

**MÉTODO DE SELECCIÓN Y
RESULTADOS ACADÉMICOS:
ESCUELA DE INGENIERÍA DE
LA UNIVERSIDAD DE CHILE***

Ronald Fischer y Andrea Repetto

En este trabajo se estudia la capacidad predictiva de rendimiento académico que tienen los distintos componentes del sistema de admisión a la Escuela de Ingeniería de la Universidad de Chile que se utilizó durante el período 1994 y 2002. Los resultados del estudio indican que las notas de educación media y las pruebas específicas de física y matemáticas tienen una capacidad predictiva importante y estable a lo largo de los estudios de ingeniería, y que el efecto conjunto de las pruebas específicas es aproximadamente el doble del efecto de las notas. La prueba de aptitud matemática (PAM) no exhibe un poder predictivo independiente, lo que aparentemente se puede explicar por el restringido rango de puntajes entre quienes ingresan a la Escuela de Ingeniería. Por otro lado, los resultados en la prueba SIMCE obtenidos por el colegio de origen son significativos, lo que parece indicar que la capacidad predictiva de las notas de la enseñanza media depende de los resultados del colegio en la prueba SIMCE.

RONALD FISCHER. Ph. D. en Economía, University of Pennsylvania. Ingeniero Civil Matemático, Universidad de Chile. Profesor-Investigador del Centro de Economía Aplicada (CEA), Depto. Ingeniería Industrial de la Universidad de Chile.

ANDREA REPETTO. Ph. D. en Economía, MIT. Profesor asistente del Departamento de Ingeniería Industrial de la Universidad de Chile.

* Agradecemos los datos suministrados por Patricio Poblete y la ayuda de Loreto Arenas. También agradecemos a Carmen Le Foulon por los datos a nivel nacional utilizados en las comparaciones. Harald Beyer y los participantes en un seminario del CEA (Centro de Economía Aplicada) nos dieron valiosas sugerencias. Por último, Martín Osorio y David Rappaport fueron excelentes ayudantes de investigación. Este trabajo contó con el apoyo de una donación de la Fundación Hewlett.

1. Introducción

Las recientes controversias sobre el mecanismo de selección apropiado para ingresar a la universidad han dejado al descubierto la notoria ausencia de estudios sobre el tema, a pesar de la rica experiencia de treinta años de un sistema de selección que casi no ha sufrido cambios. Es por lo tanto urgente disponer de estudios que permitan tomar decisiones informadas acerca de los efectos del actual sistema, y así decidir si un cambio de estrategia es deseable. Este trabajo contribuye a esta línea con una descripción del poder explicativo de los distintos componentes del sistema de admisión a la Escuela de Ingeniería de la Universidad de Chile.

Desde el punto de vista de estudiar el efecto de los puntajes de entrada sobre el rendimiento, el hecho de ser un estudio restringido a un grupo de alumnos sometidos a exigencias similares es una ventaja. Durante los primeros dos años y medio, los alumnos toman los mismos cursos, por lo que las comparaciones de rendimiento son más fáciles.

Los resultados se refieren a todos los alumnos ingresados en el período 1994-2002, y son un análisis preliminar, en el sentido que no se ha aprovechado toda la riqueza contenida en los datos, sino que sólo se trata de responder a preguntas acotadas sobre la relación entre las capacidades del postulante, medidas en base a los mecanismos de admisión a la Escuela de Ingeniería, y los resultados que los alumnos obtienen en ella. Es claro que los resultados están censurados, porque sólo se dispone de los resultados de quienes fueron admitidos y no los de toda la población. Si todos los interesados hubieran podido ingresar, habría sido posible estimar mejor el efecto de mayores puntajes. Este problema es especialmente grave en lo que se refiere a la prueba de aptitud matemática. En el año 2002, por ejemplo, el puntaje promedio de entrada en esta prueba fue de 791 puntos. Esto significa que los alumnos que ingresaron se encontraban en el 2% superior de la población de alumnos y que la varianza disponible en la muestra es menos de un 25% de la varianza en los datos originales¹.

En este sentido, es natural esperar que la influencia de los distintos indicadores usados en la selección sea relativamente baja, pues buena parte del poder explicativo de los puntajes se pierde en el umbral de entrada a la Escuela de Ingeniería, y que sean otros factores los que expliquen gran parte de la varianza de los resultados de los alumnos. En otras palabras,

¹ Los alumnos que ingresan pertenecen al 5% superior de los estudiantes que toman la prueba de conocimientos específicos de matemática. Este es su grupo de referencia, pues son quienes pueden postular a carreras de ingeniería.

mientras más severo sea el proceso de selección, se hacen relativamente más importantes en el desempeño posterior al ingreso las variables que no forman parte del proceso de selección. A pesar de esta grave desventaja, los datos son interesantes pues todos los alumnos en la muestra se ven sometidos a exigencias similares, a diferencia del caso de un estudio basado en todos los alumnos admitidos a una universidad, los que enfrentan una gran varianza en la dificultad de los cursos en las distintas carreras².

Existen distintas formas de medir el desempeño, y en este informe utilizamos varias de ellas. A pesar de esto, los resultados son bastante uniformes, en el sentido de no afectar las conclusiones en forma sustancial, lo que permite tener confianza en la robustez de las estimaciones. Consideramos medidas tales como el porcentaje de repeticiones de cursos, las notas ponderadas en el primer año y a lo largo de la carrera, y el porcentaje de avance respecto a los hitos en cada año de la carrera³.

Nuestros resultados indican que el puntaje de ingreso es un buen indicador del rendimiento de los alumnos. Dentro de las componentes del puntaje de ingreso que explican el desempeño, son claves los efectos de las notas de la educación media (NEM), de la prueba específica de matemáticas (PEM) y de la prueba específica de física (PEF), ya que son significativas y con un efecto importante en el rendimiento en cada caso. En cambio, los demás componentes del puntaje, la prueba de aptitud matemática (PAM), la prueba de aptitud verbal (PAV) y la prueba de historia y geografía (HYG) son significativas en regresiones con algunas variables de desempeño y no en otras, y en general tienen menos impacto sobre los resultados finales. El bajo poder explicativo de la PAM se debe, probablemente, a su escasa varianza en los datos. Se debe notar que su correlación con los indicadores de rendimiento, corregida por la restricción de rango, es bastante alta, y es la mayor de todas las correlaciones de componentes del mecanismo de selección y las variables de resultados de los alumnos. Esto no es demasiado relevante, sin embargo, pues el efecto de la PAM desaparece en

² Este problema puede resolverse en cierta medida usando variables *dummy*.

Para algunos resultados y controversias en esta área, véase R. Aravena, G. del Pino y E. San Martín, "Capacidad Predictiva de la PAA" (2003), pp. 203-217; así como B. Vial y R. Soto, "¿Predice la PAA el Rendimiento o Éxito en la Universidad?" (2002), pp. 24-27.

