

## Efectos de la colaboración externa y de las redes de investigación como capital social en los resultados de investigación: el caso de los ingenieros mexicanos

RODRÍGUEZ-MIRAMONTES, Jorge\*† y MAQUEDA-RODRÍGUEZ, Gabriela

Recibido Abril 4, 2016; Aceptado Junio 16, 2016

### Resumen

Este trabajo analiza cómo las redes de colaboración influyen en los resultados de la investigación realizada por científicos mexicanos en el campo de la ingeniería. Para ello se elaboró una base de datos longitudinal de todas las publicaciones y citas en el Institute of Scientific Information (ISI) de 1980 a 2007, en la que participa al menos un ingeniero en el Sistema Nacional de Investigadores (SNI) de México. En este análisis se considerando tres puntos importantes: 1) los resultados de investigación (cantidad y calidad) se consideraron con base en el número de publicaciones en un periodo de dos años así como el número de citas recibidas a estos artículos en los cuatro años posteriores a su publicación; 2) los efectos de la colaboración externa se clasificaron en cinco tipos (interinstitucional, intrainstitucional, internacional, academia-empresa y academia gobierno), mientras que los artículos de un solo autor se consideraron como variable de control; 3) los efectos de la red se midieron en función de cuatro medidas de centralidad (grado, *closeness*, *betweenness* y *eigenvector*) y dos variables del tipo ego (*structural holes* y densidad). Nuestros resultados muestran que prácticamente todos los tipos de colaboración tienen una fuerte relación con los resultados de investigación a excepción de la colaboración con la empresa la cual presenta impacto alguno. Con respecto a las redes de colaboración, las variables de centralidad que afectan los resultados de investigación son: el grado de centralidad, la cercanía y de forma negativa la intermediación; respecto las variables de ego, la densidad y structural holes tienen un fuerte impacto.

**Red de Investigación, Colaboración Externa, Capital Social, Centralidad, Egonetworks.**

### Abstract

This paper analyzes how collaboration networks influence the results of research carried out by Mexican scientists in the field of engineering. For this purpose, a longitudinal database of all the publications and quotations was elaborated at the Institute of Scientific Information (ISI) from 1980 to 2007, in which at least one engineer participates in the National System of Investigators (SNI) of Mexico. In this analysis three important points were considered: 1) the research results (quantity and quality) were considered based on the number of publications in a period of two years as well as the number of citations received to these articles in the four following years To its publication; 2) the effects of external collaboration were classified into five types (interinstitutional, intrainstitutional, international, academic-business and government academy), whereas articles of a single author were considered as a control variable; 3) the effects of the network were measured in terms of four measures of centrality (degree, closeness, betweenness and eigenvector) and two ego-type variables (structural holes and density). Our results show that virtually all types of collaboration have a strong relationship with research results, except for collaboration with the company which has an impact. With respect to collaborative networks, the centrality variables that affect the research results are: the degree of centrality, proximity and negatively the intermediation; With respect to ego variables, density and structural holes have a strong impact.

**Research Network, External Collaboration, Social Capital, Centrality, Egonetw.**

**Citación** RODRÍGUEZ-MIRAMONTES, Jorge y MAQUEDA-RODRÍGUEZ, Gabriela. Efectos de la colaboración externa y de las redes de investigación como capital social en los resultados de investigación: el caso de los ingenieros mexicanos. Revista de Investigación y Desarrollo 2016, 2-4: 24-36

\* Correspondencia al Autor (Correo Electrónico: jrodriguez@cinvestav.mx.)

† Investigador contribuyendo como primer autor.

## Introducción

La colaboración en investigación junto con las nuevas tecnologías de la informática representa un papel importante del trabajo en equipo. Así, mediante la interacción entre científicos de la misma universidad, de diferentes centros de investigación e inclusive de diferentes países, la colaboración en el ámbito de la investigación científica es un fenómeno positivo y deseable. Los incentivos para colaborar son varios, por ejemplo, el acceso a recursos (Katz & Martin 1997), la acumulación de experiencia o pericia (Beaver 2001, Birnholtz 2007), el aprendizaje de nuevas habilidades (Heinze & kuhlman 2008), mayor productividad (Beaver & Rosen 1978), mayor calidad en los resultados (Rigby & Eder 2005), acceso a fondos (Beaver 2001) y prestigio (Beaver & Rosen 1978), entre otros.

Los beneficios en relación a la colaboración han sido plenamente documentados. Por ejemplo, estudios pioneros en el área como realizado por Lotka (1926), muestran una relación entre productividad y colaboración. Price y Beaver (1966) documentaron que los autores más productivos corresponden a los más colaborativos, lo cual fue confirmado por Zuckerman (1967). Acorde a Birnholtz (2007), existe una influencia positiva entre calidad y colaboración. Por su parte Katz & Hicks (1997) identificaron una relación positiva entre coautor y el número de citas recibidas. Estos trabajos emplean datos de países desarrollados. Sin embargo, los estudios que involucran a países en vías de desarrollo han sido menos explorados y son relativamente recientes (Schofer et al., 2000; Shurum, 2005; Gonzalez-Brambila y Veloso, 2007; Gonzalez Brambila, et al. 2013).

