



ASOCIACION ARGENTINA
DE ECONOMIA POLITICA

ANALES | ASOCIACION ARGENTINA DE ECONOMIA POLITICA

LIII Reunión Anual

Noviembre de 2018

ISSN 1852-0022

ISBN 978-987-28590-6-0

Determinantes personales e institucionales del
rendimiento académico en materias de primer año
universitario análisis empírico en la Universidad
Nacional de Salta

Gonzalez Fabian Enrique



Determinantes personales e institucionales del rendimiento académico en materias de primer año universitario: análisis empírico en la Universidad Nacional de Salta

Fabián Enrique González
Facultad de Ciencias Económicas, Jurídicas y Sociales
Universidad Nacional de Salta

Resumen

En este trabajo se aplica un modelo probit para intentar explicar el rendimiento académico de una muestra de los estudiantes que cursaron Matemática I en el año 2017, en la Facultad de Ciencias Económicas, Jurídicas y Sociales de la Universidad Nacional de Salta. Las variables explicativas a emplear se refieren a determinantes personales e institucionales como el sexo, la carrera, la calidad de ingresante/recursante, la participación en un aula virtual y la aprobación de un curso de ingreso.

1. Introducción

El estudio del rendimiento académico es una preocupación de diversas disciplinas: la pedagogía, la sociología, la psicología, etc. En economía, es un tema clave de un área en particular: la economía de la educación. Numerosos artículos publicados en Argentina y en el mundo proponen formas de medirlo, ninguna de ellas perfecta, por supuesto.

El estudio de este fenómeno social es complejo debido a la multiplicidad de tipos de factores que entran en juego: psicológicos, sociales, institucionales, etc. Pero a su vez es de particular interés por las implicancias que tiene para el desarrollo de políticas educativas y metodologías pedagógicas.

Debe decirse también que numerosos economistas se preocupan por el rendimiento académico porque ellos mismos son parte de este fenómeno: muchos ejercen docencia. Esto genera un interés de incluir mediciones en los procesos de enseñanza y aprendizaje, llevando al empleo de modelos econométricos que permitan explicar y realizar predicciones del rendimiento académico.

Este trabajo tiene como objetivo estudiar la influencia de factores institucionales como el uso de herramientas pedagógicas como un aula virtual y el dictado de un curso de apoyo al ingreso universitario (de cursado no obligatorio) en el rendimiento académico de los alumnos de la asignatura Matemática I de la Facultad de Ciencias Económicas, Jurídicas y Sociales de la Universidad Nacional de Salta. Se analizarán también características de los individuos como la condición ingresante/recursante, carrera y sexo.

El artículo se organiza de la siguiente manera: en la sección 2 se expone un resumido marco teórico para el trabajo; en la sección 3 se describe a la cátedra de Matemática I; en la sección 4 se definen las variables; en la sección 5 se presenta la metodología; en la sección 6 se analiza descriptivamente las variables; en la sección 7 se muestran los resultados de las estimaciones y en la sección 8 se expone una conclusión. Así mismo, el trabajo cuenta con un anexo sobre el análisis del sesgo de selección muestral.

2. Marco teórico

2.1. El rendimiento académico

A lo largo de la literatura, podemos encontrar distintas definiciones de rendimiento académico. También existe una búsqueda constante de las variables que ayudan a explicarlo. Por ejemplo; el sexo, la edad, si el alumno vive con su familia o no, son determinantes sociales e instituciones (que definiremos más adelante). Lo habitual es identificar al rendimiento con resultados, que pueden ser las calificaciones obtenidas (rendimiento inmediato) o la productividad y eficiencia en el mercado laboral (rendimiento diferido). En el primer grupo, se encuentra la clasificación que señala Tejedor Tejedor y Valcárcel (2007):

- El rendimiento en sentido estricto, medido a través de la presentación a exámenes o éxito en las pruebas (calificaciones), que se traduce en unas determinadas tasas de promoción (superación de curso), repetición (permanencia en el mismo curso más de un año) y abandono (alumnos que dejan de matricularse en cualquiera de los cursos de la carrera)
- El rendimiento en sentido amplio, medido a través del éxito (finalización puntual en un período de tiempo determinado) o del fracaso (retraso o abandono de los estudios). También se habla de regularidad académica, cuando el concepto de rendimiento académico se presenta mediante las tasas de presentación o no a las mesas de examen.

En este trabajo se tomará como definición de rendimiento lo especificado en el primer punto. Particularmente, se trabajará con la condición final de cada individuo al cursar Matemática I. Es decir, el rendimiento académico será el resultado que tienen en la materia luego de haber cursado la misma y rendido los exámenes parciales. Así, se considerará las condiciones regular, promoción, libre y abandono.

2.2. Variables que influyen en el rendimiento

Múltiples investigaciones de los campos de la psicología, la pedagogía y de la economía de la educación (área en la que se enmarca este trabajo) señalan que los factores

que influyen en el rendimiento académico de una materia en particular y de una carrera universitaria en general son variados. Garbanzo Vargas (2007) clasifica estos determinantes en tres grupos:

- Determinantes personales: características propias del estudiante. Aquí se incluye: competencia cognitiva, motivación, autoconcepto académico, formación previa, sexo, edad, aptitudes.
- Determinantes sociales: como todo ser humano, un estudiante universitario interactúa con su ambiente, lo que determina su vida académica. Algunos vendrían a ser: diferencias sociales, entorno familiar, nivel de educación de los familiares, ingresos del hogar.
- Determinantes institucionales: son componentes no personales que intervienen en el proceso educativo. Al interactuar con los factores personales, influyen en el rendimiento académico alcanzado. Dentro de este grupo encontramos: metodologías docentes, carrera elegida (si bien es un factor personal, también se relaciona con la oferta académica), dificultad de la materia, horarios.

