

Efecto de grupo par sobre el rendimiento académico: Evidencia en la Universidad del Rosario

Camila Valencia

Tesis de la Maestría en Economía de las Políticas Públicas

Facultad de Economía

Universidad del Rosario

Asesor : Darwin Cortés

28 de Agosto de 2017

Resumen

En este trabajo se estima el efecto de grupo par sobre el rendimiento académico, utilizando una muestra de estudiantes de primer semestre de la Universidad del Rosario. Una vez construida la red completa de los pares de los estudiantes y con el propósito de resolver la endogeneidad causada por el problema de reflexión, se instrumenta el rendimiento académico de los pares con el puntaje de la prueba Saber 11 de la cohorte del colegio de cada uno de los integrantes de la red. Los resultados sugieren que un aumento de una desviación estándar en el rendimiento de la red de pares tiene un impacto de 0.8 desviaciones estándar en el rendimiento académico del estudiante.

Palabras clave: Efecto de grupo par, red social, educación, rendimiento académico

1. Introducción

Desde la publicación del Reporte de Coleman (1996), cuyo objetivo fue promover la igualdad de oportunidades educativas en Estados Unidos, se ha argumentado que la composición de los grupos pares es uno de los principales determinantes del rendimiento de los estudiantes, así como también lo son la calidad de los profesores, el tamaño de la clase y el compromiso de los padres (Sacerdote, 2011). Al momento de identificar los efectos de grupo par surgen dos principales problemas identificados por Manski (1993): el problema de reflexión, que consiste en distinguir entre el efecto de la interacción endógena¹ y el efecto de la interacción exógena², y el problema de correlación, que se ocupa de la posible correlación entre los inobservables a nivel grupal y las características individuales.

Entenderemos “efecto de grupo par” como lo define Hoxby (2000), es decir como “casi cualquier efecto que pueden tener las características del par - sus antecedentes, su comportamiento y sus resultados actuales - sobre la variable de resultado del estudiante”. También se entiende como efecto de grupo par cualquier externalidad de contagio de los pares, de la familia de los pares o las acciones actuales de los pares. Por ejemplo, existe un efecto de grupo par cuando en un salón de clase hay un estudiante que tiene mayores habilidades y demanda un ritmo más rápido por parte del profesor; otro ejemplo es cuando en un grupo de compañeros existe un estudiante sobresaliente, lo que puede incentivar a que los demás hagan un mayor esfuerzo.

Hay bastantes trabajos que muestran evidencia de la existencia y el tamaño de los efectos de grupo par. Varios de ellos se enfocan en el impacto sobre el rendimiento académico de los estudiantes a nivel de colegio, tanto en primaria como en secundaria. Los resultados de algunos de estos trabajos son modestos pero significativos (Hoxby et al. (2000); Vigdor et al. (2007), Hanushek et al. (2003)). Otros investigadores encuentran pequeños efectos (Burke y Sass (2008), Angrist y Lang (2004)). A nivel universitario, Sacerdote (2001) y Zimmerman (2003) encuentran efectos grandes y significativos pero, Foster (2006) y Lyle (2007) no encuentran ningún efecto.

La diferencia en estos resultados se debe a la metodología para manejar los problemas de reflexión y correlación. Tamayo (2016) los clasifica en los siguientes tres grupos: estudios que utilizan fuentes exógenas de variación para las características de los pares, estudios en los que los estudiantes han sido asignados aleatoriamente, y estudios que usan efectos fijos de los estudiantes y de la escuela para controlar el problema de auto-selección.

En este trabajo usé la metodología de variables instrumentales, como fuente de variación exógena, para identificar los efectos de grupo par en la Universidad del Rosario, dada la información que se tiene acerca los estudiantes de primer semestre de la Facultad de Economía en

¹ El comportamiento actual del individuo

² Los antecedentes del individuo

el semestre 2016-I. Los datos que se tienen son: desempeño académico en la universidad, rendimiento pasado dado los resultados de la prueba Saber 11 y el reporte de la red de estudiantes con la que cada individuo estudia. Mi propósito es identificar el impacto de la calidad de la red sobre el rendimiento académico de cada estudiante, y lo logro instrumentando la calidad de la red con la calidad del colegio de cada uno de los integrantes de la red.

Encuentro un efecto de 0.8 desviaciones estándar, positivo y significativo, de la calidad de la red dirigida sobre el rendimiento académico de los estudiantes. El impacto que tiene la red sobre el rendimiento, se da en las materias que tienen fuerte énfasis matemático como Cálculo y Lógica Matemática. Los resultados son robustos a diferentes especificaciones y a potenciales problemas dado el tamaño de muestra.

Este trabajo es de relevancia porque a diferencia de la limitada literatura para el caso colombiano, se enfoca en los efectos de grupo par a nivel universitario y además contó con la información de la red de estudio de cada individuo. Los resultados de este trabajo son relevantes ya que las universidades pueden entender la relevancia de los efectos de grupo par en el desempeño académico de los estudiantes.

La estructura de este trabajo es la siguiente: la sección 2 es la revisión de literatura relevante para este caso, la sección 3 describe los datos utilizados y sus fuentes, la sección 4 presenta la metodología utilizada y su aplicación empírica, la sección 5 muestra los resultados encontrados de la implementación de variables instrumentales, algunas pruebas de robustez y muestra algunas variaciones al modelo básico, y la sección 6 concluye y da algunas recomendaciones.

2. Revisión de la literatura

En esta sección mostraré los problemas que se han encontrado en las últimas décadas al identificar efectos de grupo y cómo se ha dado solución a algunos de estos problemas a nivel de colegio y de universidad. También presentaré los resultados de los trabajos más relevantes en el tema de efectos de grupo par sobre el rendimiento académico, prestando especial atención a los que han trabajado variables instrumentales, ya que es la misma especificación que usé en este trabajo.

La literatura que ha estudiado efectos de grupo par es bastante amplia y su principal objetivo es identificar estos efectos. Al momento de identificar el efecto y su magnitud surgen varios inconvenientes, An (2011) explica claramente varios de estos problemas: (1) la reflexión del efecto identificada por Manski (1993), (2) la auto selección de los individuos, (3) las variables omitidas que pueden afectar el resultado, (4) el análisis estático de las redes sociales y (5) los errores de medida.

El problema de reflexión surge porque es difícil diferenciar entre el efecto del grupo de pares sobre un individuo y el efecto del individuo sobre el grupo de pares, por esto Manski lo llama de esta manera. El segundo problema por lo cual es complicado medir el efectos de grupo par es la tendencia de los individuos a autoseleccionarse, es decir individuos similares tienden a agruparse. En este caso, los estudiantes sobresalientes pueden detectar sus pares sobresalientes y trabajar en un grupo muy similar, lo que genera un problema al tratar de diferenciar entre efectos pares y los efectos de selección (Sacerdote, 2001).

El problema de variables omitidas se debe a la existencia de otros factores de los pares, que pueden ser de contexto como por ejemplo situaciones emocionales o problemas sociales, que a pesar de estar correlacionados con los resultados del individuo de los pares no son tenidos en cuenta en el modelo. Esto puede generar un sesgo en el estimador de los efectos pares.

Entiéndase análisis estático como el estudio de un fenómeno en un momento específico en el tiempo. El problema del análisis estático en es caso se presenta al trabajar con las redes sociales, usualmente se tiene únicamente la información correspondiente a un solo momento de las redes de los estudiantes, cuando en realidad las conexiones entre individuos cambian en el tiempo.

El otro problema expuesto por An (2011) es el problema de reporte en las redes, ya que los estudiantes pueden olvidar algunas de sus conexiones al momento de la toma de datos, lo cual genera datos faltantes o también pueden hacer un sobrerreporte de conexiones, lo cual perjudica la calidad de los datos que son analizados.

