



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

EFECTOS DE PARES HETEROGÉNEOS Y RENDIMIENTO ACADÉMICO

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN
ECONOMÍA APLICADA

GONZALO MATÍAS VIVEROS ARENAS

SANTIAGO DE CHILE
NOVIEMBRE 2011



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

EFFECTOS DE PARES HETEROGÉNEOS Y RENDIMIENTO ACADÉMICO

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN
ECONOMÍA APLICADA

GONZALO MATÍAS VIVEROS ARENAS

PROFESOR GUÍA
MATTIA MAKOVEC

MIEMBROS DE LA COMISIÓN
ALEJANDRA MIZALA SALCES
BENJAMÍN VILLENA ROLDÁN
GIACOMO DE GIORGI

SANTIAGO DE CHILE
NOVIEMBRE 2011

EFFECTOS DE PARES HETEROGÉNEOS Y RENDIMIENTO ACADÉMICO

El estudio de *efectos sociales endógenos* consiste en analizar el efecto que ocurre en que el comportamiento de un individuo, de alguna manera, tiende a variar con el comportamiento del grupo que pertenezca. Este proyecto explicará el efecto social que ocurre en el alumnado de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile en el año 2010. El efecto a estudiar es denominado *efecto de pares*, que consiste en determinar los grupos de compañeros, o *pares*, con los cuales el alumno se relaciona con el fin de determinar la relevancia que tienen estos sobre el rendimiento académico.

El análisis empírico consiste en identificar la correlación entre el rendimiento de los alumnos y sus pares, por medio de un específico modelo de regresión denominado *lineal en medias*. No obstante, la correcta identificación recae en solucionar los problemas recurrentes a lo largo de la literatura. Estos son el *reflection problem* y la *endogeneidad*. Bajo la forma de capturar la información, es posible encontrar las soluciones para estos problemas y dar paso a la identificación del efecto de pares.

La manera de recopilar la información es por medio de una *encuesta* especificada en conjunto con el profesor guía. Esta es respondida por los alumnos de la Facultad con la cual se tiene especificada toda la información relevante para identificar la heterogeneidad y la *intensidad* del efecto de pares. Se destaca que con la realización de esta, se ha conformado una base de datos original dentro de la literatura.

El aporte de este proyecto hacia la literatura de efectos de pares recae en capturar la *intensidad* en la relación entre los pares, proporcionando la posibilidad de innovar los métodos usuales para la estimación de este efecto visto a lo largo de la literatura. Esta metodología propuesta se basa en la hipótesis que ha *priori* los pares más cercanos al alumno, en términos de amistad, son más influyentes en el rendimiento de este. Los resultados así lo demuestran. Un papel relevante poseen los efectos de pares sobre el rendimiento académico de los estudiantes, y que este rendimiento depende de la intensidad de la relación entre los pares.

Se recalca, además, que se especifican dos tipos distintos de grupos de pares en los cuales los alumnos interactúan a diario en la vida universitaria. El primero de ellos consiste en los compañeros con los que se junta a *estudiar* y el segundo trata sobre los compañeros con los que se juntan a hacer *vida social*. Se ha encontrado evidencia significativa que los alumnos eligen a sus pares de estudio dado por un rendimiento similar, en comparación a los pares con los cuales hace vida social que estos lo hacen de manera independiente al rendimiento. Por lo demás, son los compañeros con los que hace vida social los que perjudican el efecto positivo dado por los compañeros con los cuales se reúnen a estudiar.

Índice General

1. Introducción	1
1.1. Motivación	3
1.2. Objetivo General	4
1.2.1. Objetivos Específicos	4
1.3. Metodología	4
1.4. Resultados Esperados	5
1.5. Revisión de la Literatura	5
2. Modelo Lineal de Interacciones Sociales	10
2.1. Introducción	10
2.2. Modelo de Efectos Sociales Endógenos	11
2.2.1. Especificación del Modelo	12
2.2.2. Identificación de Parámetros	13
2.3. Modelo de Efecto de Pares con Interacción por Grupos	14
2.3.1. Superposición Parcial de Grupos de Pares	15
2.3.2. Especificación del Modelo	16
2.3.3. Identificación de Parámetros	17
2.4. Modelo en <i>Intensidad</i> del Efecto de Pares	18
2.4.1. Especificación del Modelo	19
2.4.2. Interacción por Grupo	20
2.4.3. Identificación de Parámetros	25
3. Efecto de Pares en la Red Social de la U. de Chile	26
3.1. Introducción	26
3.2. Encuesta Alumno	27
3.3. ANÁLISIS DESCRIPTIVO	28
3.3.1. Intensidad en la Relación entre Alumnos y Pares	29

3.4. RESULTADOS	30
3.4.1. Relación Lineal sobre el Rendimiento Medio	30
3.4.1.1. Intensidad en la Relación	30
3.4.2. Modelo de Efectos de Pares	32
3.4.2.1. Estudio Grupal y Vida Social	32
3.4.2.2. Modelo en Intensidad	34
Conclusión	37
Referencias	40
A. Resultados Gráficos y Tablas	42
A.1. Estadística Descriptiva	42
A.2. Relación Alumno <i>vs.</i> Pares	52
A.2.1. Intensidad en la Relación	52
A.2.2. Relación Lineal sobre el Rendimiento	55
A.3. Modelo de Efectos de Pares	59
B. Encuesta Alumno	66
B.1. Preguntas sobre sus Pares	67
B.1.1. Estudio Grupal	67
B.1.2. Vida Social	69
B.2. Preguntas sobre el Entorno Familiar	70
B.2.1. Situación Socioeconómica de la Familia	70
B.2.2. Composición de la Familia	73
B.3. Preguntas Personales	74
B.3.1. Gusto Personal o Preferencias	74
B.3.2. Contextura Física	75

CAPÍTULO 1

Introducción

Manski (1993) fue pionero en el estudio de *efectos sociales endógenos*, el cual explica el comportamiento de un individuo, de alguna manera, tiende a variar con el comportamiento del grupo que lo contenga, es decir, el comportamiento de un grupo de amigos puede afectar de alguna manera el comportamiento propio. Dependiendo del contexto, estos efectos pueden ser llamados “*normas sociales*”, “*influencias de los compañeros*”, “*efectos de vecindad*”, “*conformidad*”, “*imitación*”, “*interacciones sociales*”, “*preferencias interdependientes*”, “*comportamiento de multitud*” o *herd behavior*. Por ejemplo, Conlisk (1980) estudió el efecto de multitud y demostró que, si la toma de decisiones es costosa, puede ser óptima la opción de imitar el comportamiento de otras personas que están mejor informadas.

Esta tesis, *Efectos de Pares Heterogéneos y Rendimiento Académico*, tratará de explicar el efecto social que ocurre dentro de un establecimiento de educación superior, en particular en el alumnado de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile en el año 2010. La manera de capturar la información es por medio de una encuesta desarrollada en conjunto con el profesor guía. Los alumnos de manera voluntaria respondieron la encuesta y por esta razón no se consiguió un número relevante de encuestados. Sin embargo, este proyecto representa un piloto para realizar una encuesta de esta índole en una muestra de mayor cantidad de alumnos. La importancia radica en los aportes e innovaciones relevantes de este proyecto hacia la literatura, los que serán discutidos más adelante.

El efecto a estudiar es denominado como *efecto de pares*, que consiste en determinar los grupos de compañeros con los cuales el alumno interactúa diariamente en la universidad. En el contexto de esta tesis, los *efectos de pares heterogéneos* se refiere a la diferente intensidad de las relaciones entre los pares que se está explícitamente midiendo. En consecuencia, se establecerá en que medida la intensidad de estas interacciones afectan el rendimiento académico, en términos de esfuerzo y logro.

El estudio empírico sobre los efectos sociales consiste en identificar la correlación entre el rendimiento de los alumnos que de alguna forma interactúan entre sí. Manski (1993) distingue tres tipos de efectos que influyen en el rendimiento del individuo. El *efecto exógeno* o la influencia que tienen las características exógenas de los pares, el *efecto endógeno* como la influencia del rendimiento de los pares, y el *efecto correlacionado* donde los individuos tienden a tener rendimientos similares debido a que enfrentan un ambiente común. La manera de identificar los efectos de pares viene dado por medio de un modelo econométrico denominado *modelo lineal en medias*. Propuesto en primera instancia por este autor, ha sido fundamental en el estudio de los efectos sociales.

En los primeros estudios, la difícil identificación de los efectos de pares ha llevado a implementar condiciones restrictivas en la identificación, como por ejemplo, considerar restricciones en el modelo, en los datos e inclusive omisión de la presencia del efecto correlacionado, implicando soluciones parciales. En el último decenio se han tenido importantes avances en la especificación del modelo lineal en medias para la correcta identificación del efecto de pares. En este proyecto se han realizado importantes contribuciones hacia la literatura. Primero, corresponde ha considerar la *intensidad*¹ en la relación entre los pares para la identificación del efecto. Para llevar a cabo esto, se ha modificado el modelo lineal en medias con la finalidad de capturar la intensidad en la relación y de esta manera estimar el efecto de pares. Segundo, considerando las interacciones dentro de la vida universitaria, se analiza el efecto de pares, de manera independiente y conjunta, en dos grupos de pares de distinta índole: uno de los cuales los alumnos se juntan a *estudiar* y el otro con los cuales se juntan hacer *vida social*.

Los resultados demuestran, en primera parte, la relevancia en considerar la intensidad en la relación dado que explica de mejor manera el rendimiento de los alumnos en comparación de no considerarla. Por otra parte, se muestra que existe evidencia estadísticamente significativa para concluir la presencia de efectos de pares, especialmente en el grupo de pares con los cuales se reúnen a estudiar. Estos conforman sus grupos con compañeros de rendimiento similar con consecuencias positivas. Por último, se destaca que se encuentra evidencia para concluir que los pares con los cuales se reúnen a hacer vida social tienen una consecuencia desfavorable sobre el efecto positivo producto de los pares con los que se reúnen a estudiar.

La estructura de esta tesis es la siguiente. En las secciones que siguen se muestran la motivación, objetivos, metodología y resultados esperados que se plantean para este proyecto. Este capítulo termina con la revisión de la literatura. En el capítulo 2 se muestra la teoría

¹La intensidad en la relación consiste en que el alumno puede hacer distinción entre sus pares según el nivel de amistad que posean, y de esta forma, el alumno deberá poder *rankear* a sus pares.

relevante acerca del modelo econométrico que se utiliza para estimar el efecto de pares. En este se estudia el modelo lineal en medias (Manski, 1993), luego una extensión de este modelo en que los individuos interactúan en grupos (Bramoullé *et al.* 2008, Boucher *et al.* 2010, De Giorgi *et al.* 2010) y se finaliza con el modelo propuesto de efecto de pares que cuantifica la intensidad en la relación entre los pares. Luego, en el capítulo 3 se muestra el estudio y resultados obtenidos para el efecto de pares en la red social de la Universidad de Chile. Este proyecto culmina con el capítulo de conclusión. En el anexo A se muestran los resultados gráficos y tablas hacia la información recopilada, y en el anexo B se muestra la encuesta realizada para la captura de dicha información.

1.1. Motivación

A lo largo de la literatura se muestra que hay dos principales problemas en la identificación del efecto de pares. Primero, es difícil distinguir los reales efectos sociales (endógenos y/o exógenos) del efecto correlacionado producto de la autoselección que realizan los alumnos a los mismos grupos de estudio o amigos (*problema de endogeneidad*). Segundo, incluso ante la ausencia del efecto correlacionado, el *reflection problem* dificulta la identificación de los efectos endógenos de los efectos exógenos. El *reflection problem* se hace presente en el momento que se observa la distribución del comportamiento en la población y se trata de inferir si es el comportamiento de un individuo en algún grupo influye en el comportamiento de los individuos que componen el grupo. Considerando estos problemas, la primera etapa en la correcta identificación del efecto de pares recae en la solución de estos.

La manera de identificar los efectos sociales viene dado por medio del *modelo lineal en medias*. Al plantear este modelo es posible dar reconocimiento de ambos problemas. El modelo es el siguiente. Consideremos: y como el rendimiento del alumno; \mathbf{z} los atributos característicos del grupo; (\mathbf{x}, u) atributos que afectan a y . El investigador observa la realización aleatoria de $(y, \mathbf{x}, \mathbf{z})$, pero no observa u . Luego, el modelo se define como:

$$y = \alpha + \beta \mathbb{E}[y|\mathbf{z}] + \mathbb{E}[\mathbf{x}|\mathbf{z}]' \boldsymbol{\gamma} + \mathbf{x}' \boldsymbol{\eta} + u,$$

donde $(\alpha, \beta, \boldsymbol{\gamma}, \boldsymbol{\delta}, \boldsymbol{\eta})$ es el vector de parámetros y se considera el supuesto que el error esperado por individuo es igual a los atributos medios al cual pertenezca el individuo, *i.e.*, $\mathbb{E}[u|\mathbf{x}, \mathbf{z}] = \mathbf{z}' \boldsymbol{\delta}$. El modelo a estimar viene dado por el rendimiento esperado de un individuo, obtenido como la esperanza de y condicionado en los atributos que lo caracterizan (\mathbf{x}, \mathbf{z}) :

$$\mathbb{E}[y|\mathbf{x}, \mathbf{z}] = \alpha + \beta \mathbb{E}[y|\mathbf{z}] + \mathbb{E}[\mathbf{x}|\mathbf{z}]' \boldsymbol{\gamma} + \mathbf{x}' \boldsymbol{\eta} + \mathbf{z}' \boldsymbol{\delta},$$

la cual tiene una forma lineal. El *reflection problem* se hace presente en esta expresión. Se tiene que para explicar el rendimiento medio de un individuo ($\mathbb{E}[y|\mathbf{x}, \mathbf{z}]$) existe la presencia de $\mathbb{E}[y|\mathbf{z}]$ como regresor y ambos términos no pueden ser determinados al mismo tiempo. Este problema dificulta la identificación de los efectos endógenos de los efectos exógenos.

El segundo problema que se enfrenta ante el estudio de los efectos de pares, consiste en el *efecto correlacionado*. Los agentes que tienen algún conocimiento de ciertos individuos, tienden a elegir *endógenamente* su grupo de estudio y/o amigos. La manera que se propone a resolver este problema viene dado por Bramoullé, *et al.* (2008) y De Giorgi, *et al.* (2010) donde muestran que en el contexto de los efectos de pares los instrumentos surgen de manera natural, dada por las características exógenas de los denominados *pares excluidos* (pares de los pares).

1.2. Objetivo General

El principal objetivo de este trabajo es mostrar procedimientos alternativos a la actual literatura para *identificar la heterogeneidad e intensidad* del efecto de pares.

1.2.1. Objetivos Específicos

- Medir la heterogeneidad del efecto de pares utilizando información recopilada a través de una encuesta con el propósito de medir la intensidad de las relaciones sociales de los estudiantes, en términos de amistad y estudio.
- Establecer si los efectos de pares varían en base a la intensidad de la relación (de amistad y/o estudio) entre pares.
- Estimar el impacto de los efectos de pares sobre esfuerzo y logros académicos.

1.3. Metodología

En primera instancia, se tiene especificada la realización de una encuesta docente que será contestada por el alumnado de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile en el año 2010. Con la realización de esta encuesta, se tiene toda la información relevante para llevar a cabo satisfactoriamente este proyecto de tesis. Para este

fin, se considerarán preguntas tales como: el distinto grado de amistad entre los compañeros de estudios; con quien estudia y quienes son los amigos con quien el alumno no estudia, etc.

Luego de una acabada revisión bibliográfica, se tiene claro las principales innovaciones aplicadas en el modelo lineal en medias, lo que permite resolver los principales problemas en la identificación del efecto de pares. La clave en la solución de estos problemas consiste en la superposición parcial de los grupos de pares (De Giorgi, *et al.* 2010). El modelo propuesto para este proyecto, considera la intensidad en la relación para la identificación del efecto, siendo una *innovación* importante dentro de la literatura. Este se presenta como una extensión del enfoque propuesto por Bramoullé, *et al.* (2008) dado que es aquel que mejor se adecúa a los datos obtenidos por la encuesta.

Con la especificación del modelo propuesto, teniendo en claro las consideraciones y soluciones a los problemas de la literatura, se procede a estimar la intensidad del efecto de pares en la red en la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile.

1.4. Resultados Esperados

El primer resultado que se espera obtener es encontrar una heterogeneidad importante en los datos para así identificar grupos de pares. Se espera además otra fuente de heterogeneidad provenga de la familia.

Conjuntamente, se pretende encontrar un papel relevante de los efectos de pares sobre el rendimiento académico de los estudiantes, y que este rendimiento dependa de la intensidad de la relación entre los pares.

1.5. Revisión de la Literatura

En los últimos años, el estudio de los efectos de pares literalmente ha explotado. Numerosos estudios han presentado pruebas empíricas que documentan la existencia de efectos de pares pertinentes en muchas áreas. Ellos se han aplicado a temas tan diversos como la actividad criminal (Glaeser, *et al.* 1996), la participación de bienestar (Bertrand, *et al.* 2000), el logro académico (Sacerdote, 2001), la participación en los planes de jubilación (Sáez & Duflo, 2003), obesidad (Trogon, *et al.* 2008 y Fowler & Christakis, 2008), entre otros. Además, tanto en lo económico como en la literatura sociológica, la importancia de los efectos de pares en la configuración de la conducta individual ha sido ampliamente reconocida (Jackson, 2006).

A pesar de la amplia literatura sobre el tema, la identificación de las interacciones sociales (efectos exógenos y endógenos) sigue siendo muy problemática. Manski (1993) muestra los dos principales problemas de identificación que se plantean en el contexto de un *modelo lineal en medias*. En primer lugar, es difícil distinguir los reales efectos sociales de los efectos correlacionados, producto principalmente de la autoselección de los agentes a un grupo común (*problema de endogeneidad*). En segundo lugar, incluso ante la ausencia de los efectos correlacionados, el particular problema de simultaneidad en el comportamiento de los agentes que interactúan introduce una colinealidad perfecta entre el resultado medio esperado del grupo y sus características medias. El *reflection problem* dificulta la identificación de los efectos endógenos de los efectos exógenos (Brock & Durlauf 2001, Moffitt 2001, Lee 2007).

Este proyecto de tesis reconoce casos especiales de varios modelos estudiados en la literatura, entre los que destacan Manski (1993), Moffitt (2001) y Lee (2007). Estos autores analizan las distintas versiones del estándar modelo lineal en medias. El modelo de Manski considera que el individuo se incluye cuando se obtiene de la media de su grupo. En este caso, los efectos de pares no son identificados. En el modelo de Moffitt, el individuo se excluye de la media y todos los grupos tienen el mismo tamaño. En este estudio los efectos de pares tampoco pueden ser identificados. En contraste, el modelo de Lee considera que las interacciones en grupos de diferentes tamaños, y el individuo también es excluido de la media del grupo. Se considera que las variaciones en los distintos tamaños de grupo puede producir la identificación de estos efectos.

El reconocimiento que los individuos interactúan en grupos ha conllevado a resultados satisfactorios en la identificación de los efectos sociales. Bramoullé, *et al.* (2008) son los primeros en considerar las interacciones sociales que se estructuran a través de una red social. El análisis de las redes sociales ha sido desarrollado por los sociólogos y se ha convertido en un campo central de la investigación en la sociología. En economía, un creciente cuerpo de trabajo teórico explora cómo los incentivos individuales dan lugar a redes y, a su vez, pueden ser modelados por ellas (Jackson, 2006). A nivel empírico, algunos estudios recientes, como Boucher *et al.* (2010) y De Giorgi *et al.* (2010), se aprovechan de la valiosa información que se puede obtener sobre las relaciones entre los agentes, a fin de proporcionar evidencia más precisa sobre los efectos sociales.

El planteamiento de Boucher, *et al.* (2010) ofrece la primera aplicación del enfoque desarrollado por Lee (2007) para la identificación y la estimación de los efectos de pares. La relevancia de esta nueva aplicación es que permite resolver el problema de endogeneidad y el *reflection problem* con datos estándar y sin imponer “*restricciones de exclusión*” (planteadas por Bramoullé, *et al.* 2008). El método se basa en supuestos estructurales, lo que hace que su enfrentamiento a los datos reales sea importante. Para el análisis, han recogido datos admi-

nistrativos sobre resultados de los exámenes finales de una escuela en la provincia canadiense de Quebec. Investigan la presencia de efectos de pares en el rendimiento en matemáticas, ciencias, francés e historia aplicando el enfoque de Lee (2007). El modelo econométrico que utilizan se basa en dos hipótesis fundamentales. En primer lugar, los individuos interactúan en grupos que son conocidos por el investigador. En segundo lugar, el resultado individual es determinado por un modelo lineal en medias con efectos fijos dados por el grupo de pares del agente. Así, la calificación del examen de un estudiante se ve afectado por sus características, por la puntuación promedio de la prueba y las características de su grupo de pares. Lee (2007) muestra que los efectos de pares son totalmente identificados en dicho marco.

El trabajo publicado por De Giorgi, *et al.* (2010) propone una nueva estrategia para la identificación y estimación de los efectos de pares, y en particular de su componente endógeno (efecto correlacionado). Esta estrategia aprovecha una característica común de las redes sociales, la existencia de grupos que se superponen parcialmente de sus compañeros, por lo que es muy general y fácilmente aplicable a varias configuraciones en un análisis. Al explotar esta característica es posible resolver cuestiones relacionadas al *reflection problem* y al efecto correlacionado (o *endogeneidad*) y, por tanto, se puede identificar todos los parámetros de interés en el estándar modelo lineal en medias de interacciones sociales.

La aplicación empírica de los efectos de pares puede ser abordado en una variedad de temas. Por ejemplo, Laschever (2005) aplica un modelo de interacción social con múltiples grupos de referencia a la probabilidad de la posguerra de empleo de los veteranos de la Primera Guerra Mundial. Como se muestra, un grupo de referencia de múltiples estructuras es sólo una de muchas estructuras de interacción para que los efectos sociales puedan ser identificados. Lin (2007) utiliza datos detallados sobre los lazos de amistad para estimar la influencia de pares sobre los resultados de los estudiantes. Se pueden obtener estimaciones por separado de los efectos endógenos y exógenos sólo porque las redes de la amistad en los datos satisfacen la condición general de identificación, especificada por Bramoullé, *et al.* (2008). Estos autores muestran avances en la metodología teórica expuestas por Laschever (2005) y Lin (2007) en la estimación empírica de los efectos de pares.

Bramoullé, *et al.* (2008) utilizan una aplicación empírica sólo para ilustrar conceptos básicos del efecto de pares, debido a que el marco teórico que estudian es muy general. Consideran el consumo de actividades recreativas por estudiante de escuela secundaria, tales como la participación a deporte, arte, y organizaciones sociales. Las actividades recreativas de los alumnos se supone que no sólo dependen de sus propias características (edad, sexo, ingresos de los padres), sino también por las características de sus amigos (efectos exógenos) y actividades recreativas de los mismos (efectos endógenos). Este último efecto puede reflejar la conformidad o el placer que tiene el estudiante a participar en actividades

recreativas con amigos que también participan en estas. Además, mientras más actividades artísticas, deporte y clubes sociales estén disponibles en una escuela, es más probable que los alumnos de esta consuman actividades recreacionales (efectos correlacionados). Estos últimos efectos también pueden surgir de la formación endógena de grupos, por ejemplo, los estudiantes hiperactivos interactúan juntos y por tanto se observa que consuman el mismo (alto) nivel de actividades recreativas. Asimismo, muestran que las actividades recreativas de sus amigos tiene una influencia positiva y significativa en las actividades recreativas del estudiante. Por último, las características de los amigos, como la participación de los padres en el mercado laboral, afecta de manera positiva y significativa a las actividades recreativas de los estudiantes.