En el presente trabajo, se intentará explicar el rendimiento de los alumnos a partir de determinantes personales e institucionales. Esta elección se debe a la disponibilidad de datos y al objetivo del trabajo de aportar mediciones sobre la efectividad de las metodologías de la cátedra de Matemática I, las que se describirán en la siguiente sección.

3. La cátedra de Matemática I

La cátedra de Matemática I pertenece al primer año de las tres carreras que se ofrecen en la Facultad de Ciencias Económicas, Jurídicas y Sociales de la sede central de la Universidad Nacional de Salta, es decir Contador Público Nacional, Licenciado en Administración y Licenciado en Economía. Se dicta en el primer cuatrimestre, por lo que es una de las primeras materias en la vida universitaria de los ingresantes de dicha facultad.

Actualmente, se dictan clases teóricas y prácticas presenciales. Además, se habilita un aula virtual en la cual se disponen cuestionarios evaluativos teóricos y actividades prácticas, todos son de carácter no obligatorios y se realizan de forma periódica, pero si se aprueban otorgan puntos extras en los exámenes parciales. También se disponen espacios de consulta presenciales y en el aula virtual.

Se realizan tres exámenes parciales. Si el alumno aprueba (obtiene una nota igual o superior a 60 en una escala del 1 al 100) el primero, el segundo o ambos, puede rendir el tercero. Si desaprueba este último con una nota mayor o igual a 20 tiene derecho a rendir un recuperatorio. Cuando el estudiante desaprueba los dos primeros parciales o el recuperatorio del tercero, queda en condición de libre. Caso contrario, puede regularizar o promocionar sin rendir examen final. Para la promoción se necesita aprobar los tres parciales y tener una nota promedio de 76 o más, en una escala del 1 al 100.

Los alumnos ingresantes también pueden contar con puntos extras para el primer parcial cuando aprueban un curso de ingreso que se dicta en el mes de febrero, previo al cursado del primer cuatrimestre en marzo.

Una característica interesante de Matemática I, es que separa a ingresantes y recursantes en el cursado: tienen comisiones prácticas y aulas virtuales separadas.

4. Metodología

Se utilizarán los datos que la cátedra recolecta de los alumnos; en este caso se emplearán los que corresponden a los inscriptos a dicha asignatura en el año 2017. Las variables de las que se disponen son: carrera, notas de parciales, puntajes extras provenientes del aula virtual y del curso de ingreso y la condición final de los estudiantes. En total, se cuenta con datos de 1245 individuos.

En el presente trabajo se definirá una variable dependiente dicotómica como proxy. La denominaremos *condición_final* y tomará el valor 1 si el alumno logra un “éxito” y 0 si tiene un “fracaso”. Se entenderá por “éxito” a la regularización o promoción sin examen final y por “fracaso” a abandonar la materia (es decir, si rindiendo el primer parcial no asiste al segundo) o quedar libre según las condiciones explicadas anteriormente.

$$\text{condición_final} \begin{cases} 1 \text{ si hay un éxito (regularizar o promocionar)} \\ 0 \text{ si hay un fracaso (abandonar o quedar libre)} \end{cases}$$

Dado que la variable dependiente es dicotómica, será necesario recurrir a un modelo de regresión de variables dependientes categóricas. En el presente trabajo, se utilizará un modelo probit de probabilidad, el cual presentaremos brevemente.

Siguiendo a Greene (2012), sea Y_i^* una variable de respuesta latente que representará al rendimiento académico, y sea Z_i el vector de datos para un alumno seleccionado aleatoriamente. Definiendo Y_i^* en función de Z_i :

$$Y_i^* = \alpha + Z_i' \beta + u_i \quad (1)$$

Donde β es el vector de coeficientes de las variables independientes, α un término constante y u_i es un error estocástico. Diremos que este modelo es latente puesto que no es observable para nosotros.

Ahora definimos a la variable dicotómica Y_i como la condición final, donde sí el rendimiento académico del alumno Y_i^* alcanza y/o pasa determinado umbral (digamos a), el alumno regulariza o promociona (tiene éxito) y fracasa en otro caso, entonces:

$$Y_i = \begin{cases} 1 \text{ si } Y_i^* \geq a \\ 0 \text{ si } Y_i^* < a \end{cases} \quad (2)$$

De esta manera, podemos definir qué:

$$\begin{aligned} P_i &= P(Y_i = 1 | \alpha, a, Z_i) = P(Y_i^* \geq a | \alpha, a, Z_i) = P(\alpha + Z_i' \beta + u_i \geq a | \alpha, a, Z_i) \\ P(Y_i = 1 | \alpha, a, Z_i) &= P(u_i \geq (a - \alpha) - Z_i' \beta | \alpha, a, Z_i) \\ P(Y_i = 1 | \alpha, a, Z_i) &= P(u_i \leq (\alpha - a) + Z_i' \beta | \alpha, a, Z_i) \end{aligned} \quad (3)$$