Los modelos estáticos más populares que se han desarrollado para identificar los efectos de grupo par dado los problemas que mencionamos anteriormente son el modelo lineal en medias y el modelo de resultado binario. El modelo lineal en medias asume que el resultado de un individuo esta determinado por sus propias características y la media respectiva de las características y resultados de sus pares, por ejemplo Manski (1993), Lee (2007) y Carrell (2009). El modelo de resultado binario adopta un modelo logit o probit para predecir los efectos de grupo par, por ejemplo Brock et al. (2001a), (2001b), Bulte et al. (2001), Sorensen (2006) y Krauth (2009).

Además, se han desarrollado varias estrategias empíricas para identificar fuentes de variación exógena de las características de los pares y así solucionar el problema de endogeneidad. Hoxby (2000), Hanushek et al. (2003), McEwan (2003), Vigdor (2006) han considerado la variación del colegio o del grado, Lavy et al. (2009) ha hecho lo propio con la variación del individuo, Hoxby et al. (2005) con la reasignación de subgrupos, Sacerdote (2001), Zimmerman (2003) Duflo et al. (2008) con la asignación aleatoria y Angrist y Lang (2004), Kang (2007), Zabel (2008), Bramoullé et al (2009) con las variables instrumentales.

Algunos estudios a nivel de colegio encuentran estimadores significativos y modestos del impacto del rendimiento académico previo de los pares sobre el rendimiento académico del estudiante (Hoxby (2000), Vigdor et al. (2007), Hanushek et al. (2003)), otros encuentran pequeños efectos, por ejemplo Burke y Sass (2008), Angrist y Lang (2004). Hoxby y Weingarth (2005) proponen que el rendimiento del individuo depende de si hay cambios en la composición del grupo de pares midiéndolo por la distribución de habilidades (alta, media, baja). Encuentran que al introducir un estudiante a un grupo nuevo, los estudiantes en los percentiles más altos de calificación afectan positivamente a los estudiantes más sobresalientes de dicho grupo. Si introducen estudiantes del percentil 85 los estudiantes con menores calificaciones aumentan su rendimiento académico.

A nivel universitario se encuentran resultados contradictorios, en Estados Unidos varios de los estudios analizan los efectos en los compañeros de dormitorio pues estos son usualmente asignados aleatoriamente. Zimmerman (2003) encuentra que el rendimiento anterior y los resultados del compañero de cuarto tienen efecto en el rendimiento del individuo. Carrell et al. (2009) analizan un contexto inusual en donde tienen identificado todo el grupo de pares, además la institución impone una fuerte interacción entre pares, lo que refleja unos efectos grandes y significativos. Sacerdote (2001) usó el índice ACA, un índice académico como medida de habilidad para ver los efectos pares sobre el rendimiento académico, encuentra efectos modesto pero significativos. Mientras que Foster (2006) y Lyle (2007) no encuentran ningún efecto.

En este trabajo utilicé el método de variables instrumentales para lidiar con el problema de endogeneidad, los instrumentos que se han usado anteriormente son bastante diferentes. Angrist y Lang (2004) encuentran resultados pequeños pero significativos, al utilizar el número de los estudiantes con desventajas transferidos del programa Metco³ en Boston como instrumento para predecir el efecto sobre el rendimiento de los estudiantes de los colegios a los que ellos llegaban. Zhang (2009) usó el rendimiento académico de los nuevos pares como instrumento para identificar el efecto de grupo par en el rendimiento académico de la cohorte, usó un panel de datos de 15 colegios en un distrito de China y encontró efectos positivos pero no significativos.

Para este trabajo conté con una red de conexiones bien definidas, al igual que Bramoullé et al. (2009) que usa variables instrumentales y argumenta que cuando hay una triada intransitiva⁴. El resultado de uno de los lados de la triada, sirve como instrumento para ver el efecto entre el individuo del centro y el individuo del otro lado. En este trabajo utilicé la calidad del colegio de los pares para instrumentar la calidad de la red de cada uno de los individuos.

³Metco "Metropolitan Council for Educational Opportunity", es el programa de desagregación voluntaria en colegios de Boston

⁴ Cuando en una red hay tres individuos a , b , y c , sí hay una conexión entre a y b y entre b y c pero ninguna entre a y c .

Este análisis es de relevancia porque a diferencia de la mayoría de literatura existente cuenta con la información de la red social específica de estudio de cada uno de los estudiantes. Una red social es una estructura compuesta de nodos, generalmente son individuos u organizaciones, conectados por uno o más tipos de interdependencia, como amistad, creencias, comercio o conflicto. Bramoullé et. al (2009) explica que el análisis de las redes sociales se empezó a desarrollar gracias a los sociólogos, Jackson (2006) muestra que en los años recientes los economistas han empezado a analizar los efectos de las redes sociales en el comportamiento de los individuos.

3. Datos

En esta sección describiré los datos obtenidos de varias fuentes y analizaré, la cohorte que entró a estudiar economía y finanzas en el primer semestre del 2016 a la Universidad del Rosario, la cual está conformada por 157 estudiantes. Empezaré describiendo la Universidad del Rosario y las características que han tenido sus estudiantes en el tiempo. Después mostraré las características descriptivas de la muestra como lo son edad, genero, programa e indicador socio-económico obtenidos de la base de datos del Saber 11. Posteriormente analizaré los datos que son de interés para encontrar los efectos de grupo par.

Las variables de resultado son el desempeño de cada individuo medido por un promedio general y por la nota de cada una de las materias que los estudiantes cursaron. Las variables explicativas son la calidad académica de los estudiantes, instrumentada con la calidad del colegio del que se graduaron los estudiantes de la red, medida con los resultados de las pruebas de estado Saber 11. Por último, explicaré como se construyó el instrumento a partir de la encuesta a final de semestre en donde se pregunto, entre otras cosas por la red de estudio de cada estudiante.

Los datos provienen de la Universidad del Rosario, ubicada en Bogotá, Colombia. Esta universidad, que fue fundada en 1653, es una de las instituciones educativas de mayor tradición en Colombia y una de las que más influencia ha tenido en la historia del país. Actualmente la Universidad del Rosario cuenta con siete facultades. Por ser una universidad privada en la que el costo de la matrícula es bastante elevado para el ingreso promedio de una familia en Colombia. Sin embargo, en el semestre analizado en este trabajo, estudiantes de menores recursos ingresaron a la institución gracias al programa Ser Pilo Paga.

La muestra tenida en cuenta para este trabajo es de 157 individuos porque 22 estudiantes no contestaron una de las encuestas realizada durante el semestre. En la tabla 1 están las estadísticas descriptivas de la muestra, la edad de los estudiantes promedio es 17 años y 9 meses, el 61 % de los estudiantes son hombres. El porcentaje de los estudiantes que estudia finanzas y comercio internacional es 53 % y el 47 % restante estudia economía. El nivel socio

económico es medido por la variable estrato, la media es 3,2 con una desviación estándar de 1,09, es decir personas de clase media. Además hay un indicador socio-económico que depende de los artículos que tienen en la casa los estudiantes y de la educación de los padres⁵.