El análisis empírico desarrollado por De Giorgi, *et al.* (2010) se basa en una reciente elaboración de datos administrativos de los estudiantes de pregrado de la Universidad Bocconi. En estos, los grupos de pares están bien definidos por características institucionales. En el período abarcado, los estudiantes inicialmente se deben matricular en un programa común y al final del tercer semestre deben optar por especializarse en una de las dos principales áreas: Economía o Negocios. Durante los tres primeros semestres, todos los estudiantes toman nueve cursos obligatorios y asisten a conferencias en las clases asignadas aleatoriamente para cada curso. Este proceso de asignación aleatoria, naturalmente, define los grupos de pares que varían a nivel individual por cada estudiante y son, por tanto, inmune a la simultaneidad (*reflection problem*). Por otra parte, ya que la asignación a las clases es al azar, la endogeneidad en la formación del grupo de pares es excluida por construcción (Akerlof, 1997). Al considerar que los grupos de pares se encuentran parcialmente superpuestos, esto genera *pares excluidos* (pares de los pares), donde las características exógenas de tales pares excluidos son un conjunto natural de instrumentos para superar la potencial endogeneidad generado por la autoselección al grupo común (efecto correlacionado).

Algunos estudios recientes ya podrían implementar la estrategia de identificación planteada por De Giorgi, *et al.* (2010) sin necesidad de datos adicionales. Como Calvó-Armengol, *et al.* (2009) que se centran en la influencia de la posición de un jugador en una red, en particular en el rendimiento académico; Bayer, *et al.* (2009) estudian el comportamiento criminal donde las características de sus pares son ponderados por el tiempo que pasan en el mismo centro penitenciario. La identificación de De Giorgi, *et al.* (2010) es, por tanto, potencialmente mayor que los actuales enfoques empíricos para la estimación de los efectos de pares.

Los efectos de pares pueden explicar el comportamiento similar, o en algunas ocasiones, irracional de los agentes. Por ejemplo, Fowler & Christakis (2008) encontraron que las personas obesas forman agrupaciones en la red en todos los tiempos, y estos grupos son de hasta

tres niveles (el amigo del amigo del amigo). Boucher, *et al.* (2010) encuentran que la interacción con estudiantes de más edad suele tener un efecto negativo en su propia calificación y es independiente del efecto de los resultados de sus compañeros. De Giorgi, *et al.* (2010) ha demostrado que los compañeros pueden desviar los estudiantes en la selección de magíster en los que tienen ventaja en capacidad y/o habilidad. Los estudiantes escogen principalmente siguiendo a sus compañeros en contraste con su capacidad de graduado, con consecuencias negativas en el rendimiento académico, los salarios de entrada y la satisfacción en el trabajo.

Las interacciones sociales, y en particular la influencia endógena de los pares, es probable que sean determinantes cruciales en las opciones que los estudiantes realizan, por encima y más allá de su rendimiento observado. Al decidir si ir a la universidad, en que escuela particular asistir o en el campo que va a especializarse, los pares pueden proporcionar valiosa y, de otro modo, costosa información. Además, para muchos estudiantes, compañeros de clase también representan un grupo de referencia natural y el comportamiento conformista es muy común durante la adolescencia (Akerlof & Kranton, 2002). Alternativamente, podría ser simplemente una ganancia de utilidad de los que asisten a la escuela o cursos con los amigos.

A lo largo de la literatura, se ha concentrado en distintas metodologías en la identificación de los efectos de pares. La principal contribución que tiene este proyecto hacia la literatura recae, no sólo en identificar los efectos sociales, sino identificar (cuantificar) la *intensidad* de la relación entre los pares. De esta manera, se abre la posibilidad en identificar el efecto de pares de una manera alternativa, con mejorías notables en los resultados conforme a lo revisado en la literatura.

Una de las falencias que existe en la literatura al demostrar la evidencia de los efectos de pares consiste en tener una información completa y adecuada. Completa en el sentido de tener la totalidad de características de los individuos y adecuada en la información a analizar. Contar con una base de esta índole puede ser muy compleja de obtener. En dos publicaciones distintas, Bramoullé *et al.* (2008) y Fowler & Christakis (2008) han utilizado la información proveniente del Estudio Longitudinal Nacional de la Salud Adolescente (National Longitudinal Study of Adolescent Health o Add Health) que apunta investigar la salud física y mental adolescente, la cual proporciona información relevante para el estudio de los efectos de pares. Sin embargo, con la información presentada por Add Health hace problemática la identificación de los efectos sociales dado por la endogeneidad en la formación de los grupos. Una contribución relevante de este proyecto consiste en construir la base de datos, por medio de una encuesta (anexo B), con el propósito de estudiar el efecto de pares en los alumnos de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile. De esta manera, se tiene especificada una base de datos destinada al estudio de los efectos de pares.

CAPÍTULO 2

Modelo Lineal de Interacciones Sociales

2.1. Introducción

A pesar de la amplia literatura sobre el efecto de pares, la identificación de las interacciones sociales sigue siendo muy costosa debido a los dos problemas identificados por Manski (1993): la *endogeneidad* ya sea por sus compañeros o de la autoselección al grupo común (efecto correlacionado), y el *reflection problem* un caso particular de la simultaneidad que se plantea cuando se trata de inferir si el comportamiento de un individuo en algún grupo de estudio influye en el comportamiento de los individuos que componen el mismo. Refiérase a comportamiento como alguna acción, característica o conducta que pueda ser influenciado por sus pares, como por ejemplo, la selección de un magíster, el consumo de actividades recreacionales en un colegio o el rendimiento de los alumnos.

En los últimos años se ha innovado el estudio de los efectos de pares en que los estudiantes interactúan en grupos. La interacción por grupos consiste en que los individuos están particionados en grupos y son afectados por los individuos pertenecientes al mismo, pero no por los de afuera. La identificación de los efectos sociales por medio de interacción por grupos resuelven los dos principales problemas. Bramoullé, *et al.* (2008) son pioneros en considerar una extensión del modelo lineal en medias en que los individuos interactúan por medio de una red social. Se considera red social a una estructura social hecha de nodos (generalmente individuos u organizaciones) que están unidos por uno o más tipos específicos de interdependencia, como la amistad, valores, creencias, conflictos o comercio. Demuestran que los efectos endógenos y exógenos pueden ser identificados bajo la interacción por red. Una propuesta alternativa, De Giorgi, *et al.* (2010) trabajan con grupos de pares que se

encuentran *parcialmente superpuestos*. Por definición, los grupos de pares están parcialmente superpuestos si el set de pares de dos individuos que son pares no coinciden perfectamente¹. En el contexto que los grupos de pares no se superponen completamente, es posible identificar todos los parámetros del modelo lineal en medias de interacciones sociales.

Este capítulo consiste en mostrar la teoría relevante sobre los modelos econométricos que se han utilizado a lo largo de la literatura para estudiar el efecto de pares. Se hace hincapié en dos de ellos. El primero planteado por Manski (1993) ha servido de base para el estudio econométrico de los efectos de pares, donde formaliza la identificación de los efectos sociales y no sociales que afectan el comportamiento del individuo. Este modelo se muestra en la sección que sigue. Los modelos de interacciones sociales que consideran la interacción por grupos de pares muestran resultados relevantes, en el sentido que solucionan los principales problemas en la identificación de los efectos sociales (endogeneidad y el *reflection problem*). En la sección 2.3 se muestra una extensión de este modelo, donde se trabaja con la superposición parcial de los grupos de pares, planteado por De Giorgi, *et al.* (2010). En la sección 2.4, se muestra el modelo propuesto que cuantifica la intensidad del efecto de pares, en términos de amistad y estudio. Aquí, se hace especial énfasis en cómo se innova el modelo lineal en medias para que considere la *intensidad* en la relación.

2.2. Modelo de Efectos Sociales Endógenos

Un desafío clave en la literatura empírica sobre los efectos de pares es lo que impulsa a identificar la correlación entre los resultados de los estudiantes que interactúan entre sí. Manski (1993) trata de distinguir tres hipótesis para explicar que el comportamiento de agentes pertenecientes al mismo grupo tienden a comportarse de manera similar. Se pueden distinguir entre las que siguen.

- *Efecto Endógeno*: La propensión de un individuo a comportarse de alguna manera, varía de acuerdo con el comportamiento del grupo.
- *Efecto Exógeno*: El comportamiento de un individuo, de alguna manera, varía con las características externas del grupo.

¹Para aclarar esta definición, supongamos un set de cuatro estudiantes: A, B, C y D. Los alumnos A, B y C estudian juntos, y además B, C y D también lo hacen. Se tiene que el grupo de pares de B es {A,C,D} mientras que el grupo de pares de C es {A,B,D}. De esta manera, los grupos de pares de B y C se superponen parcialmente, ya que sólo coinciden en los alumnos A y D.

- *Efecto Correlacionado*: Los individuos en el mismo grupo tienden a comportarse de manera similar, debido a que tienen características similares o se enfrentan a un entorno común.

A manera de ejemplo, considere el rendimiento de los estudiantes en un colegio. Hay un *efecto endógeno* si el rendimiento de un alumno tiende a variar con el rendimiento medio de los otros estudiantes (compañeros de curso o amigos). Hay un *efecto exógeno* si el rendimiento de un estudiante tiende a variar según la composición socioeconómica. Hay un *efecto correlacionado* si el rendimiento de varios estudiantes es similar debido a que tienen los mismos profesores o tienen similares antecedentes familiares.

2.2.1. Especificación del Modelo

El modelo lineal de interacciones sociales propuesto por Manski (1993) se define para un valor $y \in \mathbb{R}$ de la siguiente forma:

$$y = \alpha + \beta \mathbb{E}[y|\mathbf{z}] + \mathbb{E}[\mathbf{x}|\mathbf{z}]' \boldsymbol{\gamma} + \mathbf{x}' \boldsymbol{\eta} + u, \quad (2.1)$$

el cual considera el supuesto $\mathbb{E}[u|\mathbf{x}, \mathbf{z}] = \mathbf{z}' \boldsymbol{\delta}$ y se tiene que $\boldsymbol{\theta} = (\alpha, \beta, \boldsymbol{\gamma}, \boldsymbol{\eta}, \boldsymbol{\delta})$ es el vector de parámetros. Se define:

y : Resultado u Output.

Es un valor escalar. En el contexto de los estudiantes denota el rendimiento, como por ejemplo, las notas de un curso en un determinado semestre.

\mathbf{z} : Atributos del Grupo.

Son las características que identifican el grupo del individuo, de dimensión $J \times 1$. Estas características pueden ser el curso del estudiante o grupo étnico.

\mathbf{x} : Atributos Individual.

Son las variables que afectan directamente a y , de orden $K \times 1$. Por ejemplo, se tendría la situación socioeconómica del estudiante.

u : Atributo Individual no-observable.

Son los atributos no-observables que afectan directamente a y , como por ejemplo, la habilidad individual.

Luego, el Modelo Lineal de Interacciones Sociales se caracteriza por las variables observables $(y, \mathbf{z}, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^J \times \mathbb{R}^K$, y por la variable no observable $u \in \mathbb{R}$.

El modelo a estimar viene dado por el comportamiento esperado de un individuo, obtenido como la esperanza de y condicionado a los atributos que lo caracterizan más los atributos que identifican a su grupo, dado por el vector (\mathbf{x}, \mathbf{z}) :

$$\mathbb{E}[y|\mathbf{x}, \mathbf{z}] = \alpha + \beta \mathbb{E}[y|\mathbf{z}] + \mathbb{E}[\mathbf{x}|\mathbf{z}]' \boldsymbol{\gamma} + \mathbf{x}' \boldsymbol{\eta} + \mathbf{z}' \boldsymbol{\delta}, \quad (2.2)$$

la cual tiene una forma lineal. En esta última expresión es posible identificar los tres efectos sociales:

- Si $\beta \neq 0$ hay un *efecto endógeno*. El comportamiento del individuo varía de manera directa con $\mathbb{E}[y|\mathbf{z}]$, el comportamiento medio del grupo el cual pertenece con atributos \mathbf{z} .
- Si $\boldsymbol{\gamma} \neq \mathbf{0}$ hay un *efecto exógeno*. El comportamiento del individuo varía de manera directa con $\mathbb{E}[\mathbf{x}|\mathbf{z}]$, las características medias de las personas que conforman el grupo con atributos \mathbf{z} .
- Si $\boldsymbol{\delta} \neq \mathbf{0}$ hay un *efecto correlacionado*. Los atributos \mathbf{z} , que identifican al grupo del individuo, afectan de manera directa el comportamiento de este: $\mathbb{E}[y|\mathbf{x}, \mathbf{z}]$.
- El parámetro $\boldsymbol{\eta}$ cuantifica el efecto directo de las propias características del individuo \mathbf{x} en el comportamiento y .

2.2.2. Identificación de Parámetros

El objetivo de esta sección consiste en identificar el vector de parámetros $\boldsymbol{\theta}$ en el modelo (2.2), vale decir, consiste en distinguir los tres tipos de efectos sociales entre uno del otro, y sobre los efectos no sociales. Manski (1993) ha planteado este objetivo como interrogante, dado a que se requieren de los siguientes supuestos para una correcta identificación:

- i. Las variables (\mathbf{x}, \mathbf{z}) sean discretas; o,
- ii. Las variables y y \mathbf{x} tengan varianza finita, y los regresores $(\mathbb{E}[y|\mathbf{x}, \mathbf{z}], \mathbb{E}[y|\mathbf{z}], \mathbb{E}[\mathbf{x}|\mathbf{z}])$ deben ser continuas en el soporte de (\mathbf{x}, \mathbf{z}) .

Con esto se permite que las estimaciones para cada regresor sea consistente en el soporte del vector (\mathbf{x}, \mathbf{z}) , aquel conjunto de características que identifica al individuo. Luego, podemos dar por conocidos estos regresores, y se pondrá énfasis en la estimación de parámetros.

El *reflection problem* se hace presente en (2.2). Se tiene que para explicar el comportamiento esperado de un individuo, dado por $\mathbb{E}[y|\mathbf{x}, \mathbf{z}]$, existe la presencia de $\mathbb{E}[y|\mathbf{z}]$ como regresor, y ambos términos no pueden ser determinados al mismo tiempo. Esto es producto a que el regresor $\mathbb{E}[y|\mathbf{z}]$ denota el comportamiento medio del grupo \mathbf{z} donde está contenido el individuo y . Luego, no se puede concluir si es el individuo que determina al grupo o es el grupo que determina al individuo (he aquí donde este problema obtiene su nombre). Con esto, se dificulta la identificación de los efectos endógenos de los efectos exógenos. Una solución propuesta por Manski (1993) consiste en marginalizar la expresión (2.2) respecto a la distribución de $(\mathbf{x}|\mathbf{z})$. De esta manera y teniendo que $\beta \neq 1$, es posible obtener la siguiente expresión:

$$\mathbb{E}[y|x] = \frac{\alpha}{1-\beta} + \mathbb{E}[\mathbf{x}|\mathbf{z}]' \frac{\gamma + \boldsymbol{\eta}}{1-\beta} + \mathbf{z}' \frac{\boldsymbol{\delta}}{1-\beta}, \quad (2.3)$$

resolviendo el problema y posibilitando la estimación de parámetros en este nuevo modelo. Sin embargo, esta es una solución parcial, debido a que se omite el efecto endógeno. Además se tiene que $\mathbb{E}[y|\mathbf{z}]$ es en función lineal de sus regresores y con esto, el efecto endógeno no se puede distinguir del efecto exógeno o del efecto correlacionado. En consecuencia, el vector de parámetros $\boldsymbol{\theta}$ no puede ser identificado.

En la sección que sigue, se presenta una extensión del modelo lineal en medias en donde los individuos interactúan en grupos de pares. Este planteamiento es prometedor ya que permite resolver los principales problemas: endogeneidad y *reflection problem*.

2.3. Modelo de Efecto de Pares con Interacción por Grupos

A lo largo de la literatura, el reconocimiento que los individuos interactúan en grupos ha conllevado resultados satisfactorios en la identificación de los parámetros del modelo lineal en medias. El modelo que se verá en esta sección se considera como una extensión del modelo de interacción social en donde cada individuo tiene un grupo específico de referencia, definido por todos los individuos cuya respuesta media y característica influye en su propio resultado. Relajando el supuesto de interacción por grupo, generalmente permite separar entre los efectos endógenos de los exógenos.

Los supuestos del modelo son: *lineal en medias* y los individuos interactúan en grupos, vale decir, la población está particionada en grupos. Estos grupos son conformados de tal manera que los individuos son afectados por todos los que pertenecen a su grupo y por ninguno fuera del mismo. A esta consideración de grupo de pares se denomina *grupo de*

referencia (Bramoullé *et al.* 2008, Boucher *et al.* 2010, De Giorgi *et al.* 2010).

Los principales problemas que se analizaron en la sección 2.2 sobre la identificación de los parámetros en el modelo lineal en medias de interacciones sociales, vienen dados por la endogeneidad y el *reflection problem*. La endogeneidad nace por la propia selección que hacen los pares, los que ya se conocen se eligen entre ellos, y el *reflection problem*, el particular caso de simultaneidad analizado por Manski (1993). En la sección que sigue se expone la manera propuesta en resolver estos problemas.

De aquí en adelante, se centra las ideas de las interacciones sociales en el contexto de estudiantes en un establecimiento educacional.

2.3.1. Superposición Parcial de Grupos de Pares

El efecto de pares nace por la imitación, y es incierto que la definición de grupos debe ser limitado a los amigos cercanos, ya que no sólo son ellos los que afectan el comportamiento del individuo. En el sistema educacional, la definición genérica de pares está dada por los estudiantes que toman los mismos cursos, y en este proyecto se quiere capturar la red donde los estudiantes interactúan académica y socialmente. Esta toma de cursos influye fuertemente en como, donde y con quién cada estudiante gasta su mayor parte del tiempo.

Formalmente, existen tres definiciones de pares basadas en la cantidad de cursos en que los alumnos realizan juntos. La primera consiste en la *looser definition*, que es considerado como par aquellos alumnos que al menos han sido compañeros en un curso. La segunda, son considerados pares aquellos alumnos que cursaron juntos n de los m cursos en total ($n \leq m$). Por último, existe la posibilidad de asignar un peso relativo a la cantidad de cursos en la que dos estudiantes son compañeros.

A manera de contrastar estas definiciones, considere el grupo de pares del individuo i denotado por \mathcal{P}_i , incluye todos los individuos j que fueron asignados a la misma clase que el individuo i por al menos n cursos que se consideran. Luego por cada $j \in \mathcal{P}_i$ se asigna un *importance weight* $\omega_{ij} \in (0, 1]$ respectivo a los cursos de j tomado junto a i , definido como:

$$\omega_{ij} = \begin{cases} 1, & \text{si } j \text{ asistió a todos los cursos junto a } i; \\ k/m, & \text{si } j \text{ asistió a } k \text{ de los } m \text{ cursos con } i. \end{cases}$$

y estos pesos son normalizados dentro de cada grupo. Con esta definición cabe la posibilidad de encontrar grupos de menor tamaño en comparación a la *looser definition*, lo que se tendría grupos de menor variabilidad y abriendo la posibilidad de encontrar estimaciones no consistentes.

La superposición parcial de los grupos de pares es un supuesto factible. Todos tenemos amigos, compañeros, colegas, que quizás coincidan algunos, pero muy raro sería que coincidan de manera perfecta. De Giorgi, *et al.* (2010) obtienen como resultado que los grupos de pares que no se superponen completamente, es posible resolver el problema de endogeneidad y el *reflection problem*, lo que permite la identificación de todos los parámetros del modelo lineal en medias.

La idea detrás de la solución a estos problemas es la siguiente. Con la superposición parcial de los grupos se generan pares de pares, denominados pares excluidos, que actúan como restricciones de exclusión en las ecuaciones simultaneas en el modelo de interacciones sociales, y con esto se resuelve el *reflection problem*. Por otra parte, teniendo grupos de pares que varían a nivel individual, garantiza la presencia de compañeros de clase excluidos, *i.e.*, estudiantes que no fueron compañeros del estudiante i , pero sí fueron compañeros de alguno de los pares de i . Las características exógenas de los pares excluidos surgen como un set natural de instrumentos, debido a que estas están correlacionadas con el comportamiento del par por medio de las interacciones sociales, y no están correlacionado con el shock individual del grupo. Estos pueden ser utilizados para resolver el problema de endogeneidad generada producto del efecto correlacionado.

Por tanto, la clave en la identificación son las especificaciones individuales de los grupos de pares y el uso de las características de los pares excluidos como instrumentos.

2.3.2. Especificación del Modelo

Hasta el momento, se ha evaluado el modelo lineal en medias para explicar el comportamiento de un alumno. Esto puede ser llevado para explicar, por ejemplo, la elección de un magíster: Economía o Negocios (De Giorgi, *et al.* 2010). El modelo de interacción por grupos es una extensión del modelo lineal en medias (2.1):

$$y_i = \alpha + \beta \mathbb{E}[y|\mathcal{P}_i] + \gamma \mathbb{E}[\mathbf{x}|\mathcal{P}_i] + \boldsymbol{\delta} \mathbf{x}_i + u_i, \quad (2.4)$$

donde se considera el supuesto que $\mathbb{E}[u_i|\mathcal{P}_i, \mathbf{x}_i] = 0$, *i.e.*, no hay efectos correlacionados producto de la propia selección a los grupos, y se tiene que $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \boldsymbol{\delta})$ es el vector de parámetros. En este modelo se define,

- y_i : Selección del Magíster (Economía o Negocios).
- \mathbf{x}_i : Set de tratamientos individuales.
- \mathcal{P}_i : Grupo de pares del individuo i .
- u_i : Variable individual no-observable.

Con la formulación (2.4) los grupos de pares son fijos por individuos. Esto quiere decir que si los respectivos grupos de pares de i y j coinciden, *i.e.*, $\mathcal{P}_i = \mathcal{P}_j$ para algún $i \neq j$, conlleva que el efecto endógeno no puede ser distinguido del efecto exógeno. En la práctica, es difícil tener que dos individuos coincidan en sus grupos de pares, sin que ellos no sean pares. Por tanto, el requerimiento $\mathcal{P}_i \neq \mathcal{P}_j$ es factible.

El modelo a estimar viene dado por la esperanza condicional a las características que definen al individuo, es decir, condicionando al grupo de pares y a las variables exógenas que caracterizan al individuo:

$$\mathbb{E}[y_i | \mathcal{P}_i, \mathbf{x}_i] = \alpha + \beta \mathbb{E}[y | \mathcal{P}_i] + \gamma \mathbb{E}[\mathbf{x} | \mathcal{P}_i] + \delta \mathbf{x}_i. \quad (2.5)$$

En la ecuación (2.5) se tiene que el término $\mathbb{E}[\mathbf{x} | \mathcal{P}_i]$ es la media de las variables exógenas en el grupo de pares de i , y el término $\mathbb{E}[y | \mathcal{P}_i]$ es la elección media en el grupo de pares de i . Es decir, γ cuantifica el *efecto exógeno* y β cuantifica el *efecto endógeno*.

2.3.3. Identificación de Parámetros

Los principales problemas en la identificación del vector de parámetros θ recae en la endogeneidad y el *reflection problem*. El primero de ellos nace por al menos dos razones:

1. Las personas escogen a sus pares endógenamente.
2. Shocks comunes no observables: Efecto de los profesores sobre los alumnos pueden crear grupos.