Este análisis contribuye a la literatura relacionada con colaboración y productividad en investigación mediante tres aproximaciones: 1) incrementa el conocimiento de la colaboración en países en desarrollo ya que se centra en los investigadores en ingeniería en México aludiendo que existe considerables diferencias entre disciplinas en los resultados de la colaboración (Cronin et al., 2003: 2004; Newman 2004); 2) el estudio se centra en un área estratégica de cualquier nación que es la ingeniería, cuyos estudios son casi nulos en este tipo de investigación; 3) incorpora el análisis de redes sociales lo que permite entender cómo las iteraciones interpersonales pueden afectar la productividad.

Este trabajo está organizado de la siguiente forma: la sección II cita algunos de los trabajos relacionados con esta investigación; la sección III presenta los datos y metodología junto con la descripción de las variables para el modelo econométrico correspondiente; la sección IV muestra los resultados de este análisis y por último la sección V contempla la discusión y conclusiones.

## Trabajos relacionados

En esta sección se presentan los trabajos relacionados a los objetivos de este análisis de acuerdo con dos enfoques: 1) impacto de la colaboración en la productividad científica y 2) análisis de redes de coautoría y su impacto en la productividad científica.

## Efectos de la colaboración en la productividad

Los estudios relacionados con las actividades de colaboración científica han demostrado los beneficios de ésta, por ejemplo: el acceso a equipos e instalaciones especiales, adquirir experiencia, acceso a materiales únicos, entre otros.

Desde varias décadas diversos análisis han documentado una fuerte relación entre productividad y colaboración (Lotka, 1926; Price & Beaver, 1966; Nudelman y Landers, 1972; Gordon, 1980; Pao, 1982; Diamon, 1985; Lawani, 1986; Narin y Whitlow, 1990; Adams G. et al., 2004; Lee y Bozeman, 2005).

En Europa Frenken et al., (2005), analizaron la producción del conocimiento europeo en biotecnología en un periodo de 1988 a 2002, enfocándose en el papel de la colaboración en investigación en dicha creación del conocimiento. Un aspecto importante de esta investigación es que diferenciaron el tipo de colaboración: de forma geográfica (nacional, europea e internacional) y entre instituciones (entre academia, fuera de academia e híbrido. Así, identificaron diversos aspectos interesantes: 1) la difusión del conocimiento académico medido como la tasa de citas es dependiente de las características inter e intra-organización de la colaboración; 2) el número de autores y el número de instituciones contribuyen positivamente y significativamente al número de citas recibidas por artículo; 3) la colaboración internacional incrementa el impacto en citas comparada con la nacional; 4) se presenta un efecto negativo en las citas cuando la colaboración es entre academia.

Por otra parte, Zi-Lin He et al., (2009) desarrollaron una base longitudinal de 65 científicos biomédicos de universidades de Nueva Zelanda con sus publicaciones en un periodo de 14 años, ellos encontraron que a nivel artículo la colaboración internacional e inter-universidad están relacionadas positivamente con la calidad de los artículos.

### **Análisis de redes de coautoría y su impacto en la productividad científica**

El análisis de redes sociales surgió bajo la idea de que los individuos están inmersos en grandes redes de relaciones sociales e iteraciones. Así, los sociólogos han desarrollado medidas como estatus sociométricos, centralidad, poder, prestigio, rank entre otras, para tratar de resolver preguntas de orden social. En años recientes el estudio de redes de coautoría se considera una forma importante de las redes sociales, misma que ha sido documentada como un método bibliométrico (Newman, 2001, Barbasí, et al., 2002). Sin embargo, existen pocos estudios que involucren el análisis de redes de coautoría en estudios de la productividad científica y aún menos que involucren modelos matemáticos para explicarla a través de variables de redes.

Existen dos nociones del capital social que emergen de la teoría de análisis de redes, los cuales son: "closure" y "structural holes". La perspectiva "closure" figura como la visión pesimista de la diversidad demográfica (Coleman 1988:1990; Portes y Sensenbrenner, 1993; McCain et al. 1983) y la perspectiva de "structural holes" como la visión optimista (Burt 1992; Ancona y Cadwell 1992; Bantel y Jackson 1989; Granovetter, 1973). Estas dos perspectivas en capital social no están en conflicto. Mientras que la perspectiva "closure" se enfoca en la presencia o ausencia de relaciones a nivel iteración local, los huecos estructurales son los responsables de los beneficios de la información y son quienes dividen al sistema social en forma global.

Por otra parte, Gonzalez-Brambila, et al., (2013) documentaron que las características dimensionales de la red afectan a la cantidad, pero no la calidad mientras que las características cognitivas tienen el efecto opuesto.

Respecto a las relaciones estructurales, identificaron que éstas tienen cierto grado de convergencia entre la calidad y cantidad. Adicionalmente, mostraron que existe un impacto positivo sobre el número de citas por parte de los huecos estructurales.

