



**Universidad del
Rosario**

Efecto de la perseverancia de la red de pares sobre el rendimiento académico

**Autor:
Mónica Alejandra Robayo González**

**Trabajo presentado como requisito para optar por el título de:
Magíster en Economía**

**Tutor:
Darwin Cortés**

**Facultad de Economía
Maestría en Economía
Universidad del Rosario**

**Bogotá - Colombia
Febrero 2023**

Efecto de la perseverancia de la red de pares sobre el rendimiento académico

Mónica A. Robayo-González ^{*†}

17 de Febrero de 2023

Resumen

En este trabajo estimamos el efecto del grupo par sobre el rendimiento académico con base en la perseverancia de la red de estudio, teniendo en cuenta información de habilidades socioemocionales y redes de pares de estudiantes de primer semestre de una de las mejores universidades a nivel académico de Colombia. Para identificar el efecto de la red, planteamos un modelo de variables instrumentales con el fin de resolver el problema de endogeneidad reflejada en la red de pares de primer grado, y empleando como instrumento la red de pares de segundo grado. Los resultados sugieren que la perseverancia de la red de pares aumenta 0.11 desviaciones estándar el rendimiento académico del estudiante. De igual forma, encontramos efectos heterogéneos importantes del estimado entre hombres y mujeres, los cinco grandes rasgos de la personalidad *Big-5*, otros factores no cognitivos como la maleabilidad de la inteligencia y el autocontrol, por áreas de conocimiento, y características de los colegios. Además, estudiamos los efectos del grupo par considerando comportamientos académicos de los estudiantes. Por último, realizamos una prueba de robustez utilizando como instrumento alternativo la red de pares en tercer grado, dada la cercanía de la red en segundo grado con la red de amigos. Estos resultados nos indican que la perseverancia de la red de estudio se comporta como un buen determinante del rendimiento académico.

Palabras clave: Efecto de grupo par, redes de estudio, perseverancia, habilidades socioemocionales, rendimiento académico, educación superior.

Clasificación JEL: D85, D91, I23

* *Universidad del Rosario*. E-mail: monica.robayo@urosario.edu.co.

† Quiero agradecer a Darwin Cortés, Ph.D, por el apoyo, enseñanzas y guía en este proceso. Agradezco los comentarios realizados por los participantes del Brown Bag Seminar y el Workshop de Microeconomía aplicada de la *Universidad del Rosario*. Dedicado a la memoria de mi tía Consuelo González Rozo y mi abuelito Fabio González Pérez.

1. Introducción

Las habilidades socioemocionales afectan la formación de habilidades académicas. Existe evidencia empírica donde los rasgos de la personalidad toman fuerza en la explicación del rendimiento académico Panayiotou et al. (2019), Alzahrani et al. (2019), Jiang et al. (2021). Uno de los principales rasgos de la personalidad que se comporta como predictor del logro académico y el éxito de una persona, es la *perseverancia*. Los estudios realizados por Duckworth et al. (2007) y Duckworth y Quinn (2009) introdujeron el concepto de *Grit*, definido como “la perseverancia y pasión por los objetivos a largo plazo”, siendo un rasgo esencial para medir el alto rendimiento de una persona en varios campos. Respecto a esto, varios estudios han encontrado relación entre *Grit* y el rendimiento académico, donde sugieren que la *perseverancia* es una habilidad importante para predecir el logro académico Pate et al. (2017), Mason (2018), Lam y Zhou (2019), Alhadabi y Karpinski (2020), Bennett et al. (2020), Nishikawa et al. (2022). En este trabajo, nos concentramos en la perseverancia (*Grit*) de las personas, dado que se destaca como una habilidad predictora del logro académico y el éxito en la vida adulta. De esta forma, queremos contribuir a esta discusión, centrándonos en que la socialización con compañeros con mayores niveles de (*Grit*), contribuye en mejores resultados académicos.

En esta investigación estudiamos el efecto de *Grit* de las red de compañeros y su relación con el rendimiento académico de una muestra de estudiantes universitaria. Para medir el efecto de los pares usamos el instrumento propuesta por Bramoullé et al. (2009), en el que los efectos de pares endógenos y contextuales son identificados bajo intransitividad, es decir los pares de mis pares que no son pares míos. De acuerdo con el autor, las características de los amigos de un estudiante que no son sus amigos sirven como restricción de exclusión, del cual se pueden formar triadas intransitivas. Una triada intransitiva es un conjunto de tres estudiantes l, m, n, tales que l estudia con m, y m estudia con n, pero l no estudia con n. En este contexto, nos centramos en resolver la endogeneidad asociada al problema de reflexión Manski (1993), donde los efectos endógenos y contextuales no pueden identificarse conjuntamente a partir de la forma reducida del modelo. Para este caso, la endogeneidad es controlada teniendo en cuenta que el nivel de *Grit* de los pares de pares del individuo que no son pares directos de el, no tienen un impacto directo sobre el rendimiento académico, y solamente afecta el resultado individual a través de su impacto en los resultados de los pares. Sin embargo, dicha endogeneidad es controlada parcialmente, teniendo en cuenta que los estudiantes eligen a sus propios compañeros.

Para estimar el efecto de los compañeros utilizamos la metodología de variables instrumentales, y el estimador de mínimos cuadrados en dos etapas para obtener una fuente de variación exógena

y solucionar parcialmente el problema de endogeneidad. Además, tomamos la aproximación de la escala de *Grit-5* propuesta por Duckworth y Quinn (2009), con el fin de tener una medida de perseverancia para las redes de compañeros. Para este análisis, la variable resultado es el rendimiento académico, la variable endógena la perseverancia de la red de pares en primer grado (red de amigos), y el instrumento la perseverancia de la red de pares en segundo grado (red de amigos de amigos que no son amigos). Lo que queremos mostrar es si la perseverancia de la red de estudio de los estudiantes tiene un efecto positivo sobre el rendimiento académico.

Es importante mencionar que la universidad analizada es una institución privada ubicada en Bogotá, Colombia, actualmente con acreditación institucional de alta calidad, destacada por ser una de las mejores universidades de Colombia y Latinoamérica según el QS Universities Rankings y por encontrarse en el top 10 de las instituciones a nivel nacional. Actualmente la universidad cuenta con nueve facultades, de las cuales utilizamos ocho dado que una de ellas fue establecida después del año 2018. Además desde el año 2017 por medio de la Facultad de Economía, la institución aplica a todos los estudiantes matriculados de primer semestre pruebas de habilidades cognitivas, socioemocionales y una encuesta de hábitos y redes, las cuales nos serán útiles para el presente estudio. Los datos corresponden a las cohortes de estudiantes que entraron a la universidad en 2017 y 2018.

Los resultados de esta investigación muestran que la perseverancia de la red de pares aumenta en 0.11 desviaciones estándar el rendimiento académico de los estudiantes, resultado que es robusto ante diferentes estimadores de variables instrumentales. De igual manera, evidenciamos diferencias del estimado por género, habilidades socioemocionales, áreas del conocimiento y características de los colegios. Además, estimamos los resultados principales tomando como instrumento alternativo la red de pares en tercer grado, encontrando diferencias con los resultados iniciales. Ahora bien, para complementar la hipótesis de que la perseverancia del grupo par aumenta el rendimiento académico, utilizamos como variables resultado comportamientos académicos de los estudiantes, tales como las horas de estudio, horas de permanencia en la universidad y la frecuencia de la preparación de exámenes.

Esta investigación es relevante para la literatura, puesto que contribuimos a la discusión de que las habilidades sociales de los compañeros contribuyen a mejorar las habilidades académicas, esto lo hacemos utilizando el nivel de *Grit* de la red de compañeros de estudio, trabajando con una muestra universitaria colombiana. Las principales implicaciones teóricas de este estudio se centran en las diferencias de las habilidades sociales entre género, teniendo en cuenta que las mujeres se apoyan más en su red con mayores niveles de *Grit*, sin embargo estos resultados están expuestos a la endogeneidad puesto que los individuos eligen a sus propios amigos. Además, los mecanismos

sociales juegan un papel importante en esta investigación, dado que se comportan como factores no cognitivos que se asocian directamente con el rendimiento en clase, llamados comportamientos académicos, los cuales están asociados en la dedicación a estudiar y completar tareas. Esto es importante porque es el modo en cómo los estudiantes desarrollan y demuestran sus conocimientos y habilidades académicas. Para esta investigación, los comportamientos son los mecanismos que actúan como mediadores del nivel de *Grit* y estos a su vez afectan el rendimiento académico.

El documento se encuentra organizado de la siguiente manera: en la sección 2 realizamos la revisión de literatura. En la sección 3 presentamos la descripción de los datos. En la sección 4 exponemos la estrategia empírica utilizada. En la sección 5 mostramos los resultados de las estimaciones principales y robustez del modelo. Por último, la discusión y conclusiones se encuentran en la sección 6.

2. Revisión de literatura

En los últimos años la ciencia económica se ha interesado cada vez más en cómo las redes sociales impactan el comportamiento individual en diferentes contextos, entre estos estudios encontramos: la estructura social del mercado laboral para la búsqueda de empleo [Montgomery \(1991\)](#), la relación entre las interacciones sociales y las tasas de criminalidad [Glaeser et al. \(1996\)](#), el rol de las redes en el bienestar social [Bertrand et al. \(2000\)](#) & [Bernard \(2006\)](#), efectos de pares en los resultados de los estudiantes [Sacerdote \(2001\)](#), el papel de las redes sociales y la jubilación [Duflo y Saez \(2003\)](#).

Para identificar los efectos de los pares han surgido algunos desafíos econométricos, en principio el modelo lineal en medias expuesto por [Manski \(1993\)](#) revela que, para medir el efecto del grupo par sobre el resultado individual se presentan dos tipos de efectos sociales, por un lado un componente endógeno que corresponde a la influencia de los resultados de los compañeros, y por el otro un componente exógeno que hace referencia a la influencia de las características exógenas de los compañeros. A esto se le llama problema de reflexión, donde puede ser imposible distinguir entre los efectos sociales endógenos y exógenos, debido a la simultaneidad en el comportamiento de los agentes que interactúan. Además de esto, se pueden evidenciar efectos correlacionados, los cuales se dan por la selección endógena de pares y características similares que presentan los miembros de la red. Por su parte, [Lee \(2007\)](#) expone otra preocupación basada en variables no observables, dado que dichas variables pueden tener un efecto directo en los resultados observados, el autor considera la identificación y estimación de un modelo con interacción grupal endógena y factores exógenos, donde permite la existencia de no observables correlacionados como un efecto fijo en un grupo.

Para resolver dichos problemas, algunos estudios han comprendido la importancia de las interacciones que pueden presentarse en las redes, [Bramoullé et al. \(2009\)](#) con el efecto de pares de pares que

no lo son, [Giorgi et al. \(2010\)](#) con los grupos de pares que se superponen parcialmente, [Lin \(2010\)](#) propuso un marco de modelo espacial autorregresivo, mientras que [Laschever \(2013\)](#) identifica que los individuos pueden pertenecer a múltiples grupos bien definidos, y estar influenciados por más de un grupo de referencia.

Adicional a esta literatura, la importancia de medir el efecto de los pares sobre los resultados educativos es relevante para esta investigación, en relación a ello [Calvó-Armengol et al. \(2009\)](#) estudiaron el impacto de la ubicación de la red en los resultados educativos, [Patacchini et al. \(2017\)](#) consideraron la influencia de diferentes tipos de compañeros o vínculos en los resultados educativos a largo plazo de los estudiantes, [Tincani \(2018\)](#) propuso un modelo econométrico que estima los efectos de los pares como funciones no paramétricas de las características de los estudiantes, mientras que [Verma y Meiselman \(2021\)](#) se enfocaron en el papel de los efectos disruptivos de los pares en las escuelas disciplinarias utilizando la superposición de pares.

A esta literatura incorporamos el interés por estudiar la influencia de las habilidades sociales de los compañeros sobre el rendimiento individual del estudiante. Conforme a ello, [DeLay et al. \(2016\)](#) sugieren que la intervención de aprendizaje socioemocional puede ser efectiva porque cambia la naturaleza de las relaciones sociales dentro del aula de clases, fomentando la influencia positiva de los compañeros. Por su parte, [Dasgupta et al. \(2020\)](#) midieron el efecto causal de inscribirse en una universidad selectiva en el logro cognitivo, las preferencias económicas y los rasgos de personalidad de los Cinco Grandes, y cómo la calidad de los pares afecta estos resultados. En estudios más recientes se aprovecha la asignación aleatoria para medir el efecto de los pares, [Golsteyn et al. \(2021\)](#) estudiaron si la personalidad de los compañeros afecta el desempeño en la educación en cuatro medidas distintas de los rasgos de personalidad como la persistencia, confianza en sí mismo, ansiedad y actitud de riesgo. Además, [Zárate \(2021\)](#) estudia el impacto causal de los pares en la formación de habilidades sociales y el rendimiento académico, utilizando dos tipos de tratamiento: (i) compañeros socialmente más centrales frente a menos socialmente centrales, y (ii) compañeros de mayor rendimiento frente a compañeros de menor rendimiento, de igual forma explotando la asignación aleatoria de grupos. De igual forma, [Shan y Zölitz \(2022\)](#) estimaron cómo los pares afectan el desarrollo de la personalidad de los estudiantes, encontrando que los estudiantes se vuelven más similares a sus compañeros en varias dimensiones como, competitividad, mente abierta y responsabilidad.

De acuerdo con la literatura expuesta, las habilidades socioemocionales se comportan como predictoras del rendimiento académico. En relación a esto, el concepto de *Grit* propuesto por [Duckworth et al. \(2007\)](#) toma fuerza como un determinante del rendimiento de una persona. En este trabajo los autores proponen una escala para medir la perseverancia, nombrada *Grit - Original* donde de-

muestran que *Grit* se comporta como un predictor del logro en dominios desafiantes más allá de las medidas de talento. En esta escala, incluyeron ítems que indican consistencia de intereses, y perseverancia en el esfuerzo. Posteriormente, Duckworth y Quinn (2009) sugieren otra escala nombrada *Grit-Short*, por el cual validan una medida más eficiente de *Grit*, esta se compone por un subconjunto de la escala *Grit-O*, donde la consistencia de interés y la perseverancia del esfuerzo se cargan en la escala *Grit-S* como un factor latente de segundo orden, demostrando que en los análisis factoriales confirmatorios el *Grit-S* se ajusta mejor a los datos que el *Grit-O*.

Para esta investigación, utilizamos la perseverancia de la red de pares con el objetivo de estimar el efecto sobre el rendimiento académico de los estudiantes, esto es relevante teniendo en cuenta que el enfoque de la literatura reciente no ha profundizado sobre el efecto de *Grit* de los pares sobre el resultado individual, por lo que la presente investigación estaría aportando un análisis valioso a la literatura.

3. Datos

Para el presente análisis contamos con datos de estudiantes de primer semestre inscritos en las facultades de: Economía, Administración, Estudios Internacionales, Políticos y Urbanos, Medicina, Jurisprudencia, Ingeniería, Ciencia y Tecnología, Ciencias Humanas y Ciencias Naturales, pertenecientes a las cohortes 2017-1, 2017-2, 2018-1 y 2018-2. Los datos provienen de 3 fuentes principales: la primer fuente es la *universidad analizada*, con las pruebas aplicadas de habilidades socioemocionales y cognitivas, la encuesta de hábitos y redes de estudio, información del rendimiento académico de primer semestre, rendimiento académico previo antes de entrar a la universidad (*saber11*) y características individuales de los estudiantes. La segunda fuente corresponde al *Instituto Colombiano para la Evaluación de la Educación (Icfes)*, con los resultados agregados del *Saber 11* del colegio para cada estudiante. La tercera y última fuente es el *Ministerio de Educación Nacional (MEN)*, del cual obtuvimos el índice sintético de la calidad educativa de cada institución, y las características de los colegios.

3.1. Descripción de los datos

En la tabla 1 del anexo G reportamos las estadísticas descriptivas de nuestras principales variables utilizadas en el modelo para las cohortes 2017 y 2018. La muestra analizada se compone de un total de 2325 estudiantes de primer semestre. Empezando con la variable resultado, que corresponde al promedio del periodo definitivo de primer semestre, observamos que la media del rendimiento académico es de 3.9, un promedio máximo de 4.9 y un mínimo de 0.4. La variable endógena es la perseverancia de la red de pares en primer grado, y como instrumento la perseverancia de la red de pares en segundo grado teniendo en cuenta la aproximación del instrumento propuesto por

Bramoullé et al. (2009). Además, para medir la perseverancia de la red académica, usamos la escala de *Grit* propuesta por Duckworth y Quinn (2009) (ver anexos A.2 y F-tabla F2). Para el análisis, estas variables las utilizamos de manera estandarizada, con el fin de interpretar el coeficiente de interés en desviaciones estándar.

Las variables que utilizamos como controles las dividimos en 5 grupos de análisis: en el primero encontramos las *medidas cognitivas* de los estudiantes, correspondientes a las pruebas de habilidades cognitivas aplicadas en la semana de inducción, donde son analizadas la *velocidad de procesamiento*, y la *memoria de trabajo* en el corto plazo (ver anexo B). En el segundo grupo, ubicamos el rendimiento académico previo de los estudiantes antes de entrar a la Universidad con el percentil de las pruebas aplicadas del *Saber 11* para cada estudiante. En este grupo encontramos que para las 5 áreas del conocimiento evaluadas (*ciencias naturales, matemáticas, ciencias sociales, lectura crítica e inglés*), los estudiantes de la muestra superaron en promedio el percentil 83, presentando mejor rendimiento en ciencias sociales e inglés. En el tercer grupo, tuvimos en cuenta el rendimiento académico del colegio para cada estudiante en cada área del conocimiento, donde observamos mejor rendimiento en inglés, matemáticas y ciencias naturales. Estos puntajes son calculados por el *Icfes* teniendo en cuenta los resultados de la institución en los últimos 3 años, donde entre más altos y homogéneos sean dichos puntajes, mayor será el índice calculado.

En el cuarto grupo encontramos otra medida de la calidad del colegio, llamado *índice sintético de la calidad educativa - ISCE* (ver anexo C). En el quinto y último grupo tuvimos en cuenta la información de las facultades, cohortes y características de los estudiantes. En este grupo, observamos que en promedio el 28.5% de los estudiantes pertenecen a la Escuela de Medicina y Ciencias de la salud, el 20.9% a Jurisprudencia, el 18.3% a Administración, el 13.6% a Estudios Internacionales, Políticos y Urbanos, el 9.7% a Economía, y el 8.9% a otras facultades donde se encuentran Ciencias Humanas, Ciencias Naturales, e Ingeniería, Ciencia y Tecnología. La cohorte con mayor porcentaje de estudiantes es 2017-1 con el 33.2%, le sigue 2018-1 con 29.1%, 2017-2 con el 19.3%, y 2018-2 con el 18.4% de estudiantes. Además, el 35% de la muestra analizada son hombres, el 40.8% de los estudiantes son de estrato bajo (1,2,3), mientras que el 59.1% de estrato alto (4,5,6), y la edad promedio de ingreso a la Universidad es de 17 años.

Adicionalmente, utilizamos 4 grupos de variables para estimar efectos heterogéneos. En el primer grupo encontramos los 5 grandes rasgos de la personalidad: *Apertura, Extroversión, Estabilidad Emocional, Responsabilidad y Amabilidad* (ver anexo A.1). En el segundo grupo, ubicamos a otros factores no cognitivos, como la maleabilidad de la inteligencia y el autocontrol (ver anexo A.2). Además, para la interpretación y validez interna de cada variable del grupo del *Big 5* y *Otros fac-*

tores no cognitivos, utilizamos el *Alpha* propuesto por Cronbach y Meehl. (1955). Este *Alpha* es un índice que sirve para medir la confiabilidad y consistencia interna de una escala, y para evaluar la correlación de los ítems de la misma. Los valores del *Alpha de Cronbach* entre 0.70 y 0.90 indican una buena consistencia interna (*ver anexo F*).

En tercer lugar, agrupamos las facultades por área del conocimiento: Ciencias Económicas y Administración (corresponde a la Facultad de Economía y Escuela de Administración), Ciencias Sociales (se encuentra la Escuela de Ciencias Humanas y Facultad de Estudios Int. Pol. y Urbanos), Ciencias de la Salud y Ciencias básicas (correspondiente a la Escuela de Medicina y Ciencias de la salud sin el programa de Medicina, Ciencias Naturales y MACC), Medicina (solo el programa de Medicina), y por último Jurisprudencia (solo la Facultad de Jurisprudencia). De acuerdo a esta agrupación, encontramos que el 28.0% de la muestra hacen parte del área del conocimiento de ciencias económicas y administración, el 20.9% a Jurisprudencia, el 19.6% a Ciencias Sociales, el 17.3% a Ciencias de la salud y básicas, mientras que el 14.1% pertenecen al programa de Medicina. En el cuarto y último grupo, utilizamos las características de los colegios, donde evidenciamos que el 90.3% de los estudiantes de la muestra son provenientes de los colegios privados, el 9.7% de colegios públicos, y el 32.5% de colegios privados con creencia religiosa.

Finalmente, utilizamos variables relacionadas con los comportamientos académicos de los estudiantes, entre estas variables se encuentran las horas de estudio por fuera del horario de clase, horas de permanencia en la universidad y la frecuencia de la preparación de los exámenes. Las variables de horas de estudio y horas de permanencia en la universidad, fueron utilizadas para el análisis de forma estandarizada para una mejor interpretación, y la variable de la frecuencia de preparación de exámenes tuvo una conversión de variable categórica a una variable dummy, donde indica 1 para las dos categorías más altas, y 0 para las tres categorías más bajas. Estas categorías son: 5=*durante todo el semestre*, 4=*con una o más semanas de antelación*, 3=*durante la semana anterior al parcial*, 2=*un día* y 1=*no estudia para los exámenes*. En relación a estas variables, encontramos que en promedio los estudiantes estudian 3 horas por fuera del horario de clases, y permanecen en promedio en la universidad 8 horas, el 22,6% preparan los exámenes durante todo el semestre y con una o más semanas de antelación, mientras que el 77.4% prepara los exámenes durante una semana anterior al parcial, un día antes, o no estudia para ello.