Como consecuencia, si se observa una correlación significativa entre el resultado individual y grupal, es difícil decir que este resultado es por el verdadero efecto de pares o por la simple formación endógena de los grupos (efecto correlacionado). El segundo problema, el particular caso de simultaneidad, nace por que en un grupo de pares el comportamiento de todos afecta a los otros del grupo, como un reflejo en un espejo. No se puede saber si la acción de uno es la causa o el efecto de las acciones de los pares.

La solución del *reflection problem* en la expresión (2.5) propuesta por Manski (1993) consiste en marginalizar esta expresión respecto a la densidad $(\mathbf{x} | \mathcal{P}_i)$, lo que lleva a obtener que $\mathbb{E}[y | \mathcal{P}_i]$ es una combinación lineal de los siguientes regresores:

$$\mathbb{E}[y | \mathcal{P}_i] = \frac{\alpha}{1 - \beta} + \frac{\gamma + \delta}{1 - \beta} \mathbb{E}[\mathbf{x} | \mathcal{P}_i], \quad (2.6)$$

siendo una solución parcial ya que esta regresión sólo considera el efecto exógeno. Otra solución propuesta dada por De Giorgi, *et al.* (2010) consiste en aprovechar la ventaja sobre que

los individuos interactúan en grupos parcialmente superpuestos. De esta manera, se puede reescribir la ecuación (2.6) permitiendo que los grupos de pares varíen a nivel individual:

$$\mathbb{E}[y_i|\mathcal{P}_i] = \alpha + \beta \mathbb{E}[\mathbb{E}[y|\mathcal{P}_j] | \mathcal{P}_i] + \gamma \mathbb{E}[\mathbb{E}[\mathbf{x}|\mathcal{P}_j] | \mathcal{P}_i] + \delta \mathbb{E}[\mathbf{x}_i|\mathcal{P}_i],$$

donde j es un miembro del grupo de pares de i . Se puede escribir el problema de esta forma dado que el grupo de pares de j nunca coincide con i , *i.e.*, se requiere que $\mathcal{P}_i \neq \mathcal{P}_j$ para todo $i \neq j$.

Con una superposición perfecta, $\mathbb{E}[y|\mathcal{P}_i]$ no varía a nivel individual, y sería una constante para todos los miembros del mismo grupo. Con esto, se impide la identificación separada de los efectos exógenos y endógenos, inclusive si los grupos fueran formados de manera aleatoria y ante la ausencia del efecto correlacionado.

2.4. Modelo en *Intensidad* del Efecto de Pares

En esta sección se muestra el modelo con el cual se propone utilizar para cuantificar la *intensidad* del efecto de pares en la red social de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile. Este modelo será presentado como una extensión del modelo lineal en medias de Bramoullé, *et al.* (2008), debido a que es el que mejor se adecúa a los datos obtenidos por la encuesta y se ha modificado dicho modelo para que considere la intensidad en la relación. Este modelo *propuesto* se basa en tres hipótesis fundamentales:

1. Los individuos interactúan en grupos conocidos por el investigador, *i.e.*, la población de estudiantes está dividida en grupos.
2. Los estudiantes se ven afectados por todos los individuos que pertenecen a su grupo, pero por ninguno fuera de este.
3. Los individuos pueden *rankear* a sus compañeros dentro de su grupo de pares, según la intensidad en la relación.

Por tanto, con estas hipótesis se tiene especificado una estructura de grupo muy particular, lo que conlleva a un modelo único en la literatura de efectos de pares. El modelo de interacciones sociales definido permite que cada estudiante elija su grupo de referencia y pueda *rankear* a sus pares dentro de este, con la finalidad de cuantificar la intensidad en la relación. Además se considera que cada alumno tiene su propio *grupo de referencia*, definido por las personas cuyo rendimiento medio y características influyen en su resultado. Este supuesto es típico

en los estudios de rendimiento académico, pero se deduce claramente de las limitaciones de datos *ad-hoc*.

El rendimiento individual y grupal hace referencia a las notas medias obtenidas en el último semestre cursado en la Universidad. Estas hipótesis inferen que el rendimiento individual es determinado por un modelo lineal en medias, en donde, el rendimiento en el último semestre cursado de un estudiante se ve afectado por sus características, por el rendimiento medio y características medias de su grupo de pares.

La estructura de un grupo múltiple de referencia es una de las tantas estructuras en donde los efectos sociales son identificados. Dada por la especificación de la base de datos de este proyecto, es posible realizar estimaciones separadas de los efectos exógenos y endógenos, ya que se cumplen las condiciones generales de identificación. Por medio de la metodología en la captura de la interacción entre los pares, cabe la posibilidad de innovar esta metodología para la captura de la intensidad en la relación. En la sección 2.4.2 se expone la manera propuesta de innovar el modelo, que es uno de los puntos relevantes de este proyecto.

2.4.1. Especificación del Modelo

El modelo de interacciones sociales definido en esta sección permite que cada estudiante elija su grupo de referencia y pueda *rankear* a los compañeros dentro de este, con el propósito de cuantificar la intensidad en la relación entre pares. El modelo se define como sigue. Supongamos que tenemos un set \mathcal{P} de estudiantes de tamaño n . Cada alumno i tiene su grupo de referencia \mathcal{P}_i de tamaño n_i , dado para todo $i = 1, \dots, n$. Para nuestros propósitos, se definen,

y_i : Rendimiento en el último semestre del alumno i , definido por la nota promedio obtenida en ese semestre.

\mathbf{x}_i : Vector de características del alumno i , de orden $1 \times k$, tales como el ingreso per cápita en la familia, educación de los padres, etc².

El grupo de referencia contiene a todos los estudiantes en que su rendimiento en el último semestre o alguna característica puede ser influido por el rendimiento del estudiante. Además, se supone que el individuo i es excluido de su grupo de referencia, *i.e.*, $i \notin \mathcal{P}_i$. De aquí podemos definir que un alumno se encuentra *aislado* si su grupo de referencia está vacío. La exclusión del individuo de su grupo de referencia y la superposición parcial de los grupos

²Revisar la sección A.1 ver el totalidad de estas características, junto a la descripción estadística de ellas.

de pares son los puntos claves para la solución del *reflection problem* (sección 2.3.1).

Para este modelo, se asume que el rendimiento de un estudiante puede ser afectado por el rendimiento medio de su grupo de referencia, y por las características medias de sus pares. Formalmente, el modelo de manera estructural se define como:

$$y_i = \alpha + \beta \frac{\sum_{j \in \mathcal{P}_i} y_j}{n_i} + \gamma \mathbf{x}_i + \delta \frac{\sum_{j \in \mathcal{P}_i} \mathbf{x}_j}{n_i} + \varepsilon_i,$$

en donde el término que acompaña a β denota el rendimiento medio de los pares en el grupo de referencia \mathcal{P}_i y el término que acompaña a δ denota las características medias de los pares de i . En este contexto,

- Parámetro β captura el *efecto endógeno*: El rendimiento del alumno i varía con el rendimiento medio del grupo de referencia \mathcal{P}_i .
- Parámetro δ captura el *efecto exógeno*: El rendimiento del alumno i varía con las características medias del grupo de referencia \mathcal{P}_i .
- El parámetro γ cuantifica el efecto directo de las características del alumno en su rendimiento.

La manera de tratar el efecto correlacionado es por medio de la estructura del grupo de pares. De manera natural surge un set de instrumentos dados por las características exógenas de los pares excluidos. Se consideran, además, los supuestos:

Exogeneidad : $\mathbb{E}[\varepsilon_i | \mathbf{x}_i] = 0$.

Requerimiento : $|\beta| < 1$.

Excepto de este último, no se impone ninguna otra restricción en el modelo. El término de error ε_i refleja las características no-observables para i , y se asume una exogeneidad estricta por parte de los regresores \mathbf{x} . No se requieren más supuestos sobre este término dentro de una red, especialmente no es necesario que el error sea *i.i.d.* Normal. En otras palabras, este es un modelo semiparamétrico.

2.4.2. Interacción por Grupo

La interacción por grupo ha conllevado a resultados satisfactorios en la identificación de los efectos de pares. La manera de capturar la interacción entre los pares es por medio de

la *Interaction Matrix*. Esta se define como la matriz que contiene los pesos relativos a la cantidad de pares que tiene cierto alumno en su grupo de referencia.

Para Bramoullé, *et al.* (2008) definen G como la *interaction matrix*, aquella que contiene los pesos relativos a la cantidad de pares que tiene cierto alumno, dada por:

$$G_{ij} = \begin{cases} 1/n_i, & \text{si } j \text{ es amigo de } i; \\ 0, & \text{en otro caso.} \end{cases}$$

A *priori*, estos autores les asignan el mismo peso relativo a todos los pares del alumno, sin ninguna distinción de por medio. En este proyecto se diferencia a los otros modelos en que se asigna un peso relativo a la intensidad en la amistad entre los pares, debido a que poseemos dicha información sobre las preferencias entre los pares. Se propone definir la *interaction matrix* F como sigue:

$$F_{ij} = \begin{cases} \frac{p_{ij}}{n_i}, & \text{si } j \text{ es amigo de } i; \\ \sum_{j=1} p_{ij} & \\ 0, & \text{en otro caso,} \end{cases}$$

en donde p_{ij} denota la intensidad en amistad sobre el alumno j por parte de i . La matriz propuesta F mantiene las propiedades esenciales de la *interaction matrix* en que cada valor sea menor o igual a 1, y la suma por cada fila sea 1.

Observación Se tiene que $p_{ij} \in \{1, 2, \dots, n_i\}, \forall i, j$. Por tanto la suma de estos valores queda como,

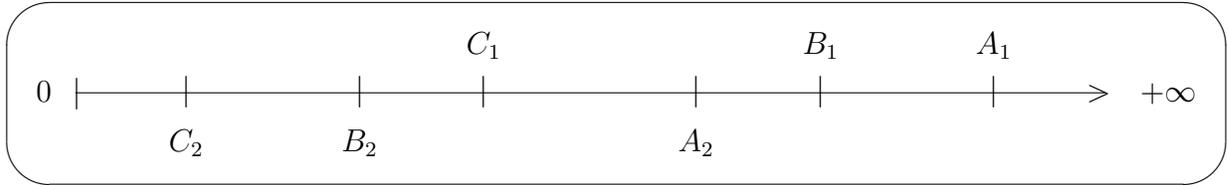
$$\sum_{j=1}^{n_i} p_{ij} = \frac{n_i(n_i + 1)}{2}.$$

Luego, la *interaction matrix* F puede ser reescrita como:

$$F_{ij} = \frac{p_{ij}}{\frac{n_i(n_i+1)}{2}} = \frac{2p_{ij}}{n_i(n_i + 1)}, \quad \text{si } j \text{ es amigo de } i.$$

Se puede ver que la *interaction matrix* es una medida de distancia normalizada entre la relación de amistad del alumno con cada uno de sus pares. Dada la especificación de la matriz G se tiene que todos los pares del alumno tienen el mismo grado de amistad y por tanto le asigna el mismo peso relativo a cada uno. Con la matriz F se trata que el amigo más lejano del alumno influye de menor forma en comparación al más amigo. Por tanto, de esta forma, el amigo más lejano en amistad tiene un peso relativo menor que en comparación

al amigo más cercano en amistad. Se debe tener en consideración que el alumno sólo puede *rankear* a sus pares según el nivel de amistad, pero no considera la *magnitud* de cuanto lo afecta. Por ejemplo, supongamos que se tienen dos alumnos con sus respectivos pares, ordenados de mejor a menor amigo: $\mathcal{P}_1 = \{A_1, B_1, C_1\}$ y $\mathcal{P}_2 = \{A_2, B_2, C_2\}$, y supongamos que hay una métrica en donde se pueden ordenar a estos pares según su influencia en el alumno (de menor a mayor):



En la parte superior se encuentran los pares del alumno 1 e inferior los pares de 2. Se puede ver que para ambos grupos se mantiene la relación de amistad, sin embargo es clara la diferencia de amistad. Se tiene que A_1 y B_1 son más amigos del alumno 1 en comparación de A_2 para el alumno 2. Además, para el alumno 1 hay una mejor relación de amistad con C_1 en comparación de B_2 y C_2 para el alumno 2. La manera de resolver esto sería por medio que los alumnos fueran capaces, por ejemplo, de evaluar a sus pares con notas de 1 a 10 según su intensidad en la amistad, donde 10 sería la máxima relación en amistad y 1 sería la mínima. De esta forma se tendría una información más detallada sobre la relación entre pares. Este punto se deja propuesto para estudios posteriores.

Una vez definida la *interaction matrix*, se procede a definir el modelo de manera estructural. Sin hacer previa selección entre G y F , digamos que $M = \{G, F\}$ es la *interaction matrix*, el modelo de manera matricial queda como:

$$Y = \alpha \mathbf{1} + \beta MY + \gamma \mathbf{X} + \delta M\mathbf{X} + \varepsilon, \quad (2.7)$$

en donde,

Y : Denota el rendimiento del alumno;

\mathbf{X} : Denota las características del alumno;

MY : Denota el rendimiento *medio* de los pares y,

$M\mathbf{X}$: Denota las características *medias* de los pares,

es decir, β cuantifica el efecto endógeno y δ el efecto exógeno. El término $\mathbf{1}$ denota un vector de unos y se considera,

$$\mathbb{E}[\varepsilon|\mathbf{X}] = \mathbf{0} \quad \text{y} \quad \mathbb{E}[\varepsilon\varepsilon'|\mathbf{X}] = \Sigma.$$

La identificación del vector de parámetros $\theta = (\alpha, \beta, \gamma, \delta)$ recae en la restricción de exogeneidad y en restricciones sobre la *interaction matrix*.

Ejemplo

A manera de clarificar todos los conceptos analizados en esta sección, consideremos el siguiente ejemplo. Supongamos que se tiene un set de cuatro estudiantes $\mathcal{P} = \{A, B, C, D\}$, donde se tienen los siguientes grupos pares con su respectiva intensidad en la relación:

$$\mathcal{P}_A = \{C, B\}, \quad \text{con } n_A = 2.$$

$$\mathcal{P}_B = \{C\}, \quad \text{con } n_B = 1.$$

$$\mathcal{P}_C = \{D, A, B\}, \quad \text{con } n_C = 3.$$

$$\mathcal{P}_D = \{B, A\}, \quad \text{con } n_D = 2.$$

Seguendo a Bramoullé, *et al.* (2008) consideran la *interaction matrix* G asignando el mismo peso a cada par respecto a la cantidad de pares que tenga el alumno:

$$G = \begin{array}{c} \begin{array}{c} A \\ B \\ C \\ D \end{array} \begin{array}{c} A \quad B \quad C \quad D \\ \left[\begin{array}{cccc} 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/2 & 1/2 & 0 & 0 \end{array} \right] \end{array} \end{array}.$$

Si consideramos la intensidad en la relación de estos pares, asumiendo *a priori* que el mejor amigo afecta el rendimiento del alumno de manera más significativa que el segundo (y así sucesivamente), se asigna un peso relativo mayor al mejor amigo, continuando con el resto. De esta forma, la *interaction matrix* F queda como:

$$F = \begin{array}{c} \begin{array}{c} A \\ B \\ C \\ D \end{array} \begin{array}{c} A \quad B \quad C \quad D \\ \left[\begin{array}{cccc} 0 & 1/3 & 2/3 & 0 \\ 0 & 0 & 1 & 0 \\ 2/6 & 1/6 & 0 & 3/6 \\ 1/3 & 2/3 & 0 & 0 \end{array} \right] \end{array} \end{array}.$$

En este punto se quiere clarificar cómo operan las características de los pares excluidos como instrumentos. Se recuerda que los *pares excluidos*, o pares de los pares, son originados por la superposición parcial de los grupos de pares. Los pares excluidos están correlacionados con el rendimiento de los pares del alumno, pero no con el rendimiento del alumno. Por tanto, las características de los pares excluidos sirven como instrumentos para el rendimiento de los pares (efecto endógeno).

Las características de los pares excluidos se obtienen de la siguiente forma. Generalizando la *interaction matrix* como $M = \{G, F\}$, se tiene en cuenta que el producto $M\mathbf{X}$ denota las características media de los pares del alumno. Luego, de manera análoga, las características medias de los pares excluidos del alumno viene dado como las características medias de los pares de los pares, es decir,

$$M(M\mathbf{X}) = M^2\mathbf{X},$$

y de esta forma se tendría que $G^2\mathbf{X}$ y $F^2\mathbf{X}$ son instrumentos válidos para las variables GY y FY , respectivamente.

Continuando con el ejemplo, se define como \mathcal{P}_j^2 los pares excluidos para cada alumno $j = A, B, C, D$. Estos vendrían dados por:

$$\mathcal{P}_A^2 = \{\mathcal{P}_C, \mathcal{P}_B\} = \{\{D, A, B\}, \{C\}\}.$$

$$\mathcal{P}_B^2 = \{\mathcal{P}_C\} = \{\{D, A, B\}\}.$$

$$\mathcal{P}_C^2 = \{\mathcal{P}_D, \mathcal{P}_A, \mathcal{P}_B\} = \{\{B, A\}, \{C, B\}, \{C\}\}.$$

$$\mathcal{P}_D^2 = \{\mathcal{P}_B, \mathcal{P}_A\} = \{\{C\}, \{C, B\}\}.$$

Recordando que se mantiene la intensidad en la relación. Por una parte, la *interaction matrix* para los pares excluidos considerando que cada par tiene la misma influencia sobre el alumno queda como:

$$G^2 = \begin{matrix} & \begin{matrix} A & B & C & D \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{bmatrix} 1/6 & 1/6 & 3/6 & 1/6 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/6 & 2/6 & 3/6 & 0 \\ 0 & 1/4 & 3/4 & 0 \end{bmatrix} \end{matrix},$$

y por otra parte, haciendo la distinción entre la intensidad en la relación, vendría dada por:

$$F^2 = \begin{matrix} & \begin{matrix} A & B & C & D \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{bmatrix} 2/9 & 1/9 & 3/9 & 3/9 \\ 2/6 & 1/6 & 0 & 3/6 \\ 3/18 & 8/18 & 7/18 & 0 \\ 0 & 1/9 & 8/9 & 0 \end{bmatrix} \end{matrix}.$$

Si analizamos los pares excluidos de B , el cual es fácil entender dado que sólo son los pares de C como pares excluidos, se mantienen las relaciones previamente definidos. Vale decir, para G_B^2 se asigna el mismo peso a cada par excluido y para F_B^2 lo hace respecto a la intensidad en la relación. Para los otros pares excluidos, es complejo el entender la dispersión de estos pesos, dada la interacción de *segundo grado* (pares de los pares) entre estos. Eso sí, al comparar estas matrices, es evidente que se tienen valores muy distintos en la interacción de

los pares excluidos. Sin embargo, hacer una distinción *a priori* de estas matrices sin analizar los resultados es irrelevante.

2.4.3. Identificación de Parámetros

La identificación de los parámetros se realizará de manera genérica en la *interaction matrix* $M = \{G, F\}$, dado que ambas mantienen las propiedades esenciales. Escribiendo el modelo (2.7) de forma reducida, considerando que $|\beta| < 1$ y que $(I - \beta M)$ es invertible, se consigue que:

$$Y = \alpha(I - \beta M)^{-1}\mathbf{1} + (I - \beta M)^{-1}(\gamma I + \delta M)\mathbf{X} + (I - \beta M)^{-1}\varepsilon. \quad (2.8)$$

en donde el intercepto toma el valor α si el estudiante se encuentra aislado y $\alpha/(1 - \beta)$ si no. La matriz de varianzas-covarianzas de este nuevo término de error viene dado por $(I - \beta M)^{-1}\Sigma(I - \beta M)^{-1}$. Los efectos sociales son identificados *si y solo si* el vector de parámetros θ del modelo estructural puede ser únicamente recuperado de los parámetros de la forma reducida en (2.8).

La identificación de parámetros recae en el valor esperado de la media grupal de los amigos. Para encontrar este valor esperado, es posible utilizar series de expansión en (2.8) dado que $|\beta| < 1$. Por tanto, es posible reescribir $(I - \beta M)^{-1} = \sum_{k=0}^{\infty} \beta^k M^k$. Asumiendo que no hay alumnos aislados, el valor esperado queda como:

$$\mathbb{E}[MY|\mathbf{X}] = \frac{\alpha}{1 - \beta}\mathbf{1} + \gamma M\mathbf{X} + (\gamma\beta + \delta) \sum_{k=0}^{\infty} \beta^k M^{k+2}\mathbf{X}.$$

La identificación del vector θ recae en las propiedades en la *interaction matrix* y en el grupo de pares. Asumiendo que $\gamma\beta + \delta \neq 0$, la proposición 1 en Bramoullé, *et al.* (2008) nos dice que si las matrices I, M y M^2 son *linealmente independiente*, entonces los efectos sociales son identificados. En la segunda proposición dice que si al menos un grupo de par tiene distinto tamaño a los otros grupos, entonces los efectos sociales son identificados. Sin embargo, si estas matrices son *linealmente dependiente* o todos los grupos de pares tienen el mismo tamaño, entonces los efectos sociales no son identificados. Ambas proposiciones se cumplen de manera satisfactoria con los datos obtenidos, con lo cual se tienen indicios en la identificación de dichos efectos en la red social de la Universidad de Chile.

CAPÍTULO 3

Efecto de Pares en la Red Social de la Universidad de Chile

3.1. Introducción

La información para llevar a cabo este proyecto consiste en los datos estudiantiles a nivel administrativo y personal de alumnos provenientes de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile en el año 2010, enfocándose principalmente en cómo se relacionan entre ellos a nivel educacional. Con esta información se tiene toda la necesaria para la estimación del efecto de pares. La forma de capturar dicha información consiste en una encuesta, la cual los alumnos la respondieron de manera voluntaria, y es por esta misma razón que la cantidad de registros conseguidos no es significativa. No está demás en recalcar que este es un estudio piloto para una próxima publicación, dada la relevancia que tiene este proyecto hacia la literatura de efectos de pares, principalmente por las nuevas metodologías proporcionadas.

Se destaca el hecho que se analiza el efecto de pares en dos tipos distintos de grupo de referencia. El primero consiste en los compañeros con los cuales se juntan a estudiar y el segundo consiste en los pares con los cuales se reúnen para hacer vida social. Al primer grupo se le denomina *Estudio Grupal* y al segundo *Vida Social*.

Este capítulo consiste en el análisis de los efectos de pares utilizando el modelo lineal en medias revisado en el capítulo anterior. Previo al análisis de estos modelos, se realiza el estudio descriptivo de los datos para tener un conocimiento más profundo acerca del problema en cuestión. La presentación de este modelo comienza con la justificación de realizar este estudio utilizando la *intensidad* en la relación y luego estimar los efectos para los pares que

se encuentran en el grupo de *Estudio Grupal y Vida Social*, de manera independiente. Este capítulo finaliza con la exposición de los resultados del modelo que considera la intensidad en la relación para la estimación del efectos de pares, de manera conjunta para ambos grupos de referencia.

3.2. Encuesta Alumno

La manera de recopilar esta información es por medio de una encuesta preparada en conjunto con el profesor guía, presentada en el anexo B. Dentro de esta, se consideran preguntas¹ tales como: el distinto grado de amistad entre los compañeros de estudios; con quien estudia y quienes son los amigos con quien el alumno no estudia.

La encuesta realizada trata de maximizar la información que se puede obtener de los alumnos y sobre sus pares. Esta consta de tres secciones:

1. Preguntas sobre sus Pares

Se enfocan principalmente en los compañeros de relación académica (amigos de estudio) y con los cuales hace vida social.

2. Preguntas sobre el Entorno Familiar

Principalmente se pregunta sobre la situación socioeconómica y la composición de la familia.

3. Preguntas Personales

Como gustos, preferencias y contextura física.