La evaluación del rendimiento de los académicos en ingeniería con el índice  $h$  fue propuesta por Hirsch (2007), en función de tratar de medir conjuntamente el número de publicaciones de un investigador (i.e. cantidad) y su impacto en otras publicaciones (i.e. calidad). Mediante dicho análisis se encontró que el grado de centralidad es estadísticamente significativo y con impacto positivo en todas las redes excepto en la de comunicación, cercanía y el Eigenvector las cuales tienen un impacto significativo en el desempeño de las citas en todas las redes. Adicionalmente, se documentó que la intermediación tiene un impacto positivo significativo únicamente en la red de publicaciones mientras que la intensidad de los lazos fue estadísticamente significativa con un impacto positivo únicamente en las redes de publicación y patentes.

### Datos y método

En este análisis se consideraron cuatro tipos de colaboración: internacional, interinstitucional, intrainstitucional, academia-empresa y academia-gobierno. Respecto a las redes sociales se consideraron las medidas de centralidad normalizadas por la ASN. Adicionalmente, retomando el debate de closure y “structural holes” se analizó el efecto de la productividad que tienen los huecos estructurales en la red de investigación. A continuación, se describen los datos empleados para realizar los análisis correspondientes.

### Datos

Básicamente se cuenta con dos fuentes de información.

La primera consiste en datos de 27,667 investigadores de todas las áreas del conocimiento que han sido parte del Sistema Nacional de Investigadores (SNI)<sup>1</sup>, por lo menos un año en el periodo 1991-2007.

Debido a que el objetivo de este trabajo es estudiar a los investigadores en el área de ingeniería<sup>2</sup>, se seleccionaron los miembros de al área VII del SNI<sup>3</sup> por lo que se obtuvo una muestra de 2150 investigadores. La segunda fuente de información empleada corresponde a una base de datos de publicaciones y citas del Science and Social Sciences Citation Index desarrollada por el *Institute of Scientific Information (ISI)*.

Finalmente, las publicaciones fueron obtenidas mediante un cruce de información entre la base de datos del SNI y la correspondiente del ISI en el periodo 1981-2007, junto con las citas correspondientes a cada artículo al 2007 (ISI, 2008). Con base en lo anterior se obtuvo una muestra de 14,275 artículos publicados en dicho periodo.

### Variables

Se consideran dos variables dependientes: número de publicaciones por un investigador y número de citas de cada publicación en los subsecuentes años. Las variables independientes se dividen en dos grupos: tipo de colaboración y características de la red de investigación.

<sup>1</sup> El sistema Nacional de Investigadores se creó en 1984 para dar estímulos económicos a los investigadores más productivos de México. El programa sigue vigente hasta hoy en día y llega a representar hasta el 30% del salario de los investigadores que pertenecen a él.

<sup>2</sup> Si bien el área de ingeniería en el SNI no es la más productiva en cuanto a artículos publicados en el ISI, fue seleccionada debido a que la ingeniería en México no ha sido lo suficientemente explorada.

<sup>3</sup> El SNI cuenta con siete áreas del conocimiento: Ciencias Físico-Matemáticas; Ciencias de la Tierra; Biología y Química; Medicina y Ciencias de la Salud; Humanidades y Ciencias de la Conducta; Ciencias Sociales; Biotecnología y Ciencias Agropecuarias e Ingenierías.

### Variables de colaboración

- Coautoría internacional: un artículo tiene direcciones reportadas en más de un país.
- Coautoría intrainstitucional: un artículo tiene una sola dirección, pero más de un autor.
- Coautoría interinstitucional: un artículo tiene más de una dirección reportada, pero un solo país.
- Coautoría academia-empresa: un artículo reporta afiliación a una o más empresas.
- Coautoría academia-gobierno: un artículo reporta como afiliación a una entidad gubernamental.

### Variables de redes

Se emplean tres medidas clásicas de centralidad (grado, cercanía e intermediación) y dos variables de redes egocéntricas (huecos estructurales y densidad)

- Grado nodal. El grado de centralidad de un nodo ( $n_i$ ) denotada por  $C_D$  corresponde al número de nodos que son adyacentes al nodo ( $n_i$ ), es decir, el número de contactos únicos ( $e_{ij}$ ) que tiene el vértice ( $n_i$ ) (Wasserman & Faust, 1994). El grado Normalizado de centralidad,  $C'_D(n_i)$  se calcula dividiendo el grado nodal  $n_i$  entre el número total de nodos,  $n$ , excluyendo  $n_i$  como  $(n-1)$ . Por tanto, el grado normalizado está representada por:

$$C'_D(n_i) = \frac{C_D(n_i)}{(n-1)} = \frac{\sum_j e_{ij}}{n-1} \quad (1)$$

- Cercanía. La cercanía de centralidad de un nodo ( $n_i$ ) denotada por  $C_C(n_i)$ , es la suma de las distancias geodésicas de todos los otros nodos en una red (Wasserman & Faust, 1994). La distancia geodésica es el camino más corto para unir a dos nodos ( $n_i$  y  $n_j$ ) o la distancia (vector) entre esos dos nodos  $d(n_i, n_j)$ . La sumatoria de la distancia geodésica es  $\sum_j^n d(n_i, n_j)$  y por tanto la cercanía, es la inversa de la suma de las distancias más cortas entre cada individuo y cualquier otra persona en la red. La cercanía normalizada se calcula multiplicando  $C_C(n_i)$  por  $(n-1)$  cuya ecuación está representada por:

$$C'_C(n_i) = C_C(n_i)(n-1) = \frac{n-1}{\sum_j^n d(n_i, n_j)} \quad (2)$$

- Intermediación. La centralidad de intermediación de un nodo ( $n_i$ ) denotada por  $C_B(n_i)$ , es la suma de la relación del número de geodésicas,  $g_{jk}(n_i)$ , que unen los nodos  $n_j$  y  $n_k$ , que contiene el nodo  $n_i$ , con el número total de geodésicas  $g_{jk}$  que une los nodos  $n_j$  y  $n_k$ , (Wasserman & Faust, 1994). También se define como el número de caminos geodésicos (distancia más corta entre todos los pares de nodos (Borgatti, 2005) que pasan a través de un nodo  $n_i$ , la ecuación de la intermediación está dada por:

$$C_B(n_i) = \sum_{i < j}^n \sum_j^n \frac{g_{jk}(n_i)}{g_{jk}} \quad (3)$$

La intermediación normalizada  $C'_B(n_i)$  (tiene valores entre 0 y 1), se calcula dividiendo a la intermediación por  $(n-1)(n-2)/2$ , lo anterior indica el número de pares no incluidos  $n_i$  por lo que la intermediación normalizada está dada por:

$$C'_B(n_i) = \frac{C_B(n_i)}{(n-1)(n-2)/2} = \frac{2 * C_B(n_i)}{n^2 - 3n + 2} \quad (4)$$

- Eigenvector. Considerando que un nodo es más central si está vinculado con nodos que están bien relacionados o que por sí mismos ya son centrales (Bonacich, 1972), este vector implica que la centralidad de un nodo no depende únicamente del número de sus nodos adyacentes (i.e. grado nodal). El Eigenvector de un nodo ( $n_i$ ) denotado por  $C_E(n_i)$  y es el múltiplo positivo de la suma de la matriz adyacente de centralidad es decir:
- Huecos Estructurales. Representan la ausencia de un vínculo entre dos contactos, los cuales están asociados a un actor. Esta variable es obtenida por la ecuación de restricción de Burt (1992), que en esencia es una medida del grado en que un nodo se relaciona con otros nodos que están conectados con otro relacionados entre sí.
- Densidad. Indica la división del número de lazos de un nodo  $n_i$  entre el número de pares que tiene la red del nodo  $n_i$ . Esta medida densidad local (ego densidad) indica el grado de colaboración o de cohesión entre los coautores que cada investigador se relaciona.

$$\lambda * C_E(n_i) = \sum_{k=1}^n (a_{ij} * C_E(n_i)) \quad \forall i \quad (5)$$

donde  $a_{ij} = 1$  si el vértice  $i$  es conectado con el vértice  $j$ , y  $a_{ij} = 0$  si ocurre lo contrario. La centralidad de todos los nodos es representada  $c = C_E(n_i) = (C_D(n_1), C_D(n_2), \dots, C_D(n_n))$  y  $\lambda$  es un escalar. Lo anterior se puede escribir mediante la notación  $\lambda c = Ac$  la cual es la ecuación característica para encontrar el eigensistema de una matriz  $A$  (Wasserman & Faust, 1994). Por convención, el eigenvector de centralidad está dado por el valor más grande de  $\lambda$  (Borgatti et al., 2002), el cual se puede medir mediante la siguiente expresión:

$$C'_E(n_i) = \frac{C_E(n_i)}{\sqrt{2}} \quad (6)$$

Un nodo que está conectado con otros nodos que a su vez están bien conectados tendrá un valor alto y por lo tanto son buenos candidatos para difundir información. Por el contrario, un nodo que tiene pocas conexiones por ende tendrá un valor bajo.

Para el cálculo de las variables de centralidad se empleó el software de análisis UCINET 6.523 (Botrgatti et al, 2002).

### Modelo

Para nuestro análisis, se empleó una extensa base de datos longitudinal (Panel Data) debido a que el comportamiento de los científicos es observable a través del tiempo. Esta base, contiene la información de diversas actividades de los investigadores mexicanos que pertenecen al área de ingeniería del SNI (área VIII) así como sus variables de redes y los tipos de coautoría.

El modelo general es:  $Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it}$ , donde  $Y_{it}$  son las variables dependientes (publicaciones y citas)  $i =$  la entidad (investigador) y  $t =$  al tiempo;  $\beta_1 X_{it}$  son las variables explicativas (tipo de colaboración y variables de redes);  $\alpha_i$  ( $i = 1 \dots n$ ) es el intercepto desconocido para cada investigador ( $n$  interceptores específicos por investigador) y  $u_{it}$  es el término error.