Tabla 1: Estadísticas descriptivas

VARIABLES	Obs	Media	Des. Stan.	Min	Max
Promedio UR	157	3.743	0.568	1.838	4.811
Puntaje Saber 11	157	69.04	6.740	52.50	91.00
Integrantes red	157	.649	7.611	1	40
Estudia Economía	157	0.477	0.501	0	1
Edad	157	17.70	0.957	16	23
Masculino	157	0.598	0.491	0	1
Estrato	157	3.267	1.094	1	6
Beneficiario Pilo Paga	157	0.312	0.464	0	1
Indicador Socioeconómico	148	0.626	0.264	0	1

Para analizar los efectos de grupo par sobre el rendimiento académico, utilizaré el promedio ponderado del semestre de todas las notas de los cursos que los estudiantes tomaron como variable de resultado. En primer semestre los estudiantes de las dos carreras siguen casi el mismo plan curricular, todos toman los cursos de Introducción a la Ciencia Económica, Cálculo Diferencial y Lógica Matemática y lo que marca la diferencia entre las dos carreras es el curso de Doctrinas Económicas que lo toman únicamente los estudiantes de Economía y el curso de Contabilidad Financiera solo lo toman los estudiantes de Finanzas. La tabla 1 muestra que el promedio de las notas de la universidad fue de 3,7 sobre 5 en toda la muestra con una desviación estándar de 0,5.

El efecto de grupo par se entenderá como el efecto que tiene la calidad de la red de cada uno de los estudiantes, es decir el promedio de las notas de la universidad de los compañeros reportados en la encuesta que se hizo en el mes de mayo del año 2016, cuatro meses después de empezar semestre. En esta encuesta los estudiantes especificaron con nombre propio cuáles eran los estudiantes con los que estudiaban por lo menos una vez en la semana por fuera del horario de la clase⁶. A partir de estos resultados se identificaron las redes de estudio con las cuales medí la calidad de la red y su instrumento.

⁵ Calculado por el método de componentes principales, el cual tiene en cuenta las diferentes características económicas relevantes. El indicador es un valor entre 0 y 1, si el indicador se acerca más a 1 quiere decir que el individuo esta en una buena situación socio económica

⁶ Pregunta del cuestionario "Seleccione las personas con las que estudia por lo menos una vez a la semana por fuera de los horarios de clase"

En total esta encuesta tenía 6 preguntas sobre redes, donde se preguntó con nombre propio cuáles eran los estudiantes con los que compartía diferentes espacios. Con esta información reconstruí las siguientes redes dirigidas⁷: (i) la red de estudio, las relaciones entre los estudiantes se pueden ver la figura 1, (ii) la red de las actividades sociales (por ejemplo, fiesta, teatro, cine y otras) de al menos una vez por semana ⁸, (iii) la red de las actividades deportivas⁹, (iv) la red de los compañeros más cercanos, ya que se les preguntaba los estudiantes a los cuáles les podrían pedir un consejo personal¹⁰, (v) la red de conocidos ex-ante¹¹ y (vi) la red de los estudiantes que son considerados líderes¹².

También analizaré la red académica, según diferentes atributos de los estudiantes como lo son el genero, el programa que estudian, el estrato socio-económico, el rendimiento académico, el ser beneficiario del programa Ser pilo paga. En la figura 1 se observa la centralidad de cada uno de los nodos, medida por su tamaño y también se ve que no hay ninguna polarización por género (los nodos morados representan a las mujeres y los nodos verdes a los hombres). Además se encontró que no hay polarización por estrato socio económico, o desempeño académico pero que los estudiantes están agrupados por la carrera que estudian según el análisis de redes por atributos. Esta agrupación era de esperarse pues cuando los estudiantes reciben su horario la facultad de economía los agrupo en seis diferentes grupos, tres de finanzas y tres de economía debido a que hay una materia diferente para cada programa.

⁷Una red dirigida es la que contiene la dirección de la relación, es decir si a reporta a b y b no reporta a a existe una flecha entre los dos nodos con una única dirección

⁸Pregunta "Seleccione las personas con las que hace planes sociales (fiesta, teatro, cine, etc.) en su tiempo libre por lo menos una vez a la semana ". Se puede observar en la figura 4 del anexo

⁹Pregunta "Seleccione a las personas con las que hace actividades deportivas por lo menos una vez a la semana". Se puede observar en la figura 5 del anexo

¹⁰ Pregunta "Seleccione a las personas a las que acudiría para pedir un consejo personal"

¹¹ Pregunta "Seleccione los estudiantes que conocía antes de ingresar a la Universidad"

¹² Pregunta "Seleccione las personas que usted considere como líderes estudiantiles"

Red de estudio fuera de clase (min 1 x semana)

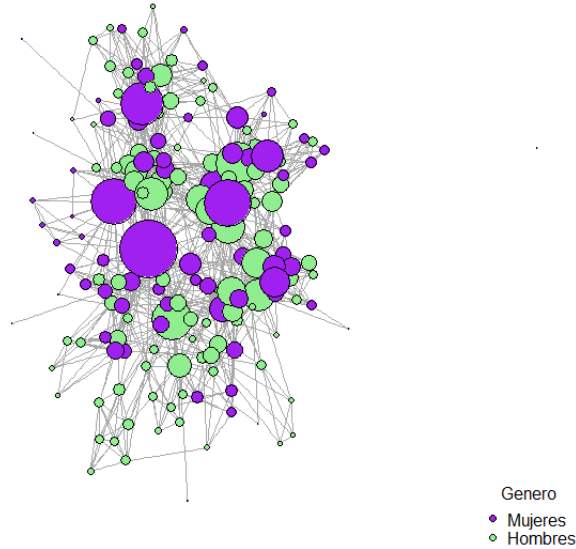


Figura 1: Red Académica por género

La información acerca de la red de estudio muestra que los estudiantes reportaron en promedio a 8 compañeros, con los que estudian por fuera de clase por lo menos una vez a la semana. El mínimo de estudiantes con los que estudia algún individuo es de 0, es decir su red es nula, a estos individuos no los tuve en cuenta para el análisis. La red más numerosa es de 40 compañeros, lo cual parece ser un problema de sobre-reporte en los outliers ya que es poco probable que un estudiante pueda estudiar con 40 estudiantes diferentes en una semana, por fuera de clase. Al ver la distribución del tamaño de la red de cada estudiante llama la atención que hay varios estudiantes que reportaron más de 20 compañeros. Es importante tener en cuenta este problema de sobre-reporte a la hora de encontrar los efectos del grupo par.

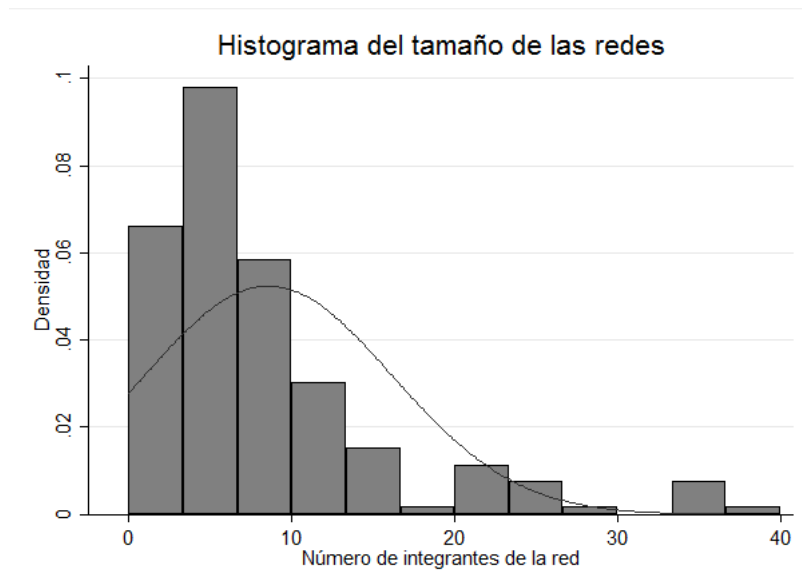


Figura 2: Histograma del tamaño de la red

La Secretaría Académica organiza los horarios de los estudiantes de primer semestre sin algún orden en específico, esto significa que no podemos garantizar el balance en los diferentes grupos en los que se reparten los estudiantes. Por esta razón puede haber una autoselección dentro de los seis grupos por características como el desempeño en los quices o primeros parciales o por características como el nivel socio-económico. Este es otro problema que también hay que tener en cuenta a la hora de analizar los resultados. Para descartar esta agrupación por desempeño académico o estrato socio-económico grafiqué la red según el desempeño del estudiante y su estrato, no encontré grupos delimitados por estas características.