Dentro del cuestionario, se consideran matrices de respuesta en las cuales el alumno encuestado tiene que ordenar de mayor a menor a los compañeros según la influencia en la relación de amistad. Luego, respecto a estos compañeros, tiene que responder preguntas respecto al estudio grupal y vida social, como por ejemplo, la cantidad de horas que dedicó al estudio con cierto compañero. De esta manera es posible tener un *ranking* según el nivel de amistad entre los compañeros influyentes en los estudios, y así cuantificar la intensidad en la relación entre pares. Con la aplicación de esta encuesta se pretende tener una base de datos original en la literatura.

¹Revisar la sección A.1 del Anexo para la completa descripción de las variables.

3.3. Análisis Descriptivo

En el Anexo A.1 se muestra la estadística descriptiva sobre el conjunto de variables obtenidos por la encuesta, dada en conjunto por tablas e histogramas presentadas en las tablas A.1-A.16. La totalidad de personas encuestadas fue de 129 donde la mayoría de estos son hombres (63.6%). El tamaño de los grupos de pares conformados, tiene un tamaño medio de 2.17 para los pares catalogados como de *Estudio Grupal* y 2.12 para los pares que hacen *Vida Social*, con una desviación estándar de 1.40 y 1.47, respectivamente.

Sobre la situación socioeconómica de la familia, se tiene que la mayoría de los padres, tanto padre como madre, poseen un nivel educacional superior completa. Además la mayoría de estos padres son profesionales con grado de licenciatura, y un porcentaje no menor sobre las madres se encuentran actualmente desempleadas. Se hace notar que la mayoría de las familias tienen un ingreso neto mayor a los tres millones de pesos chilenos, un 20.16%, seguida por las familias con un ingreso de entre los .75 - 1 millón (13.95%).

Con respecto a las preferencias de los encuestados, en ideología política no hay mucha diferencia a simple vista. Sin embargo lideran personas de izquierda, seguidos por aquellos que no tienen una ideología. Sobre las creencias religiosas, la gran mayoría son católicos y ateos/agnósticos. En tiempo libre prefieren hacer deportes e ir al cine, mientras que las actividades menos concurridas son ir al teatro e ir de compras.

Acerca del rendimiento de los alumnos, se tiene que la nota media en el último semestre fue de 5.3 con una desviación estándar de .68. Se hace notar que las notas van desde el 3.3 al 7.0. En las tablas A.12 y A.13 se muestra la distribución de las notas de los pares según la intensidad en la relación en estudio y recreación, dado por los grupos de *Estudio Grupal* y *Vida Social*, respectivamente. De manera inmediata se puede ver que hay una mayoría de individuos en el grupo de *Estudio Grupal*, pero la desviación estándar es mayor en el grupo de *Vida Social*. Acerca de la distribución de las notas, se tiene que en *Estudio Grupal* hasta el tercer compañero amigo hay una distribución simétrica en comparación al cuarto y quinto compañero. En *Vida Social* se podría decir que esta simetría se mantiene hasta el último compañero amigo. En ambos grupos se tiene que las notas medias se encuentran cerca del 5.1.

De la misma forma que las notas en el último semestre, se tiene la distribución sobre los puntajes obtenidos en la PSU. Con esta variable controlamos el rendimiento de los estudiantes *ex-ante* de ingresar a la Universidad y definir su grupo de estudio. En la tabla A.14 se muestra la distribución del puntaje para los alumnos, que no es simétrica y se tiene

una media considerablemente alta (749.0 puntos). Se hace notar que en ambos grupos de referencia (*Estudio Grupal* y *Vida Social*), y por cada intensidad en amistad, la distribución de los puntajes claramente no es simétrica, donde la mayoría se centra en los 700 puntos, siendo que la media en cada grupo es mayor.

3.3.1. Intensidad en la Relación entre Alumnos y Pares

En las figuras A.1-A.3, se muestran el rendimiento de los pares según los cinco grados de intensidad en la relación de estudio y/o recreación. En la figura A.1(a) se muestra el resultado para los pares en *Estudio Grupal*. El alumno es una influencia positiva hacia el rendimiento de sus pares, debido a que la nota del alumno es mayor a la nota de sus pares. Es sólo el cuarto compañero en amistad que tiene una influencia negativa. Acerca de los pares en *Vida Social*, en la figura A.1(b) se aprecia una tendencia decreciente respectiva a la intensidad hasta el cuarto compañero, lo cual a medida que decrece en la intensidad en la amistad, los amigos en *Vida Social* generan cada vez más una influencia negativa al rendimiento del alumno. Sin embargo, es sólo el quinto compañero que tiene una influencia positiva en el rendimiento del alumno².

Para un análisis más acabado de estos rendimientos, se han realizado este mismo estudio por quintiles. En la figura A.2 se tiene el análisis para los pares en *Estudio Grupal*. Se puede ver que en el primer quintil (aquellos alumnos con las menores notas) tiene una clara tendencia positiva respectivo a la intensidad, y la nota del alumno es menor a la de sus pares. Esto quiere decir que los pares generan una influencia positiva en el rendimiento del alumno y el amigo más cercano, en términos de amistad, es menos influyente en comparación al último amigo. Sin embargo, en el último quintil (donde los alumnos tienen los mejores rendimientos) se tiene una tendencia negativa. Luego, los amigos más cercanos en *Estudio Grupal* influyen de mayor manera en comparación al último compañero.

Sobre los pares en *Vida Social*, en la figura A.3 no se observa una clara tendencia en algún quintil. No obstante, en el primer y segundo quintil, donde los alumnos tienen el menor rendimiento, se observa que el alumno genera una influencia negativa en el rendimiento de sus pares. Por otra parte, se tiene que los alumnos generan una influencia positiva sobre los pares de alto rendimiento.

²Dentro de los pocos pares que conforman el quinto compañero, existe un alumno que sólo cursó un ramo con nota de 7.0 escapándose notablemente de la media, con lo cual se podría considerar como un *outlier* en el análisis. Omitiendo a esta persona la nota media para el quinto compañero en *Vida Social* quedaría como 5.1, teniendo un resultado más representativo.

3.4. RESULTADOS

3.4.1. Relación Lineal sobre el Rendimiento Medio

A manera de tener un conocimiento más profundo sobre la relación entre el rendimiento del alumno y sus pares, se han realizado gráficos de dispersión junto a una regresión lineal sobre las notas medias de estos compañeros. Los resultados son de gran ayuda para conocer el comportamiento de estas variables al momento de estimar el modelo de efectos de pares visto en la sección 2.4.

En las figuras A.6-A.7 se muestra la dispersión sobre las notas medias, y en las figuras A.4-A.5 se muestra la dispersión de las notas según la intensidad en la relación. Antes de discutir el estudio por medias de los grupos, se muestran los resultados sobre la distinta intensidad en la amistad dado que se quiere hacer un análisis más acabado respecto a la manera de obtener *media* grupal.

La dispersión de las notas respecto a la distinta intensidad en la amistad, en la figura A.4 se muestra una clara relación positiva y significativa para los pares en *Estudio Grupal*, hasta el tercer compañero. No obstante, para los últimos dos no se tiene una clara tendencia, y los parámetros respectivos en cada regresión no son significativos. Acerca de los pares en *Vida Social*, figura A.5, se tiene que no hay una relación explícita entre estos rendimientos en ningún grado de intensidad. Estos resultados dan indicios en concluir que los pares en *Estudio Grupal* buscan estudiar con compañeros que tengan un rendimiento similar, y en cambio a los pares que se encuentran en *Vida Social* se agrupan independiente al rendimiento.

3.4.1.1. Intensidad en la Relación

Este apartado tiene como objetivo definir el tipo de media a considerar. En la sección 2.4.2 se discutió dos formas de obtener la media respectiva a cada grupo de par. La primera consiste en la media que usualmente conocemos, aquella que le asigna el mismo peso a los pares que están en el grupo de referencia, definido por la matriz G . La segunda viene dado según la intensidad en la amistad, en donde al compañero más amigo se le asigna un peso relativo mayor que al segundo, y este mayor al tercero y así consecutivamente, determinada por la matriz F . Dado que se tiene conocimiento de la intensidad en la amistad en el grupo de referencia, es posible crear la matriz F .

El análisis sobre el rendimiento medio de los pares depende de la definición de media. Se han realizado las estimaciones según el tipo de compañero (académica o social) y respecto a cada definición de media. Los modelos de manera estructural son:

$$Y = \alpha + \beta GY + \varepsilon, \quad (3.1a)$$

$$Y = \alpha + \beta FY + \varepsilon, \quad (3.1b)$$

en donde Y denota matricialmente el rendimiento, y los productos GY y FY denotan las medias respectivas a las definiciones, recordando que el rendimiento del alumno es excluido de su grupo de referencia.

Los resultados de los modelos (3.1) se encuentran en las figuras A.6-A.7, respectivamente. A simple vista no se ve una diferencia significativa entre los resultados obtenidos por las distintas medias. Aún así, se percibe una clara relación lineal positiva entre las notas de los alumnos y sus pares en *Estudio Grupal* (figura A.6). Además se tiene que ambos modelos son estadísticamente significativos. La nota media de los pares en este grupo influye de manera positiva (significativa) sobre las notas de los alumnos. Por otra parte, en la figura A.7 no se ve ninguna relación clara de las notas de los alumnos sobre las notas medias de los pares en *Vida Social*. Con esto se concluye que los alumnos se juntan a estudiar con compañeros de rendimiento similar que otorguen un desempeño positivo y significativo, en cambio, lo que respecta a compañeros en *Vida Social* lo hacen independiente del rendimiento.

Si comparamos las estimaciones de los modelos (3.1), en la figura A.6 se aprecia que las estimaciones según la media por *intensidad* los valores se agrupan de mejor manera sobre la recta. Se tiene además que el estadístico t es mayor para este tipo de media, lo que se concluye mejores resultados para el modelo (3.1b). En lo que respecta a los pares en *Vida Social*, independiente al tipo de media, se tienen resultados no significativos. Para estas estimaciones se adjunta el estadístico de Cragg-Donald, el test \mathcal{F} para la prueba de instrumentos débiles (Stock & Yogo, 2005) relativos a la exclusión de instrumentos. Esta prueba muestra la relevancia en utilizar como instrumentos excluidos las características exógenas de los pares excluidos dados por $G^2\mathbf{X}$ y $F^2\mathbf{X}$, respectivamente al tipo de media. En las tablas A.17 y A.18 se adjuntan las primeras etapas de las estimaciones (3.1). En estas, se exponen buenos resultados en el sentido de que las características de los pares excluidos explican de manera satisfactoria las medias de los pares.

En consecuencia, las estimaciones utilizando la intensidad en la amistad de los pares tienen mejores resultados que no considerar dicha intensidad. No obstante, se recalca el hecho que se requiere tener más información para la creación de la matriz F en comparación a la matriz G . Luego, las estimaciones se realizarán utilizando la matriz F .

3.4.2. Modelo de Efectos de Pares

El modelo que se utilizará para la estimación del efecto de pares ha sido revisado en la sección 2.4. En teoría se pueden utilizar todas las variables que se tienen a disposición, sin embargo existen una gama de ellas que no proporcionan una valiosa información y en algunos casos aumentan la variabilidad en la estimación. A pesar que el conjunto de efectos exógenos corresponde exactamente a las características propias de los alumnos, el modelo puede ser estimado con distintas variables, lo cual se obtienen coeficientes de mayor significancia. Por tanto se ha llegado a la conclusión de omitir variables no significativas en el modelo. Se hace hincapié en el hecho que se poseen pocas observaciones, por tanto se tendrá flexibilidad en la estimación del modelo final.

Dado que se tiene la información de dos tipos de grupos de referencia, y con el objeto de estudiar en profundidad el efecto de pares, en primera instancia se estudiará el efecto de pares de manera independiente para cada grupo de referencia (*Estudio Grupal* y *Vida Social*) para luego analizar de manera conjunta el efecto de pares.

3.4.2.1. Estudio Grupal y Vida Social

De manera independiente, para cada grupo de referencia se ha estimado el siguiente modelo:

$$Y = \alpha + \beta FY + \gamma \mathbf{X} + \delta F\mathbf{X} + \varepsilon, \quad (3.2)$$

donde:

Y : Rendimiento del alumno, dado por las notas cursadas en el semestre.

\mathbf{X} : Vector de características de los alumnos.

Las variables multiplicadas, FY y $F\mathbf{X}$, denotan el rendimiento medio y características medias de los pares del alumno, respectivamente. Bajo este contexto, el parámetro β cuantifica el *efecto endógeno* y el vector δ el *efecto exógeno*.

El modelo que cuantifica el efecto de pares en *Vida Social* es mucho más conservador al respecto de la cantidad de variables, en comparación al modelo en *Estudio Grupal*. Esto se debe a que la información recopilada está enfocada mayoritariamente a características de los alumnos que están fuertemente correlacionados con el tipo de estudio.

Estudio Grupal

Los resultados de este modelo se encuentran en la tabla A.20. Para explicar el rendimiento de los alumnos, analizando sus propias *características*, se hace notar que la tasa de aprobación de cursos se encuentra fuertemente correlacionado. A medida que se tiene una mayor tasa, esta aumenta el rendimiento. Acerca del estudio con sus pares, si siempre estudian en las semanas de evaluación tiene un efecto contraproducente sobre el rendimiento. No obstante, el estudiar siempre en una semana normal de clases tiene un efecto positivo.

Analizando el efecto que tienen las características de los pares sobre el rendimiento de los alumnos, el *efecto exógeno*, se tienen que varias son relevantes para explicar el rendimiento, como la educación y ocupación de los padres. La tasa de aprobación de los pares también es significativo y positivo para explicar el rendimiento del alumno.

Se puede apreciar la significancia del *efecto endógeno* (notas de los pares) para explicar el rendimiento individual. Asumiendo que las demás variables se mantienen constantes, si las notas de los pares dentro del *Estudio Grupal* se incrementan en un punto, estos incrementan en un .6 puntos el rendimiento del alumno. Aunque este efecto es bastante grande, parece razonable teniendo en cuenta que se consideran sólo los pares con los cuales se junta a estudiar.

En la tabla A.19 se adjunta la primera etapa en esta estimación. Para explicar las notas medias de los pares, dentro de las características del alumno, es la tasa de aprobación que se considera relevante, y la educación de los padres con magíster significativa al 10 %. Todas las características relevantes de los pares excluidos en esta etapa, son género, tasa de aprobación, ocupación de los padres, entre otros. Se destaca el elevado valor del estadístico Cragg-Donald (de 32.74) para la prueba de instrumentos débiles, lo que recalca la significancia de las características de los pares excluidos como instrumentos.

Vida Social

En la tabla A.22 se muestran los resultados del modelo (3.2) para el rendimiento de los alumnos sólo considerando los pares con los que hacen *Vida Social*. Con respecto a las *características* del alumno, es muy significativo y positivo para explicar el rendimiento. Es interesante observar que si el alumno se reúne pocas veces con sus pares para hacer vida social, tiene un efecto positivo en el rendimiento, en contraparte a que si siempre se reúnen.

El *efecto exógeno* es significativo para la educación de los padres de los pares con grado de magíster y/o doctorado, para explicar el rendimiento de los alumnos. Cabe destacar que

hay características físicas que se consideran relevante, como la estatura de los pares que es significativo al 10 %.

En términos del *efecto endógeno*, el aumento de un punto en las notas de los pares influyen en un aumento de .2 puntos el rendimiento individual. Se destaca un incremento menor en comparación a los pares en *Estudio Grupal*, a la vez un menor nivel de significancia. Esto se explica a que se analiza el efecto de los pares cuyo objetivo es realizar actividades recreativas y/o de vida social, pero no de estudio.

La primera etapa en la estimación se muestra en la tabla A.21. Aquí, se puede ver que mayoritariamente las características de los pares del alumno y la de los pares excluidos son significativas para explicar el comportamiento de las notas medias de los pares. Son las características de los propios alumnos que no son relevantes. Esto puede corresponder a que los pares en *Vida Social* no consideran como par a los alumnos que sí los consideran como par. En otras palabras, un alumno puede considerar como par a una persona, pero no viceversa, dado que no es recíproco y se produce el efecto anteriormente expuesto. Sobre la prueba de instrumentos débiles, se consigue un estadístico de Cragg-Donald de 18.22 lo que destaca la significancia de las características de los pares excluidos.

3.4.2.2. Modelo en Intensidad

El modelo en intensidad del efecto de pares consiste en estimar de manera conjunta el efecto de pares en los grupos de referencia de *Estudio Grupal* (EG) y *Vida Social* (VS), donde se controla la intensidad en la relación utilizando la matriz F respecto a cada grupo:

$$Y = \alpha + \beta_{EG} F_{EG} Y + \beta_{VS} F_{VS} Y + \gamma \mathbf{X} + \delta_{EG} F_{EG} \mathbf{X} + \delta_{VS} F_{VS} \mathbf{X} + \varepsilon, \quad (3.3)$$

haciendo notar que se estiman ambos efectos de los grupos de referencia de manera simultánea. En este modelo, los efectos sociales vienen dados respecto al tipo de grupo de referencia, vale decir, el vector de parámetros (β_{EG}, β_{VS}) cuantifica el *efecto endógeno* y el vector $(\delta_{EG}, \delta_{VS})$ cuantifica el *efecto exógeno*, respectivo a cada grupo.

En la tabla A.25 se adjuntan los resultados del modelo (3.3). Se recalca el hecho que en teoría se pueden utilizar todas las variables obtenidas por medio de la encuesta, no obstante, se busca la “*armonía*” entre estas variables en conjunto a los grupos de referencia. Por tanto las variables escogidas para la estimación del modelo son expuestas en esta tabla.

Acerca de las *características* propias del alumno, la tasa de aprobación es relevante para explicar el rendimiento del alumno, dada la implicancia directa entre el rendimiento del

alumno y su tasa de aprobación. Agregando que los alumnos con un ingreso mayor a los 3 millones, tienden a tener un rendimiento menor. Con los pares de *Estudio Grupal* que estudian en semana normal de clases (tanto pocas veces y siempre) tienen un efecto positivo en el rendimiento, en contra parte de estudiar siempre en las semanas de evaluación. Dentro de los pares que hace *Vida Social*, el reunirse pocas veces y regularmente genera un efecto positivo en el rendimiento.

Para explicar el rendimiento de los alumnos, el *efecto exógeno* es significativo para algunas variables. Los pares que se encuentran en *Estudio Grupal*, la tasa de aprobación de estos fomenta el rendimiento positivo del alumno, siguiendo con lo obtenido anteriormente. Sin embargo, no es claro el efecto de la educación de los padres de los pares en este grupo de referencia. Acerca del efecto producto de las características de los pares en *Vida Social*, se hace notar que la tasa de aprobación no es significativa para explicar el rendimiento.

Por último, los resultados muestran que el *efecto endógeno* es significativo para los pares en *Estudio Grupal* y al 10% para el efecto de pares en *Vida Social*. El estadístico t del efecto producto de los pares en *Estudio Grupal* refleja la relevancia que tiene para explicar el rendimiento individual dado que es un valor considerablemente grande. Si comparamos este efecto por grupo de referencia, el efecto sobre el rendimiento del alumno producto de los pares pertenecientes a *Estudio Grupal* es mayor en comparación al efecto de los pares en *Vida Social*. Por tanto, cuando se trata de rendimiento, los alumnos se reúnen a estudiar con compañeros de rendimiento similar lo cual entrega resultados positivos en el rendimiento propio. En comparación a los pares con los que hace *Vida Social*, los alumnos eligen a estos de manera independiente al rendimiento que poseen.

Si comparamos el *efecto endógeno* del *Estudio Grupal* en el modelo (3.2), se consigue que es mayor si lo analizamos de manera independiente al efecto de los pares en *Vida Social*. El considerar efectos en el rendimiento individual sobre los pares que pertenecen a *Vida Social* hacen que tenga una consecuencia contraproducente en el efecto producido por los pares en *Estudio Grupal*. Sobre el *efecto endógeno* producto de los pares en *Vida Social*, no hay un cambio significativo a si se analiza de manera independiente.

Para la correcta especificación del modelo, se han utilizado algunas características de los pares excluidos como instrumentos para la variable endógena (rendimiento de los pares), de manera conjunta para los amigos de los pares en *Estudio Grupal* y *Vida Social*. De esta forma se requieren dos modelos para la primera etapa, respectivo a cada rendimiento medio de los pares según el grupo de referencia. Los resultados de estos se presentan en las tablas A.23 y A.24, respectivos al grupo. Es atrayente hacer notar que la primera etapa del modelo que evalúa el rendimiento medio de los pares en *Estudio Grupal* considera relevante las

características de los pares en este mismo grupo, pero no a las características de los pares en *Vida Social*. Ocurre el caso inverso para la primera etapa en *Vida Social*. Esto confirma la correcta especificación del modelo, y además hace notar el distinto efecto en el rendimiento producto de ambos grupos de referencia. En base de estas estimaciones, se consigue un estadístico de Cragg-Donald del orden de 8.13, siendo un valor no significativamente grande, pero aún mantiene los niveles de aceptación de los instrumentos.

Conclusión

En este proyecto de tesis, *Efectos de Pares Heterogéneos y Rendimiento Académico*, se ha abordado el análisis de los efectos de pares en la red de la Universidad de Chile en el semestre de Primavera del año 2010. La manera de medir este efecto es por modelos econométricos denominados *lineales en medias*. Luego de un acabado estudio metodológico de estos, se implementa la correcta especificación de un modelo lineal en medias a los datos conseguidos por medio de una encuesta preparada en conjunto con el profesor guía.

Cabe destacar que para llevar a cabo el estudio de los efectos de pares y los modelos lineales en medias se implementó complejos algoritmos computacionales en el software estadístico STATA (StataCorp, 2009).

Dentro del estudio de los efectos de pares se tiene el difícil acceso a información acorde al problema en cuestión. Vale decir, es muy complicado el tener acceso a datos en los cuales pueda ser implementado el modelo lineal en medias. Es por esto que de manera *parcialmente satisfactoria* se ha implementado una encuesta hacia los alumnos de la Universidad para la captura de datos acorde al problema. Se dice de manera parcialmente satisfactoria debido a que con estos datos se dio la realización de este proyecto, pero fueron pocos los alumnos que de manera voluntaria la respondieron. No obstante, este proyecto representa un piloto para realizar una encuesta del mismo estilo en una muestra más amplia de alumnos. Dentro del cuestionario, se consideran matrices de respuesta en las cuales el alumno encuestado tiene que ordenar de mayor a menor al compañero más influyente de estudio. Se destaca que de esta manera se tiene especificada una base de datos original en la literatura de los efectos de pares.

El efecto de pares puede ser utilizado para explicar una variedad de situaciones que ocurren dentro de un establecimiento educacional. Por ejemplo, estos pueden ser utilizados para estudiar la selección de un magíster (De Giorgi, *et al.* 2010) o la cantidad de horas planeadas para la realización de actividades recreacionales (Bramoullé, *et al.* 2008). En este proyecto la implementación recae en el rendimiento presentados por los alumnos, definidos por la nota media obtenida en el semestre de evaluación. Además, los pares se han definido

por dos tipos de grupo: aquellos con los que se juntan a estudiar y aquellos con los que se reúne para hacer vida social.

El poder *rankear* a los compañeros influyentes en los estudios dio paso a cuantificar de forma alternativa el *efecto endógeno* y *efecto exógeno*. A lo largo de la literatura, consideran que todos los pares dentro del grupo de referencia afectaban con la misma intensidad el comportamiento del alumno, y en base de esta hipótesis obtener dichos efectos. La manera de innovar la captura de estos efectos es como sigue. Dependiendo de este *ranking*, se le asigna un mayor peso al compañero más amigo, luego al segundo y así consecutivamente. Básicamente se dice a *priori* que son los compañeros más amigos que influyen de mayor manera en comparación al segundo y los que siguen. Los resultados son concluyentes. Al considerar la intensidad en la relación, se tienen estadísticos más significativos y mejor bondad de ajuste. Sin embargo, como se ha planteado anteriormente, el considerar el *ranking* en el grado de amistad hace que la información respectiva sea más privilegiada en obtener.