Sea $X_i' \beta$ el vector que incluye $Z_i' \beta$ y al término constante $(\alpha - a)$. Por el Teorema del Límite Central, se cumple que asintóticamente u_i sigue una distribución normal estándar, con lo que se puede afirmar:

$$\begin{aligned} P(Y_i = 1 | X_i) &= P(u_i \leq (\alpha - a) + Z_i' \beta | X_i) = P(u_i \leq X_i' \beta | X_i) \\ P(Y_i = 1 | X_i) &= F(X_i' \beta) \end{aligned} \quad (4)$$

La función $F(\cdot)$ es la función de distribución de u_i . Cuando suponemos que el error sigue aproximadamente una distribución normal estándar, con media cero y varianza 1, estamos en el modelo probit; esto es válido por el Teorema Central del Límite. De esta manera, la probabilidad de que un alumno tenga éxito en Matemática I dadas sus características es:

$$P(Y_i = 1 | X_i) = \int_{-\infty}^{X_i' \beta} \phi(X_i' \beta) dX_i' \beta = \Phi(X_i' \beta) \quad (5)$$

Donde $\phi(\cdot)$ es la función de densidad de la distribución normal estandarizada y $\Phi(\cdot)$ su función de distribución. Con esto, se puede obtener que la probabilidad de que un estudiante fracase dadas sus características es:

$$P(Y_i = 0 | X_i) = 1 - P(Y_i = 1 | X_i) = 1 - \Phi(X_i' \beta) \quad (6)$$

Para estimar este modelo, se utilizará el método de máxima verosimilitud. La función de log-verosimilitud es, siguiendo (5) y (6), y considerando que Y_i tiene una distribución binomial:

$$LL = \sum_{i=1}^n [Y_i X_i' \beta - \ln(1 + e^{X_i' \beta})] \quad (7)$$

Los estimadores de los coeficientes de las variables independientes son los que maximizan (7).

Para la interpretación de los coeficientes, derivemos (5) con respecto a una variable independiente X_k se obtiene que:

$$\frac{\partial P_i}{\partial X_k} = \phi(X_i' \beta) \cdot \beta_k \quad (8)$$

Se ve que el efecto marginal de un cambio en una variable independiente variará en cada punto, pero lo que sí es constante es el signo, el cual viene dado por el coeficiente, dado que la función de densidad es siempre positiva. Es por eso que al interpretar los coeficientes cuando utilicemos un modelo probit solo observaremos el signo para saber cuál es el sentido del efecto marginal en la probabilidad de que un alumno regularice o promocióne.

Debido a que en la muestra solo se ha considerado a los alumnos que se inscribieron y asistieron, *Yi* solo será observado si el alumno efectivamente asistió. Esto podría generar un sesgo de selección muestra, por eso se realizó un test para comprobar si existe efectivamente el sesgo, y el resultado fue negativo, por lo que solo tomamos a los 1245 individuos¹.

5. Definición de variables independientes

En este trabajo, intentaremos explicar esta variable a partir de las siguientes dimensiones:

5.1. Sexo

En trabajos como Porto y Di Gresia (2001), Giovagnoli (2002) y Di Gresia (2007) se encuentra que las mujeres tienen significativamente un mejor rendimiento. No obstante, en Dapozo et al. (2010) y Beltrán y La Serna (2009) la variable no resulta significativa.

5.2. Condición del alumno

Como se señaló anteriormente, en la cátedra de Matemática I los alumnos recurrentes tienen comisiones prácticas y un aula virtual por separado de aquellos que son ingresantes. Por otro lado, intuitivamente puede sostenerse que aquellos que están de nuevo cursando la materia conocen la dinámica de evaluación y tienen mayores conocimientos sobre la asignatura que los ingresantes. Sin embargo, el recurrente puede estar cursando materias de otros años, lo que le dificultaría el éxito en Matemática I. En definitiva, se podría decir que a priori no es posible definir el efecto que tiene la variable en la obtención de éxito. Es por estas razones que resulta interesante estudiar si la condición de ingresante o recurrente influye en el resultado obtenido en la asignatura.

5.3. Participación en el aula virtual

La participación en el aula virtual no es obligatoria, pero la realización de cuestionarios evaluativos y actividades prácticas otorga al alumno puntos extras para los exámenes parciales. A su vez son herramientas de autoevaluación, y el carácter periódico genera incentivo a que se estudien los temas dictados en clases con constancia. Por eso, a priori se espera que una mayor participación en el aula virtual mejore el rendimiento académico.

5.4. Carrera elegida

Un dato que se dispone y que también resulta interesante analizar es si la carrera elegida influye en el rendimiento académico. Recuérdese que la oferta académica de la Facultad de Ciencias Económicas, Jurídicas y Sociales está compuesta por contador público nacional, licenciado en administración y licenciado en economía.

5.5. Curso de Ingreso

Los ingresantes tienen la posibilidad de hacer un curso de ingreso preparatorio para la universidad en febrero. En este se dictan, entre otras cosas, contenidos básicos de matemática fundamentales para la comprensión de los temas de Matemática I. No es obligatorio ni eliminatorio, pero su aprobación permite obtener puntos extras para el primer parcial, 10 puntos si aprueba con una nota entre 60 y 79, o 20 si se aprueba con 80 o más, en una escala del 1 al 100. Se espera que su realización y aprobación, entonces, influyan de manera positiva en el rendimiento académico de los ingresantes.

En base a esto, definiremos las siguientes variables independientes

- *sexo*: es una variable binaria que toma los valores 1 si el individuo es mujer y 0 si es varón.
- *recursante*: variable binaria que toma los valores 1 si el individuo es recurrente y 0 si es ingresante.

¹ Véase el anexo.