³ Hemos, además, analizado el porcentaje de alumnos que egresan en seis, siete y ocho años. Los resultados no son tan robustos como los que se refieren a las otras variables de rendimiento. Es posible que esto se deba a que las exigencias del proceso de titulación en la Escuela de Ingeniería variaron durante el período de análisis. Asimismo, es posible que los mejores alumnos reciban buenas ofertas laborales antes de graduarse o que realicen memorias más difíciles, por lo que tardarían más en titularse. Los resultados de este análisis, no reportados en el trabajo, se encuentran disponibles a través de los autores.

una regresión, ya que su efecto predictivo está probablemente incorporado en los de la PEM y PEF.

Otro resultado interesante es que la combinación de las pruebas específicas de física y matemáticas tiene normalmente más poder predictivo que las notas de la enseñanza media y en el caso de algunas variables de desempeño, este poder es el doble del original. Por último, el efecto de estas pruebas y de las notas de la educación media se reduce en la medida en que los alumnos avanzan en la carrera. Otro grupo de factores que podrían ser de interés son los aspectos socioeconómicos. En un trabajo reciente se muestra que los resultados difieren entre alumnos de distintos segmentos socioeconómicos, pero que las diferencias, aunque significativas, no son importantes y son de no más de dos décimas en los promedios ponderados a lo largo de la carrera⁴.

Este resultado muestra que las componentes tradicionales del proceso de admisión son capaces de seleccionar alumnos en forma homogénea, independientemente de su condición socioeconómica.

Resulta interesante analizar en más detalle la influencia de las NEM en el rendimiento. Existen proposiciones para elevar la influencia de este indicador en la ponderación de entrada a las universidades. Por su parte, los críticos a estas propuestas han señalado que las notas de distintos colegios no son comparables y que aumentar su ponderación en el puntaje incentivará la inflación de notas. Este trabajo presenta evidencia que muestra que la calidad del colegio, medida por el promedio del colegio en las pruebas SIMCE, tiene un impacto importante en los resultados de los alumnos⁵. Tal vez más relevante aún, este efecto se mantiene a lo largo de la carrera. Dado que el SIMCE promedio mide la capacidad que tiene el colegio de entregar conocimientos a los alumnos, el hecho que su efecto no esté subsumido en el promedio de notas indica que las notas no miden lo mismo. En consecuencia, si se desea aumentar la capacidad predictiva de los puntajes de entrada, deberían corregirse los promedios de notas mediante el puntaje SIMCE del colegio de origen del alumno. Una corrección de este tipo tendría el sano efecto de reducir los incentivos a la inflación de notas.

Se debe destacar que el SIMCE no es una manera de reflejar el nivel socioeconómico de los alumnos. La correlación entre el nivel socioeconómico del alumno, de acuerdo a una clasificación realizada por el Servicio de Bienestar Estudiantil de la Escuela de Ingeniería, y el SIMCE del colegio de origen es de sólo 0,18 (aunque es significativa). Esto indica que incorpo-

⁴ Loreto Arenas, "Estudio de los Factores Explicativos del Rendimiento Académico de los Alumnos de la Escuela de Ingeniería y Ciencias de la Universidad de Chile" (2003).

⁵ Se utilizó la prueba SIMCE de 1998.

rar el SIMCE o una variable equivalente para corregir las notas de los colegios no tendría un efecto discriminador desde el punto de vista socioeconómico.

2. Los datos

Los datos representan a todos los alumnos que ingresaron a la Escuela de Ingeniería durante el período 1994-2002, es decir, más de cinco mil alumnos. De estos alumnos se dispone información sobre: i) puntaje de ingreso y sus componentes; ii) el código de cada curso tomado, su nota y el número de créditos del curso⁶; iii) la carrera en que está inscrito el alumno; iv) la situación actual del alumno (activo o inactivo), v) el año de ingreso; vi) el puntaje SIMCE del colegio del que proviene, y otras variables que no han sido utilizadas hasta ahora, tales como la procedencia, el grupo socioeconómico y el colegio donde estudió. Los puntajes asociados a NEM están reescalados a la escala que se utiliza para calcular el puntaje de ingreso.

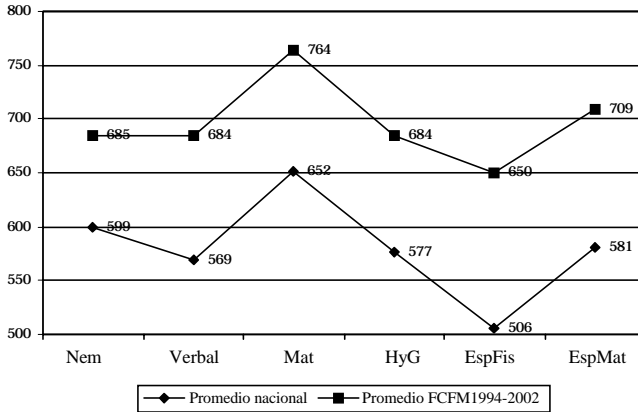
Se han eliminado los alumnos actualmente en programas de magister de la muestra, aunque esto no excluye a los alumnos que están simultáneamente siguiendo la carrera y el magister. Así, aquellos buenos alumnos que son capaces de tomar cursos de postgrado mientras siguen sus carreras no han sido excluidos. Este mecanismo tiene la ventaja de eliminar de los datos a alumnos que están en los programas de magister y ya están graduados, posiblemente de otras universidades, y que por lo tanto introducirían ruido en la muestra. Por otro lado, esta regla excluye a alumnos que han completado licenciaturas (pero no poseen título profesional) y que están tomando magísteres científicos. Su exclusión sesga nuestros resultados en la medida que se trate de buenos alumnos; sin embargo, su número debería ser pequeño.

Como se observa en la Figura 1, el perfil de los alumnos que ingresan a la Escuela de Ingeniería es bastante distinto al perfil de la población que tomó las específicas de física y matemáticas (10.011 individuos en 2001), que de por sí representa un grupo por encima del promedio nacional⁷. Es interesante observar, además, que la varianza de los puntajes de los

⁶ En aquellos casos en que el alumno repite o elimina el curso, no se le pone nota en la Escuela de Ingeniería, sino que simplemente se marca la eliminación (E) o repetición (R). Por lo tanto, hemos debido hacer análisis de sensibilidad respecto a distintos supuestos para las notas que el alumno se hubiera sacado en aquellos cursos que repitió o eliminó.

⁷ En particular, este grupo de alumnos tiene un puntaje en la PAA de matemáticas que es más de una desviación estándar superior al promedio nacional entre los alumnos que tomaron las pruebas específicas relevantes.

FIGURA 1: PUNTAJES ALUMNOS FCFM VERSUS PROMEDIOS DE ALUMNOS QUE TOMARON PRUEBAS ESPECÍFICAS



alumnos ingresados es mucho menor que las varianzas tomando el universo de todos los alumnos que dieron las pruebas. Como se observa en el Cuadro 1, en el caso de PAM, la desviación estándar en la muestra de los alumnos es un 25% de la varianza total, el caso más extremo. Este es uno de los motivos para el poco poder explicativo de la PAM, a menos que se corrija por la restricción de rango. El otro motivo es que el puntaje de la PAM y de la PEM tienen una alta correlación⁸.