El modelo de efecto de pares estimado viene dado acorde a los grupos de referencia definidos en la encuesta: *Estudio Grupal* y *Vida Social*. Para cada uno de estos, se ha estimado un modelo de efectos de pares. El modelo final consiste en analizar ambos grupos de pares de manera conjunta. Dada la construcción de los datos, es posible omitir el efecto correlacionado debido a la ausencia a nivel de grupo. Las principales dificultades en la estimación consiste en las pocas observaciones conseguidas. En teoría, es posible utilizar todas las variables, pero dada la cantidad de observaciones se decidió en eliminar las menos relevantes dentro del análisis para así obtener resultados estadísticamente significativos.

Los resultados de este proyecto muestran que existe evidencia estadística significativa para concluir la presencia de efectos de pares, especialmente en el grupo de referencia en *Estudio Grupal*. Estos pares afectan de manera positiva y significativa en el rendimiento del alumno. Sobre los pares en *Vida Social*, no hay un efecto significativo. Esto quiere decir que los alumnos conforman sus grupos de estudio con compañeros de rendimiento similar con consecuencias positivas en el rendimiento, sin embargo, lo que respecta a los compañeros en *Vida Social* estos los eligen independiente al rendimiento. En base de los modelos estimados, se encontró que el efecto endógeno producto de los pares pertenecientes al grupo de *Vida Social* genera una consecuencia contraproducente al efecto generado por los pares en *Estudio Grupal*.

Los efectos de pares son fundamentados para diseñar políticas respecto de cómo asignar los escasos recursos destinados a mejorar la educación. En base de los resultados obtenidos, es posible que alumno de alto rendimiento apoye a los que tienen más problemas (efecto endógeno positivo y significativo). Además fomentar el estudio grupal periódico en las sema-

nas normal de clases tiene un efecto positivo en el rendimiento en comparación de hacerlo en las semanas de evaluación.

Referencias

- Akerlof G.A. (1997). "Social distance and social decisions." *Econometrica*, 65(5): 1005-1027.
- Akerlof G.A. & Kranton. R.E. (2002). "Identity and schooling: Some lessons for economics education." *Journal of Economic Literature*, 40(4): 1167-1201.
- Bayer P., Pintoff R. & Pozen D. (2009). "Building criminal capital behind bars: Peer effect in juvenile corrections." *Quarterly Journal of Economics*, 124(1): 105-147.
- Bertrand M., Luttmer E.F.P. & Mullainathan S. (2000). "Network Effects and Welfare Cultures." *Quarterly Journal of Economics*, 115: 1019-1056.
- Boucher V., Bramoullé Y., Djebbari H. & Fortin B. (2010). "Do Peers Affect Student Achievement? Evidence from Canada Using Group Size Variation." *IZA Discussion Paper*, No. 4723.
- Bramoullé Y., Djebbari H. & Fortin B. (2008). "Identification of Peer Effects through Social Networks." *Journal of Econometrics*, 150(1): 41-55.
- Brock W.A. & Durlauf S.N. (2001). "Interaction-based models." In *Handbook of Econometrics*, 5: 3297-3380. Ed. J. Heckman & E. Leamer. Elsevier.
- Calvó-Armengol A., Patacchini E. & Zenou Y. (2009). "Peer Effects and Social Networks in Education." *Review of Economic Studies*, 76(4): 1239-1267.
- Conlisk J. (1980). "Costly Optimizers Versus Cheap Imitators." *Journal of Economic Behavior and Organization*, 1: 275-293.
- De Giorgi G., Pellizzari M. & Redaelli S. (2010). "Identification of Social Interactions through Partially Overlapping Peer Groups." *The American Economic Journal: Applied Economics*, forthcoming.
- Fowler J. & Christakis N. (2008). "Estimating Peer Effects on Health in Social Networks." *Journal of Health Economics*, 27(5).

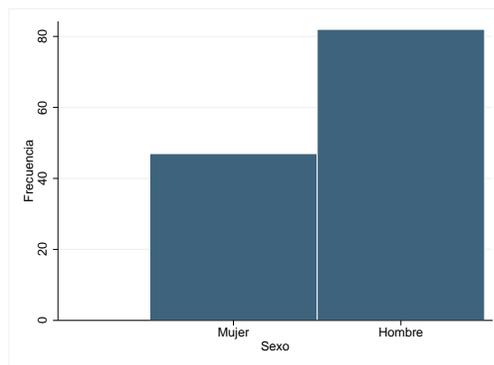
- Glaeser E., Sacerdote B. & Scheinkman J. (1996). "Crime and Social Interaction." *Quarterly Journal of Economics*, 111(2): 507-548.
- Jackson M. (2006). "The Economics of Social Networks." in *Proceedings of the 9th World Congress of the Econometric Society*, 1. Cambridge, Cambridge University Press.
- Laschever R. (2005). "The Doughboys Network: Social Interactions and Labor Market Outcomes of World War I Veterans." *Mimeo*. Northwestern University.
- Lee L.F. (2007). "Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effect." *Journal of Econometrics*, 140(2): 333-374.
- Lin X. (2005). "Peer Effects and Student Academic Achievement: an Application of Spatial Autoregressive Model with Group Unobservables." *Mimeo*. Department of Economics, Tsinghua University, Beijing.
- Manski C. (1993). "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60(3): 531-542.
- Moffitt R. (2001). "Policy Interventions, Low-Level Equilibria, and Social Interactions." *Social Dynamics*, edited by Steven Durlauf and Peyton Young. MIT press.
- Sacerdote B. (2001). "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics*, 116(2): 681-704.
- Sáez E. & Duflo E. (2003). "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence From a Randomized Experiment." *Quarterly Journal of Economics*, 118: 815-842.
- StataCorp. (2009). *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP. <http://www.stata.com/>
- Stock J.H. & 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* Cambridge, 2005: 80-108. Cambridge University Press. New York.
- Trogdon J., Nonnemaker J. & Pais J. (2008). "Peer Effects in Adolescent Overweight." *Journal of Health Economics*, 27(5): 1388-1399.

ANEXO A

Resultados Gráficos y Tablas

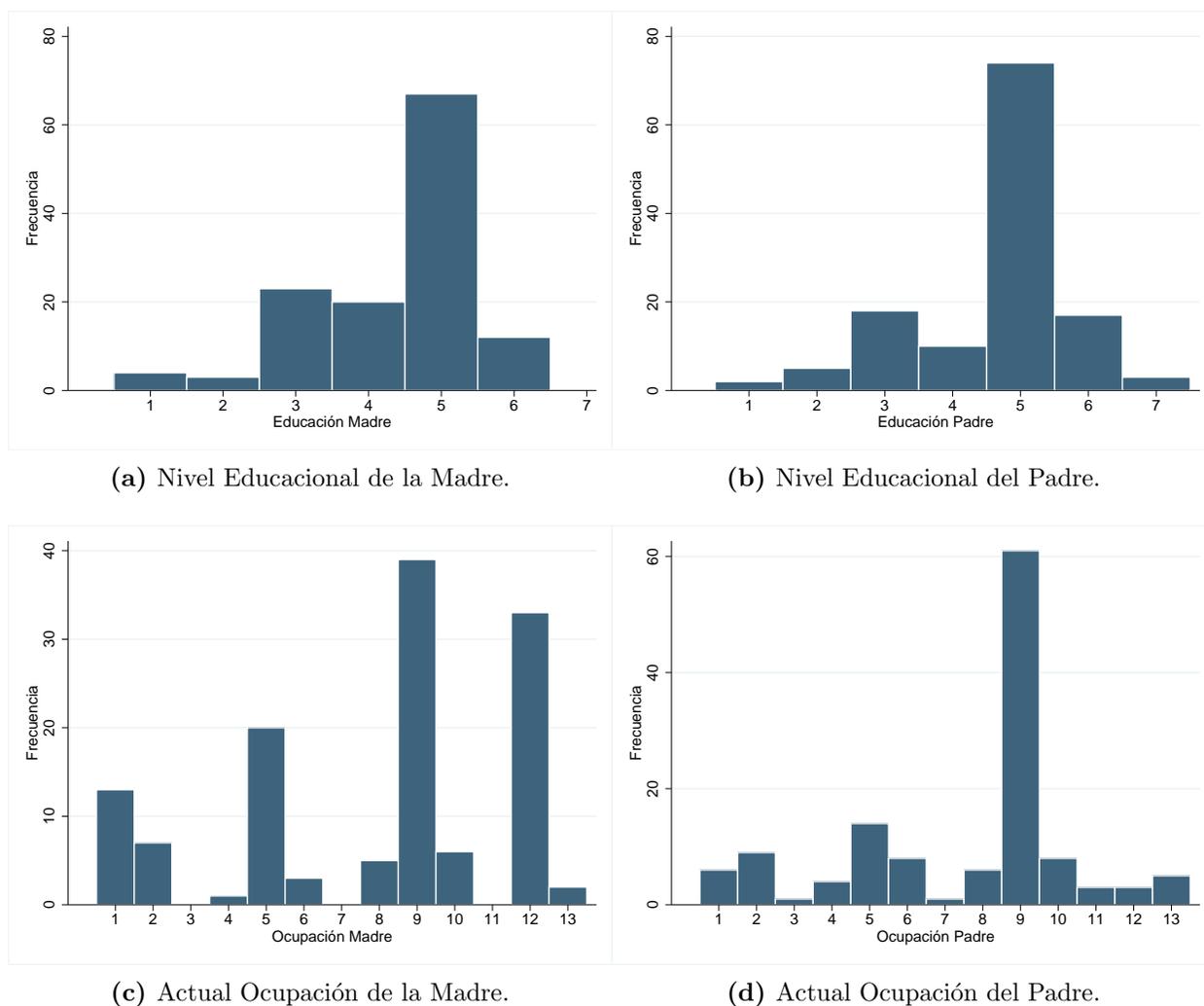
A.1. Estadística Descriptiva

Tabla A.1 Sexo



Valor	Frecuencia	Porcentaje
0. Mujer	47	36.43
1. Hombre	82	63.57
Total	129	100

Tabla A.2 Situación Socioeconómica de la Familia



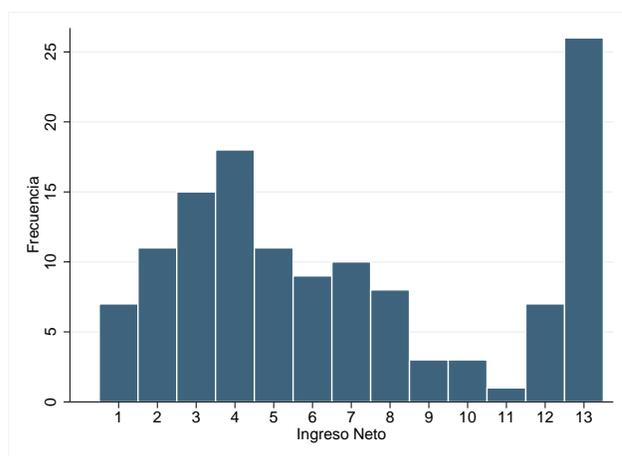
■ Figura A.2(a)-(b)

Valor	Educación Padre		Educación Madre	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
1. Básico	2	1.55	4	3.10
2. Media Incompleta	5	3.88	3	2.33
3. Media Completa	18	13.95	23	17.83
4. Superior Incompleta	10	7.75	20	15.50
5. Superior Completa	74	57.36	67	51.94
6. Magíster	17	13.18	12	9.30
7. Doctorado	3	2.33	0	0.00
Total	129	100	129	100

■ Figura A.2(c)-(d)

Valor	Ocupación Madre		Ocupación Padre	
	Frec.	Porc.	Frec.	Porc.
1. Trabajo no calificado sin contrato	13	10.08	6	4.65
2. Trabajo no calificado con contrato	7	5.43	9	6.98
3. Agricultores y trabajadores calificados	0	0.00	1	0.78
4. Operadores de instalaciones y máquinas	1	0.78	4	3.10
5. Empleado de oficina	20	15.50	14	10.85
6. Trabajadores de servicios	3	2.33	8	6.20
7. Oficiales y operarios de industrias	0	0.00	1	0.78
8. Profesional de carreras técnicas	5	3.88	6	4.65
9. Profesional con Grado de Licenciatura	39	30.23	61	47.29
10. Miembros del poder ejecutivo y legislativo	6	4.65	8	6.20
11. Fuerzas Armadas	0	0.00	3	2.33
12. Desempleado	33	25.58	3	2.33
13. Jubilado	2	1.55	5	3.88
Total	129	100	129	100

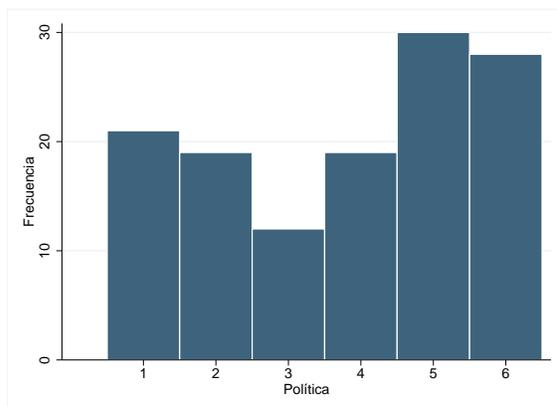
Tabla A.3 Ingreso Neto de la Familia



Valor*	Frec.	Porc.
1. Menor a 250.000	7	5.43
2. 250.001 - 500.000	11	8.53
3. 500.001 - 750.000	15	11.63
4. 750.001 - 1.000.000	18	13.95
5. 1.000.001 - 1.250.000	11	8.53
6. 1.250.001 - 1.500.000	9	6.98
7. 1.500.001 - 1.750.001	10	7.75
8. 1.750.001 - 2.000.000	8	6.20
9. 2.000.001 - 2.250.000	3	2.33
10. 2.250.001 - 2.500.000	3	2.33
11. 2.500.001 - 2.750.000	1	0.78
12. 2.750.001 - 3.000.000	7	5.43
13. Mayor a 3.000.000	26	20.16
Total	129	100

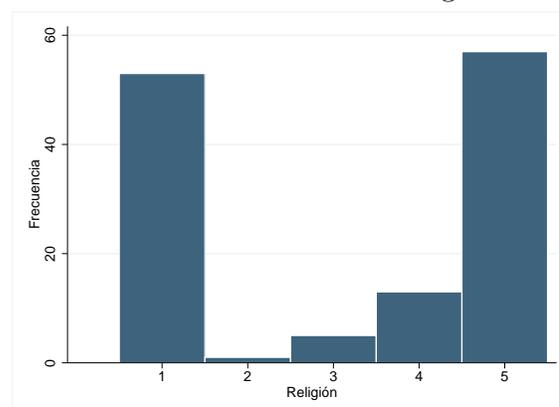
*Valores medidos en pesos Chilenos (CLP).

Tabla A.4 Ideología Política



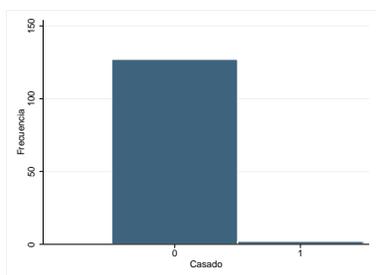
Valor	Frec.	Porc.
1. Derecha	21	16.28
2. Centro-Derecha	19	14.73
3. Centro	12	9.30
4. Centro-Izquierda	19	14.73
5. Izquierda	30	23.26
6. Ninguna	28	21.71
Total	129	100

Tabla A.5 Creencias Religiosas

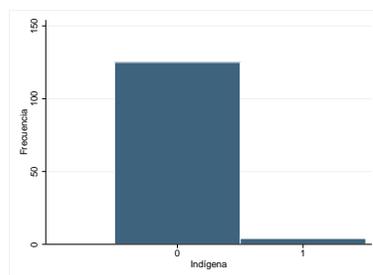


Valor	Frec.	Porc.
1. Católica	53	41.09
2. Judío	1	0.78
3. Protestante	5	3.88
4. Otra	13	10.08
5. Ateo/Agnóstico	57	44.19
Total	129	100

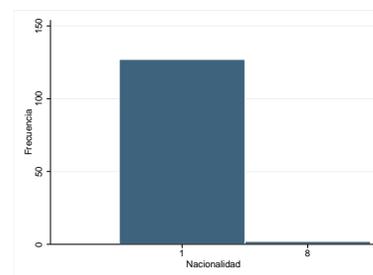
Tabla A.6 Composición de la Familia



(a) Estado Civil (1: Casado).



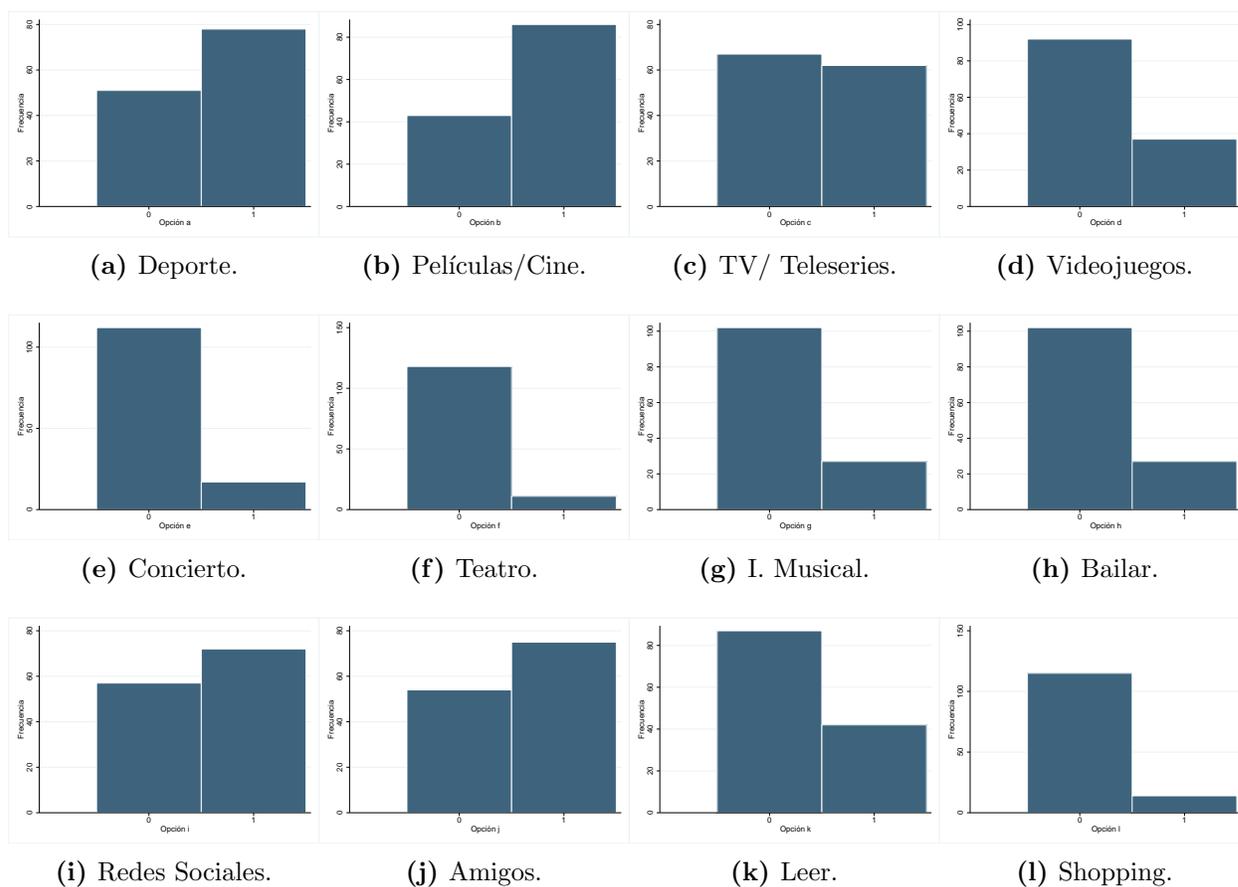
(b) Grupo Étnico Indígena (1: Pertenece).



(c) Nacionalidad (1: Chilena; 8: Otra).

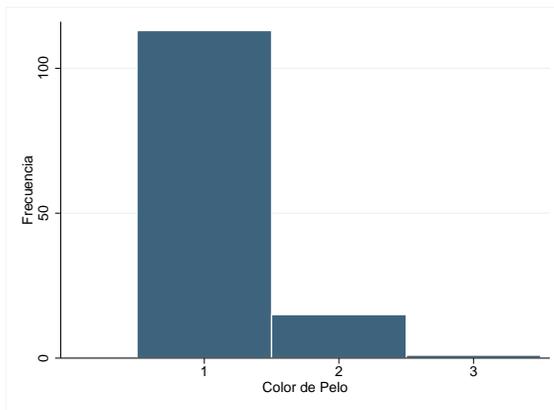
	(a) Casado		(b) Indígena		(c) Nac. Chilena	
Valor	Frecuencia	Porcentaje	Frecuencia	Porcentaje	Frecuencia	Porcentaje
0. No	127	98.45	125	96.90	2	1.55
1. Sí	2	1.55	4	3.10	127	98.45
Total	129	100	129	100	129	100

Tabla A.7 Tiempo Libre (1: Realiza)



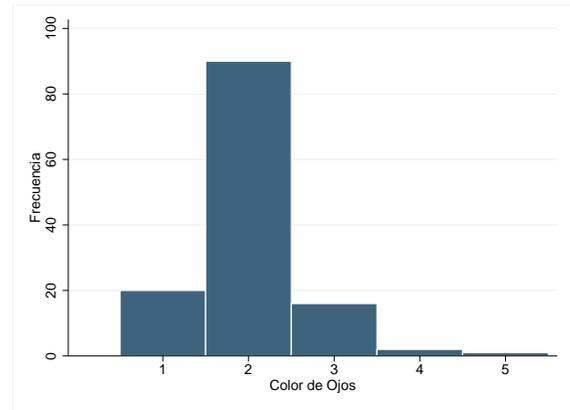
Valor	Frecuencia	Porcentaje
(a) Deporte	78	60.47
(b) Películas/Cine	86	66.67
(c) TV/Teleseries	62	48.06
(d) Videojuegos	37	28.68
(e) Concierto	17	13.18
(f) Teatro	11	8.53
(g) Instr. Musical	27	20.93
(h) Bailar	27	20.93
(i) Redes Sociales	72	55.81
(j) Amigos	75	58.14
(k) Leer	42	32.56
(l) Shopping	14	10.85
Otra	33	25.58

Tabla A.8 Color de Pelo



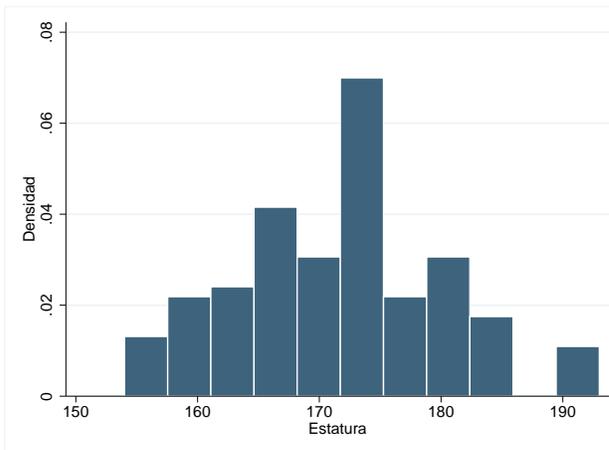
Valor	Frecuencia	Porcentaje
1. Oscuro	113	87.60
2. Claro	15	11.63
3. Pelirrojo	1	0.78
Total	129	100

Tabla A.9 Color de Ojos

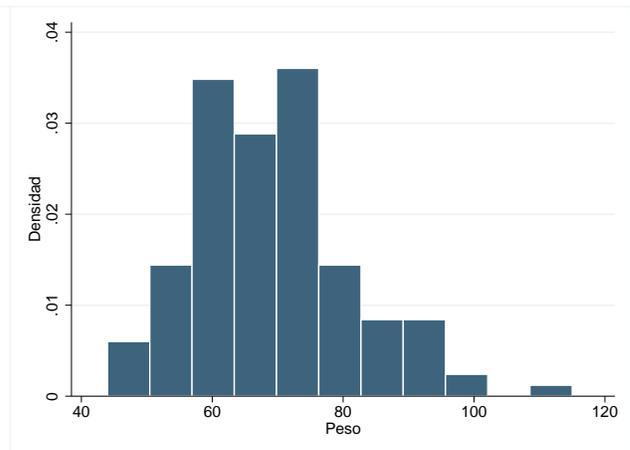


Valor	Frecuencia	Porcentaje
1. Negro	20	15.50
2. Café	90	69.77
3. Verdes	16	12.40
4. Azules	2	1.55
5. Gris	1	0.78
Total	129	100

Tabla A.10 Contextura Física



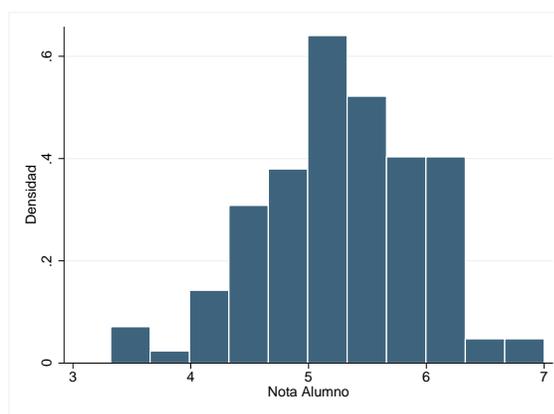
(a) Estatura (cm).