- *Tasa de participación en el aula virtual (TPAV)*²: esta es una variable cualitativa continua que se define como la cantidad de puntaje extra obtenido para el primer y segundo parcial dividido por la cantidad máxima posible de puntaje extra para dichos parciales.
- *Se inscribió a la licenciatura en administración (LA)*³: variable binaria que toma el valor 1 si el individuo está inscripto para la carrera de licenciatura en administración y 0 en otro caso.
- *Se inscribió a la licenciatura en economía (LE)*: variable binaria que toma el valor 1 si el individuo está inscripto para la carrera de licenciatura en economía y 0 en otro caso,
- *cursoingreso*: variable binaria que toma el valor 1 si el individuo cuenta con puntaje extra del curso de ingreso y 0 en otro caso.

6. Análisis descriptivo de las variables dependientes

A continuación, se expondrá descriptivamente a las variables implicadas. Para casi todas se presentarán tablas de frecuencias, debido a la naturaleza caegórica de las variables. Para el caso de la TPAV, se analizará su histograma y las medidas de posición y variabilidad.

6.1. Condición Final (*condicion_final*)

La tabla 6.1. resume a la variable binaria dependiente. De la muestra de 1245 alumnos, el 77,11% quedó libre o abandonó y el 22,89% promocionó o regularizó.

Tabla 6.1. Condición Final de los estudiantes de Matemática I, 2017

Condición Final	Frecuencia	Porcentaje
Fracaso	960	77.11
Éxito	285	22.89
Total	1,245	100.00

Fuente: inscriptos en Matemática I en 2017, datos recolectados por la cátedra.

6.2. Sexo (*sexo*)

Como se ve, se tiene una mayoría de mujeres, las que significan el 56,95% del grupo estudiado contra el 43,05% de los varones.

Tabla 6.2. Sexo de los estudiantes de Matemática I, 2017

Sexo	Frecuencia	Porcentaje
Varón	536	43.05
Mujer	709	56.95
Total	1,245	100.00

Fuente: inscriptos en Matemática I en 2017, datos recolectados por la cátedra.

² Se considera solo el puntaje extra de los dos primeros parciales puesto que corresponden a más del 60% de los cuestionarios evaluativos y actividades prácticas que se realizan en el aula virtual durante el cursado. A su vez, se divide por la cantidad máxima de puntaje extra debido a que esta magnitud difiere entre recursantes e ingresantes (18 y 15 respectivamente).

³ La categoría de comparación es el ingresante inscripto en la carrera de CPN.

6.3. Recursante (*recursante*)

En la tabla se observa que en la muestra hay una mayoría de ingresantes, con un 53,98% de las observaciones frente al 46,02% de los recursantes.

Tabla 6.3. Ingresante/Recursante; estudiantes de Matemática I, 2017

Recursante	Frecuencia	Porcentaje
Ingresante	672	53.98
Recursante	573	46.02
Total	1245	100.00

Fuente: inscriptos en Matemática I en 2017, datos recolectados por la cátedra.

6.4. Tasa de Participación en el Aula Virtual (*TPAV*)

La Tasa de Participación en el aula virtual (*TPAV*) fue creada a los fines de cumplir con el objetivo planteado por esta investigación. La tabla 6.4. y el gráfico 6.1. muestran una distribución con un claro sesgo hacia la derecha. La media es de 0,24 y el desvío es 0,25, por lo que se tiene un coeficiente de variación mayor al 100%. El TPAV tiene una variabilidad considerable debido a la existencia de que la mayor cantidad de valores se concentran en tasas de participación bajas pero existen casos atípicos de valores muy altos. Un 33% de la muestra presenta una TPAV=0, por lo que para mejor interpretación se agrega un histograma con TPAV≠0,

Tabla 6.4. TPAV para los estudiantes de Matemática I, 2017

Variable	Observaciones	Media	Desvío	Min	Max
TPAV	1245	.2400714	.25637	0	1

Gráfico 6.1. Histograma del TPAV para los estudiantes de

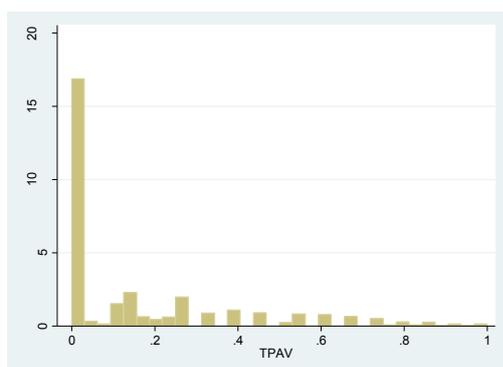
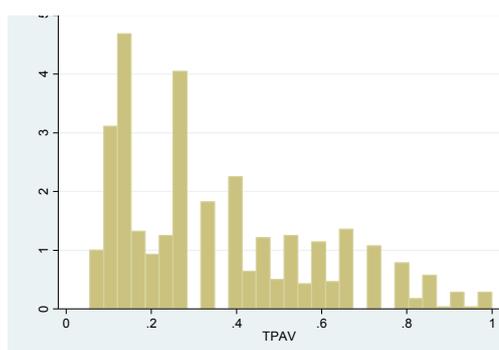


Gráfico 6.2. Histograma del TPAV para los estudiantes de Matemática I, 2017. Con TPAV≠0



Fuente: inscriptos en Matemática I en 2017, datos recolectados por la cátedra.

6.5. Curso de Ingreso (*cursoingreso*)

A la variable que indica si el alumno aprobó el curso de ingreso (por ende cuenta con puntos extras para el primer parcial) se describiera, por lo tanto, solo para el subconjunto de ingresantes de la muestra.