CUADRO 1: DISPERSIÓN PUNTAJES FCFM Y PUNTAJES PROMEDIOS NACIONALES

Prueba	s FCFM	s nacional ¹	Razón
NEM	74,73	104,95	0,71
PAV	56,28	106,51	0,53
PAM	33,64	136,09	0,25
PEM	49,31	118,95	0,42
PEF	65,06	98,60	0,66
HYG	64,76	110,55	0,59

1 : s nacional, sobre alumnos que tomaron todas las pruebas.

Fuente: Datos nacionales entregados por Carmen Le Foulon, del Centro de Estudios Públicos.

⁸ De acuerdo a datos entregados por Carmen Le Foulon (Centro de Estudios Públicos), la correlación de Pearson entre la PAM y la PEM entre los alumnos con más de 450 puntos es de 0,80.

El Cuadro 2 presenta la matriz de correlación entre los distintos componentes del sistema de admisión al sistema universitario en la Escuela de Ingeniería. Como era de esperar, la mayor parte de estas correlaciones son positivas y significativas, de modo que alumnos con buen desempeño en una prueba tienden a tener un buen desempeño en las otras. Las correlaciones más altas son entre la PEM y la PAM y entre la PEM y la PEF que alcanzan el 50%. Es interesante notar que las notas de enseñanza media están negativamente correlacionadas con la PEM y la PEF, y que no están correlacionadas con la PAM. Este último resultado puede ser explicado por la baja variación que presenta la PAM en estos datos.

CUADRO 2: CORRELACIONES ENTRE COMPONENTES SISTEMA DE SELECCIÓN

Prueba	NEM	PAV	PAM	PEM	PEF	HYG
NEM	1					
PAV	.122	1				
PAM	.047	.172	1			
PEM	-.044	.100	.494	1		
PEF	-.082	.176	.298	.512	1	
HYG	.126	.327	.150	.154	.203	1

Nota: Todas las correlaciones son significativas al 1%.

Durante el período, los puntajes de entrada han aumentado en forma significativa. Al mismo tiempo, ha aumentado la cantidad de alumnos que toma las pruebas de ingreso a la universidad (incluyendo cada vez a alumnos menos calificados), por lo que el aumento en los puntajes podría deberse a esta expansión. Sin embargo, el aumento de los puntajes también es relativo a otras facultades de ingeniería, lo que indica que la mejora en los puntajes no es un fenómeno estadístico. Otro punto a considerar es que en el año 1996 se redujo el ingreso en doscientos puestos (ver Cuadro 3). El aumento en los puntajes continuó en forma posterior, pese a que el número de alumnos se ha mantenido constante desde entonces, lo que no se explica por el cambio en el tamaño de la cohorte admitida⁹.

⁹ Desde mediados de los 90, la facultad ha enfatizado la calidad de los alumnos que entran y creó un departamento dedicado a la promoción de la facultad entre los alumnos de enseñanza media. Asimismo, en los últimos años de la muestra se redujeron los conflictos en la Universidad de Chile (paros y protestas estudiantiles), y se elevó la calidad académica de la Escuela de Ingeniería.

CUADRO 3: ENTRADA DE ALUMNOS Y PUNTAJE, 1994-2002

Año	# alumnos	Último puntaje	Puntaje prom.	Notas Año 1	Avance 1er año
1994	766	635	677	2,67	0,56
1995	783	639	675	2,66	0,55
1996	564	654	681	3,08	0,64
1997	585	670	698	3,65	0,75
1998	557	678	702	3,59	0,75
1999	563	663	708	4,32	0,86
2000	557	690	715	3,81	0,78
2001	541	696	723	3,83	0,78
2002	554	696	722	3,53	0,73

2.1. Indicadores de desempeño

Durante los tres primeros años en la Escuela de Ingeniería, los alumnos deben tomar cursos en un plan común que admite pocas elecciones de cursos. En particular, durante el primer año, los alumnos son asignados aleatoriamente a secciones, por lo que ni siquiera pueden elegir entre secciones con distinto profesor. En los años siguientes, los alumnos son ranqueados de acuerdo a su rendimiento académico, y los cursos se llenan de acuerdo a este ranking, es decir, los mejores alumnos de cada nivel pueden elegir su profesor. Los indicadores de desempeño utilizados son:

Promedio ponderado primer año o de la carrera. Definido como la suma de las notas en cursos aprobados multiplicadas por la cantidad de créditos o unidades docentes (UD) del curso, dividida por la cantidad de UD de todos los cursos que ha tomado el alumno. Esto es equivalente a suponer que el alumno se sacó un cero en los cursos que eliminó o repitió¹⁰.

Porcentaje de créditos reprobados o eliminados en la carrera. Esta variable es el número de UD reprobadas o eliminadas sobre UD totales.

Porcentaje de avance en el año i-ésimo. En el primer año los alumnos deberían aprobar 80 créditos, y en los años siguientes deberían aprobar cien créditos al año. Cumpliendo estos hitos, los alumnos se graduarían en el plazo esperado. Este indicador se define para cada alumno como el

¹⁰ Como alternativa para determinar la sensibilidad de los resultados, repetimos el proceso suponiendo que las notas en cursos eliminados o reprobados fue un 3,9. Los resultados son similares (salvo por la magnitud de los parámetros), lo que demuestra que el supuesto elegido no influye en las conclusiones.

porcentaje de logro de estas metas. Puede ser mayor que uno para alumnos que hayan hecho algunos de los programas de la Escuela de Ingeniería para alumnos de enseñanza media, o para alumnos que tomen y aprueben muchos créditos.

Probabilidad de cumplimiento de hitos. Ésta es una variable binaria que indica si el alumno ha cumplido el número de créditos definido en el plan de la facultad. Esta variable se define para cada año que el alumno está en la Escuela de Ingeniería.

El Cuadro 4 provee una descripción de los datos utilizados como indicadores de rendimiento. Se observa, por ejemplo, que el promedio de notas es de un 5,05 para los cursos aprobados, y que el máximo promedio es de 6,53¹¹. También se observa que la tasa de cumplimiento del programa aumenta desde 0,71 a 0,81 a medida que los alumnos avanzan en la carrera, lo que puede deberse a un fenómeno de selección: los alumnos que salen son los que tienen más problemas. El Cuadro 3 muestra que no existe una relación directa entre la calidad de una cohorte (medida por el puntaje de ingreso) y las notas o el grado de avance en la carrera en el primer año. Una última cifra interesante es el porcentaje de cursos repetidos o eliminados: aproximadamente un 29% de los cursos está en esa categoría.