(b) Peso (kg).

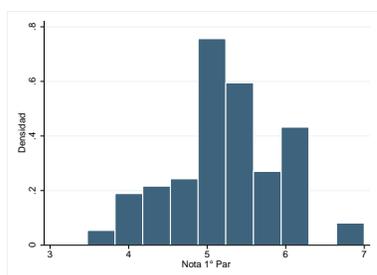
Variable	Obs.	Media	Desv. Est.	Min	Max
(a) Estatura	129	171.5	8.383	154.0	193.0
(b) Peso	129	68.9	12.066	44.0	115.0

Tabla A.11 Nota Media en el Último Semestre

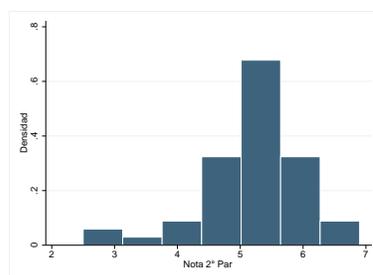


Obs.	Media	Desv. Est.	Min	Max
126	5.3	0.681	3.3	7.0

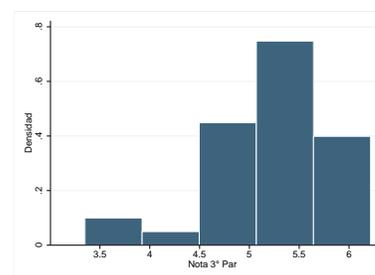
Tabla A.12 Nota Media de los Pares en *Estudio Grupal*



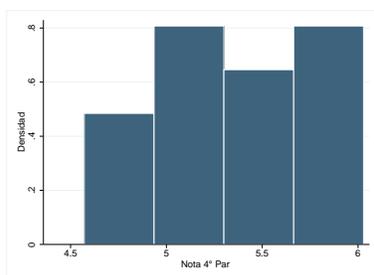
(a) Primer Compañero.



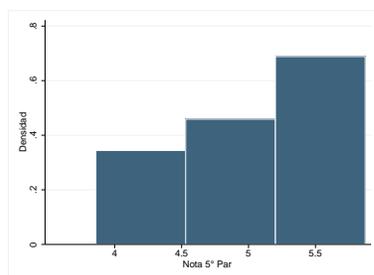
(b) Segundo Compañero.



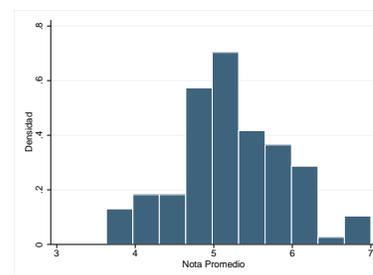
(c) Tercer Compañero.



(d) Cuarto Compañero.



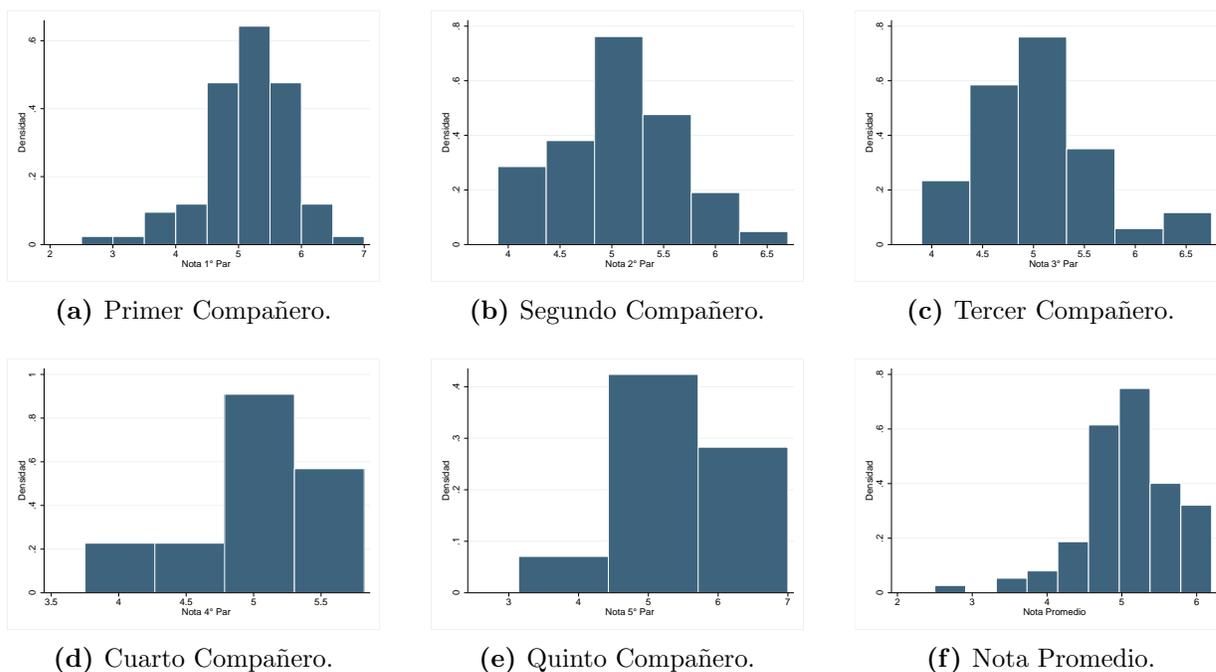
(e) Quinto Compañero.



(f) Nota Promedio.

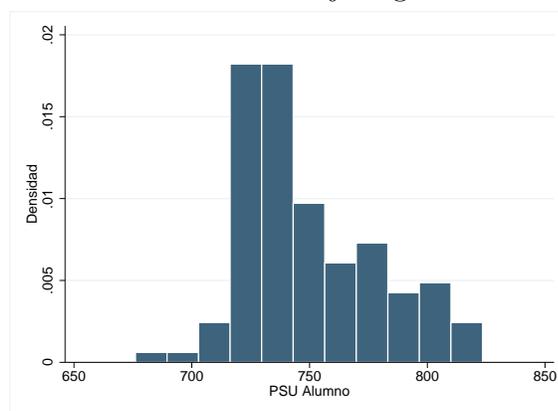
Variable	Obs.	Media	Desv. Est.	Min	Max
(a) Nota 1° Compañero	105	5.2	0.696	3.5	7.0
(b) Nota 2° Compañero	54	5.1	0.819	2.5	6.9
(c) Nota 3° Compañero	35	5.2	0.658	3.4	6.2
(d) Nota 4° Compañero	17	5.4	0.465	4.6	6.0
(e) Nota 5° Compañero	13	5.0	0.696	3.9	5.9
(f) Nota Promedio	114	5.2	0.689	3.6	7.0

Tabla A.13 Nota Media de los Pares en *Vida Social*



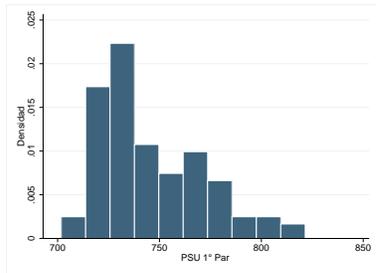
Variable	Obs.	Media	Desv. Est.	Min	Max
(a) Nota 1° Compañero	84	5.2	0.740	2.5	7.0
(b) Nota 2° Compañero	45	5.1	0.610	3.9	6.7
(c) Nota 3° Compañero	36	5.0	0.625	3.9	6.8
(d) Nota 4° Compañero	17	5.0	0.511	3.8	5.8
(e) Nota 5° Compañero	11	5.4	1.044	3.1	7.0
(f) Nota Promedio	91	5.1	0.616	2.5	6.2

Tabla A.14 Puntaje Ingreso PSU

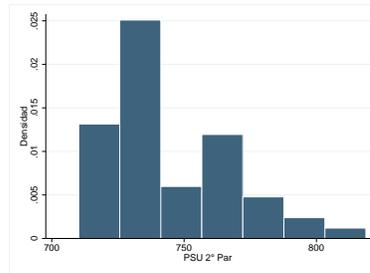


Obs.	Media	Desv. Est.	Min	Max
123	749.0	30.112	676.1	823.4

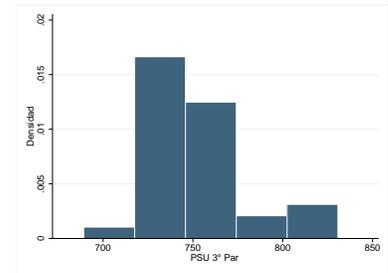
Tabla A.15 Puntaje PSU de los Pares en *Estudio Grupal*



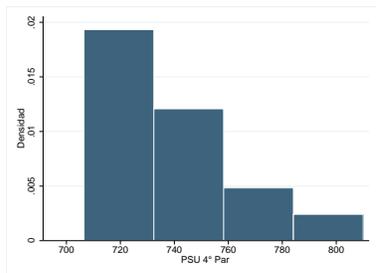
(a) Primer Compañero.



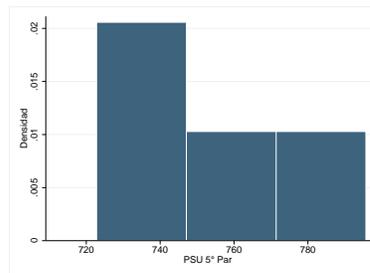
(b) Segundo Compañero.



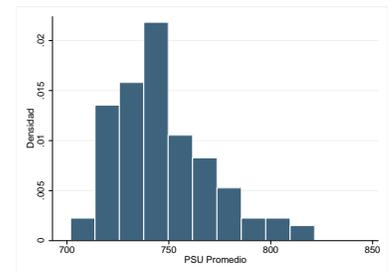
(c) Tercer Compañero.



(d) Cuarto Compañero.

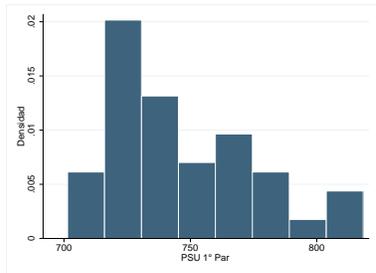


(e) Quinto Compañero.

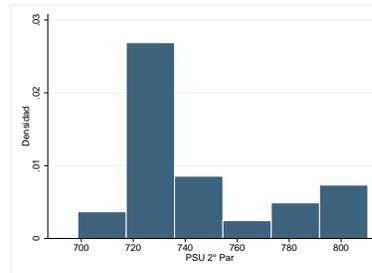


(f) Puntaje Promedio.

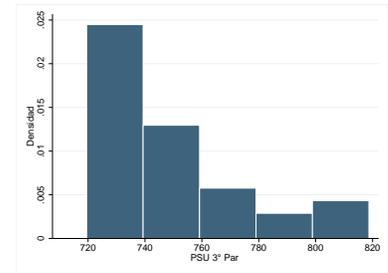
Variable	Obs.	Media	Desv. Est.	Min	Max
(a) PSU 1° Compañero	101	745.6	26.063	701.8	821.5
(b) PSU 2° Compañero	54	745.7	24.138	710.2	818.7
(c) PSU 3° Compañero	34	750.3	30.182	689.3	830.7
(d) PSU 4° Compañero	16	740.6	25.997	706.5	810.0
(e) PSU 5° Compañero	12	753.8	24.355	722.8	795.7
(f) Puntaje Promedio	111	746.4	23.496	701.8	821.5

Tabla A.16 Puntaje PSU de los Pares en *Vida Social*

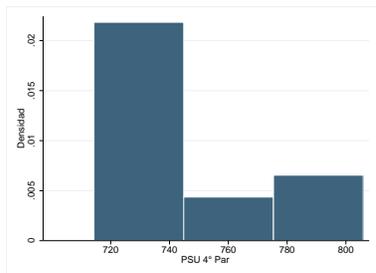
(a) Primer Compañero.



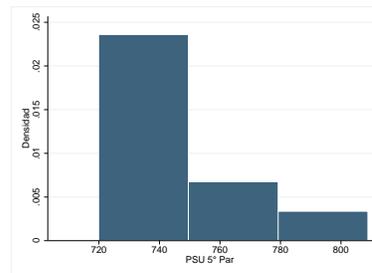
(b) Segundo Compañero.



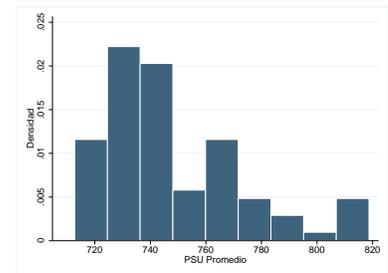
(c) Tercer Compañero.



(d) Cuarto Compañero.



(e) Quinto Compañero.



(f) Puntaje Promedio.

Variable	Obs.	Media	Desv. Est.	Min	Max
(a) PSU 1° Compañero	78	746.4	28.450	701.4	818.4
(b) PSU 2° Compañero	44	743.8	31.162	698.7	810.3
(c) PSU 3° Compañero	35	748.4	26.969	719.5	818.7
(d) PSU 4° Compañero	15	744.6	29.507	714.3	805.9
(e) PSU 5° Compañero	10	743.3	29.013	719.9	808.9
(f) Puntaje Promedio	88	748.4	25.723	712.8	818.7

A.2. Relación Alumno *vs.* Pares

A.2.1. Intensidad en la Relación

Figura A.1 Rendimiento Medio de los Pares

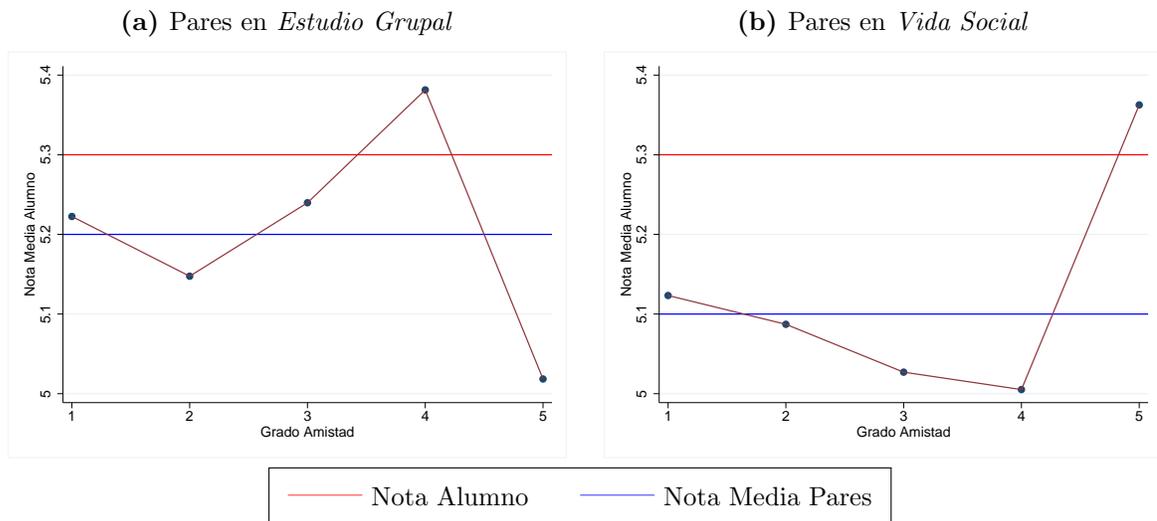


Figura A.2 Rendimiento de los Pares en *Estudio Grupal*

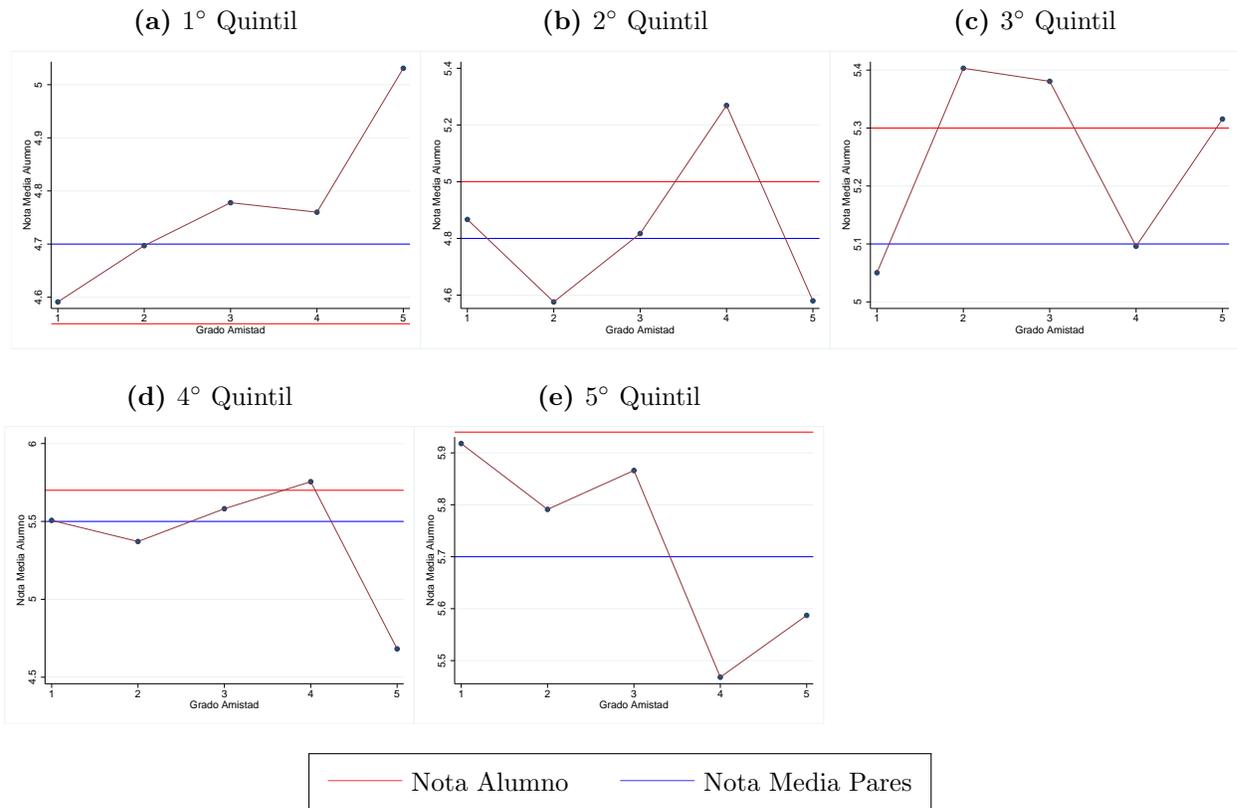
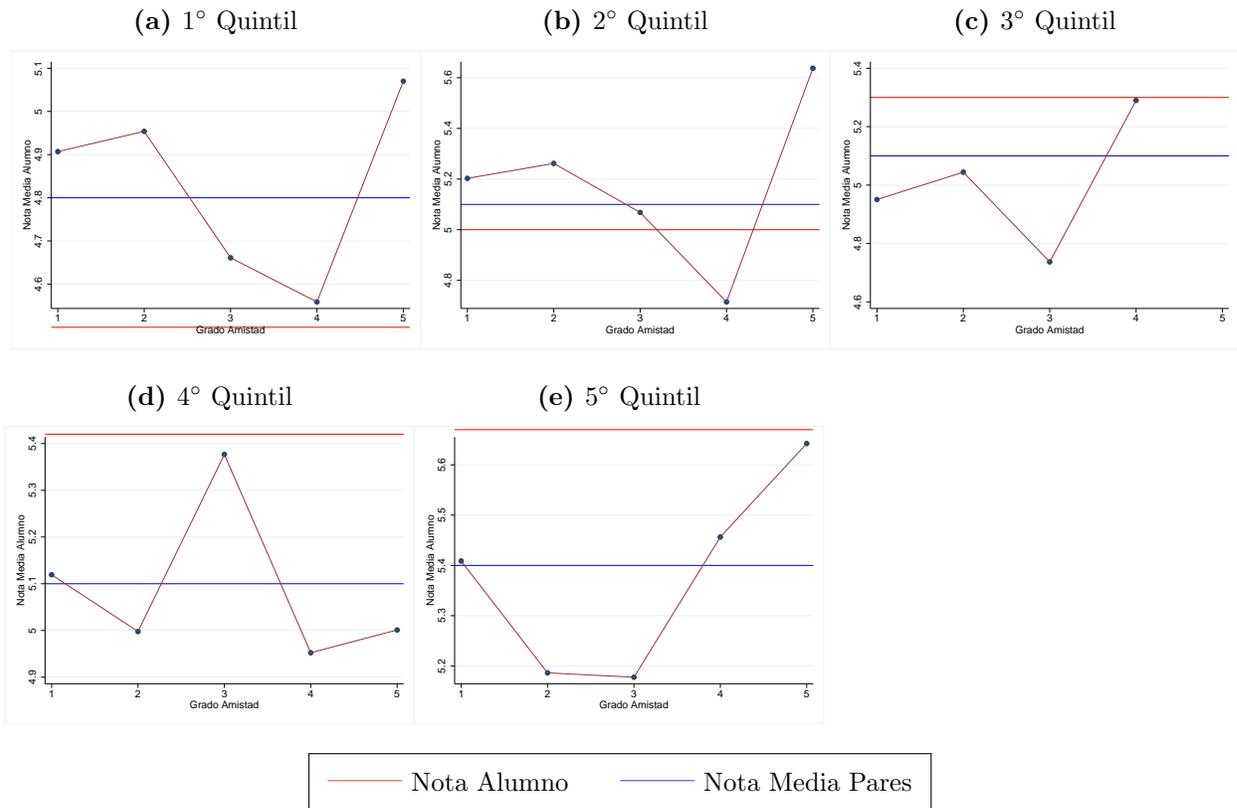
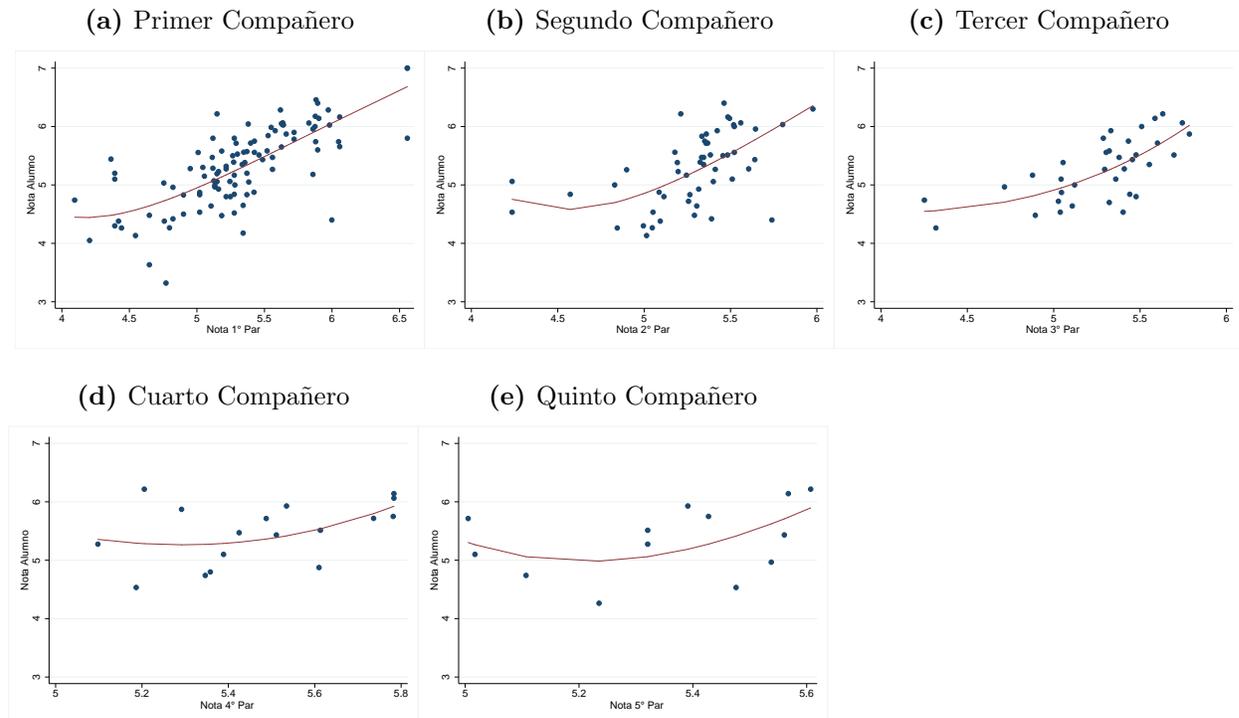