Del total de 672 ingresantes, el 20,09% cuenta con puntos del curso de ingreso para el primer parcial, frente a un 79,91% que no los consiguió o ni si quiera asistió.

Tabla 6.5. Estudiantes ingresantes que cuentan con puntos del curso de ingreso, Matemática I, 2017

Curso Ingreso	Frecuencia	Porcentaje
No tiene puntos	537	79.91
Tiene puntos	135	20.09
Total	672	100.00

Fuente: inscriptos en Matemática I en 2017, datos recolectados por la cátedra.

6.6. Carrera (LA, LE)

Un 62,33% de la muestra está inscripta en la carrera de Contador Público Nacional, un 31,73% en la de Licenciado en Administración y un 5,94% en la de Licenciado en Economía.

Tabla 6.6. Estudiantes clasificados por carrera, Matemática I, 2017

Carrera	Frecuencia	Porcentaje
LA	395	31.73
LE	74	5.94
CPN	776	62.33
Total	1245	100.00

Fuente: inscriptos en Matemática I en 2017, datos recolectados por la cátedra.

De todo el análisis, se resume que la mayoría de la muestra fracasó en el cursado de Matemática I (en el sentido de que se ha definido fracaso), hay mayoría de mujeres, ingresantes e inscriptos en la carrera de CPN, y que quienes cuentan con los puntos extras del curso de ingreso y de la participación en el aula virtual son pocos. La alta tasa de fracasos podría dar a priori indicios de que las herramientas pedagógicas de la cátedra no serían las correctas si se busca mejorar las tasas de éxitos. Sin embargo, también es posible que los pocos estudiantes que regularizan o promocionan son aquellos que tienen puntos extras del curso de ingreso y que aprueban las actividades virtuales, señalando que dichas estrategias de enseñanza influyen en el rendimiento académico del alumno de Matemática I en dicha asignatura.

7. Estimación econométrica

Aquí, como se señaló en el apartado de definición de variables independientes, se explicará la condición final del estudiante de Matemática I en dicha materia a partir de la condición de ingresante/recursante, sexo, carrera, la aprobación del curso de ingreso y los puntos extras obtenidos en actividades de la plataforma virtual. Como la variable dependiente es dicotómica, utilizaremos un modelo probit. Para realizar esto y cumplir con el objetivo de la investigación, se estimó dos modelos:

- Uno en que se utiliza toda la muestra, en el que la variable *condición_final* dependerá de *sexo*, *recursante*, *TPAV*, *LE* y *LA* (Modelo ampliado).
- Otra en la cual solo se trabajará con los datos de ingresantes, reemplazándose la variable *recursante* por *cursoingreso*. (Modelo restringido).

La intención de especificar y estimar dos modelos es estudiar, por un lado, la influencia de la variable *recursante* en el rendimiento académico de los individuos de la muestra en Matemática I y, por el otro, el efecto de *cursoingreso* en el subconjunto de ingresantes.

A continuación, están detallados los resultados.

7.1. Modelo ampliado

La tabla 7.1. muestra las estimaciones de los coeficientes de las variables explicativas.

Tabla 7.1. Modelo ampliado

Condición_final	Coeficientes	
Recursante	.3134532***	***Significativo al 1%
Sexo	-.081103	**Significativo al 5%
LE	-.174674	*Significativo al 10%
LA	-.0782733	
TPAV	3.061642***	
_cons	-1.715046***	

La log-verosimilitud es de -500.36894.

Utilizaremos como medida de bondad de ajuste el Cuanto Cuadrado, el cual asigna éxito al individuo cuando la probabilidad predicha es mayor o igual que 0.5. Luego se compara con el valor verdadero. El modelo acierta un 82,81% en sus predicciones, por lo que tiene una buena bondad de ajuste a los datos.

Las variables *sexo*, *LE* y *LA* son no significativas a los niveles habituales, con lo que tenemos evidencia de que el rendimiento académico no varía entre mujeres y varones, ni tampoco según la carrera a la cual los alumnos están inscriptos. Que el coeficiente de la variable *sexo* no haya sido significativo es un resultado que se observa en la literatura vinculada a la temática.

Por otro lado, se ve que el coeficiente de *recursante* tiene significancia, es decir que si hay diferencias entre el rendimiento académico de recursantes e ingresantes, teniendo mejor performance los primeros. Esto podría estar explicado por la intuición expresada con anterioridad: los recursantes ya conocen la dinámica de la materia y parte con más conocimientos previos sobre los contenidos que los ingresantes. Otro factor que podría estar influyendo favorablemente es la discriminación de comisiones prácticas y aula virtual. Esta metodología de enseñanza podría estar beneficiándolos.

El coeficiente de *TPAV* también es significativo, un indicio de que una mayor participación en el aula virtual mejora el rendimiento académico aumentando la probabilidad de que los alumnos promocionen o regularicen la materia. En otras palabras, el uso del aula virtual con incentivos a que los alumnos estudien con constancia mediante la habilitación de actividades no obligatorias que les permiten obtener puntos extras tiene un efecto positivo en las tasas de promoción y regularización.

Como consecuencia de los resultados obtenidos, se propone el mismo modelo pero se excluyen las variables *sexo*, *LE* y *LA* (Modelo ampliado ajustado). La tabla 7.2. muestra las estimaciones.