La Figura 2 muestra la relación entre el puntaje de ingreso y el rendimiento, medido como el promedio de notas ponderados para la muestra de todos los alumnos¹². Se aprecia que existe una fuerte relación positiva entre ambas variables, lo que indica que a mayor puntaje los resultados de los alumnos son efectivamente mejores. Sin embargo, la relación no es precisa, y existe mucha variabilidad en los resultados¹³.

Para examinar en más detalle los resultados, usamos gráficos de tipo caja¹⁴. Las Figuras 3(a)-3(d) muestran en forma gráfica la importancia de la

¹¹ Los promedios que incluyen a todos los cursos dependen de la nota que se les asigne a los cursos repetidos o eliminados. Con un cero, se tiene un 3,55, con un 3,9 se tiene un 4,72.

¹² Los puntos en 0 son alumnos que se retiran sin aprobar ninguna UD.

¹³ Aunque la figura parece indicar la existencia de una relación no lineal entre puntaje de ingreso y rendimiento, lo que ocurre es que en los puntajes más bajos aumenta la varianza. Esto se ve más claramente en la Figura 3(e), en que la línea gruesa en cada intervalo corresponde a la mediana de los datos en el intervalo, y la dispersión de los datos se muestra en el tamaño de las barras. Por último, al introducir términos de orden mayor en las regresiones de la sección 3, éstas pierden significancia.

¹⁴ En las figuras que siguen, se ha dividido el rango de las abscisas en veinte categorías de igual ancho de los puntajes, desde el menor al mayor. Cada intervalo contiene un número diferente de observaciones. La caja muestra el rango que cubre desde el 25% inferior al 75% inferior de los datos. La línea horizontal oscura es la mediana de los datos en la celda. Las líneas que salen de la caja muestran los valores adyacentes. El valor adyacente superior es la mayor observación menor o igual al percentil 75 más 1,5 veces el rango

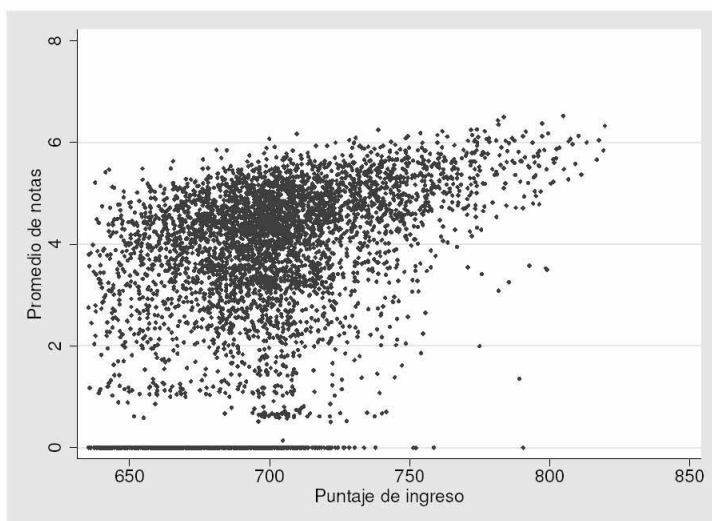
CUADRO 4: RENDIMIENTO DE LOS ALUMNOS

Variable	# obs.	Promedio	Desv. Est.	Mínimo	Máximo
Año de Ingreso a la Facultad	5062	-	-	1994	2002
Promedio, cursos aprobados	4222	5,05	0,37	4,00	6,53
Promedio ponderado ¹	4846	3,55	1,74	0,00	6,53
Promedio ponderado ²	4846	4,72	0,50	3,90	6,53
Fracción cursos R/E	4846	0,29	0,33	0,00	1,00
Fracción de avance 1er año	4846	0,71	0,38	0,00	1,98
Fracción de avance 2do año	3757	0,78	0,23	0,00	1,33
Fracción de avance 3er año	3041	0,80	0,19	0,16	1,21
Fracción de avance 4to año	2522	0,80	0,18	0,23	1,21
Fracción de avance 5to año	2020	0,81	0,19	0,19	1,29
Fracción de avance 6to año	1568	0,81	0,18	0,17	1,22
Cumpl. 1er año	5062	0,55	0,50	0,00	1,00
Cumpl. 2do año	3757	0,19	0,40	0,00	1,00
Cumpl. 3er año	3041	0,16	0,37	0,00	1,00
Cumpl. 4to año	2522	0,16	0,37	0,00	1,00
Cumpl. 5to año	2020	0,18	0,38	0,00	1,00
Cumpl. 6to año	1568	0,11	0,32	0,00	1,00

¹ Usando nota 0,0 en casos de repetición o eliminación de un curso.

² Idem, usando 3,9 en caso R/E.

FIGURA 2: PUNTAJE DE INGRESO Y PROMEDIO PONDERADO DE NOTAS



PEM y las NEM en el rendimiento de los alumnos. Aunque los datos tienen mucha dispersión, es notable lo bien que se comporta la media en cada intervalo, y cómo efectivamente se observa que a mayor puntaje NEM o PEM, el punto mediana de cada celda sigue la tendencia esperada: mayor puntaje implica menor tasa de repetición y mejores promedios de notas¹⁵.

Por último, las Figuras 3(e)-3(f) muestran el efecto del puntaje de ingreso sobre el rendimiento. Claramente, mayores puntajes están asociados a mejores rendimientos. Esto se observa también al usar como variable de rendimiento el porcentaje de avance para cada año de la carrera, como se observa en las Figuras 4(a)-4(f). En particular, en el caso de alumnos con altos puntajes de ingreso, virtualmente todos completan las 80 UD correspondientes al primer año, y los casos que no lo hacen son tan pocos que no es posible calcular la dispersión de los datos (Figura 4(a))¹⁶ el primer año. Es decir, la probabilidad de cumplimiento de hitos está positivamente relacionada con el puntaje de ingreso a la facultad. Además, es importante notar que el efecto del puntaje de entrada sobre el rendimiento persiste a lo largo de la carrera¹⁷.

Los resultados de las figuras pueden también ser resumidos en matrices de correlación entre los distintos indicadores de rendimiento y las pruebas del sistema de selección. El Cuadro 5 muestra la correlación simple entre estas variables y la correlación corregida de acuerdo al criterio de Pearson-Lawley¹⁸. No importa cómo se mida el desempeño de los alumnos, la correlación entre ellas y las pruebas de selección y las notas de enseñanza media son significativas (al 1%) y poseen el signo esperado. Las correla-

intercuantil (la diferencia entre el percentil 75 y el 25). El valor adyacente inferior es la menor observación mayor o igual al percentil 25 menos 1,5 veces el rango intercuantil. Los círculos representan puntos mayores o menores a los respectivos valores adyacentes y corresponden a *outliers* según la definición de Tukey (1977), *Exploratory Data Analysis*, Addison-Wesley.