Figura A.3 Rendimiento de los Pares en *Vida Social*

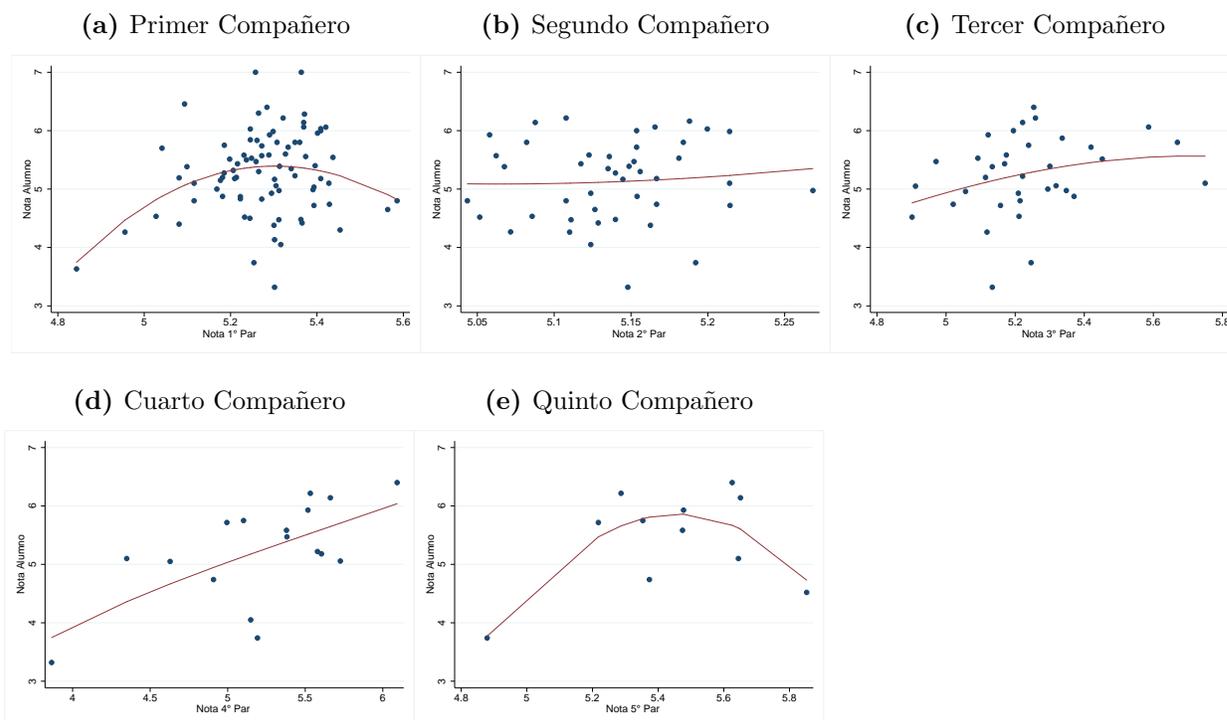


A.2.2. Relación Lineal sobre el Rendimiento

Figura A.4 Dispersión de las Notas de los Pares en *Estudio Grupal*



Modelo	Estimación	Desv. Est.	Est. t	p -valor	Intervalo de Confianza	
(a) Nota 1° Par	0.699	0.066	10.520	0.000	0.567	0.830
Intercepto	1.667	0.350	4.770	0.000	0.973	2.361
(b) Nota 2° Par	0.396	0.089	4.460	0.000	0.218	0.574
Intercepto	3.247	0.462	7.030	0.000	2.319	4.174
(c) Nota 3° Par	0.535	0.104	5.150	0.000	0.324	0.746
Intercepto	2.458	0.549	4.480	0.000	1.342	3.575
(d) Nota 4° Par	0.470	0.261	1.800	0.092	-0.087	1.028
Intercepto	2.947	1.412	2.090	0.054	-0.062	5.957
(e) Nota 5° Par	0.299	0.249	1.200	0.254	-0.248	0.846
Intercepto	3.850	1.258	3.060	0.011	1.081	6.620

Figura A.5 Dispersión de las Notas de los Pares en *Vida Social*

Modelo	Estimación	Desv. Est.	Est. t	p -valor	Intervalo de Confianza	
(a) Nota 1° Par	0.165	0.105	1.570	0.121	-0.044	0.374
Intercepto	4.431	0.545	8.120	0.000	3.346	5.516
(b) Nota 2° Par	0.080	0.176	0.460	0.651	-0.275	0.435
Intercepto	4.730	0.900	5.260	0.000	2.915	6.546
(c) Nota 3° Par	0.298	0.181	1.640	0.110	-0.071	0.667
Intercepto	3.741	0.916	4.080	0.000	1.877	5.605
(d) Nota 4° Par	1.076	0.336	3.200	0.006	0.359	1.793
Intercepto	-0.171	1.691	-0.100	0.921	-3.776	3.435
(e) Nota 5° Par	0.252	0.249	1.010	0.338	-0.312	0.816
Intercepto	4.087	1.360	3.010	0.015	1.011	7.164

Figura A.6 Dispersión sobre las Notas Medias de los Pares en *Estudio Grupal*

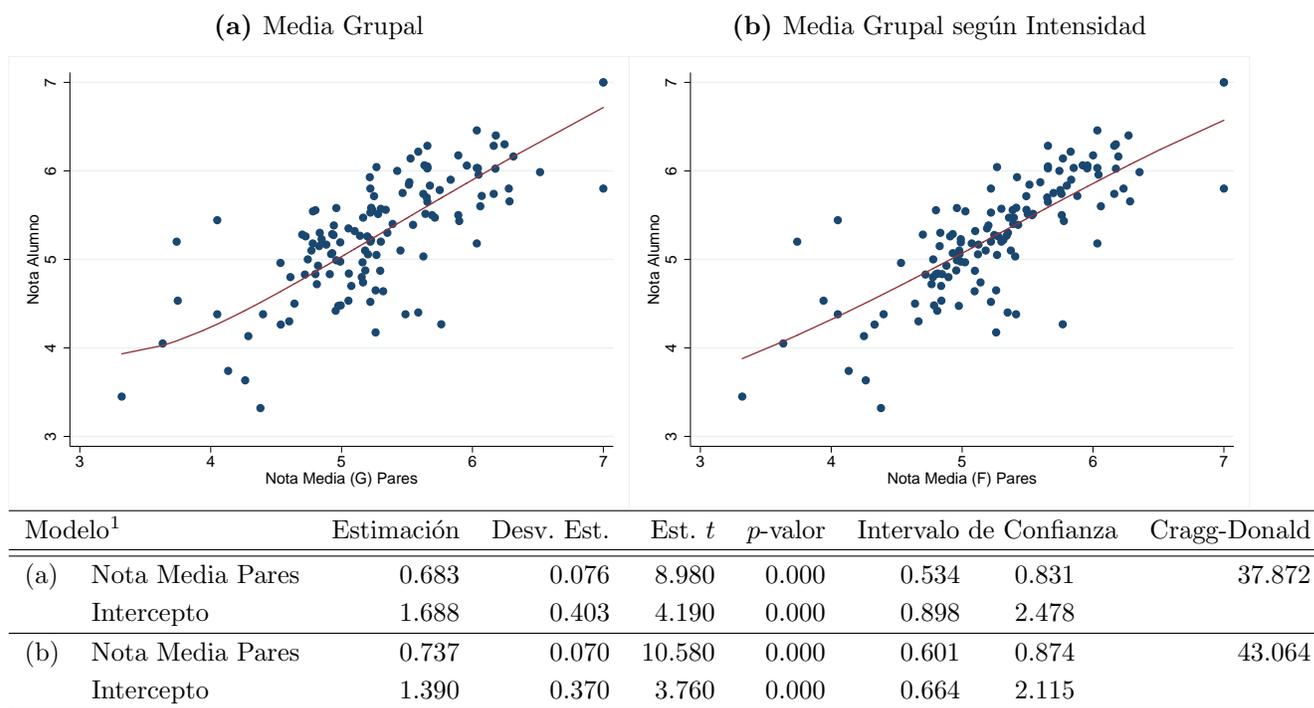
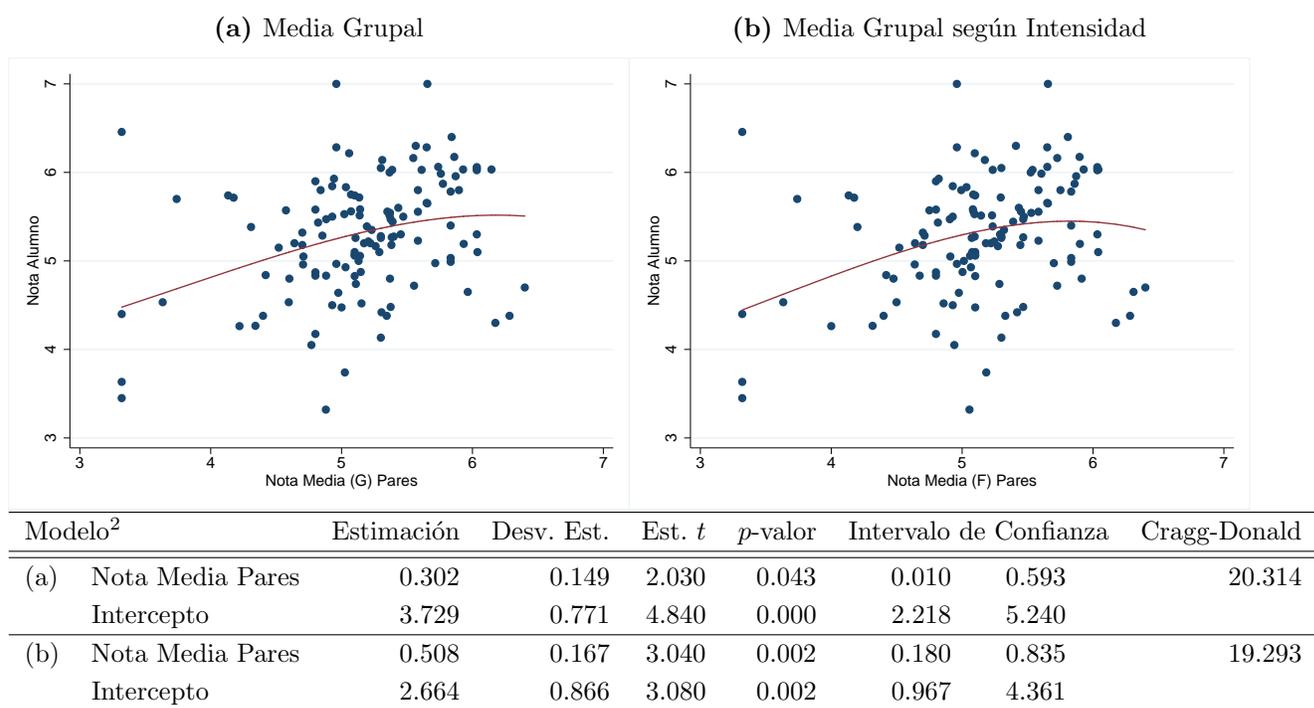


Figura A.7 Dispersión sobre las Notas Medias de los Pares en *Vida Social*



¹Primera Etapa de los modelos de pares en *Estudio Grupal* adjuntos en la tabla A.17.

²Primera Etapa de los modelos de pares en *Vida Social* adjuntos en la tabla A.18.

Tabla A.17 PRIMERA ETAPA: Dispersión sobre las Notas Medias en *Estudio Grupal*

(a) Media Grupal				(b) Media Grupal según Intensidad			
Variable [†]	Val.Est.	Desv.E.	Est. <i>t</i>	Variable [†]	Val.Est.	Desv.E.	Est. <i>t</i>
Género (1: Hombre)	-0.389	0.173	-2.250	Índice Aprobación	5.182	0.495	10.460
Índice Aprobación	5.124	0.472	10.870	Educ. Padre (1: Superior Completa)	0.418	0.112	3.730
Educ. Padre (1: Superior Completa)	0.634	0.121	5.260	Ocupación Padre (1: Profesional con Grado)	-0.599	0.107	-5.590
Ocupación Padre (1: Profesional con Grado)	-0.504	0.114	-4.430	Ingreso (0.75m - 1.5m)	0.409	0.118	3.470
Ingreso (0.75m - 1.5m)	0.330	0.124	2.670	Color de Ojos (1: Verdes o Azules)	0.419	0.128	3.270
Color de Ojos (1: Verdes o Azules)	0.496	0.126	3.940	Peso	0.012	0.004	3.270
Peso	0.024	0.006	4.180	Intercepto	-0.345	0.570	-0.610
Intercepto	-1.133	0.651	-1.740				

[†] Características de los pares excluidos.

[†] Características de los pares excluidos.

Tabla A.18 PRIMERA ETAPA: Dispersión sobre las Notas Medias en *Vida Social*

(a) Media Grupal				(b) Media Grupal según Intensidad			
Variable [†]	Val.Est.	Desv.E.	Est. <i>t</i>	Variable [†]	Val.Est.	Desv.E.	Est. <i>t</i>
Educ. Padre (1: Magíster y/o Doctorado)	-1.572	0.378	-4.160	Ocupación Madre (1: Profesional con Grado)	-1.068	0.158	-6.760
Educ. Madre (1: Magíster y/o Doctorado)	1.143	0.278	4.100	Ingreso (1.5m - 3.0m)	-0.427	0.195	-2.190
Ocupación Madre (1: Profesional con Grado)	-1.210	0.187	-6.470	Estatura	0.029	0.011	2.760
Color de Ojos (1: Verdes o Azules)	0.599	0.315	1.900	Intercepto	0.668	1.822	0.370
Intercepto	5.626	0.086	65.740				

[†] Características de los pares excluidos.

[†] Características de los pares excluidos.

A.3. Modelo de Efectos de Pares

Tabla A.19 PRIMERA ETAPA: Rendimiento de los Alumnos en *Estudio Grupal*

	Variable	Estimación	Desv. Est.	Est. <i>t</i>
<i>Características Propias</i>	Género (1: Hombre)	-0.156	0.079	-1.960
	Índice Aprobación	1.158	0.333	3.470
	Educ. Madre (1: Magíster y/o Doctorado)	0.177	0.092	1.920
	Ocupación Padre (1: Profesional con Grado)	0.049	0.066	0.750
	Ingreso (1.5m - 3.0m)	-0.094	0.069	-1.360
	Ingreso (Mayor a 3.0m)	-0.034	0.077	-0.440
	Estatura	0.009	0.005	1.860
	Peso	-0.002	0.003	-0.660
	Estudio en Semana Normal (Pocas Veces)	0.004	0.067	0.070
	Estudio en Semana Normal (Siempre)	0.143	0.190	0.750
	Estudio en Semana Evaluación (Siempre)	-0.118	0.101	-1.160
<i>Características Pares</i>	Género (1: Hombre)	-0.060	0.104	-0.580
	Índice Aprobación	3.110	0.378	8.230
	Educ. Madre (1: Superior Completa)	0.158	0.079	1.990
	Educ. Padre (1: Superior Completa)	-0.131	0.088	-1.490
	Ocupación Madre (1: Profesional con Grado)	-0.167	0.074	-2.260
	Ocupación Padre (1: Profesional con Grado)	-0.150	0.077	-1.940
	Ingreso (0.75m - 1.5m)	0.224	0.083	2.690
	Ingreso (Mayor a 3.0m)	0.068	0.084	0.810
	Color de Ojos (1: Verdes o Azules)	0.260	0.111	2.330
Estatura	0.013	0.007	1.930	
<i>Características Pares Excluidos</i>	Género (1: Hombre)	0.458	0.091	5.020
	Índice Aprobación	3.844	0.402	9.550
	Ocupación Padre (1: Profesional con Grado)	-0.467	0.106	-4.400
	Color de Ojos (1: Verdes o Azules)	0.395	0.106	3.740
	Intercepto	-5.742	1.250	-4.590

Tabla A.20 Rendimiento de los Alumnos sobre los Pares en *Estudio Grupal*[†]

	Variable	Estimación	Desv. Est.	Est. <i>t</i>
<i>Efecto Endógeno</i>	Nota Pares	0.584	0.102	5.750
<i>Características</i>	Género (1: Hombre)	-0.171	0.090	-1.910
	Índice Aprobación	2.191	0.421	5.200
	Educ. Madre (1: Magíster y/o Doctorado)	0.043	0.110	0.390
	Ocupación Padre (1: Profesional con Grado)	0.079	0.075	1.060
	Ingreso (1.5m - 3.0m)	0.079	0.079	1.000
	Ingreso (Mayor a 3.0m)	-0.164	0.088	-1.860
	Estatura	0.006	0.006	1.040
	Peso	0.008	0.004	1.990
	Estudio en Semana Normal (Pocas Veces)	0.145	0.073	1.970
	Estudio en Semana Normal (Siempre)	0.644	0.217	2.960
	Estudio en Semana Evaluación (Siempre)	-0.334	0.115	-2.890
<i>Efecto Exógeno</i>	Género (1: Hombre)	-0.120	0.110	-1.090
	Índice Aprobación	1.568	0.549	2.850
	Educ. Madre (1: Superior Completa)	0.008	0.086	0.090
	Educ. Padre (1: Superior Completa)	-0.254	0.095	-2.680
	Ocupación Madre (1: Profesional con Grado)	0.081	0.085	0.950
	Ocupación Padre (1: Profesional con Grado)	-0.208	0.091	-2.300
	Ingreso (0.75m - 1.5m)	-0.019	0.101	-0.180
	Ingreso (Mayor a 3.0m)	0.172	0.092	1.860
	Color de Ojos (1: Verdes o Azules)	-0.161	0.128	-1.260
	Estatura	0.016	0.007	2.220
	Intercepto	-5.240	1.395	-3.760
Estadístico Cragg-Donald: 32.735.				

[†] Primera Etapa del modelo en la tabla A.19.

Tabla A.21 PRIMERA ETAPA: Rendimiento de los Alumnos en *Vida Social*

	Variable	Estimación	Desv. Est.	Est. <i>t</i>
<i>Características Propias</i>	Índice Aprobación	0.208	0.336	0.620
	Educ. Madre (1: Magíster y/o Doctorado)	0.017	0.095	0.180
	Vida Social (Pocas Veces)	0.032	0.064	0.500
	Vida Social (Siempre)	-0.132	0.101	-1.300
<i>Características Pares</i>	Género (1: Hombre)	-0.538	0.117	-4.610
	Índice Aprobación	6.276	0.488	12.850
	Educ. Madre (1: Superior Completa)	0.530	0.100	5.290
	Educ. Madre (1: Magíster y/o Doctorado)	1.417	0.241	5.890
	Educ. Padre (1: Magíster y/o Doctorado)	0.036	0.197	0.180
	Ocupación Madre (1: Profesional con Grado)	-0.668	0.093	-7.170
	Estatura	0.035	0.008	4.470
<i>Características Pares Excluidos</i>	Género (1: Hombre)	0.758	0.141	5.390
	Índice Aprobación	3.639	0.907	4.010
	Educ. Madre (1: Superior Completa)	0.927	0.165	5.610
	Ocupación Madre (1: Profesional con Grado)	0.593	0.174	3.410
	Ingreso (0.75m - 1.5m)	-1.190	0.223	-5.320
	Ingreso (1.5m - 3.0m)	-1.729	0.197	-8.780
	Ingreso (Mayor a 3.0m)	-0.983	0.161	-6.120
	Color de Ojos (1: Verdes o Azules)	0.936	0.260	3.600
	Intercepto	-10.446	1.917	-5.450

Tabla A.22 Rendimiento de los Alumnos sobre los Pares en *Vida Social*[†]

	Variable	Estimación	Desv. Est.	Est. <i>t</i>
<i>Efecto Endógeno</i>	Nota Pares	0.256	0.130	1.980
<i>Características</i>	Índice Aprobación	4.591	0.495	9.270
	Educ. Madre (1: Magíster y/o Doctorado)	0.377	0.141	2.680
	Vida Social (Pocas Veces)	0.170	0.099	1.720
	Vida Social (Siempre)	-0.186	0.154	-1.210
<i>Efecto Exógeno</i>	Género (1: Hombre)	0.201	0.163	1.240
	Índice Aprobación	-0.901	0.746	-1.210
	Educ. Madre (1: Superior Completa)	0.207	0.122	1.700
	Educ. Madre (1: Magíster y/o Doctorado)	-0.816	0.334	-2.450
	Educ. Padre (1: Magíster y/o Doctorado)	0.651	0.277	2.350
	Ocupación Madre (1: Profesional con Grado)	-0.186	0.137	-1.360
	Estatura	-0.018	0.011	-1.680
	Intercepto	3.277	1.897	1.730

Estadístico Cragg-Donald: 18.218.

[†] Primera Etapa del modelo en la tabla A.21.