3.2. Descripción de las redes académicas

Para medir el efecto del grupo par utilizamos la encuesta de redes, la cual es aplicada a todos los estudiantes de primer semestre cuatro meses después de iniciar las actividades académicas. Es importante resaltar que en específico para las cohortes de los años 2017 y 2018, fue contratada una empresa para aplicar estas encuestas con el objetivo de disminuir el sesgo por ser una encuesta

auto-reportada. En esta encuesta, los estudiantes especifican y escogen a los compañeros con los que estudian al menos una vez a la semana por fuera del horario de clase. Contando con esta información, logramos identificar la red académica para cada uno de los estudiantes de la muestra, en primer lugar la red de compañeros en primer grado, es decir la red de pares de amigos, y en segundo lugar la red de compañeros en segundo grado, correspondiente a la red de pares de pares que no son amigos.

Las redes académicas por cada cohorte pueden visualizarse en el anexo H. Estas fueron ilustradas de acuerdo al género (los nodos rosados representan a las mujeres y los verdes a los hombres) y con base en el grado del nodo, definido como el número de conexiones de un nodo con otros nodos. El grado se compone por grados de salida y de entrada, es decir por enlaces entrantes y salientes de los nodos, en otras palabras por el número de conexiones asociadas a un estudiante, lo cual indica que entre más grande sea el nodo, mayor número de conexiones presenta. Es importante aclarar que ningún estudiante se repitió entre cohortes, por lo que solo se observan una vez. De manera general, para las cuatro cohortes los estudiantes reportan en promedio 6 compañeros con los que estudian al menos una vez a la semana por fuera del horario de clase, el mínimo de pares reportados es de 1 individuo, y un máximo de 44 compañeros. Para solucionar este problema de sobre reporte de pares, en la sección 5.5 estimamos el modelo principal sin tener en cuenta dicho sobre reporte. Además, es importante analizar las estadísticas de cada red académica tales como: la densidad y las medidas de centralidad descritas en la tabla 2 del anexo G.

Las estadísticas descriptivas de las redes son presentadas teniendo en cuenta la red académica completa y discriminada por género, con el fin de mostrar diferencias entre hombres y mujeres. En primer lugar, observamos que para cada cohorte la proporción de mujeres es mayor respecto a los hombres oscilando entre 78 % y 84.5 % para las mujeres, mientras que la proporción de hombres se sitúa entre 15 % y 21.6 %. En segundo lugar, encontramos la densidad de la red, definida como el número de relaciones existentes en relación con el número posible de conexiones, donde se puede identificar si una red está o no conectada como un todo. El valor de esta medida varía de 0 a 1, donde el límite inferior corresponde a redes sin relaciones y el límite superior representa redes con mayores relaciones posibles, un valor cercano a 1, significa que la red es más densa y por ende los nodos se encuentran más unificados. En este sentido, la información en redes densas puede fluir más fácilmente que la información en redes dispersas. Al revisar la densidad de las redes, en promedio presentaron un valor entre 0.003 y 0.005, lo cual indica que las redes no están conectadas como un todo, y por lo tanto las redes académicas pueden interpretarse como dispersas. Además, no evidenciamos diferencias notorias de la densidad de la red entre hombres y mujeres. En tercer lugar, analizamos las medidas de centralidad de los nodos, la cual puede medirse de distintas maneras, entre estas por el grado medio, la longitud promedio del camino más corto, y por medidas de cercanía.

Empezando con el grado medio de cada red, definido como el número de conexiones en media de cada nodo, refleja en media el número de vecinos o conexiones a otros nodos de modo que puede verse la popularidad de cada nodo. En general para las cuatro cohortes, los estudiantes interactúan en media entre 2.99 y 3.87 compañeros. Las mujeres presentan en media más interacciones que los hombres, en promedio 2 interacciones, mientras que los hombres en promedio 1 interacción. Esto nos sugiere que los nodos con bordes más altos, en este caso las mujeres, son vistas como un centro de actividad, y por ende presentan mayor popularidad. Por su parte, la longitud promedio del camino más corto, es definida como el número promedio de pasos de recorrido a través del camino más corto para todos los posibles pares de nodos de una red. De acuerdo a esta definición, en promedio las distancias más largas se presentaron en la cohorte 2018-1 y 2018-2, donde para pasar de un nodo a otro se necesitaron en promedio 5 pasos, mientras que para las otras cohortes, entre 3 y 4 pasos. Sin embargo, cuando analizamos la red por género en especial para la cohorte 2018-1, la cual gráficamente presenta mayor número de aristas respecto a las demás, el número promedio de pasos de un nodo a otro aumenta de manera notoria para ambos géneros, siendo mayor para las mujeres. Para las otras cohortes, el promedio de pasos entre los caminos más cortos es entre 2 y 3 pasos para las mujeres, y para los hombres entre 1 y 2 pasos. Esto refleja que, las mujeres deben de recorrer en promedio una mayor longitud respecto a los hombres entre todos los caminos más cortos posibles entre dos pares de nodos.

Para complementar el análisis y con el fin de identificar diferencias entre género, construimos un efecto heterogéneo con base en las medidas del grado y centralidad de la red. Este respectivo análisis lo exponemos en la sección 5.2.

4. Metodología

Para analizar el efecto de la perseverancia de la red de pares sobre el rendimiento académico, utilizamos la metodología de variables instrumentales para solucionar el problema de endogeneidad que existe en la red de compañeros asociada al problema de reflexión [Manski \(1993\)](#). Es importante mencionar que estamos abordando parcialmente el problema de reflexión, teniendo en cuenta que existe un margen donde se viola la restricción de exclusión, debido a que los estudiantes escogen a sus amigos, de manera que no podemos controlar la selección de amigos de la red de cada individuo. De acuerdo con lo anterior, la estrategia empírica utilizada corresponde al estimador de mínimos cuadrados en dos etapas. La primera etapa consiste en realizar la estimación entre la variable endógena y el instrumento, y en la segunda etapa la estimación entre la variable endógena estimada y la variable resultado.

Debido a la existencia de endogeneidad en el modelo, construimos un instrumento a partir de la perseverancia de la red medida por la escala de *Grit* propuesta por Duckworth y Quinn (2009), y la aproximación del instrumento propuesto por Bramoullé et al. (2009). En este sentido, en la primera etapa estimamos la perseverancia de la red en primer grado contra la perseverancia de la red en segundo grado. Y la segunda etapa corresponde al modelo con los coeficientes estimados de la primera etapa, es decir el efecto entre la variable dependiente que corresponde al promedio del periodo de primer semestre y la red de pares en primer grado instrumentada. De acuerdo a esta estrategia, a continuación exponemos las ecuaciones de la primera y segunda etapa.

Primera Etapa

$$D_{ipt} = \alpha_0 + \alpha_1 Z_{ipt} + \alpha_2 \mathbf{X}_{ipt} + u_{ipt} \quad (4.1)$$

Segunda Etapa

$$Y_{ipt} = \beta_0 + \beta_1 \widehat{D}_{ipt} + \beta_2 \mathbf{X}_{ipt} + \varepsilon_{ipt} \quad (4.2)$$

Los subíndices representan al individuo i , en el programa p , en el periodo t . Y_{ipt} es la variable de interés correspondiente al promedio del periodo de primer semestre del estudiante, D_{ipt} corresponde a la variable endógena como la red de pares en primer grado, Z_{ipt} es el instrumento como la red de pares en segundo grado, \mathbf{X}_{ipt} es el conjunto de variables controles del individuo i , entre las cuales encontramos: variables que miden el factor cognitivo de los estudiantes, resultados *Saber 11* individual, *Saber 11* del colegio, otra medida de la calidad educativa *ISCE*, facultad y cohorte del estudiante, género, estrato y edad de ingreso a la universidad.

De manera adicional, estimamos correlaciones con características de la red con el fin de examinar si el efecto de la perseverancia de la red de pares presenta algún tipo de variación de acuerdo a estas características. Entre estas incluimos: género, habilidades socioemocionales, áreas del conocimiento, y características de los colegios, las cuales incluimos en la ecuación principal como K_{ipt} . De manera concreta estudiamos las siguientes relaciones:

- *Género*: En primer lugar analizamos las diferencias del estimado entre hombres y mujeres, con base en la proporción de hombres y mujeres en la red de cada género. En este sentido, para cada género creamos una variable dummy de acuerdo con la mediana de la proporción de hombres y mujeres en la red, donde los compañeros por encima de la mediana presentan mayor proporción de pares hombres o mujeres, mientras que los ubicados por debajo una menor proporción de pares, y de manera adicional creamos una variable indicadora del género para cada individuo. En este punto, es importante tener en cuenta que para medir el efecto de la perseverancia de la red, fue necesario incluir en la ecuación principal dos características K_{ipt} : el género y la proporción de pares hombres y mujeres en la red.

- *Habilidades socioemocionales:* En segundo lugar, estudiamos los cinco grandes rasgos de la personalidad *Big 5* y su relación con el nivel de la perseverancia de toda la muestra. En este caso, por cada habilidad socioemocional del *Big 5* construimos una variable dummy teniendo en cuenta la mediana de cada habilidad, donde los valores por encima de la mediana indican tener más de dicha habilidad, mientras que los valores de por debajo lo contrario. Del mismo modo, para la perseverancia de toda la muestra construimos una variable dummy con base en la mediana, donde los valores por encima señalan a los más perseverantes de la muestra, y los valores de por debajo de la mediana a los menos perseverantes. De manera adicional, realizamos el mismo análisis pero tomando otros factores no cognitivos, como el autocontrol y la maleabilidad de la inteligencia. En este punto también presentamos dos características K_{ipt} : la habilidad socioemocional y el nivel de la perseverancia.
- *Áreas del conocimiento:* En tercer lugar, diferenciamos el efecto por área del conocimiento y su relación con el nivel de la perseverancia pero al interior de cada área. Las áreas que analizamos son: ciencias económicas y administrativas, ciencias sociales, ciencias de la salud y ciencias básicas, medicina, y jurisprudencia. Para ello, construimos una variable dummy para indicar cada área del conocimiento, y adicionalmente fue necesario una variable dummy con base en la mediana de la perseverancia de cada área estudiada, de manera que fue posible identificar para cada área del conocimiento su nivel de perseverancia. Del mismo modo, presentamos dos características K_{ipt} : el área del conocimiento y el nivel de perseverancia al interior de cada una.
- *Características de los colegios:* En cuarto y último lugar, analizamos el efecto por las características de los colegios, y su relación con el nivel de perseverancia de la red de los pares. Estas fueron: el sector (público y privado) y la creencia (religioso y no religioso). En este ejercicio, solo fue necesario construir una variable dummy por cada característica del colegio, representadas por K_{ipt} .

En tal sentido, para realizar la estimación con efectos heterogéneos es necesario incluir en la ecuación la interacción entre la variable endógena D_{ipt} y la característica K_{ipt} que define los efectos heterogéneos $D_{ipt} \times K_{ipt}$. Cabe resaltar que la estimación por variables instrumentales requiere al menos un instrumento por cada variable endógena de la ecuación principal, que para el caso de los efectos heterogéneos por género, habilidades socioemocionales y áreas del conocimiento, fueron tomadas las dos características K_{ipt} interactuadas con la variable endógena D_{ipt} por separado, y la triple interacción de la variable endógena con las dos características $D_{ipt} \times K_{1ipt} \times K_{2ipt}$. De manera análoga, al definir un instrumento Z_{ipt} , incluimos en la estimación las respectivas interacciones del instrumento Z_{ipt} con las dos características K_{ipt} por separado, y la triple interacción $Z_{ipt} \times K_{1ipt} \times K_{2ipt}$. Para el

caso de los efectos heterogéneos por las características de los colegios, tomamos la característica K_{ipt} del colegio interactuada con la variable endógena $D_{ipt} \times K_{1ipt}$ y de manera análoga la interacción con el instrumento $Z_{ipt} \times K_{1ipt}$. A continuación exponemos las ecuaciones de la primera y segunda etapa, teniendo en cuenta las características nombradas anteriormente:

Primera etapa:

$$D_{ipt} = \alpha_0 + \alpha_1 Z_{ipt} + \alpha_2 \mathbf{X}_{ipt} + u_{ipt} \quad (4.3)$$

$$D_{ipt} \times K_{1ipt} = \gamma_0 + \gamma_1 (Z_{ipt} \times K_{1ipt}) + \gamma_2 \mathbf{X}_{ipt} + \epsilon_{ipt} \quad (4.4)$$

$$D_{ipt} \times K_{2ipt} = \theta_0 + \theta_1 (Z_{ipt} \times K_{2ipt}) + \theta_2 \mathbf{X}_{ipt} + \nu_{ipt} \quad (4.5)$$

$$D_{ipt} \times K_{1ipt} \times K_{2ipt} = \delta_0 + \delta_1 (Z_{ipt} \times K_{1ipt} \times K_{2ipt}) + \delta_2 \mathbf{X}_{ipt} + \mu_{ipt} \quad (4.6)$$

Segunda etapa:

$$Y_{ipt} = \sigma_0 + \sigma_1 \widehat{D}_{ipt} + \sigma_2 (\widehat{D_{ipt} \times K_{1ipt}}) + \sigma_3 (\widehat{D_{ipt} \times K_{2ipt}}) + \sigma_4 (\widehat{D_{ipt} \times K_{1ipt} \times K_{2ipt}}) + \sigma_5 \mathbf{X}_{ipt} + v_{ipt} \quad (4.7)$$

En la primera etapa, por cada característica K_{ipt} realizamos la interacción con la variable endógena D_{ipt} y el instrumento Z_{ipt} . La segunda etapa, consiste en los coeficientes interactuados de la primera etapa, con el fin de identificar si el efecto de la perseverancia de la red de pares sobre el rendimiento académico, presenta variaciones respecto a las características nombradas. Este análisis lo presentamos de manera detallada en la sección 5.2. Es importante resaltar que un estimador consistente e insesgado es posible obtenerlo por el método de variables instrumentales, siempre y cuando se defina una variable instrumental Z_{ipt} que cumpla con las siguientes condiciones: relevancia del instrumento y la restricción de exclusión.

4.1. Relevancia del instrumento

La primera condición descrita por el enunciado 4.8 implica que la variable instrumento Z_{ipt} se encuentre correlacionada con la variable endógena D_{ipt} .

$$Cov(D_{ipt}, Z_{ipt}) \neq 0 \quad (4.8)$$

Si esta condición se cumple, el instrumento presenta buen poder predictivo sobre la variable endógena. En este caso observamos que la perseverancia de la red de pares en segundo grado explica la perseverancia de la red de pares en primer grado, y evidenciamos una correlación positiva y significativa. Estos resultados se pueden observar más adelante en la explicación del modelo principal

en la sección 5.1. Cabe señalar que no basta con tan solo encontrar una correlación fuerte entre el instrumento y el regresor endógeno, es necesario realizar otras pruebas de instrumento débil para corroborar dicha relación.

4.1.1. Pruebas de instrumento débil

Utilizamos tres test para comprobar la relevancia del instrumento. Empezando con la prueba AR de Anderson y Rubin (1949), testeamos la hipótesis nula de que los instrumentos son débiles, si la prueba arroja un estadístico F significativo rechazamos la hipótesis nula de instrumentos débiles. Además siguiendo a Andrews y Stock (2005), la prueba es buena cuando tenemos un solo regresor endógeno, pero el poder de la prueba disminuye cuando están presentes dos o más regresores endógenos. En segundo lugar, Stock y Yogo (2005) tabularon los valores críticos del mínimo valor propio del estadístico propuesto por Cragg y Donald (1993), si este estadístico es mayor que los valores críticos calculados por Stock y Yogo, rechazamos la hipótesis nula de instrumentos débiles. Para tabular los valores críticos, los autores definen un valor de sesgo relativo del estimador IV como una fracción o porcentaje del sesgo MCO, donde los valores críticos varían entre los valores del sesgo. Por último, el test propuesto por Montiel-Olea y Pflueger (2013) es recomendable usarlo cuando trabajamos con un único regresor endógeno. Esta prueba arroja un estadístico F efectivo (F_{eff}), junto con un nivel de confianza α y unos valores críticos con un umbral de sesgo deseado llamado τ . Si el estadístico de prueba F_{eff} es mayor que el valor crítico, se rechaza la hipótesis nula de que los instrumentos son débiles, dependiendo del valor de sesgo en el peor de los casos que el investigador esta dispuesto a tolerar.

En este escenario, al presentar un sistema de ecuaciones exactamente identificado es decir con un solo regresor endógeno, el estadístico F recomendado para estos casos es el propuesto por Montiel-Olea y Pflueger (2013) el cual es robusto bajo heterocedasticidad. Para múltiples variables endógenas, el estadístico F de Cragg y Donald (1993) es útil para evaluar que tan fuertes son los instrumentos, junto con los valores críticos propuestos por Stock y Yogo (2005) probamos la debilidad de los instrumentos. Así las cosas, el estadístico F de nuestra preferencia es el propuesto por Montiel-Olea y Pflueger (2013), dado que es aplicable para una sola variable endógena y no asume homocedasticidad. Los demás test nos pueden resultar útiles para comprobar la robustez del estadístico F .

Para el modelo propuesto, concluimos que el instrumento es relevante para los tres test utilizados. Por un lado los estadísticos F calculados por medio de la metodología propuesta por Cragg y Donald (1993) - Stock y Yogo (2005), y Montiel-Olea y Pflueger (2013) superan los valores críticos para las estimaciones realizadas. Por el otro lado, siguiendo a Anderson y Rubin (1949), el estadístico F es significativo en todos los casos. Los estadísticos F pueden visualizarse en el anexo D en la tabla D1 con los resultados en la versión extendida, y en la tabla D5 con los valores críticos.

4.2. Restricción de exclusión

La segunda condición descrita en el enunciado 4.9 es conocida como la exogeneidad del instrumento, la cual implica que el instrumento Z_{ipt} no debe estar correlacionado con el término de error u_{ipt} , es decir que el instrumento no debe explicar los determinantes no observados de la variable resultado, en otras palabras Z_{ipt} no debe tener influencia directa sobre Y_{ipt} .

$$Cov(u_{ipt}, Z_{ipt}) = 0 \quad (4.9)$$

Teniendo en cuenta la aproximación del instrumento propuesto por [Bramoullé et al. \(2009\)](#), la perseverancia de la red de pares de pares que no son amigos, no debe explicar el promedio del periodo del estudiante ni estar correlacionado con los determinantes no observados del rendimiento académico del individuo. Por consiguiente, la red de pares en segundo grado debe afectar el rendimiento académico del estudiante solo a través de la perseverancia de la red de compañeros. En este contexto, los pares de pares que no son pares actúan como restricciones de exclusión y como una fuente de variación exógena en D_{ipt} , y por ende no habría correlación con el efecto grupal de la red a nivel de individuo. Es otras palabras, utilizamos la exogeneidad de las características de los amigos, y las triadas intransitivas para estimar el efecto del grupo par. Esto permite separar el efecto de los resultados de los pares (efecto endógeno) del efecto directo de la perturbación de los pares (efecto exógeno). Por lo tanto, la presencia de triadas intransitivas, es decir pares que no están conectados pero comparten amigos en común, es una condición para poder identificar el efecto del grupo par.

No obstante, el instrumento propuesto por [Bramoullé et al. \(2009\)](#) ha recibido críticas respecto a la red de pares de pares, teniendo en cuenta que dicha red sigue siendo cercana a la red de pares en primer grado. De acuerdo con [Santavirta y Sarzosa \(2019\)](#) una preocupación sobre la restricción de exclusión cuando se identifican los efectos de los pares a través de triadas intransitivas, es la existencia de un canal directo entre las características de los amigos de los amigos y las propias acciones, y no solo un canal indirecto a través de las acciones de los amigos de uno. Esto es una preocupación, porque posiblemente puede existir interacción directa no observada entre un individuo y los amigos de sus amigos, por ejemplo compartir el aula de clase y diferentes actividades extras que no se están observando. Además, en un estudio reciente [Bramoullé et al. \(2020\)](#) argumentaron que el muestreo incompleto puede ser un problema, dado que la especificación incorrecta de la red puede invalidar los efectos de pares basados en la estructura de la red, lo cual genera que los amigos de los amigos tengan algún tipo de relación directa. Por tanto, para asegurar la ausencia de conexiones directas entre agentes conectados indirectamente, una solución potencial es combinar información de diferentes redes [Nicoletti et al. \(2018\)](#), [Giorgi et al. \(2020\)](#).

Para esta investigación no contamos con información de otras redes, sin embargo como ejercicio de robustez utilizamos las conexiones de pares en tercer grado como instrumento alternativo, con el fin de identificar si el efecto de la red puede mantenerse en magnitud y nivel de significancia. Estos resultados y la discusión la describimos con más detalle en la sección 5.5.1.

5. Resultados

En esta sección presentamos los resultados principales utilizando el estimador de mínimos cuadrados en dos etapas. En seguida comparamos el resultado principal con otros estimadores de variables instrumentales. Luego, reportamos las estimaciones aplicando efectos heterogéneos por género, habilidades socioemocionales, áreas del conocimiento y características de los colegios. Además, realizamos el análisis del grupo par teniendo en cuenta comportamientos académicos. Por último, como ejercicio de robustez encontramos el modelo principal sin tener en cuenta el sobre reporte de compañeros, y el análisis de la red en tercer grado.