Además, por la naturaleza del estudio se incurre en un problema de reflexión ya que la red puede influenciar al estudiante al igual que el estudiante puede afectar a la red. Por esto, para identificar de que manera la calidad de la red afecta el rendimiento del estudiante, recurrí al desempeño pasado de los estudiantes por medio de los resultados de las pruebas de Saber 11. Para graduarse del colegio los estudiantes en Colombia tienen que presentar un examen de estado llamado pruebas Saber 11, además es requisito indispensable ingresar a una institución de educación superior.

Las pruebas Saber 11 es un examen que evalúa el conocimiento de los estudiantes en las áreas de matemáticas, inglés, sociales, ciencias naturales y competencias ciudadanas. El promedio de los resultados de las pruebas Saber 11 de los estudiantes de la muestra fue de 69.04, con una desviación estándar de 6.7 el mínimo puntaje fue de 52,2 y el máximo de 91,0. Los resultados del examen del colegio de cada uno de los miembros de la red, es decir el desempeño del colegio, fue usado para instrumentar la variable del rendimiento académico de los miembros de la red. La variable independiente es la calidad académica de la red, medida como el promedio de las

notas de la universidad de los integrantes de esta. Debido al problema de endogeneidad, fue necesario construir el instrumento como el promedio del desempeño del colegio en las pruebas Saber 11 de cada uno de los integrantes de la red de estudio de los estudiantes de la muestra.

4. Metodología

Segun Manski (1993) y Angrist y Lang (2004) medir el efecto de grupo par implica endogeneidad por el problema de selección y el problema de reflexión. La endogeneidad genera dificultades similares a las que se pueden producir por variable omitida o errores de medida de la variable y por esto los estimadores por mínimos cuadrados ordinarios son sesgados e inconsistentes. Para solucionar este problema usé el estimador de variables instrumentales, o también conocido como estimador de mínimos cuadrados en dos etapas. La primera etapa es la regresión entre el instrumento y la variable endógena y la segunda etapa es la regresión entre la variable endógena estimada y la variable de resultados.

La existencia de endogeneidad del modelo requiere un instrumento para la variable independiente. En este caso la variable a instrumentar es la calidad académica de la red de pares de la universidad, medida por el promedio de las notas de la universidad de la red, su instrumento es el promedio de los resultados del Saber 11 del colegio de cada uno de los integrantes de la red. Por lo tanto en la primera etapa estimamé la relación entre el resultado del colegio en las pruebas Saber 11 de cada integrante de la red y el promedio de las notas de la universidad del primer semestre del 2016.

La segunda etapa es el modelo con los coeficientes estimados de la primera etapa, es decir el impacto que tiene la calidad de la red instrumentada sobre el desempeño académico del estudiante. A continuación presento el modelo que me interesa estimar, seguido por las dos etapas que llevé a cabo.

$$y_{ir} = \alpha + \beta X_{ir} + E_{ir}(y_{-ir}) + \gamma E_{ir}(X_{-ir}) + \epsilon_{ir},$$

Primera Etapa

$$E_{ir}(y_{-ir}) = \alpha + \phi E_{ir}(z_{-ir}) + u_{ir},$$

Segunda Etapa

$$y_{ir} = \alpha + \beta X_{ir} + \widehat{E_{ir}(y_{-ir})} + \gamma E_{ir}(X_{-ir}) + \epsilon_{ir}$$

Donde y_{ir} es el rendimiento académico de cada individuo medido por el promedio académico en la universidad al culminar el primer semestre, X_{ir} son los covariados a nivel individuo que son tenidos en cuenta como edad, genero, indicador socio económico, programa que estudia y el

resultado del individuo en las pruebas Saber 11. El rendimiento académico promedio de la red de cada uno de los individuos $E_{ir}(y_{-ir})$, medido por el promedio de las notas en la universidad, $E_{ir}(X_{-ir})$ son los mismos controles individuales pero en este caso de la red y finalmente $E_{ir}(z_{-ir})$ es el instrumento del rendimiento de la red dado por la calidad del colegio al que pertenecen los integrantes de la red, medido por el desempeño del colegio en las pruebas Saber 11.

A diferencia del estimador de mínimos cuadrados ordinarios, el estimador de variables instrumentales, lida con el problema de reflexión, sí es insesgado y consistente siempre y cuando cumpla con las dos siguientes condiciones:

I. La condición de relevancia

$$Cov(E_{ir}(z_{-ir}), E_{ir}(y_{-ir})) \neq 0$$

II. La restricción de exclusión

$$Cov(E_{ir}(z_{-ir}), u) = 0$$

Para observar la relevancia del instrumento se analiza su relación con la variable independiente a instrumentar haciendo la regresión de primera etapa. Si la covarianza entre el instrumento y la variable es muy pequeño el instrumento es débil lo que producirá será sobreestimación del estimador por variables instrumentales. En este caso, la covarianza no es baja, lo cual indica que la relación entre los resultados del Saber 11 y el promedio acumulado de la universidad tienen una relación positiva, lo cual era de esperarse ya que las dos variables miden el desempeño académico de los estudiantes en diferentes momentos del tiempo.

Se puede inferir entonces que hay una relación muy estrecha entre la calidad del colegio de donde los estudiantes se graduaron y su desempeño actual en la universidad. La calidad del colegio es una proxy del desempeño académico del estudiante, ya que es en el colegio en donde el individuo aprende las bases para desempeñarse adecuadamente en la universidad. Los resultados que mostraremos mas adelante corroboraran esta afirmación.

Para la restricción de exclusión, es necesario que el instrumento no esté correlacionado con el error para poder inferir causalidad. En otras palabras la condición de exclusión implica que el único canal por el cual los resultados de las pruebas Saber 11 de los compañeros del colegio del par estén relacionados con el rendimiento académico del individuo es el rendimiento académico de la red de compañeros de la universidad del individuo.

El promedio de los resultados del desempeño del colegio en las pruebas Saber 11 de cada uno de los estudiantes de la red solo esta relacionado con el rendimiento académico del estudiante por medio de la calidad actual de la red, el único otro mecanismo que podría estar siendo

omitido es que toda la red de algún estudiante perteneciera al mismo colegio y en este caso habría una relación directa con el desempeño del estudiante en la universidad y el instrumento. Sin embargo, el máximo de estudiantes que entro del mismo colegio en ese semestre fue tres. Además los estudiantes de un mismo colegio, se relacionaron con compañeros egresados de otros colegios al ingresar a la Universidad del Rosario, por lo tanto no habría un problema de correlación con el resultado en la Universidad.

5. Resultados

En esta sección mostraré los resultados de aplicar la metodología de variables instrumentales a los datos obtenidos de la Universidad del Rosario. Empezaré con los resultados de la regresión de mínimos cuadrados ordinarios para tener una línea de base, luego los resultados de mínimos cuadrados en dos etapas introduciendo el instrumento al modelo. Seguido, haré diferentes pruebas de robustez dado el tamaño de muestra que tengo. Después de esto, haré algunas modificaciones para analizar cuál es el impacto de la red sobre cada una de las materias que ven los estudiantes y por último modificaré la red que los estudiantes reportaron conservando únicamente los enlaces recíprocos.