Tabla 7.2. Modelo ampliado ajustado

Condición_final	Coeficientes	
Recursante	.3051197***	***Significativo al 1%
TPAV	3.046857***	**Significativo al 5%
_cons	-1.787288***	*Significativo al 10%

Para comparar las bondades de ajuste de ambos modelos, utilizamos un test utilizando el likelihood-ratio.

La log-verosimilitud es de -501.31554.

El modelo acierta un 82,57% en sus predicciones, por lo que tiene una buena bondad de ajuste a los datos.

Ahora todos los coeficientes son significativos y tienen la misma interpretación expuesta para el modelo sin ajustar.

Las hipótesis son:

- H0=los coeficientes de las variables excluidas no son significativamente distintas de cero.
- H1=al menos uno de ellos es significativamente distinta de cero.

El estadístico de prueba es el likelihood-ratio que se define como:

$$LR = 2(L_u - L_r)$$

Donde L_u es la log-verosimilitud del modelo con las variables incluidas y L_r es la del modelo restringido.

LR sigue asintóticamente una distribución chi-cuadrado con 3 grados de libertad.

Consideraremos un nivel de significancia del 5%. Así, el valor crítico es 7,81. El LR en este caso es 1.89. Con esto, no se rechaza la hipótesis nula; hay evidencia de que el modelo ampliado ajustado tiene una mayor bondad de ajuste.

7.2. Modelo restringido

En este caso estimaremos un modelo probit solo con los 672 ingresantes, debido a que estamos interesados en analizar el efecto de poseer los puntos extras del curso de ingreso en el rendimiento académico.

Los coeficientes que se obtienen son los que se presentan en la tabla 7.3.

Tabla 7.3. Modelo restringido

Condición_final	Coefficientes	
TPAV	2.621546***	***Significativo al 1%
LA	-.0671631	**Significativo al 5%
LE	.3369716	*Significativo al 10%
Sexo	-.0714567	
Cursoingreso	.7368701***	
_cons	-1.779026***	

La log-verosimilitud es -256.63841.

El modelo predice correctamente en un 83,93% de los casos, por lo que tiene buena bondad de ajuste.

Nuevamente se observa que no hay diferencias entre varones y mujeres y las tres carreras en lo que se refiere a rendimiento académico en Matemática I. A su vez, en el subgrupo de ingresantes sigue verificándose un efecto significativo y positivo de una mayor participación en el aula virtual.

Analizando el coeficiente de la variable *cursoingreso*, es significativamente distinto de cero y tiene un signo positivo. Es decir hay razones estadísticas para afirmar que aquellos alumnos que aprobaron el curso de ingreso, y que por ende tienen un puntaje extra en el primer parcial, tienen una mayor probabilidad de regularizar o promocionar la materia. El signo coincide con lo esperado en la sección de especificación de variables.

Se propone un modelo ajustado en el cual se excluyen las variables *sexo*, *la* y *le*. Las estimaciones correspondientes estas presentadas en la tabla 7.4.

Tabla 7.4. Modelo restringido ajustado

Condición_final	Coefficiente	
TPAV	2.56189***	***Significativo al 1%
Cursoingreso	.7631086***	**Significativo al 5%
_cons	-1.803971***	*Significativo al 10%

La log-verosimilitud es -257.78937

El modelo predice correctamente en un 83,93% de los casos, por lo que tiene una buena bondad de ajuste.

Todos los coeficientes son significativos, y tienen la misma interpretación antes enunciada.

Para comparar las bondades de ajuste de ambos modelos, vamos a utilizar el test al que recurrimos con anterioridad.

- H0=los coeficientes de las variables excluidas no son significativamente distintas de cero.
- H1=por lo menos uno de ellos es significativamente distinta de cero.

El LR= 2.30. Se distribuye chi cuadrado con 3 grados de libertad. El Valor crítico correspondiente al 5% de significancia es 7.81. No se rechaza la hipótesis nula: el segundo modelo tiene una mejor bondad de ajuste.

7.3 Predicciones de probabilidad y efectos marginales

A continuación se presentan las predicciones de la probabilidad de tener éxito en el cursado de Matemática I y los efectos marginales para distintos individuos hipotéticos con determinadas características. Utilizaremos ambos modelos ajustados.

Se expondrá un cuadro de doble entrada. Dentro de las celdas, los números resaltados son las predicciones de probabilidad y están expuestos los efectos marginales para cada una de las variables. En el caso del modelo ampliado, las estimaciones son las siguientes:

Tabla 7.5. Predicciones de probabilidad y efectos marginales en el modelo ampliado ajustado

	<i>Ingresante</i>	<i>Recursante</i>
TPAV=0	.03694549	.06914778
	TPAV .2460977***	TPAV .4052549***
	recursante .0322023***	recursante .0322023***
TPAV=.5	.39594422	.5164559
	TPAV 1.173935***	TPAV 1.214486***
	recursante .105117***	recursante .1205117***
TPAV=1	.89608764	.9411721
	TPAV .5498633***	TPAV .3573806***
	recursante .0450845***	recursante .0450845***

Observamos el efecto positivo que tiene las variables *TPAV* y *recursante*. El efecto marginal de *recursante*, expuesto en el cuadro, es la estimación de la diferencia positiva en la probabilidad de éxito entre recursantes e ingresantes. Nótese que si un individuo es recursante, en el caso de que *TPAV=0* y *TPAV=.5*, el efecto marginal de la participación en el aula virtual es más alto. Para ambas variables, la tabla sugiere que dicha variable tiene en su efecto marginal un movimiento en U invertida cuando *TPAV* aumenta.