5.1. Resultados principales

En la tabla 1 exponemos los resultados del modelo principal, en la primera fila reportamos el estimado de la segunda etapa, mientras que en la segunda fila la primera etapa. Podemos observar un resultado positivo y significativo del efecto de la perseverancia de la red sobre el rendimiento académico a un nivel de confianza estadística del 95% (*ver columna 3*). Este resultado nos indica que por un aumento en una desviación estándar en la perseverancia de la red de pares, aumenta el rendimiento académico de los estudiantes en 0.11 desviaciones estándar. En la primera etapa del modelo observamos una fuerte correlación entre el instrumento y la variable endógena a un nivel de confianza estadística del 99%, donde corroboramos la relevancia del instrumento de acuerdo a los estadísticos F reportados. Estos resultados también los estimamos incluyendo una a una el conjunto de variables controles (*ver anexo D - Tabla D1*). De acuerdo a estos resultados en forma extendida, el estimador por 2SLS no presenta mayor variación a medida que incluimos las variables controles, los errores estándar se mantienen constantes, y los resultados de la primera etapa son robustos con la condición de relevancia y con los estadísticos F calculados. Adicionalmente, estimamos la forma reducida del modelo observando que el instrumento no explica gran parte de la variable resultado, por lo que el verdadero efecto se evidencia cuando sólo la parte de la variable endógena que es explicada por el instrumento se utiliza para explicar Y .

Es importante mencionar la diferencia en magnitud del coeficiente estimado por OLS frente al estimado por IV. En los resultados observamos que la estimación por OLS es mucho menor que la estimación IV, esto puede suceder por el sesgo de atenuación que se presenta en la estimación por OLS. En este sentido, dado que la estimación IV no se ve afectada por el error de medición, tiende

a ser más grande que una estimación OLS. Además, es posible que la estimación IV sea mayor que la estimación OLS dado que IV estima el efecto local, es decir sobre un subconjunto de individuos cuyo comportamiento se ve afectado por el instrumento Z , lo cual indica que no toda la variación disponible en la variable endógena D es utilizada para explicar la variable resultado Y , mientras que con el estimador OLS se tiene en cuenta el efecto sobre toda la población. De acuerdo con este efecto local, sólo la parte de D que es explicada por Z se utiliza para explicar Y . En este caso, para el modelo principal en la medida en que la proporción de *compliers* es pequeña, el estimador de IV aumenta, lo cual se ve reflejado en la estimación de efectos heterogéneos importantes que pueden explicar el sesgo en el estimador OLS. Por lo tanto, la estimación IV es mayor que la calculada por OLS debido a la heterogeneidad en la población estudiada.

Tabla 1: **Modelo principal**

	OLS (1)	Forma reducida (IV) (2)	2SLS (IV) (3)
2da Etapa: Perseverancia de la red de pares	0.024 (0.016)		0.110** (0.052)
1era Etapa: Perseverancia de la red de pares de pares		0.033** (0.015)	0.297*** (0.031)
Número de Observaciones	2,325	2,325	2,325
Anderson & Rubin F estadístico (1949)			18,939
Cragg-Donald Wald F estadístico (1993) - Stock-Yogo (2005)			244,453
Montiel & Pflueger F estadístico efectivo (2013)			92,086
Características personales de los individuos	✓	✓	✓
Medidas Cognitivas	✓	✓	✓
Rendimiento académico previo	✓	✓	✓
Rendimiento académico del colegio	✓	✓	✓
Otras medidas de la calidad del colegio	✓	✓	✓

Notas: Resultados generales del modelo de variables instrumentales para medir el efecto del grupo par sobre el rendimiento académico con base en la perseverancia de la red. La primera fila reporta el estimado de la segunda etapa, y la segunda fila los resultados de la primera etapa del modelo. La columna (1) reporta el estimado utilizando OLS sin tener en cuenta la medida exógena del instrumento, la segunda columna la relación de la variable resultado (promedio del periodo estandarizado) con el instrumento, y la columna (3) reporta por un lado el efecto de la perseverancia de la red de pares sobre el rendimiento académico, y por el otro lado la relación entre la variable endógena y el instrumento. En cada estimación se tiene en cuenta todo el conjunto de controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo del individuo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios. Por último, se incluye en la columna (3) los estadísticos F de las siguientes pruebas de relevancia del instrumento: Anderson & Rubin (1949), Cragg-Donald (1993) - Stock-Yogo (2005), y Montiel & Pflueger (2013).

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

De manera adicional y como ejercicio de robustez, el coeficiente de interés lo calculamos teniendo en cuenta distintos estimadores de variables instrumentales (*ver anexo D - Tabla D2*), con el fin de verificar la robustez del estimador por 2SLS dado la varianza grande que puede presentar. Empezando con el estimador de *Máxima verosimilitud con información limitada (LIML)*, introducido por primera vez por [Anderson y Rubin \(1949\)](#), se usa cuando estamos interesados en estimar una sola ecuación estructural a la vez, y se diferencia del estimador por 2SLS debido a que se muestra

consistente bajo el supuesto de muchos instrumentos. En este caso, como la ecuación del modelo principal esta exactamente identificada, el estimador por *LIML* es igual al de *2SLS*, y no presenta ninguna variación en magnitud ni en los errores estándar. Así mismo, utilizamos el estimador de *LIML* de acuerdo a las modificaciones realizadas por Fuller (1977), este estimador resuelve los problemas de momentos finitos asociados a *LIML* y es asintóticamente eficiente bajo instrumentos débiles. El estimador de Fuller utiliza un parámetro a que minimiza el error cuadrático medio, cuando el parámetro $a = 1$ produce un estimador con sesgo medio igual a cero, mientras que cuando $a = 4$ produce un sesgo distinto de cero, pero minimiza el error cuadrático medio. Siguiendo esta lógica, ante las dos variaciones del parámetro a no evidenciamos un cambio significativo en la magnitud del coeficiente de interés, solamente observamos una variación mínima pero no de gran tamaño. Del mismo modo, siguiendo a Angrist et al. (1999) los autores proponen un estimador que intenta eliminar el sesgo de muestra finita de 2SLS, el cual se basa en el enfoque de "Leave one out" donde la primera etapa es replicada N veces dejando una observación por fuera de la estimación, este estimador se conoce por el nombre de *Jackknife*. Con base en este estimador el efecto de la perseverancia de la red de pares aumenta en términos de magnitud a 0.12 desviaciones estándar, los errores mostraron un mínimo incremento, y la significancia estadística se mantuvo al 95% de confianza. De acuerdo a estos resultados, concluimos que el estimador por 2SLS muestra un comportamiento robusto ante diferentes estimadores de variables instrumentales.

Por último, como ejercicio de robustez estimamos el modelo principal sin incluir la red de la cohorte 2018-1 (ver anexo D - Tabla 3), dado el número de aristas presentado en las conexiones entre pares (ver anexo H - Figura H3). En esta versión de los resultados, el efecto de la perseverancia de la red aumenta a 0.13 desviaciones estándar, los coeficientes de la primera etapa se muestran robustos con la condición de relevancia, junto con los estadísticos F reportados, y los errores estándar se mantienen constantes. Por tanto, el efecto de la perseverancia de la red de amigos no se ve afectado de manera significativa por el número de conexiones entre pares de la cohorte 2018-1.

5.2. Algunas correlaciones con características de la red

En esta sección presentamos las estimaciones de acuerdo a las siguientes características: género, habilidades socioemocionales, áreas del conocimiento y características de los colegios. Es importante mencionar que algunas de estas variables analizadas no son estrictamente exógenas, por lo que no podemos darle un sentido de causalidad, si no de correlación. En este sentido, para la red por género sabemos que los estudiantes escogen a sus amigos, además existe correlación entre el *Big 5* y la escala de *Grit*, mientras que por áreas de conocimiento se entiende que los estudiantes comparten mismos intereses y preferencias por la misma carrera. Además, vale aclarar que las regresiones fueron estimadas utilizando toda la muestra. Cada una de estas características, se exponen a continuación.

5.2.1. Género

Existen varios estudios sobre las brechas de género en las habilidades no cognitivas, mostrando que las niñas presentan mejor desempeño en dichas dimensiones. Empezando por el factor biológico, algunos autores como [McClure \(2003\)](#), [Keenan y Shaw \(1997\)](#) resaltan que las diferencias entre el cerebro masculino y femenino se dan en áreas relacionadas con el estado de ánimo y la regulación de las emociones, entre estas diferencias se encuentran un desarrollo más rápido en las niñas de la corteza frontal y el lóbulo temporal del cerebro. Adicional a esto, de acuerdo con [Knickmeyer et al. \(2005\)](#) las diferencias cerebrales estructurales y funcionales pueden deberse a la exposición intrauterina de las hormonas sexuales, donde a mayor exposición a la testosterona, tiende a una maduración más lenta del lóbulo temporal y puede expresarse en niveles más bajos de empatía y menor calidad de las relaciones sociales.

Otros autores como [Bertrand y Pan \(2013\)](#), consideran que los aportes de los padres y la calidad de la familia en la crianza también son importantes. En este estudio la mayor probabilidad de que los niños experimenten una suspensión escolar, se debe a niños criados por madres solteras, así como por madres adolescentes. Esta probabilidad se reduce cuando se enfrentan a mayores y mejores aportes de los padres. En tal sentido, estos hallazgos nos informan que la cantidad de tiempo de los padres es beneficioso para el desarrollo no cognitivo de los niños, y las brechas de género pueden presentarse por aportes diferenciales del tiempo que dedican los padres a sus hijos e hijas.

Siguiendo a [Ready et al. \(2005\)](#), las brechas de género en habilidades no cognitivas también se presentan en entornos de aprendizaje, niveles de autocontrol, habilidades interpersonales y problemas de conducta. Entre estos estudios, ([Chiu, 2001](#)) considera importante la percepción de los maestros sobre los niños y niñas en los entornos de aprendizaje, debido a que los maestros califican de manera más positiva a la niñas, dado sus niveles de atención y persistencia con las tareas, relacionado con un mayor rendimiento académico. De acuerdo con el autocontrol, los niños presentan comportamientos más directos e impulsivos y responden menos a las sugerencias de los demás ([Maccoby, 1990](#)). Estos comportamiento limitan las posibilidades para los niños en adquirir nuevas habilidades sociales ([Gresham y Elliott, 1989](#)). Respecto a la habilidades interpersonales y de conducta, los maestros tienden a calificar a las niñas con comportamientos más prosociales, mientras que califican a los niños con actitudes más agresivas y antisociales ([Maccoby y Jacklin, 1980](#)).

Otros estudios [Brophy \(1985\)](#), [Entwisle et al. \(2007\)](#), [Beaman et al. \(2007\)](#) hacen énfasis en el ambiente escolar temprano, donde el trato diferencial por parte de los maestros puede tener repercusiones duraderas en el comportamiento de los niños y niñas, en la medida en que los maestros califican a las niñas mejor que a los niños en comportamiento, y en general el contacto suele ser

negativo y disciplinario respecto a la niñas. Esto refuerza la creencia temprana en los niños y niñas de adoptar conductas específicas en el aula de clase, lo cual genera un patrón de estudiante ideal. De acuerdo con (Myhill, 2002), el patrón de interacción de las niñas de alto rendimiento sobre las calificaciones se caracteriza por tener un alto grado de concentración, voluntad de comprometerse con el maestro, y mantenerse persistentes con la tarea. Por su parte, Matthews et al. (2009) Duckworth y Seligman (2006), Duckworth et al. (2015), Carvalho (2016) y Mohoric y Taksic (2016) respaldan la hipótesis de que las habilidades de las mujeres para controlar los impulsos, seguir normas de manera más adaptativa, la auto-disciplina, auto-regulación y la comprensión de las emociones es un indicador del por qué las niñas superan en habilidades sociales a los niños.

Con base en esta literatura, las mujeres tienden a presentar mayores habilidades sociales que los hombres, no solo es un tema biológico, si no de patrones de aprendizaje que se forman en las familias y en el ambiente escolar temprano. Considerando este contexto, estimamos el efecto de la perseverancia de la red de pares por género, teniendo en cuenta la mediana de la proporción de amigos hombres y mujeres de cada individuo. Siguiendo esta lógica, la muestra ubicada por encima de la mediana indica mayor proporción de pares hombres y mujeres, mientras que los ubicados por debajo de la mediana lo contrario.

Tabla 2: **Efecto de la perseverancia de la red: con base en la proporción de pares mujeres**

	2SLS (IV)			
	Panel A: Hombres		Panel B: Mujeres	
	<Mediana Menor % de pares mujeres (1)	>Mediana Mayor % de pares mujeres (2)	<Mediana Menor % de pares mujeres (1)	>Mediana Mayor % de pares mujeres (2)
Efecto de la perseverancia de la red de pares	0.073 (0.109)	-0.164 (0.211)	0.023 (0.093)	0.173*** (0.064)
Número de Observaciones	814		1511	
Controles	✓		✓	
Test de diferencia de coeficientes (<i>p-valor</i>)	0.3088		0.1655	

Notas: Los resultados corresponden al efecto de la perseverancia de la red por género, tomando como referencia la proporción de pares mujeres en la red. En el panel A fue estimado el efecto para los hombres, mientras que en el panel B el efecto para las mujeres. Para medir el efecto heterogéneo, fue tomada como variable dummy la mediana de la proporción de pares mujeres. Por encima de la mediana, encontramos mayor proporción de pares, y por debajo de la mediana menor proporción de pares. El coeficiente de interés de la ecuación 4.7 por encima de la mediana de los hombres correspondió a la sumatoria de $\sigma_1 + \sigma_2 + \sigma_3 + \sigma_4$, y por debajo de la mediana a $\sigma_1 + \sigma_2$. Para las mujeres, el estimador por encima de la mediana fue $\sigma_1 + \sigma_3$, mientras que por debajo de la mediana a σ_1 . En cada estimación se tiene en cuenta todo el conjunto de controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios. Por último se agrega un test de diferencias de coeficientes, el cual prueba la hipótesis de nula si los coeficientes son iguales entre sí, y las hipótesis alterna de que los coeficientes son diferentes.

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

En tabla 2 presentamos el efecto de la perseverancia de la red por género con base en la proporción de pares mujeres. Para el caso de los hombres (panel A) no observamos efectos significativos, si la red se compone en mayor o menor proporción por pares mujeres. Para el caso de las mujeres (panel B), encontramos efectos significativos si la red se compone por mayor proporción de pares mujeres, mientras que no hay efecto cuando la red presenta menor proporción de mujeres. No observamos diferencias estadísticamente significativas entre los coeficientes estimados, de manera que los estudiantes ubicados por encima de la mediana no son diferentes de los ubicados por debajo de la mediana.

Para complementar esta hipótesis, construimos un efecto heterogéneo por género con base en la mediana del coeficiente de agrupación de la red. Estos resultados pueden visualizarse en el anexo E. Este coeficiente hace referencia a una medida del grado, que cuantifica qué tanto está agrupado o interconectado el nodo con sus vecinos, el cual es útil para medir el nivel de agrupación de los nodos. En tal sentido, los nodos ubicados por encima de la mediana del coeficiente de agrupación presentan mayor agrupación, respecto a los nodos ubicados por debajo de la mediana. En la tabla E.1 del anexo E, no observamos un efecto significativo para los hombres ubicados por encima y por debajo de la mediana del coeficiente de agrupación, mientras que encontramos un efecto solo para las mujeres ubicadas por encima de la mediana. Estos resultados nos indican que el efecto de la perseverancia de la red de pares es robusto para las mujeres más agrupadas de la red, mientras que para los hombres más agrupados el efecto es muy pequeño y no es significativo. No se presentan diferencias estadísticamente significativas entre los coeficientes estimados.

Adicionalmente, estimamos como efecto heterogéneo la centralidad de cercanía de la red por género con base en la mediana. Esta centralidad es definida como la inversa de la suma de todos los caminos más cortos a otros vértices, la cual es considerada como una medida de rapidez con la que se difunde la información a todos los nodos. En este sentido, si un nodo presenta un alto grado de centralidad de cercanía, difunde la información rápidamente teniendo en cuenta sus conexiones. Así las cosas, los nodos ubicados por encima de la mediana de la centralidad de cercanía presentan mayor centralidad, respecto a los nodos ubicados por debajo de la mediana. En la tabla E.2 del anexo E podemos visualizar los resultados. Para el caso de los hombres, no observamos ningún efecto de la perseverancia de la red de pares por encima o por debajo de la mediana de la centralidad de cercanía. Para el caso de la mujeres, solo observamos efecto para aquellas ubicadas por encima de la mediana. Estos resultados sugieren que en la medida en que las mujeres son más centrales en la red, el efecto de la perseverancia de la red de pares sobre el rendimiento académico es positivo, mientras que para el caso de los hombres más centrales el efecto es pequeño y no significativo. Esto revela el papel importante que juegan las mujeres en la red, teniendo en cuenta que se presentan como

importantes influenciadoras en la red de pares, respecto a los hombres. Además, no observamos diferencias estadísticamente significativas entre los coeficientes estimados para los hombres, mientras que para el caso de las mujeres sí se presentan diferencias.

Estos resultados son robustos con la literatura anteriormente expuesta, donde parte del coeficiente estimado para el género femenino, es explicado por las posibles habilidades sociales que presentan las mujeres respecto a los hombres. Además, estos resultados son oportunos con la hipótesis de que las mujeres se comportaron como importantes influenciadoras en la red de pares, dado su nivel de centralidad y nivel de agrupamiento respecto a los hombres, lo cual también puede ser el reflejo de dichas habilidades sociales dadas en el género femenino. En este sentido, al ser las mujeres un centro de actividad importante, con un mayor nivel de agrupación, y posiblemente con un mayor grado de características sociales, la perseverancia de la red de pares tiene un impacto mayor sobre el rendimiento académico.

5.2.2. Habilidades socioemocionales

5.2.2.1. Big 5

Los rasgos de la personalidad se comportan como buenos predictores del éxito académico (Mihaela, 2015), estos son definidos por los cinco grandes rasgos de la personalidad *Big 5* (Hakimi et al., 2011), (Mitrofanu y Iona, 2013), (John et al., 2020). Estos cinco grandes están correlacionados con la escala de *Grit-S*, de acuerdo con Duckworth y Quinn (2009) la responsabilidad es el rasgo que más se relaciona con la perseverancia, enseguida del neuroticismo, amabilidad, extroversión y apertura a la experiencia, lo cual puede traducirse en que estas cinco características pueden ser buenos predictores del logro educativo, pero además es necesario la medida de la perseverancia. De manera complementaria Duckworth et al. (2007) sugieren que algunos rasgos pueden ser esenciales para el éxito sin importar el dominio, donde los líderes más destacados comparten una cualidad personal llamada *Grit*. De acuerdo con los autores, la perseverancia implica trabajar arduamente para enfrentar los desafíos, lo cual requiere mantener el esfuerzo y el interés, así como de experimentar el fracaso o la adversidad. Por tanto, si bien es cierto que el talento es importante para el logro de objetivos, también es clave la constancia en la aplicación de ese talento a lo largo del tiempo.

Conforme a la literatura, el factor de la personalidad definida por el *Big 5* es buen predictor del rendimiento académico, pero para poder llegar a la meta también es necesario la aplicación del factor de la perseverancia para lograr el éxito académico. En este contexto, para encontrar complementariedad entre los cinco grandes rasgos de la personalidad y la escala de *Grit-S*, estimamos el efecto de la perseverancia de la red de pares por cada rasgo de la personalidad del *Big 5*, tomando como punto de referencia la mediana de la perseverancia de la muestra analizada.

En la tabla 3 exponemos los resultados por cada habilidad socioemocional del *Big 5*, cada fila representa una estimación y dichos resultados se deben leer de izquierda a derecha. En este sentido, por cada componente del *Big 5* tomamos la mediana de cada habilidad, donde la muestra por encima de la mediana representa a aquellos con más de dicha habilidad, mientras que los ubicados por debajo de la mediana lo contrario. Del mismo modo, tomamos la mediana de la perseverancia pero de toda la muestra analizada, aquellos por encima de la misma representan a los más perseverantes, mientras que los ubicados por debajo a los menos perseverantes.