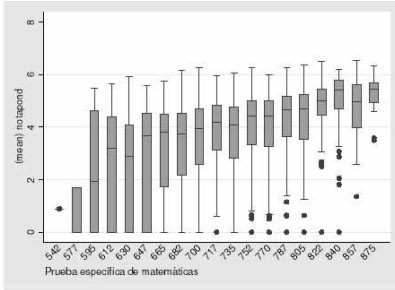
¹⁵ Al hacer figuras similares para la PAM, aunque el promedio de cada intervalo predice el rendimiento de los alumnos, la varianza es tan grande en relación al poder predictivo del promedio en cada intervalo de puntajes, que la PAM no posee capacidad explicativa en este grupo de alumnos.

¹⁶ En la Figura 4a, casi todos los alumnos de altos puntajes están en la mediana, la que está ubicada justamente en el valor 1, por lo que no se distingue fácilmente. Los demás puntos corresponden a los *outliers*, o sea, esporádicos casos extraños.

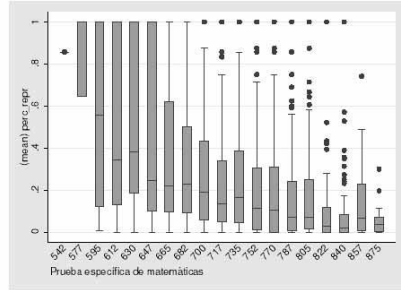
¹⁷ El sexto año los alumnos se encuentran realizando la memoria. Este proceso solía tener una duración muy variable (los buenos alumnos tomaban temas de memoria más complejos, por ejemplo), lo que tal vez explica la menor predictibilidad de los resultados para ese año.

¹⁸ Cuando una variable está restringida en su rango, como la PAM en la muestra de la Escuela de Ingeniería, su correlación con otras variables está atenuada. La corrección permite obtener una mejor medida de la correlación poblacional. Usamos los datos del universo de personas que dieron las PEM y PEF y obtuvieron al menos 450 en todas las pruebas para hacer la corrección.

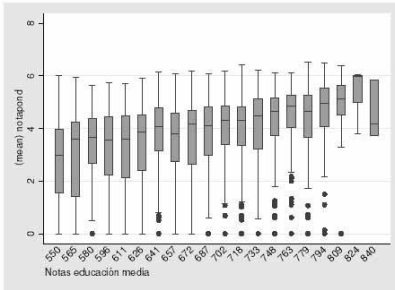
FIGURA 3: PREDECIBILIDAD DEL RENDIMIENTO DE LOS ALUMNOS



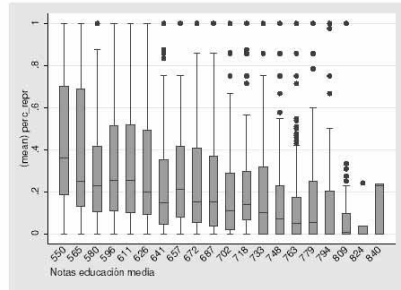
(a) Notas ponderadas y PEM



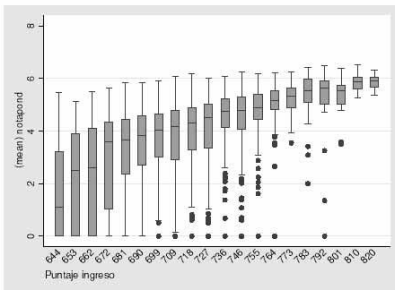
(b) Porcentaje R/E y PEM



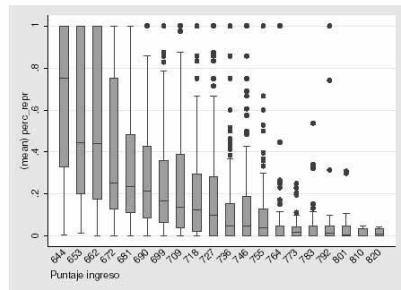
(c) Notas ponderadas y NEM



(d) Porcentaje R/E y NEM

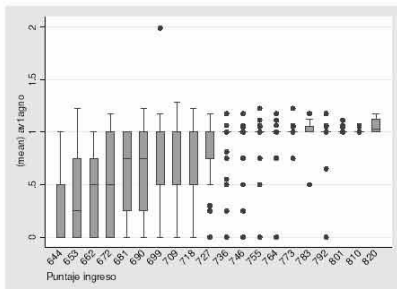


(e) Promedio ponderado y puntaje de ingreso

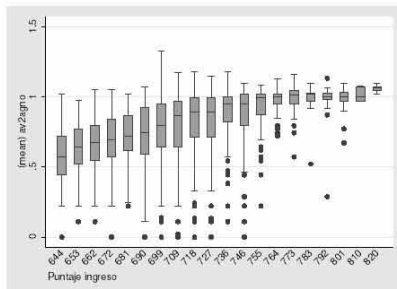


(f) Porcentaje R/E y puntaje de ingreso

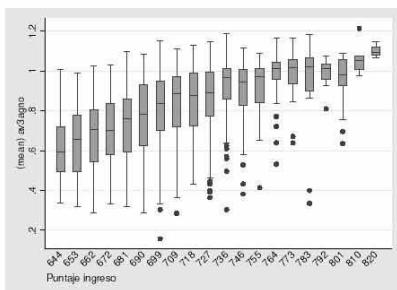
FIGURA 4: AVANCE EN LA CARRERA Y PUNTAJE DE ENTRADA



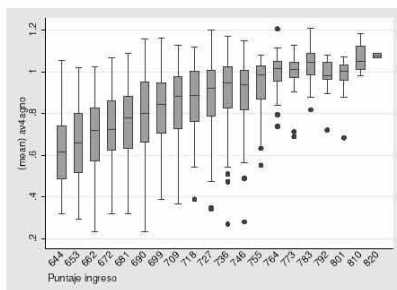
(a) Porcentaje avance 1er año y puntaje de ingreso



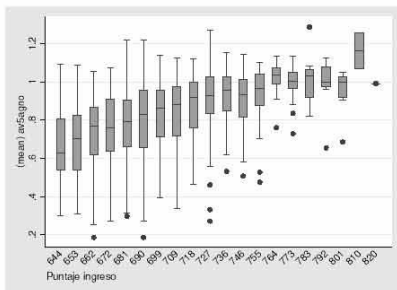
(b) Porcentaje avance 2o año y puntaje de ingreso



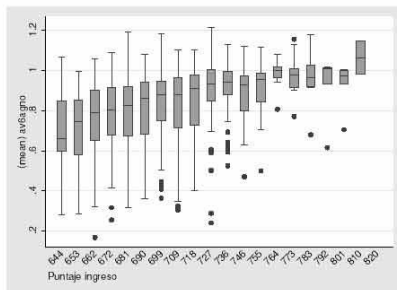
(c) Porcentaje avance 3er año y puntaje de ingreso