Según Murray (2010) una buena práctica cuando se usan variables instrumentales con experimentos no aleatorios es seguir los siguientes pasos: (i) chequear la significancia y el signo estimado del instrumento, (ii) hacer un testeo de instrumentos débiles, (iii) obtener las estimaciones usando Fuller¹³ en vez de mínimos cuadrados en dos etapas y (iv) interpretar los resultados de variables instrumentales.

Las tres primeras columnas de la tabla 2 muestran las estimaciones de los coeficientes por mínimos cuadrados ordinarios. El efecto de la calidad de la red sobre el rendimiento del estudiante, es grande y significativo a un nivel de confianza del 99%. Analizo el efecto sin ninguna variable de control en la primera columna, en la segunda agrego los controles del individuo (edad, programa, género, resultado de las pruebas Saber 11 y el indicador socioeconómico), en la tercera agrego los mismos controles individuales más los controles de la red. Dado un problema de reflexión, la calidad de la red se ve impactada por el rendimiento del individuo y viceversa, surge el problema de endogeneidad y por esta razón es probable que los coeficientes estén sesgados.

Para solucionar el problema de endogeneidad utilizo el método de variables instrumentales como estrategia de identificación. Primero, hago un análisis de la relevancia del instrumento, en

¹³Variaciones de LIML que hizo Fuller (1977). En la variación (1) cuando el parámetro a es igual a 1, el estimador en este caso es aproximadamente insesgado. Cuando el parámetro a es igual a 4, el estimador es sesgado pero tiene un menor error cuadrático medio

Tabla 2: Modelo Base

	OLS			Primera Etapa			2SLS		
	Notas UR	Notas UR	Notas UR	Notas UR red	Notas UR red	Notas UR red	Notas UR	Notas UR	Notas UR
Notas UR Red	0.881*** (0.124)	0.641*** (0.151)	0.738*** (0.180)				0.968*** (0.225)	0.782** (0.317)	0.806** (0.319)
Icfes Colegio				0.420*** (0.0552)	0.359*** (0.0732)	0.408*** (0.0662)			
Promedio Icfes		0.380*** (0.064)	0.393*** (0.075)		0.0305 (0.0434)	0.0420 (0.0437)		0.362*** (0.066)	0.374*** (0.075)
Características individuales		✓	✓		✓	✓		✓	✓
Características de la red			✓			✓			✓
Constante	-0.077 (0.067)	0.519 (1.032)	-1.592 (3.267)	-0.404*** (0.0791)	0.112 (0.602)	-0.775 (2.517)	-0.092 (0.066)	0.560 (1.006)	-1.135 (3.084)
F-Wald	157	149	130	49.912	30.827	34.367	149	148	129
Número de observaciones	0.250	0.424	0.469	0.253	0.397	0.504	0.256	0.421	0.470

nota: *** $p \leq 0,01$, ** $p < 0,05$, * $p < 0,1$

Características individuales: genero, edad, programa, indicador socioeconómico

Características de la red: genero, edad, programa, indicador socioeconómico

las tres siguientes columnas de la Tabla 2 muestran los resultados de la primera etapa. Estimo la calidad académica de la red de compañeros de estudio por medio del promedio del resultado de las pruebas Saber 11 del colegio, las tres especificaciones muestran un efecto significativo del instrumento sobre la calidad de la red. El test de Wald es usado para analizar la relevancia, un instrumento es relevante cuando la F-Wald es mayor que 10, en este caso el test de relevancia de Wald es de 49.912 en el modelo que no tiene ninguna variable de control. En el modelo con los controles del individuo el F estadístico de Wald es 30.827, cuando se incluyen los controles de la red el F de Wald es de 34.367. Por lo tanto, el instrumento es relevante en todas las especificaciones del modelo.

En la misma tabla están los resultados de la segunda etapa de variables instrumentales, en el modelo sin controles encontramos un coeficiente de 0.968 a un nivel de significancia del 99 %. En el modelo con controles del individuo el coeficiente estimado es significativo a un nivel de significancia del 95 %, por un aumento en una desviación estándar en el rendimiento académico de la red instrumentada el rendimiento académico del estudiante aumenta en 0.782 desviaciones estándar, también se ve que el efecto del desempeño en las pruebas Saber 11 sobre el promedio académico es significativo a un nivel de confianza del 99 %, por cada desviación que aumenta el resultado de las pruebas Saber 11 el promedio de la universidad aumenta en 0.362 desviaciones.

En el modelo con todos los controles, por un aumento en una desviación estándar en el rendimiento académico de la red el rendimiento académico del estudiante aumenta en 0.806 desviaciones estándar a un nivel de significancia del 95 %, de lo cual se puede inferir que tener un grupo de pares más aplicado tiene efectos positivos en el desempeño de los estudiantes.

Además de que los resultados de la F de Wald muestran que el instrumento es relevante, quise hacer un chequeo de robustez ya que la muestra con la que conté para hacer la inferencia es pequeña, en la Tabla 3 están los resultados de la estimación de los coeficientes por medio de "Limited information maximum likelihood" (LIML). LIML se basa en el supuesto de normalidad de los errores en la ecuación estructural y en la ecuación de la primera etapa, se diferencia de 2SLS en las muestras finitas dada las diferencias en los pesos de los instrumentos, la literatura reciente ha demostrado que LIML tiene un sesgo menor que el estimador de 2SLS. En este caso, es el mismo coeficiente de 0.806 del modelo base con una diferencia en los niveles de significancia de 95 % a 99 % por LIML, encuentro el mismo efecto de grupo par que en la estimación por 2SLS.

En la tabla 3 también muestran las variaciones de LIML que hizo Fuller (1977). En la variación (1) cuando el parámetro α es igual a 1, el estimador en este caso es aproximadamente insesgado. Cuando el parámetro α es igual a 4, el estimador es sesgado pero tiene un menor error cuadrático

medio. En las dos modificaciones no se observa ningún cambio significativo en el tamaño del estimador pero si en la significancia cambia de 95 % a 99 %. Por último, usé la alternativa que propone Angrist (1999), que es independiente a las perturbaciones que pueden tener las muestras pequeñas, la independencia se logra replicando la primera etapa N veces, pero cada vez se va dejando una observación por fuera de la regresión. Este estimador es conocido como Jackknife de variables instrumentales, en este caso el coeficiente muestra robustez pues es igual al de 2SLS.

Tabla 3: Chequeo de robustez

	LIML	Fuller (1)	Fuller (4)	Jackknife
	Notas UR	Notas UR	Notas UR	Notas UR
Notas Universidad Red	0.806*** (0.308)	0.803*** (0.304)	0.797*** (0.292)	0.806** (0.367)
Promedio Icfes	0.374*** (0.081)	0.374*** (0.081)	0.374*** (0.081)	0.374*** (0.084)
Características individuales	✓	✓	✓	✓
Características de la red	✓	✓	✓	✓
Constante	-1.135 (3.655)	-1.143 (3.649)	-1.166 (3.634)	-1.135 (3.450)
F-Wald	34.367	34.367	34.367	
Número de observaciones	129	129	129	129
R2	0.470	0.471	0.471	0.470

nota: *** $p \leq 0,01$, ** $p < 0,05$, * $p < 0,1$

Características individuales: genero, edad, programa, indicador socioeconómico

Características de la red: genero, edad, programa, indicador socioeconómico

Dado que la formación de las redes es endógena, el efecto que estimo puede contener un efecto de auto-selección por ciertas características de los estudiantes que no observo. Así la interpretación del efecto de la calidad de la red sobre el desempeño académico puede deberse a que los estudiantes se agrupan en redes por características no observadas. Grafiqué las redes diferenciando a los estudiantes por el color de los nodos según su estrato socio-económico y su desempeño académico. Como se observa en las gráficas ¹⁴ no hay ninguna agrupación por estas características.