Analizando el modelo en el cual solo se consideran a los alumnos ingresantes:

Tabla 7.6. Predicciones de probabilidad y efectos marginales en el modelo restringido ajustado

	<i>Tiene puntos</i>	<i>No tiene puntos</i>
TPAV=0	.14896971	.03561793
	TPAV .5945856***	TPAV .2008195***
	cursoingreso .1133518***	cursoingreso .1133518***
TPAV=.5	.59486687	.30047808
	TPAV .9930114***	TPAV .8913916***
	cursoingreso .2943888***	cursoingreso .2943888***

TPAV=1	.93587353	.77575025
	<i>TPAV</i>	<i>TPAV</i>
	.3214364***	.7668885***
	<i>cursoingreso</i> .1601233***	<i>cursoingreso</i> .1601233***

Están expuestos los efectos positivos que tienen *TPAV* y *cursoingreso*. El efecto marginal de *cursoingreso*, expuesto en el cuadro, es la estimación de la diferencia positiva en la probabilidad de éxito entre ingresantes que aprobaron el curso de ingreso y aquellos que no. Obsérvese que el efecto marginal de *TPAV* es mayor para aquellos individuos que cuentan con puntos extras del curso de ingreso en el caso en que esta variable sea igual a 0 y 0.5. Ambas variables tienen efectos marginales con forma de U invertida a medida que aumenta la participación en el aula virtual.

8. Conclusión

A lo largo del trabajo se estudió el rendimiento académico de alumnos que cursaron Matemática I en 2017. Resulta interesante destacar los resultados obtenidos.

Primero, se encontró que la participación en el aula virtual es un factor que influye favorablemente en el rendimiento académico del alumno. Para medir esta participación se utilizó una Tasa de Participación en el Aula Virtual (TPAV) y al plantear el modelo probit se encontró una relación positiva entre la probabilidad de promocionar o regularizar la materia y valores más altos de la tasa. En otras palabras, este trabajo evidencia que la aplicación de herramientas pedagógicas como las del aula virtual de Matemática I tiene impacto positivo en las tasas de promoción y regularidad.

Segundo, se observó que los recursantes tienen más probabilidad de regularizar o promocionar la materia. Se señaló que esto se puede deber a los conocimientos previos del estudiante y/o al trato diferencial entre ingresantes y recursantes que se realiza en el cursado de la asignatura. Los efectos de la participación en el aula virtual se ven potenciados para los recursantes.

Tercero, estudiando solo a los ingresantes se halló que la aprobación del curso de ingreso favorece a la performance en la cursada. Los conocimientos básicos dictados en este y los puntos extras obtenidos influyen positivamente. Los efectos de la participación en el aula virtual se ven potenciados para los que aprobaron dicho curso.

Por último, en la muestra no hay evidencia de diferencias en el desempeño entre varones y mujeres, y entre carreras.

La evidencia aquí presentada nos sugiere que para mejorar las tasas de regularidad y promoción, particularmente en una asignatura de primer año universitario con características similares a la de Matemática I, se deben aplicar y perfeccionar herramientas pedagógicas virtuales y el dictado de temas básicos en los cursos de ingreso.

9. Anexo: Análisis de la existencia del sesgo de selección muestral.

Como se dijo anteriormente, es probable que exista un sesgo de selección muestral puesto que se excluyó a los alumnos que se inscribieron a la materia pero no asistieron. Para ver si existe un problema de esta naturaleza, se hizo un procedimiento que sigue la lógica del Heckit, pero con una ecuación de interés no lineal que se estima mediante un probit. Este método lo podemos encontrar expuesto en Greene(2012).

En términos econométricos, la primera ecuación es la ecuación de selección, que representa el proceso por el cual los individuos forman parte de la muestra o no. La segunda, es nuestro modelo de interés.

1. Ecuación de selección (E.S.):

$$Y_{2i}^* = Wi'\alpha + u_{2i} \quad (9)$$

Donde Y_{2i}^* es una variable latente que determina si el individuo, habiéndose inscrito, asiste o no a clases. Así, definimos la variable Y_{2i}^* que toma el valor 1 si el individuo asiste y el valor 0 si no lo hace. Esta ecuación puede estimarse mediante el método probit antes enunciado.

2. El modelo de interés (M.I.)⁴:

$$Y_i^* = Xi'\beta + u_i \quad (10)$$

Es el modelo antes enunciado, con Y_i^* =condición final y Xi el vector de valores de las variables del individuo i , en el cual solo se observará ($Y_i^* Xi$) cuando $Y_{2i}^* = 1$.

Cabe aclarar que Wi debe incluir a Xi , e incluso puede darse el caso en que $Wi = Xi$ según el criterio del investigador.

Se define a ρ como el coeficiente de correlación entre u_i y u_{2i}

$$\text{corr}(u_i; u_{2i}) = \rho \quad (11)$$

Si este coeficiente es significativamente distinto de cero, entonces ambos errores están correlacionados, con lo que se diría que existe un problema de selección muestral. Caso contrario, solo se estimaría la ecuación (10) sin tener en cuenta (9).

Aquí es necesario utilizar una extensión del modelo probit bivariado, el cual se especifica de la siguiente manera:

$$P(Y_i = 1 | Y_{2i} = 1) = \frac{\Phi_2(Xi'\beta, Wi'\alpha, \rho)}{\Phi(Wi'\alpha)} \quad (12)$$

Donde $\Phi_2(\cdot)$ es la función de distribución de la normal estandarizada bivariada.