(d) Porcentaje avance 4o año y puntaje de ingreso



(e) Porcentaje avance 5o año y puntaje de ingreso



(f) Porcentaje avance 6o año y puntaje de ingreso

CUADRO 5: MATRIZ DE CORRELACIONES ENTRE VARIABLES DE SELECCIÓN Y RESULTADOS

	NEM	PAV	PAM	PEM	PEF	HYG	Ingreso
Prom. Pond. 1er año (R/ E = 0,0)	0.2357	0.1380	0.2591	0.3822	0.4307	0.1642	0.4970
	0.2589	0.1713	0.5179	0.5087	0.4530	0.2169	0.7341
Prom. Pond. (R/ E = 0,0)	0.2731	0.1204	0.1942	0.2940	0.3362	0.1713	0.4369
	0.2993	0.1497	0.4079	0.4023	0.3553	0.2261	0.6758
Fracción cursos R/ E	-0.2400	-0.1003	-0.1840	-0.2658	-0.3072	-0.1607	-0.3941
	-0.2635	-0.1248	-0.3892	-0.3665	-0.3251	-0.2124	-0.6292
Avance 1er año	0.2034	0.1119	0.2282	0.3274	0.3752	0.1472	0.4296
	0.2237	0.1392	0.4676	0.4437	0.3958	0.1948	0.6682
Avance 2o año	0.2352	0.0964	0.1063	0.2342	0.2643	0.1207	0.3526
	0.2583	0.1200	0.2346	0.3254	0.2801	0.1602	0.5796
Avance 3er año	0.3319	0.1196	0.1849	0.2477	0.2406	0.1931	0.4386
	0.3624	0.1487	0.3909	0.3431	0.2552	0.2541	0.6776
Avance 4o año	0.3280	0.1123	0.1485	0.2151	0.2407	0.1918	0.4203
	0.3582	0.1397	0.3210	0.3002	0.2553	0.2524	0.6582
Avance 5o año	0.3188	0.1074	0.1143	0.1907	0.1995	0.1870	0.3837
	0.3484	0.1336	0.2513	0.2674	0.2119	0.2462	0.6171
Avance 6o año	0.2884	0.0809	0.1043	0.1690	0.1733	0.1627	0.3343
	0.3158	0.1008	0.2303	0.2379	0.1842	0.2149	0.5564

Nota: La primera fila de cada entrada presenta el coeficiente de correlación simple; la segunda, la correlación corregida por restricción de rango.

ciones no corregidas por restricción de rango entre las medidas del rendimiento de los alumnos y las variables PEM, PEF y NEM, son más altas que la correlación con la prueba de aptitud académica en matemáticas (PAM). Sin embargo, una vez corregidas por restricción en rango, las correlaciones entre el rendimiento y la PAM se vuelven igualmente importantes, si no más, que las correlaciones con las otras pruebas así como con las NEM. Si no existiera la prueba específica de física, y si no hubiera restricción de rango, la prueba de aptitud matemática sería un excelente predictor del rendimiento de los alumnos en Escuela de Ingeniería.

Es interesante notar que en la medida en que los alumnos avanzan en la carrera, la correlación con las NEM tiende a subir, mientras que la correlación con las PAV, PAM, PEM y PEF tiende a bajar. Es decir, estas pruebas tendrían un poder predictivo decreciente a lo largo de la carrera, mientras que las NEM tendrían un poder predictivo creciente o constante. Esto podría indicar que la constancia en el esfuerzo, expresado en mejores notas en educación media, se reflejan en un avance más consistente a lo largo de la carrera.

3. Resultados de regresiones

Hemos realizado una serie de regresiones para explicar el rendimiento en base a los distintos componentes del puntaje de ingreso. Los principales regresores son NEM, PAM, PAV, PEF, PEM y HyG y el SIMCE promedio en matemáticas del colegio de origen en 1998¹⁹. Para controlar por la elevación en los puntajes de entrada a la Escuela de Ingeniería —el puntaje de ingreso promedio sube de 677 en 1994 a 722 en 2002— que podría indicar un cambio en el tipo de alumnos que ingresa a la facultad, hemos incluido una serie de variables *dummy* correspondiente al año de entrada de los alumnos²⁰. También hemos incorporado una serie de *dummies* correspondientes a la carrera elegida por el alumno, ya que existe la posibilidad que las distintas carreras tengan distintos grados de dificultad²¹. Es probable que la existencia de muchas variables mudas sea lo que explique los elevados R^2 de las regresiones (véase Cuadro 6). Por último, hemos incluido el SIMCE de matemáticas de segundo medio de 1998 del

¹⁹ Se usa el SIMCE de matemáticas debido a que la alta correlación de este promedio con el de castellano implicaba colinealidad de los regresores.

²⁰ Por ejemplo, el puntaje mínimo de ingreso sube de 635 a 696 en el período.

²¹ Hemos usado las *dummies* por carreras pese a que los alumnos están en plan común durante el primer año. Si hay programas más difíciles que otros, existe un efecto de autoselección en la elección de carreras.

CUADRO 6: EFECTO DE 100 PUNTOS MÁS EN CADA PRUEBA SOBRE EL RENDIMIENTO

	NEM	PAV	PAM	PEM	PEF	HYG	SIMCE	R ²	R ² P-L							
Promedio ponderado 1er año (0.0 cursos R/E)	0.598	0.029	0.048	0.039	0.082	0.079	0.646	0.050	0.610	0.038	0.010	0.035	0.269	0.090	0.29	0.71
Promedio ponderado (0.0 cursos R/E)	0.566	0.023	-0.017	0.030	0.009	0.055	0.293	0.039	0.293	0.029	0.088	0.027	0.431	0.066	0.38	0.65
Fración de Cursos R/E carrera	-0.079	0.004	0.009	0.005	0.010	0.010	-0.035	0.007	-0.041	0.005	-0.010	0.005	-0.074	0.012	0.33	0.51
Avance en primer año	0.083	0.006	0.002	0.008	0.009	0.016	0.086	0.010	0.094	0.008	-0.007	0.007	0.055	0.019	0.24	0.55
Avance en segundo año	0.075	0.004	-0.011	0.006	0.020	0.011	0.046	0.007	0.056	0.006	0.012	0.005	0.051	0.012	0.50	0.59
Avance en tercer año	0.080	0.004	-0.009	0.006	0.012	0.011	0.050	0.007	0.040	0.006	0.016	0.005	0.078	0.013	0.37	0.62
Avance en cuarto año	0.087	0.005	-0.012	0.006	0.011	0.013	0.040	0.008	0.048	0.006	0.023	0.006	0.087	0.015	0.33	0.63
Avance en quinto año	0.090	0.006	-0.014	0.007	0.018	0.015	0.030	0.010	0.041	0.007	0.028	0.007	0.094	0.018	0.29	0.60
Avance en sexto año	0.079	0.006	-0.017	0.008	0.018	0.016	0.024	0.011	0.034	0.008	0.022	0.007	0.076	0.020	0.25	0.55
Cumplir meta primer año	0.201	0.013	0.026	0.016	0.025	0.030	0.207	0.023	0.162	0.016	0.006	0.014	0.117	0.034	0.15	
Cumplir meta primer año	0.128	0.010	-0.028	0.013	0.031	0.026	0.078	0.017	0.082	0.013	0.059	0.012	0.044	0.029	0.22	
Cumplir meta primer año	0.089	0.010	-0.001	0.013	0.035	0.026	0.064	0.017	0.059	0.013	0.039	0.012	0.051	0.028	0.21	
Cumplir meta primer año	0.092	0.011	-0.009	0.014	0.066	0.028	0.075	0.019	0.060	0.015	0.032	0.013	0.065	0.033	0.16	
Cumplir meta primer año	0.117	0.013	-0.027	0.017	0.074	0.034	0.057	0.023	0.073	0.017	0.030	0.015	0.147	0.042	0.14	
Cumplir meta primer año	0.070	0.012	-0.036	0.015	0.021	0.030	0.043	0.020	0.044	0.015	0.023	0.013	0.088	0.037	0.16	