Debido a que las variables dependientes son números enteros positivos, se asume un modelo de distribución binomial negativo empleando estimadores de máxima verosimilitud con efectos fijos. Se decidió emplear una distribución binomial negativa en lugar de una distribución de Poisson debido a que la varianza estimada es superior a la media y no es constante. Por otro lado, elegimos el modelo de efectos fijos debido a que permite analizar las variaciones temporales de las variables, así como controlar por las características no observadas de las a través del tiempo, y por tanto, los resultados se consideran más precisos que si se usara efectos aleatorios o un análisis transversal.

### Construcción del Modelo

En este análisis se incluyen dos variables dependientes: Publicaciones (mide el conteo lineal de las publicaciones que el investigador publica en ISI en un periodo de dos años) y Citas (infiere la calidad de dichas publicaciones al añadir el número de citas que las publicaciones han recibido en los subsecuentes 4 años de su publicación). Las cinco variables del tipo de colaboración corresponden al conteo lineal de los artículos publicados según su categoría por el investigador en ISI, en los tres años posteriores. Como variable de control en se incluyó el número de artículos que un investigador publicó como autor único. Para el cálculo de las variables relevantes de red de cada investigador se evaluó cómo es que las características de la red en periodos anteriores afectan el desempeño en la investigación.

La tabla 1 muestra los 8 periodos a estudiar. En el periodo 1, se consideran las publicaciones que el investigador realizó en una ventana de dos años (1992 y 1993). Para el caso de las citas, se suma en número de las citas que recibieron durante los cuatro años posteriores a la publicación (1992-1995).

A fin de evaluar los efectos de publicaciones previas se analizaron los artículos acordes al tipo de colaboración publicados tres años atrás (1989-1991). Finalmente, los efectos de la inserción a la red de colaboración se evaluaron en función de características de tres años previos a las publicaciones (1989-1991).

Periodo	T (Pubs)	citas t+4	t-1 (redes)
8	2007/2006	2006/2007/2008/2009	2005/2004/2003
7	2005/2004	2004/2005/2006/2007	2003/2002/2001
6	2003/2002	2002/2003/2004/2005	2001/2000/1999
5	2001/2000	2000/2001/2002/2003	1999/1998/1997
4	1999/1998	1998/1999/2000/2001	1997/1996/1995
3	1997/1996	1996/1997/1998/1999	1995/1994/1993
2	1993/1995	1994/1995/1996/1997	1993/1992/1991
1	1993/1992	1992/1993/1994/1995	1991/1990/1989

Tabla 1 Periodos

### Resultados

La tabla 2 muestra la estadística descriptiva de las variables descritas anteriormente. La muestra de investigadores tiene una media 0.98 publicaciones por dos años con una desviación estándar de aproximadamente 2.5 publicaciones. Se registra un promedio de 12 coautores por artículo y cada investigador recibió en promedio 3 citas en los próximos cuatro años. Estos datos son similares a los que reporta Lee y Bozman (2005) quienes documentaron un número de coautores por artículo en ingeniería (incluyendo ciencias de la computación) es de 13. En relación al número de contactos directos, aunque es en un estudio de biotecnología y biología aplicada, Singh (2007) reporta para su muestra un promedio de 4.8.

En cuanto al tipo de colaboración, el 31% de los artículos involucra colaboración interinstitucional (entre la misma institución), el 25% está relacionada con la colaboración interinstitucional (entre distintas instituciones), el 35% implica colaboración internacional, el 1% involucra colaboración con la industria (academia-empresa), el 0.8% implica colaboración con instituciones del gobierno (academia-gobierno) y solo el 5% son artículos de un solo autor.

Para comprobar que nuestro modelo no tiene problemas de autocorrelación entre variables, fue sometido al test de correlación serial de Wooldrige (2002), cuyo resultado descarta problemas de este tipo.

VARIABLES	Mean	Std.De	Min	Max
pubs	0.986	2.425	0	38
citas	3.020	14.13	0	657
interinst	0.427	1.183	0	29
intrainst	0.450	1.452	0	35
intern	0.447	1.417	0	40
ind	0.0176	0.195	0	10
gob	0.0126	0.148	0	6
simcol	0.0499	0.417	0	30
grado	0.0721	0.203	0	4.062
eigenvector	0.0587	1.110	0	52.13
closeness	0.027	0.042	0	0.187
betweenness	0.04	0.214	0	6.691
structuralholes	0.212	0.329	0	1.620
directies	3.029	6.587	0	110
densit	21.81	35.66	0	100
Observaciones 17200				
n = 2150				
T= 8				

Tabla 2 Estadística descriptiva

La tabla 3 muestra la correlación de Spearman entre las 14 variables analizadas. Se observa que existe correlación alta entre algunas variables, especialmente entre la variable dependiente (publicaciones) y las medidas de centralidad normalizadas (grado y cercanía). Para explicar la calidad (i. e. citas) estas medidas tienen una baja correlación.