Tabla 3: **Efecto de la perseverancia de la red de pares: Big 5**

		2SLS (IV) - Big 5			N
		Perseverancia		Test de diferencia	
		>mediana	<mediana	de coeficientes (p-valor)	
Apertura a la experiencia	>mediana	0.197** (0.081)	-0.001 (0.117)	0.1491	2325
	<mediana	0.148 (0.090)	0.097 (0.084)	0.2980	
Estabilidad emocional	>mediana	0.228** (0.100)	0.069 (0.078)	0.1797	2325
	<mediana	0.151** (0.076)	0.026 (0.150)	0.4315	
Extroversión	>mediana	0.216** (0.085)	0.005 (0.134)	0.1548	2325
	<mediana	0.118 (0.083)	0.082 (0.080)	0.7468	
Responsabilidad	>mediana	0.205*** (0.078)	-0.013 (0.082)	0.0469	2325
	<mediana	0.108 (0.090)	0.086 (0.091)	0.8555	
Amabilidad	>mediana	0.382** (0.182)	0.066 (0.089)	0.1075	2325
	<mediana	0.109* (0.060)	0.037 (0.095)	0.5078	
Controles		✓			

Notas: Los resultados corresponden al efecto de la perseverancia de la red de pares de acuerdo a los cinco grandes rasgos de la personalidad Big5, donde cada fila representa una estimación. Para medir el efecto heterogéneo, se tomó como variable dummy la mediana de la perseverancia de toda la muestra, donde los valores por encima señalaron a los más perseverantes de la muestra, y los valores de por debajo de la mediana a los menos perseverantes. Además, por cada habilidad socioemocional del Big 5 se construyó una variable dummy teniendo en cuenta la mediana de cada habilidad, donde los valores por encima de la mediana indicaron tener más de dicha habilidad, mientras que los valores de por debajo indicaron lo contrario. Con base en la ecuación 4.7, el coeficiente de interés por encima de la mediana de la perseverancia y la habilidad del *Big 5* correspondió a $\sigma_1 + \sigma_2 + \sigma_3 + \sigma_4$, mientras que por debajo de la mediana de la perseverancia $\sigma_1 + \sigma_2$. Por su parte, cuando nos encontramos por debajo de la mediana de la habilidad del *Big 5*, y por encima de la mediana de la perseverancia el coeficiente de interés es $\sigma_1 + \sigma_3$, mientras que cuando nos ubicamos por debajo de la mediana de la perseverancia es σ_1 . En cada estimación se tiene en cuenta todo el conjunto de controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios. Por último se agrega un test de diferencias de coeficientes, el cual prueba la hipótesis de nula si los coeficientes son iguales entre sí, y las hipótesis alterna de que los coeficientes son diferentes. Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Los estimados calculados muestran un efecto positivo y significativo para los cinco grandes rasgos de la personalidad para la muestra ubicada por encima de la mediana de cada habilidad socioemocional, y además para aquellos ubicados por encima de la mediana de la perseverancia, salvo para

estabilidad emocional y amabilidad, donde obtuvimos de igual forma un resultado robusto pero con la muestra ubicada por debajo de la mediana. Adicionalmente, calculamos un test de diferencia de coeficientes con el fin de observar diferencias de aquellos ubicados por encima de la mediana de la perseverancia, respecto a los ubicados por debajo. De acuerdo a ello, no observamos diferencias estadísticamente significativas entre los coeficientes estimados, salvo para el factor de la responsabilidad, resultado que es robusto con el estudio realizado por [Duckworth y Quinn \(2009\)](#) y [Duckworth et al. \(2007\)](#), en el que la responsabilidad es el componente que más tiene correlación con la escala de *Grit*. En tal sentido, a mayores niveles de perseverancia y responsabilidad, la perseverancia de la red de pares aumenta el rendimiento académico de los estudiantes. Además, aquellos estudiantes con mayores niveles de estos dos componentes socioemocionales son diferentes estadísticamente hablando, de aquellos ubicados por debajo de la mediana. De acuerdo con [John et al. \(2008\)](#), la responsabilidad facilita los comportamientos dirigidos al logro de tareas y metas, descripción que es coherente con la definición de *Grit*.

Por lo tanto, estos resultados son evidencia de la complementariedad entre las características socioemocionales del *Big 5* y el nivel de perseverancia *Grit-S*, y dicha complementariedad asociada para ciertos niveles de *Grit*. En tal sentido, la perseverancia de la red de pares tiene un impacto positivo y significativo sobre el rendimiento académico, para aquellos estudiantes con un nivel mayor de cada componente del *Big 5* y de *Grit*.

5.2.2.2. Otros factores no cognitivos

Otros componentes socioemocionales como la mentalidad de crecimiento y el autocontrol, también presentan relación con el rendimiento académico y el nivel de perseverancia. Empezando con la mentalidad de crecimiento, dicho concepto es introducido inicialmente como la teoría implícita de la inteligencia de una persona, definida como un predictor consistente de la orientación hacia las metas. En este sentido, fueron definidas dos tipos de teorías sobre la inteligencia, por un lado una teoría incremental, y por el otro una teoría de entidad. La primera cree que la inteligencia es una cualidad maleable, incrementable y controlable, mientras que la segunda cree que la inteligencia es un rasgo fijo o incontrolable. Por tanto, en la medida en que se tiene la creencia de una inteligencia maleable, los individuos persiguen el objetivo de aprendizaje de aumentar su competencia, mientras que aquellos que creen que la inteligencia es fija tienden a no desarrollar el talento ([Dweck y Leggett, 1988](#)). Este concepto de *mindset* fue propuesto oficialmente en la investigación de [Dweck \(2006\)](#), donde la teoría de la mentalidad fue presentada con un enfoque de la psicología positiva, y la inteligencia fue entendida como una cualidad maleable.

Adicionalmente, la mentalidad de crecimiento también puede ser el resultado del logro académico [Zhang et al. \(2017\)](#). Con base en ello, estudios recientes de [Nrumah et al. \(2021\)](#) y [Kotrotsiou et al.](#)

(2020) encontraron que un nivel de mentalidad de crecimiento y perseverancia pueden estar relacionados con el rendimiento académico de los estudiantes. Además, Karlen et al. (2019) comprobaron que la teoría incremental de la mentalidad esta correlacionada positivamente con los niveles de perseverancia de los estudiantes.

Por su parte, el autocontrol también se comporta como un determinante del logro académico, Duckworth et al. (2019) define el autocontrol como “la alineación de pensamientos, sentimientos y acciones con objetivos valorados de forma duradera frente a alternativas momentáneamente más atractivas”. Además, Duckworth y Gross (2014) demostraron que el autocontrol y el nivel de perseverancia son dos variables importantes que explican el éxito de la vida de una persona. En otros estudios Oriol et al. (2017) y Kannangara et al. (2018) mostraron que la perseverancia y el autocontrol están fuertemente relacionados entre sí, y se comportan como buenos indicadores de desempeño y del logro académico.

De acuerdo con la literatura, tanto la maleabilidad de la inteligencia como el autocontrol están relacionados con el nivel de perseverancia, lo cual repercute de forma positiva en el rendimiento académico. En este contexto, para encontrar complementariedad entre estas habilidades socioemocionales y la escala de *Grit-S*, estimamos el efecto de la perseverancia de la red de pares por cada habilidad.

En la tabla 4 exponemos los resultados para la mentalidad de crecimiento, el autocontrol en el trabajo y el autocontrol interpersonal, y utilizamos de igual forma la mediana de cada habilidad analizada junto con la mediana de la perseverancia de toda la muestra. Los estimados calculados muestran un efecto positivo y significativo para la muestra ubicada por encima de la mediana de cada habilidad socioemocional y de la perseverancia, siendo el estimado de la mentalidad de crecimiento el que más destaca en términos de impacto. Por tanto, estos resultados son evidencia de la complementariedad que existe entre el nivel de perseverancia y otros factores no cognitivos que son determinantes del logro educativo. Además, no encontramos diferencias estadísticamente significativas entre los coeficientes estimados. En tal sentido, la perseverancia de la red de pares tiene un impacto positivo y significativo sobre el rendimiento académico, para aquellos estudiantes que se destacan por su nivel nivel de mentalidad de crecimiento y autocontrol, y adicionalmente por un mayor nivel de perseverancia.

Tabla 4: Efecto de la perseverancia de la red de pares: otros factores no cognitivos

		2SLS (IV)		Test de diferencia de coeficientes (p-valor)	N
		Perseverancia >mediana	<mediana		
Mentalidad de crecimiento	>mediana	0.246*** (0.088)	0.061 (0.074)	0.0940	2325
	<mediana	0.109 (0.088)	0.054 (0.119)		
Autocontrol en el trabajo	>mediana	0.190*** (0.073)	0.011 (0.109)	0.1590	2325
	<mediana	0.148 (0.104)	0.077 (0.086)		
Autocontrol interpersonal	>mediana	0.181*** (0.072)	0.003 (0.082)	0.0897	2325
	<mediana	0.177 (0.108)	0.127 (0.121)		
Controles		✓			

Notas: Los resultados corresponden al efecto de la perseverancia de la red de pares de acuerdo a otros factores no cognitivos, como: la mentalidad de crecimiento y el autocontrol, donde cada fila corresponde a una estimación. Para medir el efecto heterogéneo, se tomó como variable dummy la mediana de la perseverancia de toda la muestra, donde los valores por encima señalaron a los más perseverantes de la muestra, y los valores de por debajo de la mediana a los menos perseverantes. Además, por cada factor no cognitivo se construyó una variable dummy teniendo en cuenta la mediana de cada habilidad, donde los valores por encima de la mediana indicaron tener más de dicha habilidad, mientras que los valores de por debajo indicaron lo contrario. Con base en la ecuación 4.7, el coeficiente de interés por encima de la mediana de la perseverancia y la habilidad socioemocional estudiada correspondió a $\sigma_1 + \sigma_2 + \sigma_3 + \sigma_4$, mientras que por debajo de la mediana de la perseverancia $\sigma_1 + \sigma_2$. Por su parte, cuando nos encontramos por debajo de la mediana de la habilidad socioemocional estudiada, y por encima de la mediana de la perseverancia el coeficiente de interés es $\sigma_1 + \sigma_3$, mientras que cuando nos ubicamos por debajo de la mediana de la perseverancia es σ_1 . En cada estimación se tiene en cuenta todo el conjunto de controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios. Por último se agrega un test de diferencias de coeficientes, el cual testea la hipótesis de nula si los coeficientes son iguales entre sí, y las hipótesis alterna de que los coeficientes son diferentes.

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5.2.3. Áreas del conocimiento

En esta investigación también estudiamos las diferencias del estimado principal de acuerdo a las áreas del conocimiento, entre estas encontramos: ciencias económicas y administración, medicina, ciencias de la salud y básicas, ciencias sociales y jurisprudencia. Para medir el efecto, tomamos la mediana de la perseverancia al interior de cada área del conocimiento, donde los ubicados por encima de la mediana son vistos como más perseverantes, mientras que los ubicados por debajo el caso contrario.

De acuerdo con la literatura, los compañeros influyen en el comportamiento individual en muchos dominios. En primera instancia, [Kiessling et al. \(2019a\)](#) exploran los determinantes individuales de la selección de pares en tres principales aspectos: el rendimiento relativo, las diferencias de personalidad y la presencia de lazos de amistad. Los autores encuentran que los lazos de amistad impulsan la selección de compañeros, y muestran rasgos de homofilia en términos de rendimiento y de perso-

alidad. En tal sentido, argumentan que la similitud en el desempeño se torna muy importante para la selección de compañeros, y que la mayoría prefiere compañeros con un rendimiento ligeramente superior pero similar. Más tarde los mismos autores [Kiessling et al. \(2019b\)](#), estudian cómo los compañeros auto-seleccionados afectan el rendimiento en contraste con los pares asignados al azar. Los autores argumentan que la auto-selección de pares conduce a mejorar el desempeño, destacando que el par auto-seleccionado dentro de un grupo no es igual a un compañero aleatorio. Por su parte [Cicala et al. \(2017\)](#), proponen una teoría de las interacciones sociales basada en la auto-selección y la ventaja comparativa, en el que resaltan que los estudiantes eligen grupos de pares en función de su ventaja comparativa dentro de un entorno social.

Tabla 5: **Efecto de la perseverancia de la red de pares: áreas del conocimiento**

	Perseverancia		2SLS (IV)	Mediana de la perseverancia	N
	>mediana	<mediana	Test de diferencia de coeficientes (p-valor)		
Ciencias Económicas y Adm.	0.667*** (0.229)	0.210 (0.168)	0,004	-0,020	652
Medicina	-0.137* (0.075)	-0.126 (0.135)	0,145	0,412	327
Ciencias de la Salud y básicas	0.229 (0.216)	0.139 (0.152)	0,165	0,027	403
Ciencias Sociales	-0.018 (0.175)	-0.008 (0.099)	0,902	-0,105	456
Jurisprudencia	0.559 (0.979)	0.798 (1.984)	0,639	0,234	487
Controles	✓				

Notas: Los resultados corresponden al efecto de la perseverancia de la red de pares de acuerdo a las áreas del conocimiento, dentro de las cuales se encuentran: ciencias económicas y administrativas, medicina, ciencias de la salud y ciencias básicas, ciencias sociales y jurisprudencia, donde cada fila representa una estimación. Para medir el efecto heterogéneo, se tomó como variable dummy la mediana de la perseverancia al interior de cada área, donde los valores por encima señalaron a los más perseverantes por cada área del conocimiento, y los valores por debajo de la mediana a los menos perseverantes. Con base en la ecuación 4.7, el coeficiente de interés por encima de la mediana de la perseverancia correspondió a $\sigma_1 + \sigma_2 + \sigma_3 + \sigma_4$, mientras que por debajo de la mediana de la perseverancia a $\sigma_1 + \sigma_2$. En cada estimación se tiene en cuenta todo el conjunto de controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios. Adicionalmente, se agrega un test de diferencias de coeficientes, el cual testea la hipótesis de nula si los coeficientes son iguales entre sí, y las hipótesis alterna de que los coeficientes son diferentes. Por último, por cada área del conocimiento fue calculada la mediana de la perseverancia de la red de pares.

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

De acuerdo a lo anterior, en este estudio encontramos que la red de pares por áreas de conocimiento muestra un comportamiento interesante en relación a la auto-selección. En la tabla 5, observamos que el área de ciencias económicas y administración presenta un nivel de perseverancia bajo, mientras que los estudiantes de medicina un nivel de perseverancia alto. En tal sentido, los estudiantes de ciencias económicas y administración se apoyan más en la perseverancia de la red de pares, dado

su nivel de perseverancia. Con ello, es posible argumentar que al presentar un nivel bajo de perseverancia, los pares posiblemente se auto seleccionan al apoyarse en la perseverancia de sus pares. En contraste, los estudiantes de medicina al presentar una perseverancia muy alta, no se apoyan en la perseverancia de la red de los pares, razón por la cual el estimado da con signo negativo. Además, no evidenciamos diferencias estadísticamente significativas entre los coeficientes estimados, salvo para el área de ciencias económicas y administración, resultado que muestra mayor robustez respecto a los demás. En relación a las demás áreas del conocimiento, no encontramos efectos estadísticamente significativos.

A la luz de la literatura, la auto-selección indica elegir compañeros con desempeños específicos, y dichos compañeros seleccionados pueden ejercer efectos más fuertes que otros. En este caso, destacamos la perseverancia del grupo par como complemento para mejorar el desempeño académico individual. Además, de acuerdo con el concepto de homofilia en las redes sociales, los pares tienden a auto-seleccionarse de acuerdo a preferencias similares, donde pueden existir ideas afines, tales como en las áreas del conocimiento expuestas, donde comparten el gusto por la misma carrera, y pueden desarrollar comportamientos parecidos en función de la carrera estudiada. En este caso, la perseverancia del grupo par aumenta el rendimiento de los estudiantes de las áreas de ciencias económicas y administración, mientras que para los estudiantes de medicina no es así, dado que dichos estudiantes presentan un nivel de perseverancia alto. Con estos resultados, damos a conocer evidencia de que la auto-selección de compañeros también puede darse para niveles diferentes de *Grit*.

5.2.4. Características de los colegios

Teniendo en cuenta las características de los colegios encontramos efectos de la perseverancia de la red de pares por las siguientes categorías: el sector (público y privado), y por la creencia (colegio religioso y no religioso), donde en esta última categoría estimamos el efecto por la creencia de los colegios privados.

En este contexto, es importante indagar sobre el rol de la producción educativa en las instituciones de educación, y el papel de los efectos de los compañeros en el aula de clases. En este sentido, de acuerdo con Lazear (2001) la educación en el aula tiene aspectos de bien público, donde juega un papel clave el tamaño de la clase. En este estudio se argumenta que el aprendizaje en el aula tiene efectos de congestión, definidas como externalidades negativas creadas cuando un estudiante impide el aprendizaje de todos los demás compañeros de clase. Ante esto, existen estudiantes con una probabilidad de ser aprendices no disruptivos, y aquellos que lo son. Esta diferenciación, marca un mejor rendimiento educativo en las clases, cuando en el salón hay más estudiantes no disruptivos. En esta línea, el autor argumenta que segregar a los estudiantes por capacidad académica maximiza

el rendimiento educativo total, es decir se maximiza emparejando en lugar de mezclar los tipos de estudiantes en las clases.

Además de estos aspectos, existen diferencias en la producción educativa de acuerdo a la creencia. [Coleman y Hoffer \(1987\)](#) afirmaron que el rendimiento es más alto en las escuelas privadas y en las escuelas religiosas, que en las escuelas públicas. La razón se da porque las escuelas privadas y religiosas funcionan como comunidades funcionales y de valores, que encajan con las creencias inculcadas en el hogar, lo cual es más acentuado en los colegios de tipo religioso. Siguiendo de nuevo a [Lazear \(2001\)](#) las escuelas católicas y privadas pueden seleccionar a los estudiantes más atentos, y expulsar a aquellos más disruptivos, dejando una población de estudiantes que son más fáciles de enseñar que los de las escuelas públicas. Adicionalmente, el autor afirma que lo que caracteriza a una escuela religiosa es la estricta disciplina, lo cual puede atribuir el éxito en los programas educativos. Además, en esta línea [Evans y Schwab \(1995\)](#) argumentaron que los adolescentes matriculados en escuelas religiosas tienen una probabilidad significativamente más alta de completar la escuela secundaria y comenzar la universidad. De igual forma, [Hoffer et al. \(1985\)](#) encuentran que las escuelas religiosas producen estudiantes de mayor rendimiento, dado que exigen más tareas y cursos avanzados. Otros estudios de [Thapa \(2015\)](#) y [Opore \(1999\)](#), argumentan la asociación entre asistir a una escuela privada y tener mejores resultados de rendimiento, esto lo justifican porque las escuelas privadas están mejor equipadas, mejor administradas y más apoyadas por los padres, argumentos que van en línea por lo expuesto por [Coleman y Hoffer \(1987\)](#) sobre la formación de comunidades funcionales de las escuelas privadas y religiosas.

De acuerdo con esta literatura existen diferencias en la producción de la educación, en específico en los colegios privados y religiosos respecto a los colegios de carácter público. Las diferencias pueden darse en que los colegios privados y religiosos pueden seleccionar y admitir a los estudiantes menos disruptivos y presentar ciertos criterios de admisión, en comparación con los colegios públicos donde estos criterios de admisión no se dan. Además, el rol de la disciplina, la exigencia de más tareas y cursos avanzados, parece ser que marca un diferencial. Esta literatura va en línea con los efectos encontrados en la tabla 6, donde el efecto de la perseverancia de la red de pares es robusto para aquellos estudiantes provenientes de colegios privados, y para aquellos de colegios religiosos y a su vez privados. Con estos resultados mostramos diferencias entre los colegios privados, religiosos, y públicos, pero adicionalmente complementamos esta literatura con el nivel de perseverancia de los estudiantes. En este sentido, si bien es cierto que los colegios privados y religiosos producen estudiantes de mayor rendimiento académico, dado el nivel de disciplina y exigencia, también es cierto que el factor de la perseverancia de los pares refuerza un mejor rendimiento académico.

Tabla 6: **Efecto de la perseverancia de la red de pares: características de los colegios**

	2SLS (IV)			
	Sector		Creencia: colegio privado	
	Público	Privado	No religioso	Religioso
	(1)	(2)	(3)	(4)
Efecto de la perseverancia de la red de pares	0.018 (0.120)	0.119** (0.055)	0.084 (0.065)	0.171** (0.083)
Número de Observaciones	2,325		2,325	
Controles	✓		✓	

Notas: Los resultados corresponden al efecto de la perseverancia de la red de pares de acuerdo a las características de los colegios, entre las cuales se encuentran: el sector y la creencia de los colegios privados. Para medir el efecto heterogéneo, se realizó la interacción entre la variable endógena y la característica del colegio. Con base en la ecuación 4.7, basta solo con tomar las dobles interacciones, para el panel A el coeficiente de interés para los colegios privados es $\sigma_1 + \sigma_2$, mientras que para los colegios públicos es σ_1 . Para el panel B, el coeficiente de interés para los colegios privados-religiosos es $\sigma_1 + \sigma_2$, mientras que para los colegios privados-no religiosos es σ_1 . En cada estimación se tiene en cuenta todo el conjunto de controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios.

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5.3. Mecanismos de mediación: comportamientos académicos

Las calificaciones de los estudiantes son el reflejo del rendimiento académico dentro de las distintas clases que toman, estas pueden medir los conocimientos adquiridos y las habilidades académicas como el pensamiento crítico, trabajo en equipo, iniciativa, comunicación, entre otras. Conley (2007) y Farkas (2003) argumentan que las calificaciones también reflejan comportamientos y estrategias académicas que son clave para obtener un buen rendimiento, por ejemplo técnicas de estudio, gestión del tiempo, persistencia, hábitos de trabajo, resolución de problemas, y la capacidad de utilizar grupos de estudio.

De acuerdo con Farrington et al. (2012) existen factores no cognitivos que son clave para el éxito a largo plazo de los estudiantes que se asocian directamente con el rendimiento en clase. Entre estos se encuentran los comportamientos académicos, los cuales se relacionan con ser un buen estudiante, por ejemplo asistir regularmente a clase, prestar atención, llegar preparado a la clase y dedicar tiempo extra a estudiar. El autor argumenta que los comportamientos académicos se comportan como los factores no cognitivos más próximos al rendimiento académico de los estudiantes, y los otros factores no cognitivos tales como las habilidades académicas, las estrategias de aprendizaje, habilidades blandas y habilidades sociales, ejercen su efecto a través de los comportamientos académicos de los estudiantes. Con base en lo anterior, estudiamos el efecto de la perseverancia de la red sobre los comportamientos académicos y la relación con el rendimiento académico, esto quiere decir que la perseverancia de la red de los pares debe actuar a través de los comportamientos académicos y esto afecta el rendimiento académico.