Notas:

- (1) La primera columna para cada variable muestra el efecto estimado; la segunda su error estándar.
- (2) El cumplimiento de metas se estimó usando probit. El R² en estos casos es el pseudo-R².
- (3) R² P-L es el coeficiente de correlación múltiple corregido por restricción de rango de acuerdo al método de Pearson-Lawley.
- (4) No se han incluido los parámetros correspondientes a las *dummies* de año y carrera.

colegio de procedencia del alumno, para controlar por la calidad de la educación adquirida antes de ingresar a la Escuela de Ingeniería. Se ha utilizado la matriz de White para obtener errores robustos en los parámetros²².

Los resultados de las regresiones son interesantes, y se describen en el Cuadro 6²³. En primer lugar, 100 puntos más en las notas de educación media tienen un efecto significativo (al 5%) e importante en el rendimiento de los alumnos, no importando cómo se mida el rendimiento. El promedio de un alumno con 100 puntos más que el promedio de los alumnos subiría 0,57, su fracción de créditos reprobados bajaría de 29% a 21% (una caída de más de un 25%), y el porcentaje de las UD que cumpliría en el primer, segundo y tercer año subiría en aproximadamente un 10%. En particular, el porcentaje promedio de UD aprobadas el primer año subiría de 71% a 79%. Estos efectos son importantes.

El efecto combinado de la PEM y la PEF es en general más importante que el efecto de las NEM. Para cada medida de rendimiento, la PEM y la PEF son significativas al 5%. Un alumno con 100 puntos más que el promedio en la PEM y la PEF tendría una nota promedio ponderada 1,2 puntos mayor en el primer año y su promedio ponderado a lo largo de la carrera pasaría de 3,6 a 4,2. Su fracción de cursos repetidos o eliminados bajaría en casi lo mismo que con 100 puntos más en la NEM. La tasa de avance en la carrera el primer año sería de 89%, comparada con la tasa de 71% del promedio. A lo largo de la carrera, el efecto asociado a mayor puntaje en las pruebas específicas iría disminuyendo, pero incluso en su quinto año, el efecto combinado de cien puntos más en estas pruebas elevaría la fracción de cumplimiento del hito (480 UD) a 88%, versus el 81% del alumno promedio.

También se ha testeado el cumplimiento de hitos en la carrera, definido como la variable binaria que denota si el alumno ha cumplido con las UD que la Escuela de Ingeniería recomienda para cada año de la carrera. Como se observa en las últimas líneas del Cuadro 6, a pesar que los efectos se debilitan a lo largo de la carrera, los efectos importantes provienen de las NEM, PEF y PEM, con estas dos últimas teniendo un efecto mayor que el efecto de las NEM por sí solas. Incluso en el sexto año, los efectos son fuertes. El hecho que la PEM y la PEF sean casi dos veces más importantes que las notas de educación media el primer año puede explicarse por el

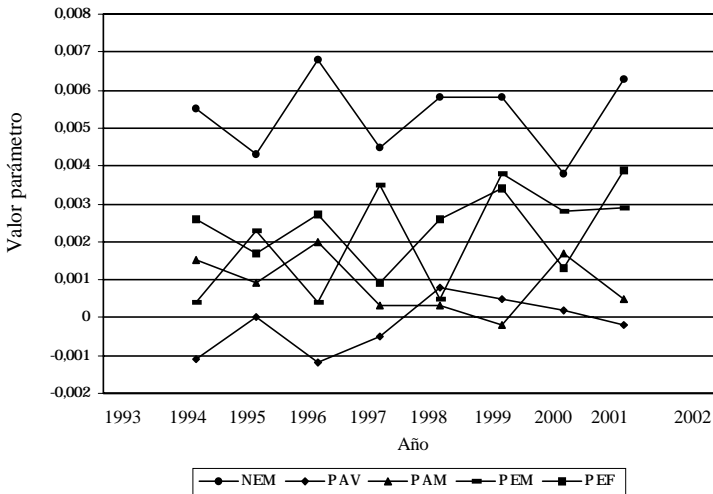
²² Dado que los alumnos tienen restricciones al número mínimo de cursos que deben aprobar para continuar en el programa, los datos están censurados. Al correr las regresiones utilizando un modelo tobit para los primeros años (cuando existe censura), los resultados no modifican nuestras conclusiones.

²³ Los resultados no dependen de si se usa un 0,0 o un 3,9 como nota en los cursos repetidos o eliminados. Se presentan sólo los resultados que usan un cero en este caso.

hecho que los conocimientos requeridos para un buen puntaje en la PEM y PEF son en cierta medida similares a los que se enseñan el primer año. Sin embargo, esto no explica por qué los efectos de la PEM y PEF se mantienen en el tiempo²⁴.

Los otros componentes del puntaje de ingreso son menos importantes, y aunque hay casos en que tienen efectos significativos sobre el rendimiento, éstos son erráticos y siempre menos importantes que la PEF, PEM y NEM. Por último, el poder predictivo del rendimiento de estas variables cae en la medida en que los alumnos avanzan en la carrera.

FIGURA 5: ESTABILIDAD DE LOS PARÁMETROS ESTIMADOS EN LA REGRESIÓN DE NOTAS PONDERADAS



3.1. Efecto de la calidad de los colegios de origen

Las regresiones incorporan una variable correspondiente al promedio en SIMCE de matemáticas de segundo año medio de 1998 del colegio de origen para examinar si la calidad de los colegios de origen tiene un impacto independiente sobre el rendimiento de los estudiantes²⁵. La desvia-

²⁴ Hemos estudiado la estabilidad de las regresiones, repitiéndolas para cada cohorte en forma individual. Los valores de los parámetros son relativamente estables, como lo muestra la Figura 5.

