEESIE, J. y FLAP, H. (eds.): *Social Networks Through Time*, Isor, Utrech.

MELIN, G., y O. PERSSON (1996), «Studying research collaboration using co-authorships», *Scientometrics*, 36: 363-377.

MERTON, R. K. (1973), *The sociology of science: Theoretical and empirical investigations*, University of Chicago Press, Chicago.

MILGRAM, S. (1967), «The small-world problem», *Psychology Today*, 1: 61-67.

MOFFIT, R. A. (2001), «Policy interventions, low-level equilibria and social interactions», en DURLAUF, S., y YOUNG, P. (eds.): *Social Dynamics*, MIT Press.

MOODY, J. (2004), «The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999», *American Sociological Review*, 69: 213-238.

NEWMAN, M. E. J. (2001a), «The structure of scientific collaboration networks», *PNAS*, 98: 404-409.

— (2001b), «Who is the best connected scientist? A study of scientific coauthorship networks», *Physical Review E* 67(arXiv:cond-mat/0011144v2).

— (2004), «Coauthorship networks and patterns of scientific collaboration», *PNAS*, 101 (S1), 5200-5205.

PALMA DE, A.; GINSBURGH, V.; PAPAGEORGIOU, Y. Y., y J.-F. THISE (1985), «The principle of minimum differentiation holds under sufficient heterogeneity», *Econometrica*, 53: 767-781.

PRICE DE SOLLA, D. (1965), «Networks of scientific papers. The pattern of bibliographic references indicates the nature of the scientific research front», *Science*, 149: 510-515.

RAUBER, M., y H. W. URSPRUNG (2008), «Life cycle and cohort productivity in economic research: The case of Germany», *German Economic Review*, 9: 431-456.

STEPHAN, P. E. (2010), «The economics of science», en B. HALL, H., y ROSENBERG, N. (Eds.): *Handbook of the economics of innovation*, Volume 1, pp. 217-273. Amsterdam, North Holland.

STRANG, D., y N. B. TUMA (1993), «Spatial and temporal heterogeneity in diffusion», *American Journal of Sociology*, 3: 614-639.

WATTS, D., y S. STROGATZ (1998), «Collective dynamics of 'small-world' networks», *Nature*, 393: 440-442.

WHITE, H. D.; WELLMAN, B., y N. NAZER (2004), «Does citation reflect social structure? Longitudinal evidence from the 'Globenet' interdisciplinary research group», *Journal of the American Society for Information Science and Technology*, 55: 111-126.