Con este resultado, podemos construir la función de verosimilitud para estimar los parámetros de ambas ecuaciones. Existen tres tipos de observaciones posibles en la muestra:

$$\begin{aligned} Y_{2i} = 0: P(Y_{2i} = 0 | Xi, Wi) &= 1 - \Phi(Wi'\alpha) \\ Y_i = 0; Y_{2i} = 1: P(Y_i = 0, Y_{2i} = 1 | Xi, Wi) &= \Phi_2(-Xi'\beta, Wi'\alpha, -\rho) \\ Y_i = 1; Y_{2i} = 1: P(Y_i = 1, Y_{2i} = 1 | Xi, Wi) &= \Phi_2(Xi'\beta, Wi'\alpha, \rho) \end{aligned} \quad (13)$$

Así, la función de verosimilitud será:

$$\prod_{i=1}^n \Phi_2[(2Y_i - 1)Xi'\beta; Wi'\alpha; (2Y_i - 1)\rho][1 - Y_{2i} + Y_{2i}(1 - \Phi(Wi'\alpha))] \quad (14)$$

Puede aplicarse logaritmo natural para obtener la log-verosimilitud. Se maximizará esta última o la función presentada.

Con este método, estimaremos tanto el modelo restringido como el ampliado (ver la sección de Estimación econométrica) con selección muestral. Consideraremos que $Xi = Wi$ en ambos casos.

Estimando el modelo ampliado con selección muestral, los resultados son los siguientes.

⁴ Sin pérdida de la generalidad, suponemos que $\alpha = 0$ y $\alpha + Zi'\beta = Xi'\beta$.

Cuadro 9.1.: estimación del modelo ampliado con selección muestral

	Coeficiente	Error Estandar	Valor-p
condicion_final(M.I.)			
TPAV	3.337.979	.2828585	0.000
LA	-.1007285	.0983577	0.306
LE	-.1664467	.2056744	0.418
recursante	.3157917	.0932484	0.001
Sexo	-.0815086	.0907954	0.369
_cons	-1.859.085	.153271	0.000
asistio(E.S.)			
TPAV	9.342.855	.6205454	0.000
Sexo	-.0258755	.0742057	0.727
recursante	.0996303	.0746288	0.182
LE	.1612985	.1549421	0.298
LA	-.2667959	.0779365	0.001
_cons	-.0684785	.0779778	0.380
Rho	.3380442	.3458701	

Para probar si rho es significativamente distinta de cero, utilizaremos un test de Wald:

H0: $\rho=0$ (No existe sesgo de selección)

H1: $\rho \neq 0$ (Existe sesgo de selección)

El estadístico de prueba es una variable que se distribuye asintóticamente chi cuadrado con 1 grado de libertad, y para este caso su valor es 1.61. Esto determina que su valor p será de 0.20, por lo que no se rechaza la hipótesis nula. Rho no es significativamente distinta de cero.

En el caso del modelo restringido en que se estudió solo ingresantes, los resultados son los siguientes:

Cuadro 9.2.: Estimación del modelo restringido con selección muestral

	Coeficientes	Error Estandar	Valor P
condicion_final (M.I.)			
TPAV	301.528	.2342231	0.000
LA	-.0928307	.1327669	0.484
LE	.2966078	.2510233	0.237
cursoingreso	.6916804	.1437523	0.000
Sexo	-.0854911	.1246407	0.493
_cons	-1.976.383	.1295622	0.000

asistio (E.S.)			
TPAV	8.557.421	.7895286	0.000
Sexo	-.0355981	.1065574	0.738
cursoingreso	-.3724585	.2244314	0.097
LE	.1088202	.210861	0.606
LA	-.4244172	.1123554	0.000
_cons	.0560727	.096441	0.561
Rho	.9892528	.9645433	

Aplicamos nuevamente el mismo test sobre rho. El estadístico de prueba es 3.93 y su valor p 0,0475. En este caso, se rechaza la hipótesis nula al 1% de significancia. Con esto, afirmamos que no hay sesgo de selección.

10. Bibliografía

- Beltrán, A. y La Serna Studzinski, K. (2009): ¿Qué explica la evolución del rendimiento académico universitario? Un estudio de caso en la Universidad del Pacífico. Universidad Pacífico.
- Dapozo, G., López, M. y Porcel, E. (2010) Predicción del rendimiento académico de alumnos de primer año de la FACENA (UNNE) en función de su caracterización socioeducativa. Universidad Nacional del Nordeste.
- Di Gresia, L. (2007) Rendimiento académico universitario. Universidad Nacional de la Plata.
- Garbanzo Vargas, G. (2007) Factores asociados al rendimiento académico en estudiantes universitarios, una reflexión desde la calidad de la educación superior pública.
- Giovagnoli, P. (2002) Determinantes de la deserción y graduación universitaria: Una aplicación utilizando modelos de duración. Universidad Nacional de la Plata.
- Green, W. H. (2012) Econometric Analysis. Editorial Pearson. Páginas 778-795 y 920-923.
- Novales Cinca, A. (1993) Econometría. Editorial McGraw-Hill.
- Tejedor Tejedor, F. y García-Valcárcel Muñoz-Repiso, A. (2007) Causas del bajo rendimiento del estudiante universitario (en opinión de profesores y alumnos). Propuestas de mejora en el marco del EEES. Universidad de la Salamanca.
- Universidad Nacional de Salta. Facultad de Ciencias Económicas, Jurídicas y Sociales. Base de datos de la cátedra de Matemática I.