La multicolinealidad ocurre cuando existe una alta correlación entre dos o más variables independientes en regresiones múltiples, por lo que una variable independiente se puede predecir a partir de otras. Este problema puede ser más común cuando medidas de redes sociales son empleadas como predictores (Cimenler et. Al., 2014).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) pubs	1													
(2) citas	0.644*	1												
(3) interinstitucional	0.270*	0.328*	1											
(4) intrainstitucional	0.582*	0.358*	0.362*	1										
(5) internacional	0.266*	0.354*	0.293*	0.323*	1									
(6) academia-empresa	0.134*	0.064*	0.077*	0.060*	0.060*	1								
(7) academia-gobierno	0.068*	0.046*	0.059*	0.077*	0.030*	-0.003	1							
(8) single author	0.130*	0.114*	0.142*	0.060*	0.060*	0.038*	0.020*	1						
(9) degree	0.066*	0.019*	0.268*	0.277*	0.200*	0.148*	0.133*	0.118*	1					
(10) eigenvector	0.308*	0.304*	0.241*	0.309*	0.328*	0.108*	0.060*	0.012	0.458*	1				
(11) closeness	0.009*	0.427*	0.382*	0.549*	0.230*	0.142*	0.116*	0.120*	0.933*	0.400*	1			
(12) betweenness	0.247*	0.421*	0.465*	0.236*	0.448*	0.141*	0.080*	0.063*	0.274*	0.400*	0.590*	1		
(13) structuralholes	0.229*	0.241*	0.484*	0.388*	0.434*	0.094*	0.109*	0.103*	0.720*	0.256*	0.590*	0.241*	1	
(14) density	0.022*	0.161*	0.511*	0.219*	0.220*	0.118*	0.111*	0.056*	0.709*	0.149*	0.780*	0.380*	0.801*	1

Tabla 3 Correlación de spearman

La tabla 4 muestra los resultados de las regresiones considerando publicaciones y citas como variables dependientes. Se presentan 8 modelos de los cuales seis introducen sólo una variable de redes para medir el efecto de ésta en la productividad y el tipo de colaboración. Los dos modelos restantes analizan la perspectiva de *closure* y *structural holes* dónde todas las variables de redes centralidad se miden en primera instancia con la variable densidad (perspectiva *closure*) y posteriormente se evalúan las mismas variables para la perspectiva *structural holes*.

Prácticamente en todos los modelos especificados la variable control (single autor), es estadísticamente significativa ( $p > 0.01$ ), lo cual indica que ésta afecta positivamente la futura productividad y la calidad (citas) de los investigadores en ingeniería, este resultado es consistente con Gonzalez-Barambila et al. 2013.

En todos los modelos las publicaciones interinstitucionales afectan positivamente el número de publicaciones y citas futuras. Respecto a la colaboración intrainstitucional, nuestros resultados muestran que es significativa solamente para efectos de citas.

La colaboración con la empresa en relación con las citas prácticamente no tiene ningún impacto, únicamente se presenta una significancia ( $p < 0.1$ ) con el variable grado la cual es negativa, lo cual indica que este tipo de colaboración tiene una relación negativa cuando se relaciona sólo con la variable grado; relación significativa positiva sólo con la variable de *structural holes* y relación con el eigenvector. Respecto a la colaboración con el gobierno, ésta tiene una relación con las publicaciones respecto a la variables *structural holes*, *eigenvector*, *closeness* y *betweenness*.

Nuestros resultados muestran que prácticamente todas las variables de redes son estadísticamente significativas y están relacionadas con la productividad futura. En el modelo 7 y 8 las dos variables que defienden el punto de vista de *closure* (densidad) y de los *structural holes* son estadísticamente significativas ( $p > 0.01$ ). En estos dos modelos las variables de tipo de colaboración son significativas, caso contrario con las variables de redes. En la tabla 5, se observa que los modelos 1 al 6 incluyen las variables de red por separado, únicamente *betweenness* reporta significancia estadística para las citas mientras que las demás tienen relación positiva con el número de citas. Para los modelos 7 y 8 (perspectiva del capital social), el *eigenvector* no tiene efecto en la calidad, mientras que la variable *betweenness* sólo es significativa para la perspectiva de los huecos estructurales, aunque de forma negativa.

VARIABLES	(1) citas	(2) citas	(3) citas	(4) citas	(5) citas	(6) citas	(7) citas	(8) citas	(9) citas
degree	0.841***								
eigenvector		0.015**							
closeness			7.349***						
betweenness				0.068					
structuralholes					0.654***				
density						0.005***			
interinst	0.055***	0.078***	0.060***	0.074***	0.078***	0.081***	0.055***	0.0021***	0.0003
intrainst	0.017**	0.038***	0.023***	0.021***	0.042***	0.041***	0.025***	0.028***	0.056***
intem	0.039***	0.049***	0.039***	0.053***	0.054***	0.061***	0.034***	0.037***	0.035***
academia-empresa	-0.086*	0.038	0.0057	0.023	0.046	0.035	-0.045	-0.046	-0.045
academia-gobierno	0.112	0.168**	0.131*	0.172**	0.168**	0.158**	0.099	0.099	0.099
single author	0.047***	0.044***	0.051***	0.041***	0.041***	0.050***	0.048***	0.051***	0.048***
Constant	-1.964***	-1.900***	-2.212***	-1.896***	-2.156***	-2.172***	-2.311***	-2.248***	-2.313***
Standard errors	emom		in		quercilicos		***		p=0.01
									p=0.05