Además del chequeo de robustez, analicé en qué materias en específico hay efecto de grupo par. En la tabla 4 se muestran las estadísticas descriptivas de las diferentes materias que los estudiantes de primer semestre cursan. Todos los estudiantes están inscritos a los cursos de Cálculo, Introducción a la Ciencia Económica y Lógica, pero la materia de contabilidad financiera es únicamente para los estudiantes del programa de Finanzas y Comercio Internacional

¹⁴Se puede observar en la figura 6 y 7 del anexo

y la materia de Doctrinas Económicas es únicamente para los del programa de Economía. Los cursos de Cálculo y Lógica son cursos puramente matemáticos a diferencia de los otros cursos, su media es 3.37 y 3.40 respectivamente.

Tabla 4: Estadísticas descriptivas por materia

Variables	Obs	Media	Des. Stan.	Min	Max
Nota de Cálculo	167	3.372	0.918	0	5
Nota de ICE	177	3.716	0.749	0	5
Nota de Lógica	176	3.407	0.818	0	5
Nota de Contabilidad	92	3.584	0.748	0	4.7
Nota de Doctrinas	85	4.411	0.656	0	4.9

En la tabla 5 esta el efecto de la calidad de la red en cada una de las materias por separado, se evidencia que solo hay efecto en los cursos de Cálculo y de Lógica con unos estimadores de 1.305 y de 0.999 a un nivel de significancia del 99 %. En cambio no hay ningún efecto en las materias Introducción de la Ciencia Económica, Contabilidad Financiera o Doctrinas Económicas.

Esto se puede deber a que son las materias del área de matemáticas, los estudiantes usualmente vienen con unas bases muy débiles en esta área. Dado esto, estas son las materias en las que necesitan un mayor apoyo, recurriendo a su grupo de compañeros más que en las otras materias.

Tabla 5: Efecto por materia

	Calculo	ICE	Logica	Conta	Doctrinas
Notas Universidad Red	1.350*** (0.407)	0.376 (0.342)	0.999*** (0.356)	0.565 (0.372)	0.078 (0.415)
Promedio Icfes	0.400*** (0.097)	0.434*** (0.080)	0.424*** (0.091)	0.476*** (0.124)	-0.005 (0.079)
Características individuales	✓	✓	✓	✓	✓
Características de la red	✓	✓	✓	✓	✓
Constante	-0.780 (3.584)	-2.496 (3.866)	-2.071 (3.881)	4.887 (5.657)	-2.957 (3.283)
F-Wald	36.608	34.367	33.797	32.906	5.744
Numero de observaciones	122	129	128	70	59
R2	0.372	0.337	0.330	0.375	0.346

nota: *** $p \leq 0,01$, ** $p < 0,05$, * $p < 0,1$

Características individuales: genero, edad, programa, indicador socio económico

Características de la red: genero, edad, programa, indicador socio económico

Como tenía una red específica quise replicar el trabajo de Bramoullé et al (2009), en el cual se utiliza como instrumento el resultado de los estudiantes que no están en la red del estudiante pero que están en la red de los integrantes de la red del mismo, es decir las conexiones de segundo orden. Los resultados de la replicación están en el cuadro 1 del anexo. Los resultados muestran que no hay ningún efecto de la calidad de la red sobre el rendimiento académico utilizando este instrumento. Esto puede darse a que la red dirigida reportada por los estudiantes es muy densa por lo cual la red de segundo nivel no tiene muchas interacciones.

Al hacer esto también quise analizar el problema de sobre-reporte que se vio en los datos descriptivos del tamaño de la red de cada uno de los individuos. Por eso modelé la red reciproca, manteniendo solo los enlaces que fueron reportados por ambos estudiantes. En la Figura 3 se puede observar cómo es esta nueva red, la media del tamaño de esta red es de 3.5 integrantes, la desviación estándar es de 2.8, el mínimo es 0 y su máximo es 17.

Red de estudio simétrica (min 1 x semana)



Figura 3: Red Académica por género

En la Tabla 6¹⁵ se muestran las regresiones de mínimos cuadrados ordinarios, la primera etapa y la segunda etapa del modelo base pero esta vez con la red no dirigida. Los resultados de la segunda etapa son significativos cuando el modelo no tiene ningún control pero a medida que incluimos los controles la significancia se pierde y el tamaño del efecto disminuye. Esto se puede deber a que la relaciones verticales entre estudiantes, es decir los estudiantes que tienen

¹⁵Ver el anexo

mayor conocimiento pueden estar siendo reportados o reportan en un solo sentido, son las que están generando los efectos de grupo par.

6. Conclusiones

En este trabajo encontré efectos de grupo par grandes y significativos con la red dirigida que reportaron los estudiantes. El aumento en una desviación estándar en el rendimiento de la red de pares tiene un impacto de 0.8 desviaciones estándar en el rendimiento académico del estudiante. Este resultado es robusto a las estimaciones por LIML, Fuller y Jackknife que son estimaciones más precisas con tamaño de muestra pequeño.

Los resultados mostraron que el efecto de grupo par se da en las materias del área de matemáticas, a saber, Cálculo y Lógica Matemática. Lo anterior puede darse por el mayor nivel de dificultad de estas materias y por la clase de apoyo que buscan los estudiantes como el trabajo en grupo o los grupos de estudios que brinda la Universidad.

El impacto de la red puede estar influido por que la red se forma de manera endógena. Al comienzo del semestre los grupos que se formaron no fueron aleatorizados ni balanceados, los estudiantes se auto-seleccionaron de alguna manera. La formación de la red se pudo dar por características que no observé. Al graficar la red para encontrar alguna clase de autoselección por desempeño académico y por estrato socioeconómico no encontré ninguna agrupación en especial.

Para lidiar con estos problemas Branas et al. (2017) proponen hacer una reconstrucción de manera incentivada, es decir por medio de incentivos poder medir la importancia de cada uno de los lazos que tiene cada estudiante con los demás. Al igual que hacer un análisis del lado de las relaciones que los estudiantes declaran y del lado de los estudiantes que son reportados por sus compañeros.

Los estudiantes de primer semestre tienden a reportar una red más grande de lo que en realidad es. Los resultados de la red no dirigida no son significativos. Esto sugiere que los enlaces que están generando el efecto son los enlaces no recíprocos. Estos hallazgos dan soporte al uso de estrategias pedagógicas o didácticas que promuevan el aprendizaje de las matemáticas entre pares en la educación superior. En todo caso es necesario analizar con más detalle la naturaleza de los enlaces no recíprocos. Estos análisis se dejan para la investigación futura.

7. Bibliografía

An, W., (2011). “Models and Methods to Identify Peer Effects,” In: Scott, John, Carrington, Peter J. (Eds.), *The Sage Handbook of Social Network Analysis*, The Sage Publications, London, 514-532.

An, W., (2015). “Instrumental Variables Estimates of Peer Effects in Social Networks,” *Social Science*

Angrist, J.D., Lang, K., (2004). Does school integration generate peer effects? Evidence from Boston’s Metco program. *Am. Econ. Rev.* 94, 1613–1634.

Angrist, J., and V. Lavy (1999), “Using Maimonides’ Rule to Estimate the Effect of Class Size on Student Achievement,” *Quarterly Journal of Economics*.