Tabla A.23 PRIMERA ETAPA: Rendimiento de los Alumnos en *Estudio Grupal*

	Grupo de Ref.	Variable	Estimación	Desv. Est.	Est. t	
<i>Características Propias</i>	<i>Propias</i>	Género (1: Hombre)	-0.067	0.075	-0.900	
		Índice Aprobación	1.373	0.397	3.460	
		Educ. Madre (1: Magíster y/o Doctorado)	0.138	0.104	1.330	
		Educ. Padre (1: Magíster y/o Doctorado)	-0.009	0.091	-0.100	
		Ocupación Padre (1: Profesional con Grado)	0.034	0.070	0.490	
		Ingreso (1.5m - 3.0m)	-0.166	0.071	-2.350	
		Ingreso (Mayor a 3.0m)	-0.082	0.075	-1.090	
		Color de Ojos (1: Verdes o Azules)	0.109	0.074	1.470	
		Peso	0.002	0.003	0.530	
	<i>Estudio Grupal</i>	Reuniones Sociales (Pocas Veces)	-0.029	0.071	-0.410	
		Reuniones Sociales (Regularmente)	-0.098	0.086	-1.140	
		Estudio en Semana Normal (Pocas Veces)	0.181	0.068	2.660	
		Estudio en Semana Normal (Siempre)	0.383	0.201	1.910	
		Estudio en Semana Evaluación (Siempre)	-0.046	0.104	-0.440	
	<i>Vida Social</i>	Reuniones Sociales (Pocas Veces)	0.088	0.072	1.230	
		Reuniones Sociales (Regularmente)	0.043	0.096	0.450	
	<i>Características Pares</i>	<i>Estudio Grupal</i>	Índice Aprobación	3.078	0.403	7.650
			Educ. Madre (1: Superior Completa)	0.349	0.096	3.640
			Educ. Madre (1: Magíster y/o Doctorado)	0.716	0.148	4.830
Educ. Padre (1: Superior Completa)			-0.182	0.150	-1.220	
Educ. Padre (1: Magíster y/o Doctorado)			-0.168	0.182	-0.930	
Ocupación Madre (1: Profesional con Grado)			-0.194	0.087	-2.240	
Ocupación Padre (1: Profesional con Grado)			-0.389	0.106	-3.680	
Ingreso (0.75m - 1.5m)			0.137	0.084	1.630	
Color de Ojos (1: Verdes o Azules)			0.350	0.113	3.090	
Estatura			0.013	0.004	2.910	
<i>Vida Social</i>			Género (1: Hombre)	0.413	0.144	2.860
			Índice Aprobación	0.314	0.633	0.500
		Educ. Madre (1: Superior Completa)	-0.098	0.108	-0.910	
		Educ. Madre (1: Magíster y/o Doctorado)	0.025	0.245	0.100	
		Educ. Padre (1: Magíster y/o Doctorado)	-0.145	0.255	-0.570	
		Ocupación Madre (1: Profesional con Grado)	0.297	0.124	2.390	
		Ocupación Padre (1: Profesional con Grado)	-0.302	0.126	-2.400	
Ingreso (Mayor a 3.0m)		-0.039	0.103	-0.380		
Estatura		-0.019	0.009	-2.120		
<i>Características Pares Excluidos</i>	<i>Estudio Grupal</i>	Género (1: Hombre)	0.424	0.081	5.250	
		Índice Aprobación	1.943	0.519	3.740	
		Educ. Madre (1: Superior Completa)	0.320	0.140	2.290	
		Educ. Madre (1: Magíster y/o Doctorado)	0.615	0.187	3.290	
		Educ. Padre (1: Superior Completa)	0.498	0.138	3.620	
		Ocupación Madre (1: Profesional con Grado)	-0.225	0.119	-1.900	
		Ocupación Padre (1: Profesional con Grado)	-0.427	0.114	-3.740	
		Ingreso (1.5m - 3.0m)	-0.503	0.161	-3.120	
		Color de Ojos (1: Verdes o Azules)	0.518	0.138	3.760	
	<i>Vida Social</i>	Género (1: Hombre)	-0.241	0.256	-0.940	
		Índice Aprobación	0.825	1.091	0.760	
		Educ. Madre (1: Superior Completa)	0.520	0.212	2.450	
		Educ. Madre (1: Magíster y/o Doctorado)	0.339	0.349	0.970	
		Educ. Padre (1: Superior Completa)	-0.653	0.213	-3.070	
		Ocupación Madre (1: Profesional con Grado)	-0.493	0.255	-1.940	
		Ocupación Padre (1: Profesional con Grado)	0.777	0.282	2.750	
		Ingreso (0.75m - 1.5m)	-0.067	0.291	-0.230	
		Ingreso (1.5m - 3.0m)	-0.178	0.284	-0.630	
		Ingreso (Mayor a 3.0m)	-0.058	0.210	-0.280	
Color de Ojos (1: Verdes o Azules)	-0.658	0.291	-2.260			
Estatura	0.007	0.015	0.460			
Intercepto	-2.316	3.583	-0.650			

Tabla A.24 PRIMERA ETAPA: Rendimiento de los Alumnos en *Vida Social*

	Grupo de Ref.	Variable	Estimación	Desv. Est.	Est. <i>t</i>	
<i>Características Propias</i>	<i>Propias</i>	Género (1: Hombre)	-0.049	0.088	-0.560	
		Índice Aprobación	0.007	0.464	0.020	
		Educ. Madre (1: Magíster y/o Doctorado)	0.087	0.121	0.720	
		Educ. Padre (1: Magíster y/o Doctorado)	-0.106	0.107	-1.000	
		Ocupación Padre (1: Profesional con Grado)	0.171	0.082	2.080	
		Ingreso (1.5m - 3.0m)	-0.023	0.083	-0.280	
		Ingreso (Mayor a 3.0m)	0.084	0.088	0.950	
		Color de Ojos (1: Verdes o Azules)	-0.026	0.087	-0.310	
		Peso	-0.002	0.004	-0.460	
	<i>Estudio Grupal</i>	Reuniones Sociales (Pocas Veces)	-0.030	0.083	-0.360	
		Reuniones Sociales (Regularmente)	-0.043	0.101	-0.420	
		Estudio en Semana Normal (Pocas Veces)	-0.141	0.080	-1.770	
		Estudio en Semana Normal (Siempre)	0.472	0.235	2.010	
		Estudio en Semana Evaluación (Siempre)	-0.292	0.122	-2.390	
	<i>Vida Social</i>	Reuniones Sociales (Pocas Veces)	0.055	0.084	0.650	
		Reuniones Sociales (Regularmente)	0.157	0.113	1.390	
	<i>Características Pares</i>	<i>Estudio Grupal</i>	Índice Aprobación	0.284	0.471	0.600
			Educ. Madre (1: Superior Completa)	0.111	0.112	0.990
			Educ. Madre (1: Magíster y/o Doctorado)	0.100	0.174	0.580
Educ. Padre (1: Superior Completa)			-0.262	0.175	-1.500	
Educ. Padre (1: Magíster y/o Doctorado)			-0.192	0.213	-0.900	
Ocupación Madre (1: Profesional con Grado)			-0.034	0.102	-0.330	
Ocupación Padre (1: Profesional con Grado)			0.181	0.124	1.470	
Ingreso (0.75m - 1.5m)			0.095	0.099	0.970	
Color de Ojos (1: Verdes o Azules)			-0.098	0.133	-0.730	
Estatura		-0.005	0.005	-1.020		
<i>Vida Social</i>		Género (1: Hombre)	-0.974	0.169	-5.760	
		Índice Aprobación	6.835	0.741	9.230	
		Educ. Madre (1: Superior Completa)	0.658	0.127	5.180	
		Educ. Madre (1: Magíster y/o Doctorado)	1.280	0.286	4.470	
		Educ. Padre (1: Magíster y/o Doctorado)	0.773	0.298	2.590	
		Ocupación Madre (1: Profesional con Grado)	-1.143	0.145	-7.860	
		Ocupación Padre (1: Profesional con Grado)	0.189	0.147	1.280	
		Ingreso (Mayor a 3.0m)	-0.284	0.120	-2.370	
		Estatura	0.062	0.011	5.830	
<i>Características Pares Excluidos</i>	<i>Estudio Grupal</i>	Género (1: Hombre)	0.047	0.094	0.490	
		Índice Aprobación	0.757	0.607	1.250	
		Educ. Madre (1: Superior Completa)	-0.178	0.164	-1.090	
		Educ. Madre (1: Magíster y/o Doctorado)	-0.399	0.219	-1.830	
		Educ. Padre (1: Superior Completa)	-0.286	0.161	-1.770	
		Ocupación Madre (1: Profesional con Grado)	0.162	0.139	1.170	
		Ocupación Padre (1: Profesional con Grado)	0.062	0.134	0.460	
		Ingreso (1.5m - 3.0m)	0.272	0.189	1.440	
		Color de Ojos (1: Verdes o Azules)	0.103	0.161	0.640	
	<i>Vida Social</i>	Género (1: Hombre)	1.658	0.300	5.520	
		Índice Aprobación	4.379	1.277	3.430	
		Educ. Madre (1: Superior Completa)	1.161	0.249	4.670	
		Educ. Madre (1: Magíster y/o Doctorado)	0.828	0.408	2.030	
		Educ. Padre (1: Superior Completa)	1.029	0.249	4.130	
		Ocupación Madre (1: Profesional con Grado)	1.247	0.298	4.190	
		Ocupación Padre (1: Profesional con Grado)	-1.172	0.330	-3.550	
		Ingreso (0.75m - 1.5m)	-1.886	0.340	-5.540	
		Ingreso (1.5m - 3.0m)	-2.235	0.332	-6.720	
		Ingreso (Mayor a 3.0m)	-1.387	0.246	-5.640	
		Color de Ojos (1: Verdes o Azules)	1.917	0.341	5.620	
		Estatura	-0.035	0.018	-1.960	
Intercepto	-10.272	4.193	-2.450			

Tabla A.25 Modelo en Intensidad del Efecto de Pares

	Grupo de Ref.	Variable	Estimación	Desv. Est.	Est. <i>t</i>	
<i>Efecto Endógeno</i>	<i>Estudio Grupal</i>	Nota de Pares	0.491	0.095	6.240	
	<i>Vida Social</i>	Nota de Pares	0.239	0.086	1.620	
<i>Características</i>	<i>Propias</i>	Género (1: <i>Hombre</i>)	-0.110	0.081	-1.370	
		Índice Aprobación	2.278	0.433	5.260	
		Educ. Madre (1: <i>Magíster y/o Doctorado</i>)	0.205	0.113	1.810	
		Educ. Padre (1: <i>Magíster y/o Doctorado</i>)	-0.235	0.100	-2.340	
		Ocupación Padre (1: <i>Profesional con Grado</i>)	0.197	0.081	2.440	
		Ingreso (1.5m - 3.0m)	-0.074	0.083	-0.900	
		Ingreso (Mayor a 3.0m)	-0.187	0.086	-2.170	
		Color de Ojos (1: <i>Verdes o Azules</i>)	-0.170	0.088	-1.920	
		Peso	0.012	0.004	3.430	
		<i>Estudio Grupal</i>	Reuniones Sociales (<i>Pocas Veces</i>)	-0.082	0.084	-0.970
	Reuniones Sociales (<i>Regularmente</i>)		-0.075	0.097	-0.780	
	Estudio en Semana Normal (<i>Pocas Veces</i>)		0.112	0.077	1.460	
	Estudio en Semana Normal (<i>Siempre</i>)		0.482	0.240	2.010	
	Estudio en Semana Evaluación (<i>Siempre</i>)		-0.420	0.120	-3.490	
	<i>Vida Social</i>	Reuniones Sociales (<i>Pocas Veces</i>)	0.155	0.082	1.880	
		Reuniones Sociales (<i>Regularmente</i>)	0.066	0.103	0.640	
	<i>Efecto Exógeno</i>	<i>Estudio Grupal</i>	Índice Aprobación	1.486	0.513	2.900
			Educ. Madre (1: <i>Superior Completa</i>)	0.048	0.109	0.440
			Educ. Madre (1: <i>Magíster y/o Doctorado</i>)	-0.144	0.190	-0.760
			Educ. Padre (1: <i>Superior Completa</i>)	-0.216	0.142	-1.520
Educ. Padre (1: <i>Magíster y/o Doctorado</i>)			0.210	0.191	1.100	
Ocupación Madre (1: <i>Profesional con Grado</i>)			0.116	0.092	1.250	
Ocupación Padre (1: <i>Profesional con Grado</i>)			-0.145	0.105	-1.380	
Ingreso (0.75m - 1.5m)			-0.060	0.091	-0.660	
Color de Ojos (1: <i>Verdes o Azules</i>)			-0.312	0.130	-2.400	
Estatura			0.008	0.005	1.660	
<i>Vida Social</i>		Género (1: <i>Hombre</i>)	0.146	0.126	1.160	
		Índice Aprobación	-0.789	0.550	-1.430	
		Educ. Madre (1: <i>Superior Completa</i>)	0.143	0.087	1.640	
		Educ. Madre (1: <i>Magíster y/o Doctorado</i>)	-0.481	0.236	-2.040	
		Educ. Padre (1: <i>Magíster y/o Doctorado</i>)	0.423	0.264	1.600	
		Ocupación Madre (1: <i>Profesional con Grado</i>)	0.062	0.138	0.450	
		Ocupación Padre (1: <i>Profesional con Grado</i>)	-0.298	0.132	-2.260	
		Ingreso (Mayor a 3.0m)	0.166	0.107	1.550	
		Estatura	-0.010	0.008	-1.330	
		Intercepto	-1.668	1.645	-1.010	

Estadístico Cragg-Donald: 8.126

[†] Primeras Etapas del modelo dado por los pares en EG y VS en las tablas A.23 y A.24, respectivamente.

ANEXO B

Encuesta Alumno

Estudio de Pares y Rendimiento Académico

Encuesta Alumno

Profesor : Mattia Makovec
Alejandra Mizala

Semestre : Primavera 2010

Alumno : Gonzalo Viveros A.

El propósito de esta encuesta consiste en estudiar la relevancia de las redes sociales sobre el rendimiento académico. La información recopilada será procesada con el fin de estimar el impacto de los “*efecto de pares*” sobre los resultados académicos y otros comportamientos sociales de los alumnos de la Facultad, en el ámbito de un proyecto de investigación a cargo de los profesores del Departamento de Ingeniería Industrial Mattia Makovec y Alejandra Mizala, y del alumno tesista Gonzalo Viveros. Los datos se tratarán con confidencialidad, no serán difundidos a terceros y serán utilizados sólo para objetivos de investigación.

B.1. Preguntas sobre sus Pares

B.1.1. Estudio Grupal

En las matrices que se muestra a continuación, se pide que responda las preguntas según el curso que se muestra en cada matriz. En la columna de compañero, por *orden de amistad* nombre a lo más cinco compañeros de universidad, donde 1 es más amigo y 5 es menos amigo, con quien se junta a *estudiar*. Luego, en base de estos compañeros, responder las preguntas 1-7 que se muestran en cada matriz.

	Compañero	P. 1	P. 2	P. 3	P. 4	P. 5	P. 6	P. 7
1								
2								
3								
4								
5								

1. ¿Es su pareja? (pololo(a), esposo(a), etc.)
 - a) Sí
 - b) No

2. ¿Desde cuándo se conocen? (*Mes/Año*, por ejemplo 03/2005)

3. Durante el semestre pasado, ¿Cuántas veces se juntan al mes para hacer vida social?

4. ¿Cómo se conocieron (por primera vez)?
 - a) Compañeros de curso en la Universidad
 - b) Compañeros de Colegio
 - c) Los padres se conocían
 - d) Amigo(a) de los hermanos(as)
 - e) Amigos de vecindario
 - f) Amigos de amigos
 - g) Compañeros de Piso (Departamento, Casa, etc.)
 - h) Amigos(as) de los compañeros de Piso
 - i) Centro de Estudiantes
 - j) Compañeros en actividad deportiva de la Universidad
 - k) Compañeros en actividad deportiva externa a la Universidad
 - l) Practicando alguna actividad recreacional o de voluntariado de la Universidad
 - m) Practicando alguna actividad recreacional o de voluntariado externa a la Universidad
 - n) Actividades Parroquiales
 - ñ) Otra

5. Durante el semestre pasado, ¿Cuántas horas a la semana estudió con su compañero en una semana de estudio normal (sin evaluaciones)?
6. Durante el semestre pasado, ¿Cuántas horas a la semana estudió con su compañero en una semana de evaluación (controles, exámenes, etc.)?
7. ¿Qué hacen cuando estudian juntos? (puede escoger más de una alternativa)
 - a) Estudian a la par
 - b) Enseña a sus compañeros
 - c) Su compañero le enseña
 - d) Presta la solución a las tareas y/o ejercicios
 - e) Se consigue la solución de las tareas y/o ejercicios

B.1.2. Vida Social

A continuación, se pide que por *orden de amistad* nombre a lo más cinco compañeros de Universidad (1 es más amigo y 5 es menos amigo) con quien sólo sale y hace vida social (*no se junta a estudiar*). Luego, en base de estos compañeros, responder las preguntas 8-11 que se muestran a continuación.

	Compañero	P. 8	P. 9	P. 10	P. 11
1					
2					
3					
4					
5					

8. ¿Es su pareja? (pololo(a), esposo(a), etc.)
 - a) Sí
 - b) No
9. ¿Desde cuándo se conocen? (*Mes/Año*, por ejemplo 03/2005)
10. Durante el semestre pasado, ¿Cuántas veces se juntan al mes para hacer vida social?

11. ¿Cómo se conocieron (por primera vez)?
- a) Compañeros de curso en la Universidad
 - b) Compañeros de Colegio
 - c) Los padres se conocían
 - d) Amigo(a) de los hermanos(as)
 - e) Amigos de vecindario
 - f) Amigos de amigos
 - g) Compañeros de Piso (Departamento, Casa, etc.)
 - h) Amigos(as) de los compañeros de Piso
 - i) Centro de Estudiantes
 - j) Compañeros en actividad deportiva de la Universidad
 - k) Compañeros en actividad deportiva externa a la Universidad
 - l) Practicando alguna actividad recreacional o de voluntariado de la Universidad
 - m) Practicando alguna actividad recreacional o de voluntariado externa a la Universidad
 - n) Actividades Parroquiales
 - ñ) Otra

B.2. Preguntas sobre el Entorno Familiar

B.2.1. Situación Socioeconómica de la Familia

12. Nivel Educativo de la Madre
- a) Básico
 - b) Media Incompleta
 - c) Media Completa
 - d) Superior Incompleta
 - e) Superior Completa
 - f) Magíster
 - g) Doctorado

13. Nivel Educativo del Padre

- a) Básico
- b) Media Incompleta
- c) Media Completa
- d) Superior Incompleta
- e) Superior Completa
- f) Magíster
- g) Doctorado

14. Actual Ocupación de la Madre

- a) Trabajo no calificado (vendedor ambulante, conserje, etc)
- b) Trabajo no calificado con contrato, servicio doméstico con contrato, junior o microempresaria (kiosko, taxi, etc)
- c) Agricultores y trabajadores calificados agropecuarios y pesqueros
- d) Operadores de instalaciones y máquinas
- e) Empleada de oficina
- f) Trabajadores de servicios, jefe de sección o técnico especializado (vendedor, etc)
- g) Oficiales y operarios de industrias u oficios (mineros, operarios de la construcción, etc)
- h) Profesional de carreras técnicas o paramédicas
- i) Profesional con Grado de Licenciatura
- j) Miembros del poder ejecutivo y legislativo, personal directivo de la administración pública, dirigentes de organizaciones especializadas, directores de empresa y gerentes de empresa
- k) Fuerzas Armadas
- l) Desempleada
- m) Jubilada

15. Actual Ocupación del Padre

- a) Trabajo no calificado (vendedor ambulante, conserje, etc)
- b) Trabajo no calificado con contrato, servicio doméstico con contrato, junior o microempresario (kiosko, taxi, etc)
- c) Agricultores y trabajadores calificados agropecuarios y pesqueros
- d) Operadores de instalaciones y máquinas
- e) Empleado de oficina
- f) Trabajadores de servicios, jefe de sección o técnico especializado (vendedor, etc)
- g) Oficiales y operarios de industrias u oficios (mineros, operarios de la construcción, etc)
- h) Profesional de carreras técnicas o paramédicas
- i) Profesional con Grado de Licenciatura
- j) Miembros del poder ejecutivo y legislativo, personal directivo de la administración pública, dirigentes de organizaciones especializadas, directores de empresa y gerentes de empresa
- k) Fuerzas Armadas
- l) Desempleado
- m) Jubilado

16. Colegio y Universidad donde su Madre terminó sus estudios

	<i>Nombre</i>	<i>Comuna</i>	<i>Ciudad</i>
Colegio	: _____	_____	_____
Universidad	: _____	_____	_____

17. Colegio y Universidad donde su Padre terminó sus estudios

	<i>Nombre</i>	<i>Comuna</i>	<i>Ciudad</i>
Colegio	: _____	_____	_____
Universidad	: _____	_____	_____

18. El ingreso neto mensual de su hogar de origen es:

- a) Menor a 250.000 CLP
- b) 250.001 - 500.000 CLP
- c) 500.001 - 750.000 CLP
- d) 750.001 - 1.000.000 CLP
- e) 1.000.001 - 1.250.000 CLP
- f) 1.250.001 - 1.500.000 CLP
- g) 1.500.001 - 1.750.001 CLP
- h) 1.750.001 - 2.000.000 CLP
- i) 2.000.001 - 2.250.000 CLP
- j) 2.250.001 - 2.500.000 CLP
- k) 2.500.001 - 2.750.000 CLP
- l) 2.750.001 - 3.000.000 CLP
- m) Mayor a 3.000.000 CLP

B.2.2. Composición de la Familia

19. ¿Está casado?

- a) Sí
- b) No

20. ¿Tiene hijos?

- a) Sí
- b) No

21. (Responder sólo si en la P.20 su respuesta es a) ¿Cuántos hijos tiene?

22. ¿Tiene hermano(s)?

- a) Sí
- b) No

23. (Responder sólo si en la P.22 su respuesta es a) Edad de sus hermanos:

24. ¿Usted pertenece a algún grupo étnico indígena?

- a) Sí
- b) No

25. Su nacionalidad es:

- a) Chilena
- b) Argentina
- c) Boliviana
- d) Brasileña
- e) Colombiana
- f) Peruana
- g) Uruguayaya
- h) Otra

B.3. Preguntas Personales

B.3.1. Gusto Personal o Preferencias

26. Ideología Política

- a) Derecha
- b) Centro-Derecha
- c) Centro
- d) Centro-Izquierda
- e) Izquierda
- f) Ninguna

27. Creencias Religiosas

- a) Católica
- b) Judío
- c) Protestante
- d) Otra
- e) Ateo/Agnóstico

28. ¿Qué hace más en su tiempo libre? (Rankear las primeras 3, donde 1 es de mayor preferencia)

- a) Deporte : _____
- b) Ver Películas/Cine : _____
- c) Ver Programas de TV/Teleseries : _____
- d) Videojuegos/Juegos on-line : _____
- e) Concierto : _____
- f) Teatro : _____
- g) Tocar algún Instrumento Musical : _____
- h) Bailar : _____
- i) Redes Sociales, como Facebook : _____
- j) Reuniones en casa de amigos : _____
- k) Leer : _____
- l) Ir de Shooping : _____
- m) Otra

B.3.2. Contextura Física

29. ¿Color de Pelo?

- a) Oscuro (Negro o Castaño)
- b) Claro (Rubio)
- c) Pelirrojo

30. ¿Color de Ojos?

a) Negro

b) Café

c) Verdes

d) Azules

e) Gris

31. Estatura (Centímetros)

32. Peso (Kilogramos)