Para comprobar este mecanismo tomamos como variable de interés los comportamientos académicos registrados en la encuesta de hábitos y redes utilizada en este estudio, tomando las horas de estudio, horas de permanencia en la universidad y la frecuencia de preparación de exámenes. Los resultados los registramos en la tabla 7 donde observamos que para los tres casos la perseverancia del grupo par aumenta los comportamientos académicos, y dichos comportamientos se asocian al buen rendimiento académico. Adicionalmente realizamos el mismo ejercicio con aquellos estudiantes con algún tipo de beca, los resultados se pueden visualizar en la tabla 6 del anexo D. Para este tipo de estudiantes, el efecto de la perseverancia de la red se muestra robusto para las horas de permanencia en la universidad y la preparación de las exámenes, mientras que no encontramos efecto para las horas de estudio.

Tabla 7: **Comportamientos académicos**

	2SLS (IV)		
	Horas de estudio (1)	Horas U (2)	Preparación exámenes (3)
Efecto de la perseverancia de la red de pares	0.104* (0.059)	0.196*** (0.066)	0.090*** (0.031)
Número de Observaciones	2,272	2,292	2,320
Controles	✓	✓	✓

Notas: Resultados generales del modelo de variables instrumentales para medir el efecto del grupo par con base en la perseverancia de la red, tomando como resultado variables relacionadas con los comportamientos académicos de los estudiantes, entre estas: horas de estudio, horas de permanencia en la universidad y preparación de exámenes. La fila reporta el estimado para cada uno de los comportamientos académicos nombrados. Para una mejor interpretación del estimado principal, las horas de estudio y horas de permanencia en la universidad se tomaron de manera estandarizada, mientras que el outcome de preparación de exámenes fue transformado en una variable dummy que indica 1 para las dos categorías más altas de la variable inicial (5=Durante todo el semestre, 4=Con una o más semanas de antelación), y 0 para las demás (3=Durante la semana anterior al parcial, 2= Un día, 1=No estudia para los exámenes). En cada estimación se tiene en cuenta todo el conjunto de controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios.

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Para complementar este análisis, utilizamos los efectos de mediación causal donde un tratamiento (D) y un resultado (M) al cual llamamos mediador causan conjuntamente un resultado de interés (Y), siguiendo la metodología propuesta por Imai et al. (2010), Pinto et al. (2019), Dippel et al. (2020) (ver anexo D - tabla 5). El análisis de mediación permite usar un solo instrumento para estimar el efecto causal por medio de una variable intermedia en el resultado final. Para ello, se descompone el efecto total sobre (Y) en un efecto indirecto mediado por (D) sobre el outcome (Y) que opera a través de (M), y un efecto directo que no opera a través de (M). En nuestro caso, esto implica que la perseverancia de la red (D) afecta el rendimiento académico (Y) indirectamente a través de los comportamientos académicos (M) de los estudiantes. Una vez estimado el efecto total, directo e indirecto, el efecto de mediación como porcentaje del efecto total es la relación entre el

efecto indirecto dividido el efecto total. Los resultados los reportamos en la tabla 5 del anexo D sin incluir la frecuencia de preparación de exámenes, debido a que es una variable dummy y los resultados no se pueden interpretar de forma clara para este caso. Con base en ello, observamos que las horas de estudio explican el 66.56 % del efecto total, mientras que las horas de permanencia en la universidad explican el 88.95 %. Con esto mostramos evidencia de cómo la perseverancia de la red actúa por medio de los comportamientos académicos, que a su vez tienen una relación con el rendimiento académico.

5.4. Robustez

5.5. Resultados sin outliers

En primer lugar, como ejercicio de robustez en la tabla 8 presentamos los resultados de la estimación principal sin tener en cuenta el sobre reporte de amigos, por el cual utilizamos los percentiles del total de compañeros reportados. En tal sentido, en el percentil 99 de la muestra los estudiantes reportaron más de 30 pares de compañeros, mientras que en el percentil 95 más de 18 pares. Con base en estas referencias estimamos el efecto sin tener en cuenta el sobre reporte de 30 pares, y sin incluir el reporte de más de 18 compañeros.

Tabla 8: Modelo sin el sobre reporte de pares

	2SLS (IV)		
	Estimación Principal (1)	Percentil 99 % >30 pares (2)	Percentil 95 % >18 pares (3)
2da Etapa: Perseverancia de la red de pares	0.110** (0.052)	0.112** (0.052)	0.107** (0.053)
1era Etapa: Perseverancia de la red de pares de pares	0.297*** (0.031)	0.296*** (0.031)	0.295*** (0.031)
Número de Observaciones	2,325	2,297	2,204
Anderson-Rubin F estadístico (1949)	18,939	19,293	18,036
Cragg-Donald Wald F estadístico (1993) - Stock-Yogo (2005)	244,453	240,078	227,298
Montiel-Pflueger F estadístico efectivo (2013)	92,086	91,273	88,624
Controles	✓	✓	✓

Notas: Resultados generales del modelo de variables instrumentales para medir el efecto del grupo par sobre el rendimiento académico con base en la perseverancia de la red, sin tener en cuenta el sobre reporte de pares. La primera columna toma el estimado de la regresión principal, en la segunda columna se encuentra el estimado sin tener en cuenta aquellos individuos que reportaron más de 30 pares, y en la tercera columna se encuentra el estimado sin tener en cuenta el sobre reporte de más de 18 pares. La primera fila corresponde a los estimados de la segunda etapa, y en la segunda fila se encuentran los estimados de la primera etapa. En ambas etapas se agrega todo el conjunto de controles: características personales de los individuos, medidas cognitivas, rendimiento académico previo, y otras medidas de la calidad del colegio. Por último, se incluyen los estadísticos F de las siguientes pruebas de relevancia del instrumento: Anderson & Rubin (1949), Cragg-Donald (1993) - Stock-Yogo (2005), y Montiel & Pflueger (2013).

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

En la columna 1 reportamos la estimación principal, mientras que en la columna 2 el efecto de la perseverancia de la red de pares sin tener en cuenta el sobre reporte de más de 30 compañeros, y en la columna 3 el estimado sin el sobre reporte de más de 18 pares. Siguiendo esta lógica, el coeficiente de interés no muestra una variación significativa en términos de magnitud, se mantiene la

significancia estadística, los errores estándar se mostraron constantes, y los resultados de la primera etapa se mostraron robustos con la condición de relevancia, por lo cual concluimos que el efecto del grupo par no está determinado por el sobre reporte de compañeros. De manera adicional, estudiamos la posibilidad de realizar la estimación con base en la red bilateral, sin embargo no encontramos resultados significativos al condicionar la regresión a las conexiones bilaterales.

5.5.1. Red en tercer grado

En segundo lugar, como ejercicio de robustez tomamos como instrumento alternativo la red de pares en tercer grado, con el fin de identificar si dicha red es robusta con el efecto encontrado en la red en segundo grado. Los resultados los encontramos en la tabla 9, la primera fila corresponde a la segunda etapa mientras que la segunda fila a la primera etapa, en cada estimación incluimos todo el conjunto de covariables utilizado.

Tabla 9: **Modelo principal: Red en tercer grado**

	OLS (1)	Forma reducida (IV) (2)	2SLS (IV) (3)
2da Etapa: Perseverancia de la red de pares	0.027 (0.016)		0.053* (0.029)
1era Etapa: Perseverancia de la red de pares de pares de pares		0.031* (0.017)	0.585*** (0.018)
Número de Observaciones	2,311	2,311	2,311
Anderson & Rubin F estadístico (1949)			16,53
Cragg-Donald Wald F estadístico (1993) - Stock-Yogo (2005)			1.335,35
Montiel & Pflueger F estadístico efectivo (2013)			1.014,60
Características personales de los individuos	✓	✓	✓
Medidas Cognitivas	✓	✓	✓
Rendimiento académico previo	✓	✓	✓
Rendimiento académico del colegio	✓	✓	✓
Otras medidas de la calidad del colegio	✓	✓	✓

Notas: Resultados generales del modelo de variables instrumentales para medir el efecto del grupo par sobre el rendimiento académico con base en la perseverancia de la red, tomando como instrumento la red de pares en tercer grado. La primera fila reporta el estimado de la segunda etapa, y la segunda fila los resultados de la primera etapa del modelo. La columna (1) reporta el estimado utilizando OLS sin tener en cuenta la medida exógena del instrumento, la segunda columna la relación de la variable resultado (promedio del periodo estandarizado) con el instrumento, y la columna (3) reporta por un lado el efecto de la perseverancia de la red de pares sobre el rendimiento académico, y por el otro lado la relación entre la variable endógena y el instrumento. En cada estimación se tiene en cuenta todo el conjunto de controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo del individuo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios. Por último, se incluye en la columna (3) los estadísticos F de las siguientes pruebas de relevancia del instrumento: Anderson & Rubin (1949), Cragg-Donald (1993) - Stock-Yogo (2005), y Montiel & Pflueger (2013).

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Con base en la primera etapa del modelo (*ver fila 2 - modelo 3*), observamos correlación entre el instrumento y la variable endógena a un nivel de confianza estadística del 99%, corroborando la relevancia del instrumento de acuerdo a los estadísticos F reportados. Respecto a la segunda etapa, encontramos un resultado positivo y significativo del efecto de la perseverancia de la red sobre el rendimiento académico a un nivel de confianza estadística del 90% (*ver fila 1 - modelo*

3). Sin embargo, el efecto de la perseverancia presenta una disminución considerada, y en términos de significancia estadística no se muestra muy robusto. Estos resultados también fueron estimados incluyendo una a una el conjunto de variables controles (*ver anexo D - Tabla D4*). De acuerdo a estos resultados en forma extendida, el estimador por 2SLS presenta variaciones tanto en magnitud como en significancia estadística a medida que se incluyen las variables controles, en especial cuando agregamos el rendimiento académico y la calidad de los colegios. De igual forma los errores estándar presentan variaciones pero no considerables, mientras que los resultados de la primera etapa se muestran robustos con la condición de relevancia y con los estadísticos F calculados.

De acuerdo con estos resultados, el resultado principal no se muestra robusto al utilizar el instrumento en tercer grado, lo cual pone sobre la mesa si en realidad el efecto del grupo par es posible mantenerlo al tomar una red lejana, y no tan similar. De acuerdo con [Ma et al. \(2015\)](#) las personas con características similares son más propensas a formar lazos, concepto denominado como homofilia, esto significa que los individuos que tienen vínculos directos tienden a tener rasgos similares. Con base en la literatura de redes, [Khanam et al. \(2020\)](#) argumentan que la homofilia puede verse como un concepto social donde las redes de los individuos tienden a ser más homogéneas que heterogéneas, de modo que la interacción con pares similares ocurre con más frecuencia que con pares diferentes. Por su parte, [Bramoullé et al. \(2012\)](#) discuten que la homofilia también afecta el comportamiento y oportunidades, lo cual impacta en el bienestar de los individuos conectados en una red social. Además, [Easley y Kleinberg \(2010\)](#), exponen que la homofilia proporciona una primera ilustración de cómo los contextos que rodean una red pueden impulsar la formación de sus vínculos. Con base en lo anterior, los entornos compartidos juegan un papel importante teniendo en cuenta que son espacios que con frecuencia son más homogéneos, lo cual puede generar patrones de homofilia. En tal sentido, los individuos conectados por un camino corto en la red o que comparten un enfoque social, son más similares que aquellos que son distantes, lo cual lleva a que las personas similares prefieren convertirse en amigos [Kossinets y Watts \(2009\)](#).

La homofilia en el contexto de estudiantes universitarios puede darse por diferentes características, [Mayer y Puller \(2008\)](#) encuentran que la raza está fuertemente relacionada con los lazos sociales. De acuerdo con [Godley \(2008\)](#), las amistades parecen estar determinadas por las preferencias de género, además las actividades conjuntas se comportan como predictoras significativas de la amistad, y la especialización académica compartida toma un rol importante. Por su parte, [Pilbeam y Denyer \(2009\)](#) argumentan que los lazos por la amistad se forman por modalidad o programa, teniendo en cuenta el año de ingreso a la universidad, y pueden ser vistos como predictores de la formación de vínculos entre estudiantes.

Respecto al rendimiento académico, [Smirnov y Thurner \(2017\)](#) encuentran que la homofilia académica es el resultado de la selección, es decir los individuos prefieren reorganizar gradualmente sus redes sociales de acuerdo con sus niveles de rendimiento, en lugar de adaptar su rendimiento al nivel de su grupo local, este hallazgo es posible verlo en redes dinámicas, donde se pueda evidenciar la evolución de la red a lo largo del tiempo. Además, [Santavirta y Sarzosa \(2019\)](#) reconocen que las redes en las aulas de clase son el resultado de elecciones activas de los estudiantes, por lo tanto son endógenas y la formación de vínculos se basa en la homofilia. Estos autores argumentan que la homofilia puede ser medida teniendo en cuenta la distancia entre dos nodos, con base en las características que los componen. El agente i se hace amigo de j si hay un valor positivo de hacerlo, entonces la homofilia implica que las diadas en las que la distancia es baja tienen un mayor excedente. Por lo tanto, el tamaño del efecto del grupo par depende potencialmente de la longitud del camino entre los dos agentes, es decir de lo cerca que estén socialmente.

Con base en esta literatura, podemos sustentar la posible homofilia que se presenta en la red analizada. En tal sentido, al presentar intereses y espacios compartidos como la misma facultad y programa académico en pro de mejorar el rendimiento académico, posiblemente la red en segundo grado tiene un comportamiento más homogéneo en función de la perseverancia de la red de pares en primer grado, mientras que la red en tercer grado se aleja de la medida de perseverancia del grupo par, a pesar de que los individuos mantienen un espacio compartido y homogéneo. Estos resultados indican que el efecto de la perseverancia de la red en tercer grado, no es robusto dado la posible heterogeneidad de los pares. Esto es una contribución interesante, debido a que la homofilia también puede presentarse de acuerdo al nivel de *Grit* del grupo par, y es importante para mantener el efecto del grupo. Por lo tanto, ser perseverante conduce a un buen rendimiento académico y además las habilidades blandas de los compañeros influyen en este, por lo que posiblemente la red en tercer grado al ser lejana, la perseverancia del grupo par a dicho nivel no es suficiente para mantener el rendimiento académico. Sin embargo, una posible limitación para estudiar en más detalle la homofilia es que la red presentada es estática, y no es posible ver su evolución a través del tiempo.

En este orden de ideas, la red en segundo y tercer grado son instrumentos válidos puesto que cumplen con la condición de relevancia. Sin embargo, los posibles patrones de homofilia de acuerdo a *Grit* influyen en el efecto encontrado, e informan sobre la naturaleza de los pares de auto-seleccionarse de acuerdo a su nivel de perseverancia, donde el rol de la distancia entre los individuos dado su nivel de *Grit* revela un papel importante.

6. Discusión y Conclusiones

En el presente trabajo encontramos efectos positivos y significativos de la perseverancia de la red de los pares sobre el rendimiento académico, resultados que son consistentes con la relación entre el éxito académico y la escala de *Grit-S* propuesta por Duckworth y Quinn (2009) y con la influencia de los pares Bramoullé et al. (2009).

Los resultados principales informan que por un aumento en una desviación estándar en la perseverancia de la red de pares, el rendimiento académico aumenta en 0.11 desviaciones estándar. Este resultado es robusto con los efectos encontrados por diferentes estimadores de variables instrumentales en específico por *LIML*, *Fuller* y *Jackknife*. Además, el efecto del grupo par encontrado en la estimación principal no está determinado por el sobre reporte de amigos, ni por el número de conexiones presentadas en la cohorte 2018-1.

Adicionalmente evidenciamos diferencias del estimado entre género, habilidades socioemocionales, áreas de conocimiento y características de los colegios. Para las mujeres, el efecto de la perseverancia de la red de pares es mucho más robusto respecto a los hombres, lo cual es explicado a la luz de la literatura por las habilidades sociales que presentan las mujeres. Esto implica teóricamente que las mujeres aprovechan más el *Grit* de su red de compañeras. Además, las mujeres se muestran más centrales en la red, lo cual indica que son un centro de actividad importante entre los pares de amigos. Para las habilidades socioemocionales, entre estos el *Big 5* y otros factores no cognitivos como la maleabilidad de la inteligencia y el autocontrol, se correlacionan fuertemente con la perseverancia de la red de los pares, lo cual es explicado por la complementariedad que puede haber entre dichas habilidades. Para las áreas del conocimiento, encontramos patrones de auto-selección, de acuerdo al nivel de *Grit* de cada una. En cuanto a las características de los colegios, a la luz de la literatura los colegios privados y religiosos muestran mejores resultados en cuanto al rendimiento académico de los estudiantes. Esto implica que los colegios privados y religiosos aprovechan mejor el *Grit* y su relación con el rendimiento académico.

De acuerdo con la restricción de exclusión y el instrumento propuesto por Bramoullé et al. (2009), para esta investigación la perseverancia de los pares de pares del estudiante que no son pares directos del estudiante afecta el rendimiento académico solo a través de la perseverancia de los pares. No obstante, el instrumento propuesto puede tener una limitación por posibles factores de confusión, los cuales se dan por los efectos directos del estudiante con la red de pares de pares que no se están observando, por lo cual se puede llegar a pensar en posibles interacciones en clases o actividades dentro de la universidad. Para tratar esto y dada la cercanía de las redes en primer y segundo grado, utilizamos como instrumento alternativo la red en tercer grado. Como resultado de este ejercicio,

evidenciamos posibles patrones de homofilia en las conexiones, dado que los estudiantes se ven afectados de manera más directa por sus compañeros más cercanos. Por lo tanto, la disminución del coeficiente de interés con base en la red de tercer grado sugiere que a medida que aumenta la distancia entre las conexiones, la perseverancia de la red de los pares es muy lejana, por lo que el efecto del grupo par disminuye. Con base en ello, ambos instrumentos son válidos de acuerdo a la condición de relevancia y se muestran robustos con los test aplicados, la diferencia entre ambos posiblemente se debe a las variaciones del nivel *Grit* del grupo par, lo cual se ve reflejado en la magnitud del coeficiente de interés.

Para complementar este análisis, utilizamos los comportamientos académicos de los estudiantes para justificar que la perseverancia de la red actúa a través de ellos, y esto a su vez impacta de forma positiva el rendimiento académico. Esto nos sugiere que el rendimiento académico no solamente depende de las habilidades académicas, también depende de mediadores académicos como las horas de estudio, las horas de permanencia en la universidad, y la preparación de los exámenes para mantener un buen rendimiento académico.

Es importante mencionar que existen limitaciones a la hora de resolver la endogeneidad. En primer lugar, los individuos de la red escogen a sus propios amigos con quienes estudian, de manera que no se puede controlar dicha elección. Y en segundo lugar, el problema de variables no observables dado el entorno compartido en relación a las clases, actividades y mismos intereses de acuerdo a cada carrera. Estas dos limitaciones hacen que no se cumpla en su totalidad la restricción de exclusión, de manera que abordamos parcialmente el problema de reflexión, por lo que recomendamos realizar análisis de sensibilidad.

En general estos resultados muestran que la perseverancia del grupo par se comporta como un buen determinante del rendimiento académico, lo cual es compatible con la literatura encontrada. Además, es importante reconocer el rol de la influencia de los compañeros, dado el impacto que puede tener en los estudiantes hacia el éxito académico, dado que los pares académicos pueden actuar como modelos a seguir positivos. En términos de la educación superior esta investigación invita a las universidades a desarrollar programas de acompañamiento a los estudiantes, del cual se puedan identificar estudiantes vulnerables, como con niveles bajos de *Grit*, autocontrol, responsabilidad, entre otros. Además, promover espacios de aprendizaje como las tutorías ayuda a la interacción constante con los pares, lo cual puede ser clave para el desarrollo de las habilidades socioemocionales. En todo caso, es necesaria una muestra más amplia donde se puedan estudiar otro tipo de redes de estudio que no están siendo observadas, y de esta forma analizar de manera más profunda los resultados.

Referencias

- Alhadabi, A. y Karpinski, A. C. (2020). Grit, self-efficacy, achievement orientation goals, and academic performance in university students. *International Journal of Adolescence and Youth*, 25(1):519–535.
- Alzahrani, M., Alharbi, M., y Alodwani, A. (2019). The effect of social-emotional competence on children academic achievement and behavioral development. *International Education Studies*, 12(12):141–149.
- Anderson, T. W. y Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *The annals of Mathematical Statistics*, 20(1):46–63.
- Andrews, D. W. y Stock, J. H. (2005). Inference with weak instruments. *Cowles Foundation Discussion Papers*.
- Angrist, J. D., Imbens, G. W., y Krueger, A. B. (1999). Jackknife instrumental variables estimation. *Journal of Applied Econometrics*, 14(1):57–67.
- Beaman, R., Wheldall, K., y Kemp, C. (2007). Differential teacher attention to boys and girls in the classroom. *Educational Review*, 58(3):339—366.
- Benedict, R. H., Amato, M. P., Boringa, J., Brochet, B., Foley, F., Fredrikson, S., Hamalainen, P., Hartung, H., Krupp, L., Penner, I., Reder, A. T., y Langdon, D. (2012). Brief international cognitive assessment for ms (bicams): international standards for validation. *BMC neurology*, 12(1):55.
- Benet-Martinez, V. y John, O. P. (1998). Los cinco grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the big five in spanish and english. *Journal Personality and social psychology*, 75(3):729–750.
- Bennett, D., McCarty, C., y Carter, S. (2020). The relationship between grit and academic performance in the classroom. *Journal for Economic Educators*, 20(2).
- Bernard, M. E. (2006). It's time we teach social-emotional competence as well as we teach academic competence. *Reading & Writing Quarterly*, 22(2):103–119.
- Bertrand, M., Luttmer, E., y Mullainathan, S. (2000). Network effects and welfare cultures. *The Quarterly Journal of Economics*, 115(3):1019–1055.
- Bertrand, M. y Pan, J. (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics*, 5(1):32–64.