Table 5 resultado de la regresión para citas

Tabla 5 resultado de la regresión para citas

Discusión y conclusiones

El objetivo de este trabajo es analizar cómo la colaboración externa en conjunto con las redes de colaboración afecta los resultados de investigación de científicos mexicanos en el campo de la ingeniería. Mediante una base longitudinal de datos, se concluye lo siguiente:

Colaboración Externa

Los resultados indican que, a pesar del bajo índice de artículos de un solo autor, éstos tienen un impacto positivo tanto en términos de la productividad futura cómo en el número de las citas recibidas. Adicionalmente, se muestra que los investigadores que colaboran más son a su vez los que publican en un alto índice a nivel individual. Este resultado es similar al reportado en científicos mexicanos del área de ciencias exactas por Gonzalez-Brambila et al 2013, lo cual sugiere que los científicos mexicanos de todos los campos de la ciencia tienen este comportamiento.

Respecto a los investigadores del área VII del SNI de México, nuestros resultados muestran que la colaboración con otros miembros tiene un impacto positivo en términos de citas y de publicaciones futuras, aunque la colaboración interinstitucional resultó ser la más significativa. Cabe señalar que este tipo de colaboración incluye la vinculación entre profesores y estudiantes.

Uno de los principales hallazgos de esta investigación es que para los investigadores en ingeniería el colaborar con la industria no es significativo, lo cual contradice los resultados obtenidos por Singht 2007, Frenken et al. 2005.

Tabla 4 Resultados de la regresión para publicaciones

Por otra parte, acorde con los resultados obtenidos la colaboración con entidades gubernamentales tiene una ligera importancia para los investigadores mexicanos.

### Redes de Colaboración

Nuestros resultados muestran que el tipo de colaboración se mantiene constante, lo que nos permite identificar que variables de red es la más adecuadas para medir el efecto de la red en los resultados de la investigación de los científicos del área de ingeniería en México.

Respecto a la variable de centralidad de grado nodal, los resultados muestran que tiene un fuerte impacto en la productividad futura y las citas; en términos de red esto significa que si un investigador en ingeniería tiene una gran cantidad de contactos directos figura como un investigador estrella o un líder en el área del conocimiento y por tanto podrá tener influencia sobre los demás, de tal forma que será un investigador con quien los demás querrán participar y por ende sus trabajos serán más conocidos y citados.

Adicionalmente, acorde con los resultados se muestra que la centralidad *Closeness*, el vector propio de centralidad o *eigenvector* son importantes para los científicos en términos de redes de colaboración, a diferencia de la centralidad *Betweenness* que no tiene ningún efecto sobre las variables dependientes.

Una de las conclusiones principales de este análisis es que las variables apropiadas de centralidad para medir la influencia de la red son: el grado nodal y la centralidad. Adicionalmente, se recomienda considerar las dos variables de "ego" en la red, densidad y huecos estructurales.

Finalmente, acorde al modelo número 9 de este trabajo, se puede inferir que las dos perspectivas en capital social no están en conflicto debido a que las dos variables se involucran con las variables de centralidad y de colaboración externa manteniendo resultados consistentes.

### Referencias

- Andrews, S. Fastqc, (2010). A quality control tool for high throughput sequence data.
- Adams, J. Black, G. C., Clemmons, J.R., Stephan, P.E. (2004). Scientific teams and Institutional collaborations: evidence from US universities, 1981-1999. NBER Working Paper, 10640.
- Adams, J. Black, G. C., Clemmons, J.R., Stephan, P.E. (2004). Scientific teams and Institutional collaborations: evidence from US universities, 1981-1999. NBER Working Paper, 10640.
- Ancona, D.G., Caldwell, D.F. (1992). Demography and design: Predictors of new product team productivity. *Organization Science*. 3, 321–341.
- Bantel, K.J., Jackson, E.S. (1989). Top Management and innovations in banking: Does the composition of the top team make the difference?. *Strategic Management Journal*. 10(1), 107-124.
- Barbasi, A.L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A. & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physics A*. 311(3-4), 590-614.
- Beaver, D., & Rosen, R. (1978). Studies in scientific collaboration –part i. the professional origins of scientific co-authorship. *Scientometrics*. 1(1), 65-84.
- Beaver, D.D (2001). Reflections on scientific collaboration (and study): Past, present, and future. *Scientometrics*. 52(3), 365-377.