Bramoullé, Yann, Djebbari, Habiba, Fortin, Bernard, (2009). Identification of peer effects through social networks. *J. Econ.* 150 (1), 41–55.

Branas, P., Jimenes, N., y Ponti, G. (2017). Eliciting real-life social networks: a guided tour. *Journal of Behavioral Economics for Policy*, 1(1), 33-39.

Brock, W.A. and Durlauf, S.N. (2001a) Interactions-based models. Chapter 54 of *Handbook of Econometrics V*, edited by J.J. Heckman and E. Leamer. Elsevier Science.

Brock, W.A. and Durlauf, S.N. (2001b) Discrete choice with social interactions. *The Review of Economic Studies*, 68(2): 235–60.

Bulte, C.V. d. and Lilien, G.L. (2001) Medical innovation revisited: Social contagion versus marketing effort. *American Journal of Sociology*, 106(5): 1409–35

Burke, M.A., Sass, T.R., (2008). Classroom peer effects and student achievement. *Calder Urban Institute Working Paper* 18.

Carrell, S. E., Fullerton, R. L., West, J. E., (2009). “Does Your Cohort Matter? Measuring Peer Effects in College Achievement,” *Journal of Labor Economics*, 27(3), 439-464.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., York, R. L., (1966). *Equality of Educational Opportunity*, U. S. Department of Health, Education, and Welfare, Office of Education, Government Printing Office.

Duflo, E., Dupas, P., Kremer, M. (2008). Peer effects and the impact of tracking: Evidence from a randomized evaluation in Kenya. Unpublished manuscript

Foster, G., (2006). It’s not your peers, and it’s not your friends: Some progress toward understanding the educational peer effect mechanism. *J. Public Econ.* 90, 1455–1475

Fuller, W.A. (1977): “Some Properties of a Modification of the Limited Information Estimator”. *Econometrica* 45, 939-954.

Hanushek, E.A., Kain, J.F., Markman, J.M., Rivkin, S.G., (2003). Does peer ability affect student achievement? *J. Appl. Econom.* 18, 527–544.

Hasan, S., Bagde, S., (2013). “The Mechanics of Social Capital and Academic Performance in an Indian College,” *American Sociological Review*, 78(6), 1009-1032.

- Hoxby, C.M., (2000). Peer effects in the classroom: Learning from gender and race variation. NBER Working Paper No. 7867.
- Hoxby, C.M., Weingarth, G., (2005). Taking race out of the equation: School reassignment and the structure of peer effects. Mimeograph.
- Kang, C. (2007). Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea. *J. Urban Econ.* 61, 458–495.
- Jackson, M. (2006). The Economics of Social Networks. In: Proceedings of the 9th World Congress of the Econometric Society, vol. 1. Cambridge University Press, Cambridge.
- Krauth, B.V. (2009) Simulation-based estimation of peer effects. *Journal of Econometrics*, 1: 243–71
- Lavy, V., Silva, O., Weinhardt, F. (2009). The good, the bad and the average: Evidence on the scale and nature of ability peer effects in schools. NBER Working Paper Series No. 15600.
- Lee, L., 2007, Identification and estimation of econometric models with group interactions, contextual factors and fixed effects, *Journal of Econometrics*, 140, 333-374
- Lyle, D.S., (2007). Estimating and interpreting peer and role model effects from randomly assigned social groups at West Point. *Rev. Econ. Stat.* 89, 289–299.
- Manski, C.F., 1993, Identification of endogenous effects: the reflection problem, *Review of Economic Studies*, 60, 531-542.
- McEwan, P.J. (2003). Peer effects on student achievement: Evidence from Chile. *Econ. Educ. Rev.* 22, 131–141
- Murray, Michael (2010). “The Bad, the Weak, and the Ugly: Avoiding the Pitfalls of Instrumental Variables Estimation,” Mimeo, Bates College.
- Rumberger, R. W., Palardy, G. J., (2005). “Test Scores, Dropout Rates, and Transfer Rates as Alternative Indicators of High School Performance,” *American Educational Research Journal*, 42(1), 3-42.
- Sacerdote, B., (2001). “Peer Effects with Random Assignment: Results for Dartmouth Roommates,” *Quarterly Journal of Economics*, 116(2), 681-704.
- Sacerdote, B., (2011). “Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?” In: Hanushek, E., Machin, S., Woessmann, L. (Eds.), *Handbook of the Economics of Education*, First edition, 3. Elsevier.
- Sorensen, A.T. (2006). Social learning and health plan choices. *RAND Journal of Economics*, 37(4): 929–45
- Tamayo, J. A.,(2016). The Impact of Peer Effects on Student Outcomes: Evidence from Colombia.
- Vigdor, J.L., 2006. Peer effects in neighborhoods and housing. In: Dodge, K.A., Dishion, T.J., Lansford, J.E. (Eds.), *Deviant Peer Influences in Programs for Youth: Problems and Solutions*. Guilford Press.
- Vigdor, J.L., Nechyba, T.S., (2007). Peer effects in North Carolina public schools. In: Woess-

mann, L., Peterson, P.E. (Eds.), Schools and the Equal Opportunity Problem. The MIT Press, pp. 73–101.

Zabel, J.E. (2008). The impact of peer effects on student outcomes in New York City public schools. *Education Finance and Policy* 3, 197–249

Zimmerman, D.J., 2003. Peer effects in academic outcomes: Evidence from a natural experiment. *Rev. Econ. Stat.* 85, 9–23.

Zhang, Hongliang (2009) “Peer Effects on Student Achievement: Evidence from Middle School in China.” Working Paper, Department of Economics, Chinese University of Hong Kong.

8. Anexos



Figura 4: Red de actividades sociales por género

Red de actividades deportivas (min 1 x semana)

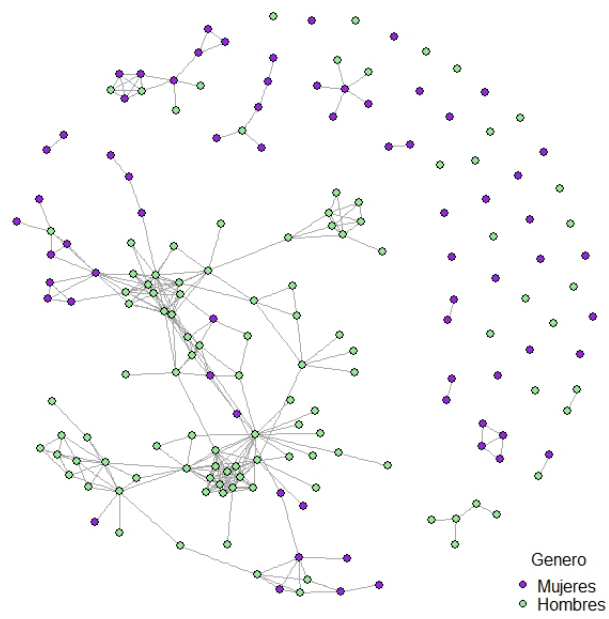


Figura 5: Red de deporte por género

Red de estudio fuera de clase (min 1 x semana)