- Bramoullé, Y., Currarini, S., Jackson, M. O., Pin, P., y Rogers, B. W. (2012). Homophily and long-run integration in social networks. *Journal of Economic Theory*, 147(5):1754–1786.
- Bramoullé, Y., Djebbari, H., y Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55.
- Bramoullé, Y., Djebbari, H., y Fortin, B. (2020). Peer effects in networks: A survey. *Annual Review of Economics*, 12:603–629.
- Brophy, J. (1985). Interactions of male and female students with male and female teachers in gender influences in classroom interaction, edited by Louise C. Wilkinson and Cora B. Marrett. *New York: Academic Press*, pages 115–42.
- Calvó-Armengol, A., Patacchini, E., y Zenou, Y. (2009). Peer effects and social networks in education. *The Review of Economic Studies*, 76(4):1239—1267.
- Carvalho, R. G. G. (2016). Gender differences in academic achievement: The mediating role of personality. *Personality and Individual Differences*, 94:54–58.
- Chiu, S. (2001). *Exploring kindergartners' social and cognitive competence: An application of ECLS-K*. Dissertation Abstracts International, 62(12), 4051. (University Microfilms No. AAT 3035879).
- Cicala, S., Fryer, R. G., y Spenkuch, J. L. (2017). Self-selection and comparative advantage in social interactions. *Journal of the European Economic Association*, 16(4):983—1020.
- Coleman, J. S. y Hoffer, T. B. (1987). *Public and Private High Schools: The Impact of Communities*.
- Conley, D. T. (2007). Toward a more comprehensive conception of college readiness. *Eugene, OR: Educational Policy Improvement Center*.
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., y Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12:769—786.
- Cragg, J. G. y Donald, S. G. (1993). Testing identifiability and specification in instrumental variables models. *Econometric Theory*, 9(2):222–240.
- Cronbach, L. J. y Meehl, P. E. (1955). Construct validity in psychological tests. *Psychol Bull*, 52(4):281–302.
- Dasgupta, U., Mani, S., Sharma, S., y Singhal, S. (2020). Effects of peers and rank on cognition, preferences, and personality. *The Review of Economics and Statistics*, 104(3):587—601.

- DeLay, D., Zhang, L., Hanish, L. D., Miller, C. F., Fabes, R. A., Martin, C. L., Kochel, K. P., y Updegraff, K. A. (2016). Peer influence on academic performance: A social network analysis of social-emotional intervention effects. *Prevention Science*, 17(8):903–913.
- Dippel, C., Ferrara, A., y Heblich, S. (2020). ivmediate: Causal mediation analysis in instrumental variables regressions. *The Stata Journal*, 20.
- Duckworth, A. L. y Gross, J. J. (2014). Self-control and grit: Related but separable determinants of success. *Curr Dir Psychol Sci*, 23:319—325.
- Duckworth, A. L., Peterson, C., Matthews, M. D., y Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6):1087—1101.
- Duckworth, A. L. y Quinn, P. D. (2009). Development and validation of the short grit scale (grit-s). *Journal of Personality Assessment*, 91(2):166–174.
- Duckworth, A. L. y Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, 98(1):198—208.
- Duckworth, A. L., Shulman, E. P., Mastrorarde, A. J., Patrick, S. D., Zhang, J., y Druckmana, J. (2015). Will not want: Self-control rather than motivation explains the female advantage in report card grades. *Learning and Individual Differences*, 39:13—23.
- Duckworth, A. L., Taxer, J. L., Eskreis-Winkler, L., Galla, B. M., y Gross, J. J. (2019). Self-control and academic achievement. *Annual Review of Psychology*, 70:373—399.
- Duflo, E. y Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics*, 118(3):815–842.
- Dweck, C. S. (1999). *Self-Theories: Their Role in Motivation, Personality, and Development. Essays in Social Psychology*.
- Dweck, C. S. (2006). *Mindset: The New Psychology of Success. New York, NY: Random House Incorporated*.
- Dweck, C. S. y Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2):256—273.
- Easley, D. y Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world. Cambridge University Press*.

- Entwisle, D. R., Alexander, K. L., y Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *Sociology of Education*, 80(1):114—138.
- Evans, W. N. y Schwab, R. M. (1995). Finishing high school and starting college: Do catholic schools make a difference? *The Quarterly Journal of Economics*, 110(4):941–974.
- Farkas, G. (2003). Cognitive skills and noncognitive traits and behaviors in stratification processes. *Annual Review of Sociology*, 29:541–562.
- Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., y Beechum, N. O. (2012). Teaching adolescents to become learners: The role of noncognitive factors in shaping school performance—a critical literature review. *International Education Studies*.
- Fuller, W. A. (1977). Some properties of a modification of the limited information estimator. *Econometrica*, 45(4):939–953.
- Giorgi, G. D., Frederiksen, A., y Pistaferri, L. (2020). Consumption network effects. *Review of Economic Studies*, 87(1):130–163.
- Giorgi, G. D., Pellizzari, M., y Redaelli, S. (2010). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, 2(2):241–275.
- Glaeser, E. L., Sacerdote, B., y Scheinkman, J. A. (1996). Crime and social interactions. *The Quarterly Journal of Economics*, 111(2):507—548.
- Godley, J. (2008). Preference or propinquity? the relative contribution of selection and opportunity to friendship homophily in college. *Department of Sociology, University of Calgary, Alberta, Canada*, 1(1):65–80.
- Golsteyn, B. H. H., Non, A., y Zölitz, U. (2021). The impact of peer personality on academic achievement. *Journal of Political Economy*, 129(4):1052–1099.
- Gresham, F. M. y Elliott, S. N. (1989). Social skills deficits as a primary learning disability. *Journal of Learning Disabilities*, 22(2):120–124.
- Hakimi, S., Hejazi, E., y Lavasani, M. G. (2011). The relationships between personality traits and students' academic achievement. *Procedia - Social and Behavioral Sciences*, 29:836–845.
- Hoffer, T., Greeley, A. M., y Coleman, J. S. (1985). Achievement growth in public and catholic schools. *Sociology of Education*, 58(2):74–97.

- Icfes (2016). Guía metodológica edición no 05. Índice sintético de la calidad educativa.
- Imai, K., Keele, L., y Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4):309—334.
- Jiang, L., Zhang, S., Li, X., y Luo, F. (2021). How grit influences high school students' academic performance and the mediation effect of academic self-efficacy and cognitive learning strategies. *Current Psychology*.
- John, O. P., Naumann, L. P., y Soto, C. J. (2008). Paradigm shift to the integrative big-five trait taxonomy: History, measurement, and conceptual issues. *Handbook of personality: Theory and research*, 3(2):114–158.
- John, R., John, R., y ur Rehman Rao, Z. (2020). The big five personality traits and academic performance. *Procedia - Social and Behavioral Sciences*, 2(1):10—19.
- Kannangara, C. S., Allen, R. E., Waugh, G., Nahar, N., Khan, S. Z. N., Rogerson, S., y Carson, J. (2018). All that glitters is not grit: Three studies of grit in university students. *Frontiers in Psychology*, 9:1539.
- Karlen, Y., Suter, F., Hirt, C., y Merki, K. M. (2019). The role of implicit theories in students' grit, achievement goals, intrinsic and extrinsic motivation, and achievement in the context of a long-term challenging task. *Learning and Individual Differences*, 74:101757.
- Keenan, K. y Shaw, D. (1997). Developmental and social influences on young girls' early problem behavior. *Psychol Bull*, 121(1):95–113.
- Khanam, K. Z., Srivastava, G., y Mago, V. K. (2020). The homophily principle in social network analysis. *ArXiv*.
- Kiessling, L., Radbruch, J., y Schaube, S. (2019a). Determinants of peer selection. *University of Bonn and University of Mannheim, Germany*.
- Kiessling, L., Radbruch, J., y Schaube, S. (2019b). Self-selection of peers and performance. *University of Bonn and University of Mannheim, Germany*.
- Knickmeyer, R., Baron-Cohen, S., Raggatt, P., y Taylor, K. (2005). Foetal testosterone, social relationships, and restricted interests in children. *Child Psychol Psychiatry*, 46(2):198–210.
- Kossinets, G. y Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, 115(2).

- Kotrotsiou, S., Barbouta, A., y Barbouta, C. (2020). Growth mindset and grit: How do university students' mindsets and grit affect their academic achievement? *International Journal of Caring Sciences*, 13:654.
- Lam, K. K. L. y Zhou, M. (2019). Examining the relationship between grit and academic achievement within k-12 and higher education: A systematic review. *Psychol Schs*, 56(10):1654–1686.
- Laschever, R. (2013). The doughboys network: Social interactions and the employment of world war i veterans. *Social Science Research Network*.
- Lazear, E. P. (2001). Educational production. *The Quarterly Journal of Economics*, 116(3):777–803.
- Lee, L.-F. (2007). Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics*, 140(2):333–374.
- Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics*, 28(4):825–860.
- Ma, L., Krishnan, R., y Montgomery, A. L. (2015). Latent homophily or social influence? an empirical analysis of purchase within a social network. 61(2):454–473.
- Maccoby, E. E. (1990). Gender and relationships: A developmental account. *American Psychologist*, 45(4):513—520.
- Maccoby, E. E. y Jacklin, C. N. (1980). Sex differences in aggression: A rejoinder and reprise. *Child Development*, 51(4):964—980.
- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542.
- Mason, H. D. (2018). Grit and academic performance among first-year university students: A brief report. *Journal of Psychology in Africa*, 28(1):66–68.
- Matthews, J. S., Ponitz, C. C., y Morrison, J. F. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology*, 101(3):689—704.
- Mayer, A. y Puller, S. L. (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of Public Economics*, 92:329–347.
- McClure, I. (2003). *The Essential Difference: Men, Women and the Extreme Male Brain*. BMJ : British Medical Journal 327(7405), 57.

- Mihaela, P. L. (2015). Psychological factors of academic success. *Procedia - Social and Behavioral Sciences*, 180:1632—1637.
- Mitrofanu, N. y Iona, A. (2013). Predictors of academic performance. the relation between the big five factors and academic performance. *Procedia - Social and Behavioral Sciences*, 78:125—129.
- Mohoric, T. y Taksic, V. (2016). Emotional understanding as a predictor of socio-emotional functioning and school achievement in adolescence. *Psihologija*, 49(4):357—374.
- Montgomery, J. D. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *The American Economic Review*, 81(5):1408–1418.
- Montiel-Olea, J. L. y Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3):358–369.
- Myhill, D. (2002). Bad boys and good girls? patterns of interaction and response in whole class teaching. *British Educational Research Journal*, 28(3):114—138.
- Nicoletti, C., Salvanes, K. G., y Tominey, E. (2018). The family peer effect on mothers' labor supply. *American Economic Journal: Applied Economics*, 10(3):206–34.
- Nishikawa, K., Kusumi, T., y Shirakawa, T. (2022). The effect of two aspects of grit on developmental change in high school students' academic performance: Findings from a five-wave longitudinal study over the course of three years. 191.
- Nrumah, E., Nkrumah, P., y Frimpong, Y. (2021). The relationship between growth mindset, grit, and academic achievement: Does goal commitment matter? *EPRA International Journal of Research and Development*, 6.
- Opare, J. A. (1999). Academic achievement in private and public schools: Management makes the difference. *Journal of Educational Management*, 2:1–12.
- Oriol, X., Miranda, R., Oyanedel, J. C., y Torres, J. (2017). The role of self-control and grit in domains of school success in students of primary and secondary school. *Frontiers in Psychology*, 8:1716.
- Panayiotou, M., Humphrey, N., y Wigelsworth, M. (2019). An empirical basis for linking social and emotional learning to academic performance. 56:193–204.
- Patacchini, E., Rainone, E., y Zenou, Y. (2017). Heterogeneous peer effects in education. *Journal of Economic Behavior and Organization*, 134:190–227.

- Pate, A. N., Payakachat, N., Harrell, T. K., Pate, K. A., J, D., Caldwell, y Franks, A. M. (2017). Measurement of grit and correlation to student pharmacist academic performance. *American Journal of Pharmaceutical Education*, 81(6).
- Pilbeam, C. y Denyer, D. (2009). Lone scholar or community member? the role of student networks in doctoral education in a uk management school. *Studies in Higher Education*, 34(3):301–318.
- Pinto, R., Dippel, C., Gold, R. S., y Heblich, S. (2019). Mediation analysis in iv settings with a single instrument. *UCLA working paper*.
- Ready, D. D., LoGerfo, L. F., Burkam, D. T., y Lee, V. E. (2005). Explaining girls' advantage in kindergarten literacy learning: Do classroom behaviors make a difference? *The Elementary School Journal*, 106(1):21–38.
- Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly Journal of Economics*, 116(2):681—704.
- Salgado, E., Vargas-Trujillo, E., Schmutzler, J., y Wills-Herrera, E. (2016). Uso del inventario de los cinco grandes en una muestra colombiana. *Avances en Psicología Latinoamericana*, 34(2):365–382.
- Santavirta, T. y Sarzosa, M. (2019). Effects of disruptive peers in endogenous social networks. *Working Paper*.
- Shan, X. y Zölitz, U. (2022). Peers affect personality development. *Working paper Shan No. 9729 - Social Science Research Network*.
- Shipstead, Z., Lindsey, D. R., Marshall, R. L., y Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, 72:116–141.
- Smirnov, I. y Thurner, S. (2017). Formation of homophily in academic performance: students prefer to change their friends rather than performance. *PLoS ONE*, 12(8).
- Stock, J. H. y Yogo, M. (2005). Testing for weak instruments in linear iv regression. In D.W.K. Andrews and J.H. Stock, eds. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg.*, 5(1):80–108.
- Thapa, A. (2015). Public and private school performance in nepal: an analysis using the slc examination. *Education Economics*, 23(1):47–62.

- Tincani, M. M. (2018). Heterogeneous peer effects in the classroom. *R&R at Journal of Political Economy*.
- Tsukayama, E., Duckworth, A. L., , y Kim, B. (2013). Domain-specific impulsivity in school-age children. *Developmental Science*, 16(6):879—893.
- Verma, A. P. y Meiselman, A. Y. (2021). Disruptive interactions: Long-run peer effects of disciplinary schools. *Job Market paper*.
- Wechsler, D., Coalson, D., y Raiford, S. (2008). *Wechsler Adult Intelligence Scale: Fourth Edition. Technical and interpretative manual*. San Antonio, TX: NCS Pearson.
- Zhang, J., Kuusisto, E., y Tirri, K. (2017). How teachers' and students' mindsets in learning have been studied: Research findings on mindset and academic achievement. *Psychology*, 8:1363–1377.
- Zárate, R. A. (2021). Uncovering peer effects in social and academic skills. *Social Science Research Network*.

Anexos

A. Pruebas de habilidades socioemocionales

A.1. Big 5

Para este test, utilizamos la versión en español propuesta por Benet-Martinez y John (1998) y los ítems sugeridos para el caso colombiano por Salgado et al. (2016). Esta metodología consiste en un cuestionario que consta de 44 frases donde el individuo debe realizar una calificación de acuerdo a una escala Likert de 5 puntos, cuyo rango va desde “totalmente en desacuerdo” (1 punto) hasta “totalmente de acuerdo” (5 puntos), donde se puede identificar los 5 grandes rasgos de la personalidad: apertura, extroversión, neuroticismo, responsabilidad (meticulosidad), y amabilidad. Siguiendo la investigación de (John et al., 2008), los 5 grandes rasgos de la personalidad se definen de la siguiente manera: (1) *apertura a la experiencia* “describe la amplitud, profundidad, originalidad y complejidad de la vida mental y aquella derivada de la experiencia de una persona”; (2) *responsabilidad* “describe el control de los impulsos que es prescrito socialmente y facilita los comportamientos dirigidos al logro de tareas y metas, como pensar antes de actuar, postergar las recompensas, seguir normas y reglas, y planear, organizar y priorizar las tareas”; (3) *extroversión* “implica una aproximación energética hacia el mundo social y material e incluye rasgos como sociabilidad, actividad, asertividad y emocionalidad positiva”; (4) *amabilidad* “contrasta una orientación prosocial y comunal hacia los otros con el antagonismo, e incluye rasgos como altruismo, ternura, confianza y modestia”, y (5) *neuroticismo* “contrasta la estabilidad emocional y un ánimo estable con la emocionalidad negativa, tal como sentirse ansioso, nervioso, triste y tenso”.

Para esta investigación, utilizamos la técnica de análisis factorial exploratorio (EFA) con el fin de construir el conjunto de variables latentes o factores comunes que explican los ítems que componen el test del *Big 5*. Con base en esta técnica, los 5 grandes rasgos de la personalidad fueron construidos aplicado a los 44 ítems de la encuesta, y luego fueron estimados como variables latentes en un sistema de ecuaciones simultáneas, a partir de los ítems con mayores cargas para los primeros 5 factores identificados en el EFA.

A.2. Otros factores no cognitivos

En otros factores no cognitivos encontramos las siguientes habilidades: *perseverancia*, *maleabilidad de la inteligencia*, y *autocontrol*.

A.2.1. Perseverancia

Los ítems los tomamos del estudio realizado por Duckworth y Quinn (2009) donde se expone una escala denominada *Grit-Short* que representa la perseverancia del esfuerzo y consistencia de interés, la cual corresponde a un subconjunto de elementos del *Grit-Original* del estudio realizado por

Duckworth et al. (2007). De acuerdo a los autores, *Grit-S* puede interpretarse como una medida más eficiente de la perseverancia y la pasión por los objetivos a largo plazo a nivel de rasgos. Esta se mide a través de un cuestionario que consta de 5 afirmaciones donde el individuo debe realizar una calificación de acuerdo a una escala Likert de 5 puntos, cuyo rango va desde “para nada cierto” (1 punto) hasta “completamente cierto” (5 puntos). El puntaje se obtiene promediando las 5 respuestas. La puntuación máxima es 5 (alta perseverancia) y el mínimo es 1 (baja perseverancia).

A.2.2. Maleabilidad de la inteligencia

Para medir la maleabilidad de la inteligencia, utilizamos la escala propuesta por Dweck (1999), por el cual se puede intuir que la inteligencia es un rasgo que no puede mejorarse denominado *Fix-Mindset*, o es un atributo que se puede desarrollar a través del esfuerzo y la experiencia, conocido como *Growth mindset*. Esta se mide a través de un cuestionario que consta de 3 afirmaciones donde el individuo debe realizar una calificación de acuerdo a una escala Likert de 6 puntos, cuyo rango va desde “totalmente en desacuerdo” (1 punto) hasta “completamente de acuerdo” (6 puntos). El puntaje se obtiene promediando las 6 respuestas. La puntuación máxima es 6 (alta mentalidad de crecimiento) y el mínimo es 1 (baja mentalidad de crecimiento).

A.2.3. Autocontrol

Siguiendo a Tsukayama et al. (2013), la impulsividad se puede distinguir en diferentes dominios, por un lado se encuentra la impulsividad interpersonal, y por el otro lado la impulsividad con el trabajo escolar. La primera hace referencia en controlar el temperamento, ser educado y entablar una conversación sin interrumpir al otro. La segunda se relaciona con un ambiente académico y consiste en tener disciplina, preparar las clases, seguir instrucciones y realizar los trabajos asignados. Estas se miden a través de un cuestionario que consta de 8 afirmaciones donde el individuo debe realizar una calificación de acuerdo a una escala Likert de 5 puntos, cuyo rango va desde “para nada cierto” (1 punto) hasta “completamente cierto” (5 puntos). Para medir el *autocontrol al trabajar*, el puntaje se obtiene promediando las 4 primeras respuestas. La puntuación máxima es 5 (alto autocontrol al trabajar) y el mínimo es 1 (bajo autocontrol al trabajar). Mientras que para medir *autocontrol interpersonal*, el puntaje se obtiene promediando las 4 últimas respuestas. La puntuación máxima es 5 (alto autocontrol interpersonal) y el mínimo es 1 (bajo autocontrol interpersonal).

B. Pruebas de habilidades cognitivas

Las pruebas de habilidades cognitivas son aplicadas de la siguiente manera: un test de *velocidad de procesamiento* (*processing speed index*), compuesta por dos pruebas que son *sustitución de símbolos* (*digit symbol substitution*) y *búsqueda de símbolos* (*symbol search*), y un test de *memoria de trabajo* (*working memory*), donde se aplica la prueba de *intervalo de simetría* (*symmetry span*). De acuerdo a estos test, se desglosan los siguientes componentes cognitivos, como se muestra a continuación.

B.1. Velocidad de procesamiento

B.1.1. Sustitución de símbolos

Este test está basado en la metodología propuesta por [Wechsler et al. \(2008\)](#) y con ella se busca medir la velocidad de procesamiento cognitivo a través de la codificación de símbolos, tal y como se muestra en la figura 1. En este ejercicio, el cuadro superior asocia un dígito a cada símbolo y el objetivo consiste en codificar la mayor cantidad de símbolos en el cuadro inferior. Esta prueba tiene una duración de dos minutos.

Figura 1: Sustitución de símbolos (Symbol Substitution)

≥	±	«	π	ж	ψ	Δ	ο	↑
1	2	3	4	5	6	7	8	9

ψ	±	π	ψ	±	ο	≥	Δ	↑	ж	±	«	±	≥	Δ
6	2	4												
ж	Δ	↑	ο	π	«	Δ	↑	ж	±	«	«	«	ж	ψ
ο	±	«	π	ж	ψ	≥	ο	±	≥	±	«	«	ψ	ο
≥	π	«	ψ	ж	±	Δ	ο	↑	ο	±	«	π	ж	«
±	±	«	π	ж	ψ	ο	±	ο	≥	±	«	π	ο	ψ
«	π	«	Δ	«	π	Δ	ο	↑	Δ	«	«	Δ	ж	ψ
≥	±	«	±	ж	«	±	ο	«	≥	±	±	π	Δ	ψ

Nota: Tomado de [Benedict et al. \(2012\)](#).