- Birnholtz J.P. (2007). When do researchers collaborate? Toward a model of collaboration propensity. *Journal of the American Society for Information Science and Technology*. 58(14), 2226-2239.
- Bonacich, P. (1972). Factoring and Weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*. 2(1), 35-38.
- Borgatti, S.P. (2005). Centrality and network flow. *Social Networks*, 27(1), 55-71.
- Borgatti, S.P., Everett, M.G., & Freeman, L.C. (2002). *Ucinet for Windows: Software for social network analysis*. Harvard, MA: Analytic Technologies.
- Burt, Ronald S. (1992). *Structural Holes*. Cambridge, MA: Harvard University Press.
- Cimenler, O., Reeves, K.A., Skvoretz, J. (2014). A regression analysis of researchers' social networks metrics on their citation performance in a college of engineering. *Journal of Infometrics*. 8(3), 667-682.
- Coleman, J.S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*. 94, 95-120.
- Coleman, J.S. (1990). *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.
- Collins, H.M. (1974). The TEA set: tacit knowledge and scientific networks. *Science Studies*. 4(2). 165-186.
- Cronin, B., Shaw, D. & La Barre, K. (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(9), 855-871.
- Cronin, B., Shaw, D. & La Barre, K. (2004). Visible, less visible, and invisible work: Patterns of collaboration in 20th century chemistry. *Journal of the American Society for Information Science and Technology*. 55(2), 160-168.
- Diamond, A.M. (1985). The money value of citations to single-authored and multiple-authored articles. *Scientometrics*. 8(5-6), 315-320.
- Frenken, K., Hözl, W., Friso de Vor. (2005). The citation impact of research collaborations: the case of European biotechnology and applied microbiology (1982-2002). *Journal of Engineering and Technology Management*. 22, 9-30.
- Gonzalez-Brambila, C. N., M. Veloso F., Krackhardt D. (2013). The impact of network embeddedness on research output. *Research Policy*. 42(9), 1555-1567.
- Gonzalez-Brambila, C.N., M. Veloso F. (2007). The determinants of research output and impact: A study of Mexican researchers. *Research Policy*. 36(7), 1035-1051.
- Gordon M.D. (1980). A critical reassessment of inferred relations between multiple authorship, scientific collaboration, the production of papers and their acceptance for publication. *Scientometrics*. 2(3), 193-210.
- Granovetter, M.S. (1973). The strength of weak ties. *American Journal of Sociology*. 78(6), 1360-1380.
- Heinze, T. & Kuhlman, S. (2008). Across institutional boundaries? –research collaboration in german public sector nanoscience. *Research Policy*. 37(5), 888-899.
- Hirsch, J.E. (2007). Does the h-index have a predictive power?. *Proceedings of the National Academy of Sciences of the United States of America*. 102(46), 16569-16572.

- Katz J.S, Martin B.R. (1997). What is research collaboration?. *Research policy* 26, 1-18.
- Katz, J.S, D. Hicks. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics*. 40(3), 541-554.
- Lawani S.M. (1986). Some bibliometric correlates of quality in scientific research. *Scientometrics*. 9(1-2), 13-25.
- Lee, S. Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*. 35(5), 673-702.
- Lotka A.J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Science*. 16(12), 317-323.
- McCain, B.E, O'Reilly., Pfeffer,J. (1983). The effects of departmental demography on turnover: The case of university. *Academy of Management Journal*. 26(4), 626-641.
- Narin F., and Whitlow E.S. (1990). *Measurements of Scientific Cooperation and Coauthorship in CEC-related Areas of Science* (Report EUR 12900, Office for Official Publications of the European Communities, Luxembourg).
- Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Science of the United States of America*. 101(1), 5200-5205.
- Newman, M.E.J. (2001a). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*. 98(2), 404-409.
- Nudelman A.E. and Landers C.E. (1972). The Failure of 100 divided by 3 to equal 33 1/3. *American Sociologist*. 7(9), 9.
- Pao, M.L. (1982). Collaboration in Computational Musicology. *Journal of the American Society for Information Science*. 33(1), 38-43.
- Portes, A., Sensenbrenner, J. (1993). Embeddedness and Immigration: Notes on the Social Determinants of Economic Action. *The American Journal of Sociology*. 98(6), 1320-1350.
- Price, D. S and D, de B. Beaver. (1966). Collaboration in an invisible college. *American Psychologist*. 21, 1011-1018.
- Schofer, E., Ramirez. F.O., Meyer, J. W. (2000). The effects of science on national economic development, 1970-1990. *American Sociological Review*. 65(6), 866-887.
- Shurum, W. (2005). Reagency of internet, or, how I became a guest for science?. *Social Studies of Science*. 35(5), 723-754.
- Singht J. (2007). *External Collaboration, Social Networks and Knowledge Creation: Evidence from Scientific Publication*. INSEAD, Mimeo.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, New York; Cambridge University Press.
- Wooldridge. J.M. (2002). *Econometric Analysis of Cross Section and Data Panel*. The MIT Press. Cambridge England. (pp. 159-230).
- Zi-Lin, H., Xue-Song, G., Campbell-Hunt,C. (2009). Research collaboration and research output: a longitudinal study of 65 biomedical scientists in a New Zealand university. *Research Policy*. 38(2), 306-317.

Zuckerman, H. (1967). Nobel laureates in science: Patterns of productivity, collaboration, and authorship. *American Sociological Review*. 32(3), 391-403.