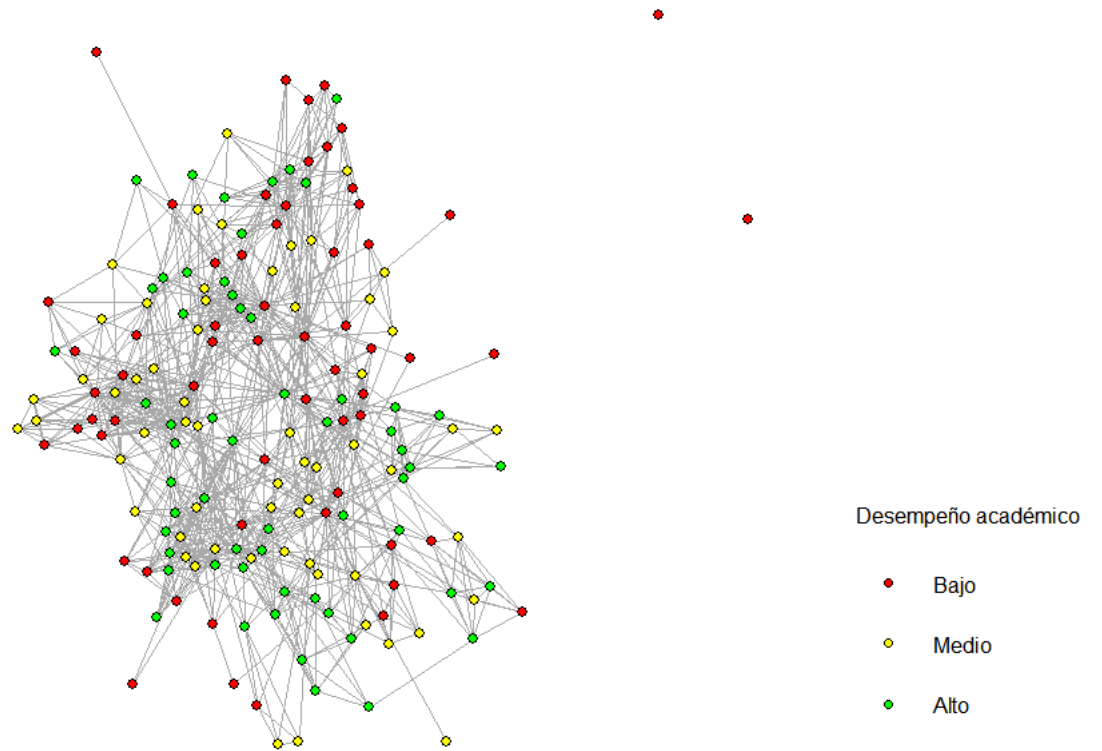


Figura 6: Red Académica según Desempeño

Red de estudio fuera de clase (min 1 x semana)

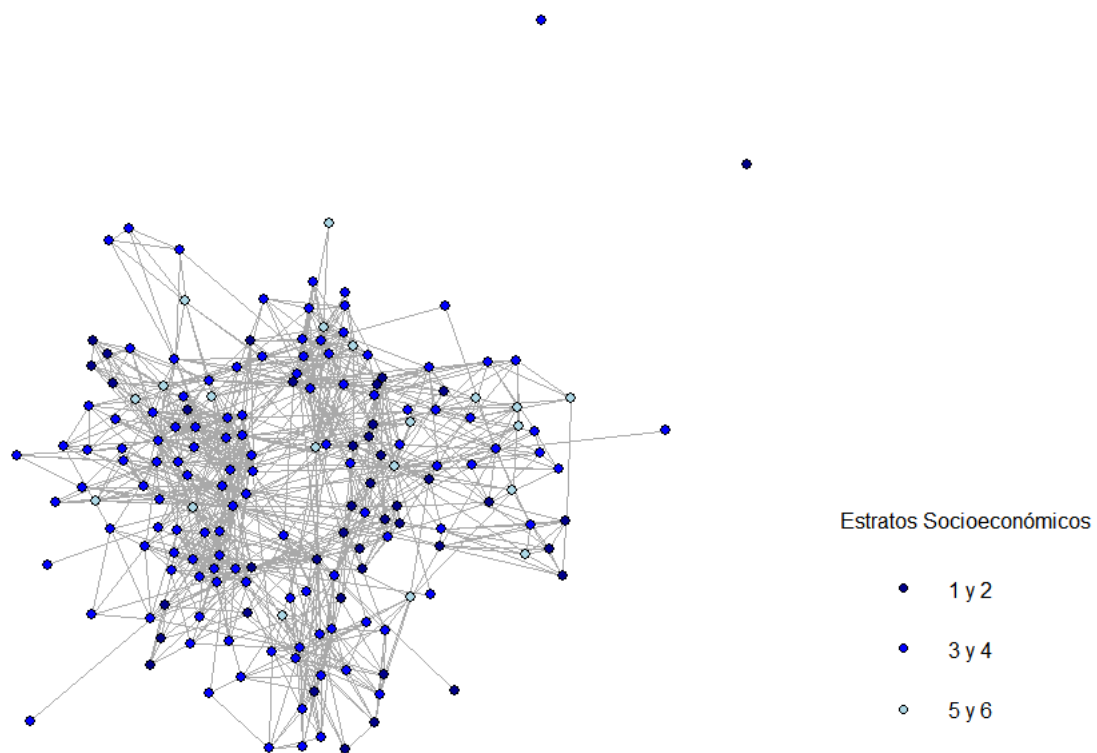


Figura 7: Red Académica según Estrato Socio económico

Tabla 6: Modelo Base Red Simetrica

	OLS		Primera Etapa		2SLS			
	Notas UR	Notas UR	Notas UR red	Notas UR red	Notas UR	Notas UR		
Notas Universidad Red	0.518*** (0.130)	0.220* (0.134)	0.203 (0.160)		0.768*** (0.186)	0.506** (0.218)	0.308 (0.354)	
Icfes Colegio				0.443*** (0.0683)	0.463*** (0.0947)			
Promedio Icfes		0.429*** (0.064)	0.416*** (0.065)	0.0537 (0.0491)	0.0485 (0.0497)	0.396*** (0.070)	0.409*** (0.068)	
Características individuales		✓	✓	✓	✓	✓	✓	
Características de la red			✓		✓		✓	
Constante	0.009 (0.061)	0.726 (1.100)	0.645 (3.211)	-0.256 (0.682)	-0.451 (2.036)	-0.049 (0.071)	0.958 (1.107)	0.898 (3.287)
F-Wald	143	143	142	15.157	8.538	79.528	45.588	28.948
Numero de observaciones				142	142	149	142	142
R2	0.114	0.397	0.411	0.351	0.521	0.111	0.370	0.408

nota: *** p<0.01, ** p<0.05, * p<0.1

Características individuales: genero, edad, programa, indicador socioeconómico

Características de la red: genero, edad, programa, indicador socioeconómico

Tabla 7: Modelo Bramoullé

	OLS		Primera Etapa		2SLS		
	Notas UR	Notas UR	Notas UR red	Notas UR red	Notas UR	Notas UR	
Notas Universidad Red	0.532** (0.224)	0.202 (0.232)	0.182 (0.249)		2.705 (2.210)	2.196 (1.771)	-0.295 (3.116)
Icfes Colegio				0.102 (0.0691)	0.120 (0.0758)	0.0663 (0.0938)	
Promedio Icfes		0.415*** (0.067)	0.417*** (0.070)	0.0308 (0.0365)	0.0390 (0.0292)	0.358*** (0.116)	0.436*** (0.152)
Características individuales		✓	✓	✓	✓	✓	✓
Características de la red			✓		✓		✓
Constante	-0.003 (0.084)	0.489 (1.154)	-2.871 (3.623)	0.0360 (0.0967)	-0.409 (0.476)	-0.711 (1.903)	-3.359 (5.469)
F-Wald	145	145	145	2.507	3.549	0.655	145
R2	0.031	0.342	0.371	0.016	0.144	0.272	-0.038

nota: *** $p \leq 0,01$, ** $p < 0,05$, * $p < 0,1$

Características individuales: genero, edad, programa, indicador socioeconómico

Características de la red: genero, edad, programa, indicador socioeconómico