B.1.2. Búsqueda de símbolos

De igual forma, este test está basado en la metodología propuesta por [Wechsler et al. \(2008\)](#), con ello se busca medir la velocidad de procesamiento con base en estímulos visuales como se puede observar en la figura 2, donde el individuo debe escoger alguno de los 2 estímulos con alguno de los 5 estímulos visuales presentados. Si no existe ningún emparejamiento, los participantes deben escoger la opción "NO". Esta prueba tiene una duración de dos minutos.

Figura 2: Búsqueda de símbolos (Symbol Search)

↙		⊕	⊞	↖	↙	↙	No
↖	↙	⊞	⊗	↖	↙	↖	No
↖	⊕	⊕	⇒	⊕	⊗	⊞	No
↖			↖	⇒	↙	↖	No

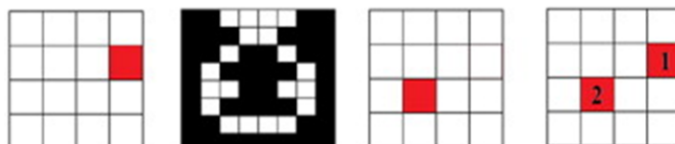
Nota: Tomado de pruebas habilidades cognitivas UR.

B.2. Memoria de trabajo

Con base en [Conway et al. \(2005\)](#), este test consiste en medir la memoria de trabajo de corto plazo de los individuos usando métodos no verbales. Como observamos en la figura 3, se presentan bloques de diferente dificultad, donde se pide recordar todas las ubicaciones de los cuadrados rojos en el

orden que fueron presentados. De forma intercalada, se realiza un juicio de simetría vertical de una imagen compleja, que corresponde a una matriz de cuadros blancos y negros. En esta prueba se obtienen tres resultados: el primero corresponde a la identificación correcta de la simetría de la imagen (*Symmetry accuracy*), el segundo a la ubicación correcta de los cuadros rojos (*Symmetry squares*), y por último el total de cuadrados rojos correctamente recordados solo si se recordaron correctamente todos los cuadrados rojos del bloque (*Symmetry sspan*).

Figura 3: Intervalo de simetría (*Symmetry span*)



Nota: Tomado de Shipstead et al. (2014).

C. Índice sintético de la calidad educativa - ISCE

De acuerdo con la guía metodológica edición No. 5 del [Icfes \(2016\)](#), el *Índice sintético de la calidad educativa - ISCE* es un índice que mide cuatro aspectos de la calidad de la educación en todos los colegios de Colombia. Este índice es calculado por el Instituto Colombiano para la Evaluación de la Educación *Icfes* con base en los resultados de las pruebas Saber y la eficiencia interna de los Establecimientos Educativos. El *ISCE* se compone de cuatro componentes: *Progreso*, *Desempeño*, *Eficiencia* y *Ambiente Escolar*, donde el nivel educativo de la educación Media no tiene este último componente. Este índice presenta una escala de 1 a 10, en donde Progreso y Desempeño pesan 40 %, cada uno, y Eficiencia y Ambiente Escolar pesan 10 %, cada uno. Para el caso de la educación Media, el componente Eficiencia pesa el 20 %. Para esta investigación, utilizamos dos componentes: el *Progreso* que indica el mejoramiento del aprendizaje de los estudiantes en las últimas dos aplicaciones de las pruebas Saber de Lenguaje y Matemáticas de cada Establecimiento Educativo, es decir qué tanto ha mejorado el colegio en relación a los resultados de las pruebas Saber del año anterior. Y la *Eficiencia* que corresponde a la tasa de aprobación escolar de la última cohorte.

C.1. Progreso y Eficiencia

El componente del progreso pesa 40 % del ISCE; es decir que el máximo valor posible de obtener es 4. Esta definido por el *Nivel de desempeño Insuficiente* y el cambio porcentual de un año a otro en el porcentaje de estudiantes en el nivel de desempeño avanzado. La eficiencia pesa el 10 % del ISCE en los niveles educativos Primaria y Secundaria; es decir que el máximo puntaje posible de obtener es 1, mientras que en la educación Media este pesa el 20 %. La Eficiencia es definida como la tasa de aprobación al siguiente año escolar de cada nivel educativo. La tasa de aprobación está en una escala de 0 a 1, y para el caso de la educación Media, se multiplica por dos la tasa de aprobación.

D. Resultados: versión extendida y ejercicios de robustez

Tabla D 1: Modelo principal: versión extendida

	OLS		2SLS (IV)				
	(0)	(1)	(2)	(3)	(4)	(5)	(6)
2da Etapa: Perseverancia de la red de pares	0.024 (0.016)	0.213*** (0.051)	0.170*** (0.054)	0.152*** (0.053)	0.121** (0.050)	0.123** (0.052)	0.110** (0.052)
1era Etapa: Perseverancia de la red de pares de pares		0.333*** (0.027)	0.306*** (0.028)	0.303*** (0.030)	0.297*** (0.030)	0.298*** (0.031)	0.297*** (0.031)
Número de Observaciones	2,325	2,998	2,977	2,610	2,576	2,353	2,325
Anderson & Rubin F estadístico (1949)		18,624	18,819	21,243	22,706	21,005	18,939
Cragg-Donald Wald F estadístico (1993) - Stock-Yogo (2005)		405,876	328,005	286,231	272,521	250,300	244,453
Montiel & Pflueger F estadístico efectivo (2013)		154,702	120,595	103,792	98,278	93,883	92,086
Características personales de los individuos	✓		✓	✓	✓	✓	✓
Medidas Cognitivas	✓			✓	✓	✓	✓
Rendimiento académico previo	✓				✓	✓	✓
Rendimiento académico del colegio	✓					✓	✓
Otras medidas de la calidad del colegio	✓						✓

Notas: Resultados generales del modelo de variables instrumentales para medir el efecto del grupo par sobre el rendimiento académico con base en la perseverancia de la red. La primera fila corresponde a los estimados de la segunda etapa desde la columna (1) hasta la (6), y la columna (0) al estimado de la regresión OLS sin tener en cuenta la medida exógena del instrumento. En la segunda fila se encuentran los estimados de la primera etapa que explican la correlación entre la variable endógena (perseverancia de la red en primer grado) con el instrumento (perseverancia de la red en segundo grado). En ambas etapas se agregan uno a uno el grupo de variables controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios. Por último, se incluyen los estadísticos F de las siguientes pruebas de relevancia del instrumento: Anderson & Rubin (1949), Cragg-Donald (1993) - Stock-Yogo (2005), y Montiel & Pflueger (2013).

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla D 2: Modelo principal: ejercicio de robustez

	2SLS	LIML	FULLER (1)	FULLER (4)	JACKKNIFE
	(1)	(2)	(3)	(4)	(5)
2da Etapa: Perseverancia de la red de pares	0.110** (0.052)	0.110** (0.052)	0.110** (0.052)	0.109** (0.051)	0.124** (0.061)
Número de Observaciones	2,325	2,325	2,325	2,325	2,325
Características personales de los individuos	✓	✓	✓	✓	✓
Medidas Cognitivas	✓	✓	✓	✓	✓
Rendimiento académico previo	✓	✓	✓	✓	✓
Rendimiento académico del colegio	✓	✓	✓	✓	✓
Otras medidas de la calidad del colegio	✓	✓	✓	✓	✓

Notas: Resultados generales del modelo de variables instrumentales para medir el efecto del grupo par sobre el rendimiento académico (variable resultado) con base en la perseverancia de la red, teniendo en cuenta distintos estimadores de IV. En cada estimación se tiene en cuenta todo el conjunto de variables controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios.

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla D 3: **Modelo principal: versión extendida sin la cohorte 2018-1**

	OLS		2SLS (IV)				
	(0)	(1)	(2)	(3)	(4)	(5)	(6)
2da Etapa: Perseverancia de la red de pares	0.037* (0.020)	0.215*** (0.045)	0.166*** (0.048)	0.156*** (0.046)	0.135*** (0.043)	0.139*** (0.044)	0.134*** (0.045)
1era Etapa: Perseverancia de la red de pares de pares		0.414*** (0.031)	0.381*** (0.033)	0.381*** (0.036)	0.375*** (0.036)	0.386*** (0.037)	0.386*** (0.037)
Número de Observaciones	1,648	2,173	2,166	1,865	1,836	1,669	1,648
Anderson & Rubin F estadístico (1949)		24,409	24,333	28,345	30,960	28,388	27,077
Cragg-Donald Wald F estadístico (1993) - Stock-Yogo (2005)		550,025	426,238	379,823	363,794	352,954	345,546
Montiel & Pflueger F estadístico efectivo (2013)		178,887	132,172	115,042	108,383	111,109	108,868
Características personales de los individuos	✓		✓	✓	✓	✓	✓
Medidas Cognitivas	✓			✓	✓	✓	✓
Rendimiento académico previo	✓				✓	✓	✓
Rendimiento académico del colegio	✓					✓	✓
Otras medidas de la calidad del colegio	✓						✓

Notas: Resultados generales del modelo de variables instrumentales para medir el efecto del grupo par sobre el rendimiento académico con base en la perseverancia de la red, sin tener en cuenta la cohorte 2018-1. La primera fila corresponde a los estimados de la segunda etapa desde la columna (1) hasta la (6), y la columna (0) al estimado de la regresión OLS sin tener en cuenta la medida exógena del instrumento. En la segunda fila se encuentran los estimados de la primera etapa que explican la correlación entre la variable endógena (perseverancia de la red en primer grado) con el instrumento (perseverancia de la red en segundo grado). En ambas etapas se agregan uno a uno el grupo de variables controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios. Por último, se incluyen los estadísticos F de las siguientes pruebas de relevancia del instrumento: Anderson & Rubin (1949), Cragg-Donald (1993) - Stock-Yogo (2005), y Montiel & Pflueger (2013).

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla D 4: **Modelo principal: versión extendida red en tercer grado**

	OLS		2SLS (IV)				
	(0)	(1)	(2)	(3)	(4)	(5)	(6)
2da Etapa: Perseverancia de la red de pares	0.027 (0.016)	0.133*** (0.029)	0.090*** (0.029)	0.097*** (0.028)	0.067** (0.027)	0.055* (0.029)	0.053* (0.029)
1era Etapa: Perseverancia de la red de pares de pares		0.589*** (0.015)	0.587*** (0.016)	0.583*** (0.017)	0.579*** (0.017)	0.584*** (0.018)	0.585*** (0.018)
Número de Observaciones	2,311	2,976	2,955	2,596	2,562	2,339	2,311
Anderson & Rubin F estadístico (1949)		20,99	20,45	21,61	22,42	17,31	16,53
Cragg-Donald Wald F estadístico (1993) - Stock-Yogo (2005)		1.850,38	1.698,81	1.534,67	1.460,61	1.354,18	1.335,35
Montiel & Pflueger F estadístico efectivo (2013)		1.537,48	1.426,60	1.242,16	1.154,48	1.029,63	1.014,60
Características personales de los individuos	✓		✓	✓	✓	✓	✓
Medidas Cognitivas	✓			✓	✓	✓	✓
Rendimiento académico previo	✓				✓	✓	✓
Rendimiento académico del colegio	✓					✓	✓
Otras medidas de la calidad del colegio	✓						✓

Notas: Resultados generales del modelo de variables instrumentales para medir el efecto del grupo par sobre el rendimiento académico con base en la perseverancia de la red, tomando como instrumento la red de pares en tercer grado. La primera fila corresponde a los estimados de la segunda etapa desde la columna (1) hasta la (6), y la columna (0) al estimado de la regresión OLS sin tener en cuenta la medida exógena del instrumento. En la segunda fila se encuentran los estimados de la primera etapa que explican la correlación entre la variable endógena (perseverancia de la red en primer grado) con el instrumento (perseverancia de la red en tercer grado). En ambas etapas se agregan uno a uno el grupo de variables controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios. Por último, se incluyen los estadísticos F de las siguientes pruebas de relevancia del instrumento: Anderson & Rubin (1949), Cragg-Donald (1993) - Stock-Yogo (2005), y Montiel & Pflueger (2013).

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla D 5: **Análisis de mediación IV: Comportamientos académicos**

	2SLS (IV)	
	Horas de estudio (1)	Horas U (2)
<i>1.) Efecto del tratamiento (D) en el resultado (Y) instrumentado con (Z)</i>		
Efecto de la perseverancia de la red sobre el rendimiento académico: Efecto total	0.101** (0.050)	0.112** (0.052)
<i>2.) Efecto del tratamiento (D) en el mediador (M) instrumentado con (Z)</i>		
Efecto de la perseverancia de la red sobre el mediador	0.134* (0.076)	0.167** (0.083)
<i>3.) Efecto del mediador (M) en el resultado (Y), controlando por el tratamiento (D) instrumentado con (Z)</i>		
3.1) Efecto directo: relación entre el resultado (Y) y el tratamiento (D)	0.034 (0.022)	0.012 (0.022)
3.2) Efecto de M sobre Y	0.504 (0.378)	0.598 (0.448)
Efecto indirecto: mediado por (D) sobre el outcome (Y) que opera a través de (M)	0.067 (0.063)	0.100 (0.090)
Efecto de mediación como porcentaje total: Efecto indirecto/Efecto total	66.56 %	88.95 %
Número de Observaciones	2,272	2,292
<i>Estadístico F Montiel & Pflueger (2013)</i>		
Primera etapa 1: (D en Z)	92,077	92,084
Primera etapa 2: (M en Z D)	5,705	4,726

Notas: Resultados de los efectos de mediación causal teniendo en cuenta el modelo de variables instrumentales (IV). El modelo informa los efectos del tratamiento de una regresión estándar IV en los efectos directos e indirectos de un tratamiento endógeno, utilizando variables mediadoras sin la necesidad de un instrumento adicional para cada variable mediadora. La tabla reporta tres modelo 2SLS: el primer modelo corresponde al efecto del tratamiento sobre la variable de interés instrumentado con (Z), el segundo corresponde al efecto del tratamiento (D) sobre el mediador (M) instrumentada con (Z), la tercera estimación corresponde al efecto del mediador (M) sobre el outcome (Y), instrumentado con (Z) y controlando por el tratamiento (D), del cual se obtienen dos coeficientes por un lado el efecto directo (relación entre el outcome (Y) y el tratamiento (D)), y por el otro el efecto del mediador (M) sobre (Y). Por tanto, el efecto indirecto es el producto del estimado de la regresión (2) con el segundo estimado de la regresión (3). Así, el efecto de mediación como porcentaje total es el efecto indirecto sobre el efecto total. Los resultados informan dos estadísticos F de la primera etapa.

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla D 6: **Resultados: Estudiantes con modalidad de beca vs. no becados**

	2SLS (IV)			
	Promedio del periodo (1)	Horas de estudio (2)	Horas U (3)	Preparación exámenes (4)
No becados	0.114** (0.056)	0.153** (0.067)	0.142* (0.074)	0.082** (0.034)
Becados	0.096 (0.101)	-0.058 (0.099)	0.359*** (0.115)	0.118* (0.061)
Número de Observaciones	2,325	2,272	2,292	2,320
Controles	✓	✓	✓	✓

Notas: Resultados generales del modelo de variables instrumentales para medir el efecto del grupo par con base en la perseverancia de la red para los estudiantes que reciben algún tipo de beca respecto a los que no, tomando como resultado el promedio del periodo y los comportamientos académicos de los estudiantes (horas de estudio, horas que permanece en la universidad y preparación de exámenes). Cada columna reporta el estimado para cada uno de los comportamientos académicos nombrados, y para la variable de interés. Para una mejor interpretación del estimado principal, el promedio del periodo, las horas de estudio y horas que permanece en la universidad se tomaron de manera estandarizada, mientras que el outcome de preparación de exámenes fue transformado en una variable dummy que indica 1 para las dos categorías más altas de la variable inicial (5=Durante todo el semestre, 4=Con una o más semanas de antelación), y 0 para las demás (3=Durante la semana anterior al parcial, 2= Un día, 1=No estudia para los exámenes). En cada estimación se tiene en cuenta todo el conjunto de controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios.

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla D 7: Estadísticos F y sus valores críticos del modelo principal

Estadístico F (Cragg y Donald, 1993)	N
405,876	2998
328,005	2977
286,231	2610
272,521	2576
250,300	2353
244,453	2325
Valores críticos (Stock y Yogo, 2005)	% nivel de sesgo
16,38	10 %
8,96	15 %
6,66	20 %
5,53	25 %
Estadístico F efectivo (Montiel-Olea y Pflueger, 2013)	N
154,702	2998
120,595	2977
103,792	2610
98,278	2576
93,883	2353
92,086	2325
Valores críticos	% de sesgo en el peor de los casos τ
37,418	5 %
23,109	10 %
15,062	20 %
12,039	30 %
nivel de confianza α	5 %
Estadístico F (Anderson y Rubin, 1949)	N
18,624	2998
18,819	2977
21,243	2610
22,706	2576
21,005	2353
18,939	2325
p - valor	0,000

Notas: La tabla corresponde a los estadísticos F y los valores críticos de cada una de las estimaciones, con base en la metodología propuesta por (Cragg y Donald, 1993), (Stock y Yogo, 2005), (Montiel-Olea y Pflueger, 2013) y (Anderson y Rubin, 1949).

E. Efectos Heterogéneos: Grado y centralidad de la red

Tabla E 1: Grado de la red

	Panel A: Hombres		Panel B: Mujeres	
	Coeficiente de agrupamiento			
	<Mediana	>Mediana	<Mediana	>Mediana
	(1)	(2)	(1)	(2)
Efecto de la perseverancia de la red de pares	0.097 (0.277)	0.013 (0.102)	0.351 (1.060)	0.139*** (0.051)
Número de Observaciones	814		1511	
Controles	✓		✓	
Test de diferencia de coeficientes (<i>p</i> -valor)	0,755		0,838	

Notas: Los resultados corresponden al efecto de la perseverancia de la red por género, teniendo en cuenta el coeficiente de agrupamiento de la red. En el panel A se encuentra el efecto sobre los hombres, y en el panel B se puede visualizar el de las mujeres. Para medir el efecto heterogéneo se tomó como variable dummy la mediana del coeficiente de agrupamiento, por encima de la mediana encontramos mayor agrupación de los nodos, mientras que por debajo de la mediana el caso contrario. En cada estimación se tiene en cuenta todo el conjunto de controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios. Por último se agrega un test de diferencias de coeficientes, el cual testea la hipótesis de nula si los coeficientes son iguales entre sí, y las hipótesis alterna de que los coeficientes son diferentes.

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tabla E 2: Centralidad de la Red

	Panel A: Hombres		Panel B: Mujeres	
	Centralidad de cercanía			
	<Mediana	>Mediana	<Mediana	>Mediana
	(1)	(2)	(1)	(2)
Efecto de la perseverancia de la red de pares	-0.007 (0.184)	0.029 (0.115)	-0.012 (0.100)	0.209*** (0.068)
Número de Observaciones	814		1511	
Controles	✓		✓	
Test de diferencia de coeficientes (<i>p</i> -valor)	0,869		0,066	

Notas: Los resultados corresponden al efecto de la perseverancia de la red por género, teniendo en cuenta la centralidad de cercanía de la red. En el panel A se encuentra el efecto sobre los hombres, y en el panel B se puede visualizar el de las mujeres. Para medir el efecto heterogéneo se tomó como variable dummy la mediana de la centralidad, por encima de la mediana encontramos mayor centralidad de cercanía, mientras que por debajo de la mediana el caso contrario. En cada estimación se tiene en cuenta todo el conjunto de controles utilizado: en características personales de los individuos se encuentran género, facultad, estrato socioeconómico, edad, y cohorte del estudiante; en las medidas cognitivas se toman variables que miden la velocidad de procesamiento y la memoria de trabajo; en el rendimiento académico previo se toma el percentil de las pruebas SABER 11 de los estudiantes, y en otras medidas de la calidad del colegio se encuentran dos indicadores del índice sintético de calidad educativa los cuales corresponden al progreso y la eficiencia de los colegios. Por último se agrega un test de diferencias de coeficientes, el cual testea la hipótesis de nula si los coeficientes son iguales entre sí, y las hipótesis alterna de que los coeficientes son diferentes.

Errores estándar robustos en paréntesis.

Niveles de significancia: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

F. Análisis Alpha de Cronbach

Tabla F 1: Alpha de Cronbach - 5 rasgos de la personalidad

Descripción - ítem	Signo	Item-test correlation	Item-rest correlation	Covarianza promedio	Alpha de Cronbach
Apertura					
Es original, se le ocurren ideas nuevas	+	0.6847	0.5568	0.3095	0.7583
Valora lo artístico, lo estético	+	0.6101	0.4630	0.3276	0.7733
Tiene una imaginación activa	+	0.7263	0.6107	0.2994	0.7495
Es inventivo	+	0.6740	0.5431	0.3121	0.7606
Es ingenioso, analítico	+	0.6353	0.4943	0.3215	0.7684
Le gusta reflexionar, jugar con las ideas	+	0.6067	0.4588	0.3285	0.7740
Es educado en arte, música o literatura	+	0.6041	0.4557	0.3291	0.7744
Tiene pocos intereses artísticos	-	0.5558	0.3971	0.3408	0.7835
Test scale				0.3211	0.7909
Estabilidad emocional					
Es depresivo, melancólico	-	0.6223	0.4399	0.4156	0.7805
Es calmado, controla bien el estrés	+	0.7458	0.6055	0.3639	0.7410
Con frecuencia se pone tenso	-	0.6817	0.5177	0.3907	0.7623
Es emocionalmente estable, difícil de alterar	+	0.7498	0.6112	0.3622	0.7396
Es temperamental, de humor cambiante	-	0.6989	0.5409	0.3835	0.7567
Mantiene la calma en situaciones difíciles	+	0.6874	0.5254	0.3883	0.7604
Test scale				0.3840	0.7891
Extroversión					
Es buen hablador	+	0.6431	0.5114	0.3713	0.8053
Es reservado	-	0.5922	0.4489	0.3843	0.8137
Está lleno de energía	+	0.5710	0.4232	0.3897	0.8172
Tiende a ser callado	-	0.7777	0.6845	0.3372	0.7807
A veces es tímido, inhibido	-	0.7320	0.6244	0.3488	0.7894
Irradia entusiasmo	+	0.6827	0.5611	0.3613	0.7984
A veces es frío y distante	-	0.5553	0.4045	0.3936	0.8196
Es extrovertido, sociable	+	0.7805	0.6882	0.3364	0.7802
Test scale				0.3653	0.8216
Responsabilidad					
Puede a veces ser algo descuidado	+	0.5837	0.4205	0.2770	0.7284
Es un trabajador cumplidor, digno de confianza	-	0.5835	0.4203	0.2770	0.7284
Tiende a ser desorganizado	+	0.5927	0.4316	0.2749	0.7263
Persevera hasta terminar el trabajo	-	0.6353	0.4847	0.2651	0.7164
Tiende a ser flojo, vago	+	0.6743	0.5344	0.2562	0.7068
Hace las cosas de manera eficiente	-	0.5921	0.4309	0.2751	0.7265
Hace planes y los sigue cuidadosamente	-	0.5505	0.3802	0.2846	0.7358
Se distrae con facilidad	+	0.6113	0.4546	0.2707	0.7220
Test scale				0.2726	0.7499
Amabilidad					
Tiende a ser crítico	-	0.5532	0.3145	0.2514	0.6267
Es generoso y ayuda a los demás	+	0.6173	0.3969	0.2283	0.5966
Inicia disputas con los demás	-	0.5292	0.2848	0.2600	0.6373
A veces es maleducado con los demás	-	0.6238	0.4054	0.2259	0.5934
Es considerado y amable con casi todo el mundo	+	0.6629	0.4580	0.2118	0.5733
Le gusta cooperar con los demás	+	0.6220	0.4031	0.2266	0.5943
Test scale				0.2340	0.6470

Notas: La tabla corresponde a los resultados del *Alpha de Cronbach* para los cinco grandes rasgos de la personalidad. Se puede observar el signo del ítem, es decir la dirección en la que una variable del ítem entró en la escala. Además, de la correlación entre el ítem-prueba y la correlación entre el ítem y la escala. Los valores del Alpha de Cronbach entre 0.70 y 0.90 indican una buena consistencia interna. Elaborado con base en los ítems teóricos de (Salgado et al., 2016), y las pruebas de habilidades socioemocionales UR.

Tabla F 2: Alpha de Cronbach - Otros factores no cognitivos

Descripción - ítem	Signo	Item-test correlation	Item-rest correlation	Covarianza promedio	Alpha de Cronbach
Perseverancia					
Termino todo aquello que inicio	+	0.6928	0.4905	0.3595	0.6919
Trabajo independiente con concentración	+	0.6062	0.3725	0.4099	0.7353
Me esfuerzo incluso después de experimentar el fracaso	+	0.6978	0.4975	0.3567	0.6892
Me mantengo comprometido con mis metas	+	0.7630	0.5930	0.3188	0.6518
Continúo trabajando duro incluso cuando deseo dejar de hacerlo	+	0.7249	0.5365	0.3409	0.6741
Test scale				0.3571	0.7353
Maleabilidad de la inteligencia					
Usted tiene cierto nivel de inteligencia, y no puede hacer mucho para cambiarlo	+	0.8247	0.5958	0.5116	0.6769
Su inteligencia es algo que usted no puede cambiar significativamente	+	0.8427	0.6308	0.4671	0.6368
Usted puede aprender cosas nuevas, pero no puede realmente cambiar su nivel inicial de inteligencia	+	0.8015	0.5525	0.5688	0.7252
Test scale				0.5159	0.7617
Autocontrol para trabajar					
Vengo a clases preparado	+	0.6857	0.4091	0.3211	0.5866
Presto atención y resisto a tentaciones durante la clase	+	0.7083	0.4431	0.3001	0.5627
Recuerdo y sigo las instrucciones que recibo	+	0.6952	0.4232	0.3123	0.5767
Cuando se me asigna una tarea o trabajo empiezo a trabajar inmediatamente en ella, en lugar de esperar hasta el último minuto	+	0.6934	0.4206	0.3140	0.5786
Test scale				0.3119	0.6445
Autocontrol interpersonal					
Permito que otros hablen sin interrumpirlos	+	0.6444	0.3543	0.3857	0.6532
Soy educado con mis compañeros de clase y otros adultos	+	0.6572	0.3724	0.3737	0.6416
Puedo controlar mi temperamento	+	0.7687	0.5433	0.2689	0.5246
Puedo permanecer calmado incluso cuando soy criticado o provocado	+	0.7489	0.5111	0.2876	0.5477
Test scale				0.3290	0.6623

Notas: La tabla corresponde a los resultados del *Alpha de Cronbach* para otros factores no cognitivos. Se puede observar el signo del ítem, es decir la dirección en la que una variable del ítem entró en la escala. Además, de la correlación entre el ítem-prueba y la correlación entre el ítem y la escala. Los valores del Alpha de Cronbach entre 0.70 y 0.90 indican una buena consistencia interna. Los ítems de la escala de *Grit-Short* fueron tomados de [Duckworth y Quinn \(2009\)](#), para maleabilidad de la inteligencia se tuvo en cuenta el estudio por [Dweck \(1999\)](#), y para medir el autocontrol se tuvo en cuenta a [Tsukayama et al. \(2013\)](#).

G. Estadísticas descriptivas

Tabla G 1: Estadísticas descriptivas

Variable	Obs.	Media	Des.Est.	Min	Max
Variable Resultado					
Promedio del periodo	2325	3,949	0,430	0,44	4,94
Variable Endógena					
Perseverancia de la red de pares de primer grado	2325	0,024	0,957	-6,873	2,968
Instrumento					
Perseverancia de la red de pares de segundo grado	2325	0,007	1,009	-7,295	3,809
Integrantes de la red					
Amigos reportados	2325	6,110	5,571	1	44
Variables Controles					
Variables Cognitivas					
<i>velocidad de procesamiento</i>					
Busqueda de símbolos (Symbol search)	2325	30,502	8,48	0	88
Sustitución de símbolos (Symbol substitution)	2325	48,475	9,78	17	99
<i>memoria de trabajo</i>					
Precisión de simetría (Symmetry accuracy)	2325	90,845	1,05	0	100
Cuadrados de simetría (Symmetry squares)	2325	26,866	8,13	0	42
Intervalo de simetría (Symmetry span)	2325	18,399	8,75	0	42
Resultados Saber 11 Individual (Percentil)					
Saber 11 Ciencias Naturales	2325	86,015	1,51	9	100
Saber 11 Matemáticas	2325	84,868	1,58	3	100
Saber 11 Ciencias Sociales	2325	87,355	1,44	12	100
Saber 11 Lectura Crítica	2325	85,834	1,52	13	100
Saber 11 Inglés	2325	89,759	1,28	7	100
Resultados Saber 11 Colegio					
Saber 11 Matemáticas	2325	0,827	0,058	0,592	0,937
Saber 11 Ciencias naturales	2325	0,824	0,051	0,602	0,921
Saber 11 Sociales y ciudadanas	2325	0,817	0,050	0,580	0,920
Saber 11 Lectura crítica	2325	0,811	0,043	0,608	0,910
Saber 11 Inglés	2325	0,865	0,076	0,597	0,969
Índice Sintético de la Calidad Educativa -ISCE					
Progreso	2325	3,198	0,581	0	4
Eficiencia	2325	1,955	0,098	0	2

Continúa en la siguiente página

Notas: La tabla corresponde a las variables utilizadas en el modelo propuesto. La variable resultado corresponde al promedio del periodo (en las estimaciones fue utilizada la versión estandarizada). La variable endógena corresponde a la perseverancia de la red de pares en primer grado, y el instrumento a la red de pares en segundo grado, ambas estandarizadas. Como variables controles fueron utilizadas el siguiente grupo: variables cognitivas que miden la velocidad de procesamiento y la memoria de trabajo, los resultados del Saber 11 de los estudiantes (percentil), los resultados del Saber 11 de los colegios, y el índice sintético de la calidad educativa.

Variable	Obs.	Media	Des.Est.	Min	Max
Otras variables UR					
<i>Características de los individuos</i>					
Género (1=Hombre)	2325	0,350	0,477	0	1
Estrato bajo (1,2,3)	2325	0,408	0,492	0	1
Estrato alto (4,5,6)	2325	0,591	0,492	0	1
Edad de ingreso a la U	2325	17,908	1,079	15,25	38,66
<i>Facultades y cohortes</i>					
Escuela de Administración	2325	0,183	0,387	0	1
Escuela de Medicina y Ciencias de la Salud	2325	0,285	0,451	0	1
Facultad de Economía	2325	0,097	0,296	0	1
Facultad de Estudios Internacionales, Políticos y Urbanos	2325	0,136	0,343	0	1
Facultad de Jurisprudencia	2325	0,209	0,407	0	1
Otras Facultades	2325	0,089	0,285	0	1
Cohorte 2017-1	2325	0,332	0,471	0	1
Cohorte 2017-2	2325	0,193	0,395	0	1
Cohorte 2018-1	2325	0,291	0,454	0	1
Cohorte 2018-2	2325	0,184	0,387	0	1
VARIABLES UTILIZADAS EN EFECTOS HETEROGÉNEOS					
Habilidades Socioemocionales					
<i>5 rasgos de la personalidad (Big 5)</i>					
Apertura	2325	-0,010	0,549	-1,930	0,960
Extroversión	2325	0,005	0,558	-1,740	1,145
Estabilidad emocional	2325	-0,002	0,490	-1,009	1,580
Responsabilidad	2325	0,018	0,407	-1,593	0,901
Amabilidad	2325	-0,009	0,313	-0,595	1,641
Otros factores no cognitivos					
Mentalidad de crecimiento	2325	0,005	0,820	-2,933	1,162
Autocontrol interpersonal	2325	0,019	0,680	-2,507	2,337
Autocontrol en el trabajo	2325	0,009	0,850	-4,204	1,901
Áreas del conocimiento					
Ciencias Económicas y Administración	2325	0,280	0,449	0	1
Ciencias Sociales	2325	0,196	0,397	0	1
Ciencias de la salud y básicas	2325	0,173	0,379	0	1
Medicina	2325	0,141	0,348	0	1
Jurisprudencia	2325	0,209	0,407	0	1
Características de los colegios					
Privado	2325	0,903	0,296	0	1
Público	2325	0,097	0,296	0	1
Creencia del colegio privado (1=Religioso)	2325	0,325	0,469	0	1
Análisis de Mediación					
Comportamientos académicos					
Horas de estudio	2272	3,729	1,914	1	12
Horas de permanencia en la U	2292	8,042	2,169	1	14
Frecuencia de preparación de exámenes	2320	0,226	0,418	0	1

Notas: La tabla corresponde a la continuación de la tabla 1. Además, como variables controles fueron utilizadas las características individuales de los estudiantes, como el género, estrato, edad de ingreso a la universidad, información de las facultades, y cohortes. Adicionalmente para el modelo propuesto fueron utilizadas variables para estimar efectos heterogéneos, en este grupo se encuentran: habilidades socioemocionales que corresponde a los 5 grandes rasgos de la personalidad como apertura, extroversión, estabilidad emocional, responsabilidad y amabilidad, otros factores no cognitivos como la mentalidad de crecimiento, perseverancia, autocontrol interpersonal y autocontrol en el trabajo. De igual forma, fueron incluidas variables por áreas del conocimiento donde se encuentran ciencias económicas y administración, ciencias sociales, ciencias de la salud y básicas, medicina y jurisprudencia. De la misma manera, encontramos las características de los colegios, como colegios públicos, privados, y la creencia del colegio privado (religioso o no religioso). Por último, son expuestas las variables de los comportamientos académicos, entre estas las horas de estudio, horas de permanencia en la universidad y la frecuencia de preparación de exámenes (1=Durante todo el semestre, con una o más semanas de antelación. 0= Durante la semana anterior al parcial, un día, no estudia para los exámenes).

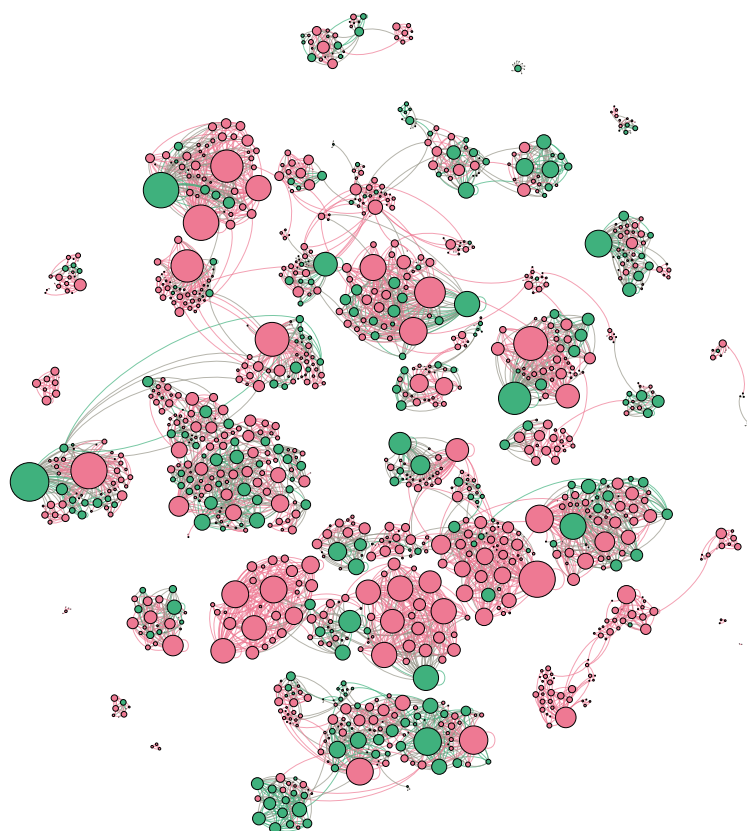
Tabla G 2: Estadísticas descriptivas: Redes

Cohorte	Red	Proporción de nodos en la red	Densidad	Grado Medio	Longitud media del camino
2017-1	Completa		0,003	3,871	4,858
	Mujeres	79,57 %	0,003	2,610	2,599
	Hombres	20,43 %	0,003	2,142	2,374
2017-2	Completa		0,003	2,966	3,212
	Mujeres	84,51 %	0,003	2,190	2,612
	Hombres	15,49 %	0,010	1,345	1,797
2018-1	Completa		0,003	3,599	5,678
	Mujeres	78,38 %	0,003	2,295	7,635
	Hombres	21,62 %	0,005	1,291	4,491
2018-2	Completa		0,005	3,423	5,104
	Mujeres	78,52 %	0,004	2,329	3,917
	Hombres	21,48 %	0,011	1,669	1,881

Notas: La tabla muestra las estadísticas descriptivas de las redes de estudio para las cuatro cohortes analizadas. Los cálculos fueron realizados en el software para el análisis de redes llamado *Gephi*. Además estos cálculos fueron presentados para la red completa, para la red de mujeres y la red de hombres. Dentro de las estadísticas fueron expuestas la proporción de nodos en la red, la densidad, el grado medio, y la longitud media del camino.

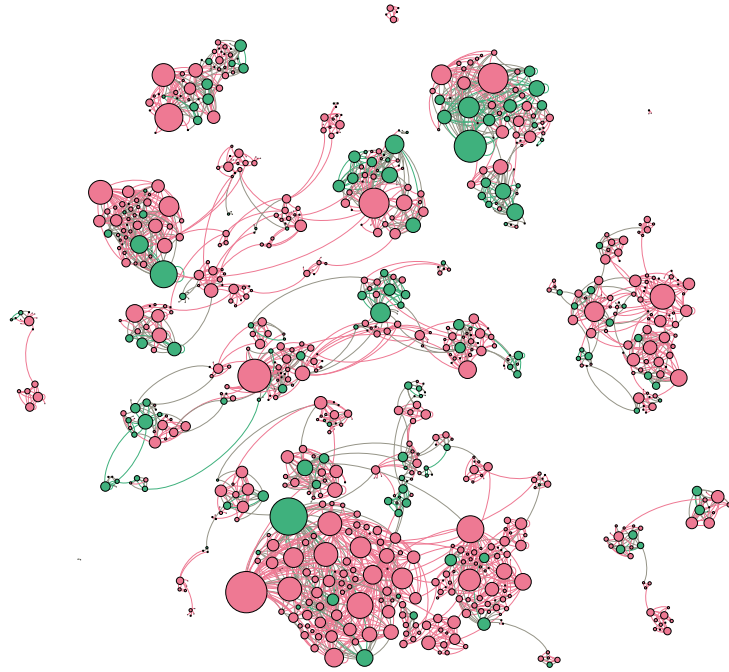
H. Redes académicas por cohorte

Figura H 1: Red de estudio cohorte 2017-1 : por género



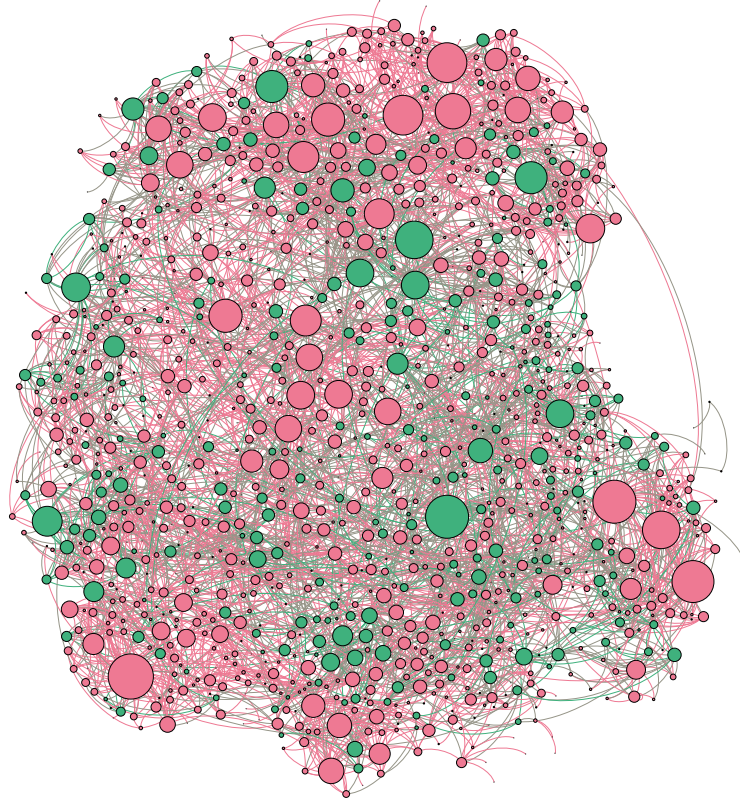
Nota: Los nodos rosados corresponden a las mujeres, y los verdes a los hombres.

Figura H 2: Red de estudio cohorte 2017-2 : por género



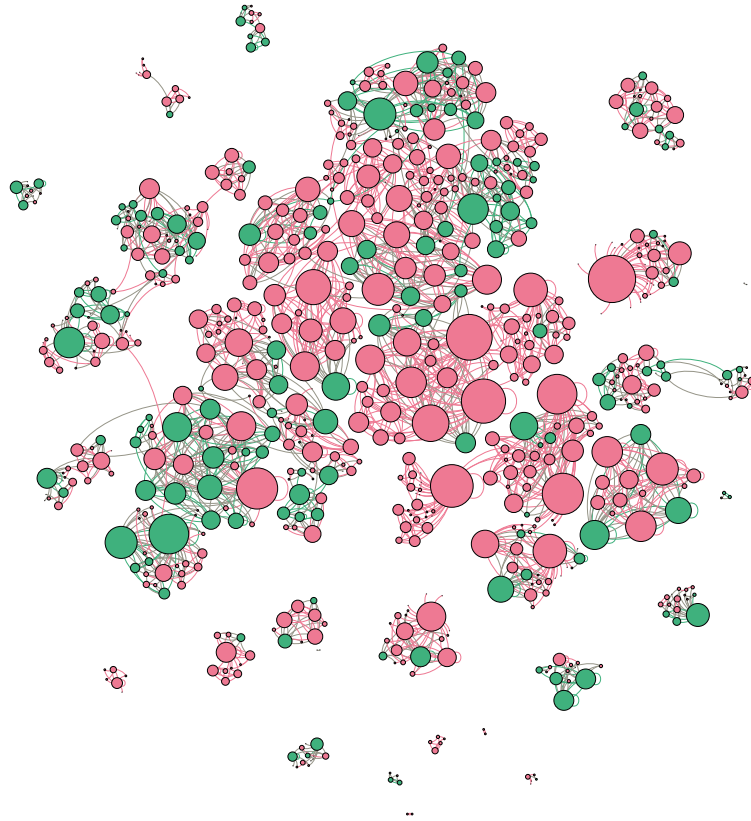
Nota: Los nodos rosados corresponden a las mujeres, y los verdes a los hombres.

Figura H 3: Red de estudio cohorte 2018-1 : por género



Nota: Los nodos rosados corresponden a las mujeres, y los verdes a los hombres.

Figura H 4: Red de estudio cohorte 2018-2 : por género



Nota: Los nodos rosados corresponden a las mujeres, y los verdes a los hombres.