²⁵ La correlación entre el Simce de lenguaje y de matemáticas en la muestra es tan alto (entre 0.8 y 0.9) que se descartó el SIMCE de lenguaje. Se eligió el año 1998 pues es el SIMCE más cercano para todos los estudiantes, ya que incluso quienes entraron en 2002 directamente de 4° medio hicieron su 2° medio en 1999.

ción estándar en el SIMCE de escuelas a nivel nacional es de 35,6 puntos²⁶. Entre quienes han tomado las específicas de matemáticas y física y que obtuvieron más de 450 puntos en las pruebas de selección (nuestro universo) la desviación estándar de las distintas pruebas de selección a la universidad es de aproximadamente 80 puntos²⁷. Por lo tanto, un aumento de 100 puntos en las componentes de la PAA equivale, aproximadamente, a un aumento de 50 puntos en el SIMCE.

Un aumento de 50 puntos en el SIMCE se traduce en un aumento en el promedio ponderado de notas desde un 3,55 a un 3,77. El porcentaje de cursos reprobados o eliminados cae desde 29% a 25,3%, es decir, en más de un quince por ciento. Aunque el efecto sobre el grado de avance durante los primeros años es relativamente menor (un 30% del efecto de la prueba específica de matemáticas durante el primer año), es un efecto que aumenta en el tiempo. En el tercer año es casi equivalente al efecto de la prueba específica de matemáticas y para el quinto año el efecto es un 50% mayor que el de la misma prueba. Este resultado es sorprendente y revela que existe un efecto de la calidad del colegio sobre el rendimiento académico en la carrera que es independiente de las notas en la educación media y de lo que se puede medir en las pruebas de selección. Este resultado es consistente con la noción de que las notas en la educación media tienen distintos significados que dependen del establecimiento educacional: un buen estudiante en un colegio selectivo puede obtener notas peores que un alumno peor en un colegio menos exigente. Es interesante observar que existe una correlación negativa e importante (-0.28) entre el SIMCE y las notas de la enseñanza media, lo que parece indicar que los colegios de peor calidad —en términos del SIMCE— ponen mejores notas.

Una explicación alternativa es que existe una asociación entre el SIMCE del colegio y el nivel socioeconómico, por lo que el SIMCE del colegio de origen actuaría como un *proxy* de éste. Sin embargo, la correlación entre el nivel socioeconómico de los alumnos en la Escuela de Ingeniería y el SIMCE promedio del colegio de procedencia es de sólo 0.18 (aunque es un valor significativo)²⁸. Al incorporar la variable que mide el nivel socioeconómico de los alumnos en las regresiones, los resultados para las demás variables no cambian sustancialmente. En las regresiones de las notas ponderadas, se replica el efecto de Arenas (2003), el efecto del nivel

²⁶ Agradecemos este cálculo a Pilar Romaguera.

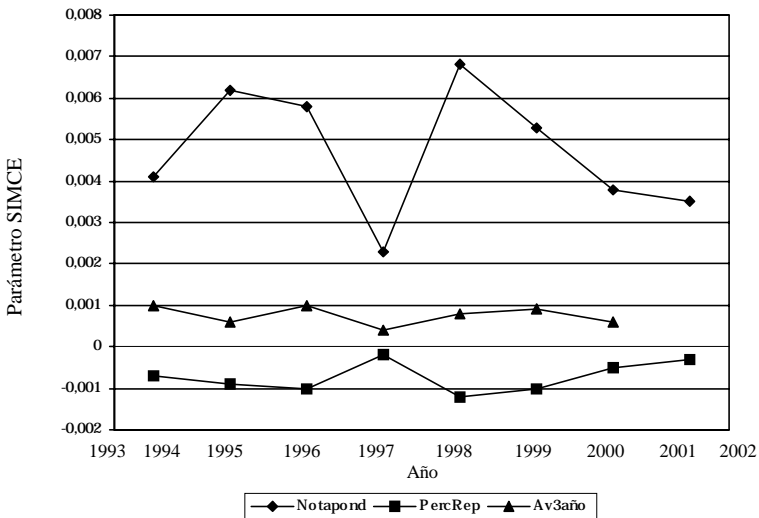
²⁷ Datos de Carmen Le Foulon.

²⁸ Al ingresar a la facultad, los alumnos se clasifican por nivel socioeconómico (NSE). El NSE resume información sobre la educación y ocupación del jefe de hogar, posesión de bienes del hogar y sector de residencia. Alumnos de clase media media y media alta representan el 62% de los alumnos.

socioeconómico sobre el promedio de notas es de -0.22 al pasar de clase alta a clase baja. La clase socioeconómica no tiene un efecto significativo sobre la tasa de repetición o eliminación de cursos. Tampoco tiene un efecto significativo sobre el grado de avance en la carrera. Este resultado revela que el nivel socioeconómico no representa un obstáculo académico en la facultad y que el filtro de entrada tiene un efecto igualador entre los alumnos en términos de esta variable.

Por lo tanto, es probable que la capacidad del SIMCE para predecir resultados se deba a que las notas de educación media no miden lo mismo en distintos colegios. Efectivamente, como recomendación de política, las notas de educación media deberían ser corregidas mediante un instrumento estándar, como el SIMCE. Al aumentar la importancia de las notas en los puntajes de admisión, no sólo se crean incentivos a la inflación de notas sino que se castiga a buenos alumnos de colegios con altas exigencias educacionales, tales como el Instituto Nacional²⁹.

FIGURA 6: ESTABILIDAD DEL PARÁMETRO ASOCIADO AL SIMCE EN DISTINTAS REGRESIONES, ESTIMADAS PARA CADA AÑO



²⁹ Los efectos del SIMCE en las regresiones son estables, en el sentido de mantenerse en regresiones por cohortes a través de las distintas regresiones, como se observa en la Figura 6.

3.2. Capacidad predictiva del modelo

El modelo lineal es capaz de explicar un porcentaje relativamente importante de la varianza de los rendimientos, pues el rango de los R^2 de nuestras regresiones está entre 0.25 y 0.50. Sin embargo ésta es una subestimación de la capacidad predictiva del modelo debido al problema de restricción de rango. Corrigiendo este indicador mediante el método de Pearson-Lawley, obtenemos los resultados que se indican en la última columna del Cuadro 6. De acuerdo a ésta, el mecanismo de selección utilizado por la facultad es capaz de explicar más del 50% de la varianza en los resultados y para el primer año, es capaz de explicar más del 70% de la varianza en los resultados³⁰.

4. Conclusiones

Los resultados muestran que el sistema de selección actual tiene un gran poder predictivo del rendimiento de los alumnos en la Escuela de Ingeniería. Tanto las NEM como las PEF y PEM explican parte considerable de la variación en el rendimiento de los alumnos. Al contrario de lo esperado, la PAM tiene poca importancia en los resultados, lo cual posiblemente se explica por el rango reducido de variación de los puntajes: tiene sólo un cuarto de la variación en la población que toma los exámenes necesarios para postular a la Escuela de Ingeniería. Alternativamente, la explicación es que mide la misma aptitud que la PEM, ya que están altamente correlacionadas.

Es importante notar que la calidad del colegio de origen también tiene un efecto significativo sobre el rendimiento, el que se mantiene a lo largo de toda la carrera. Además, la calidad del colegio no es un *proxy* para el nivel socioeconómico, el que no tiene una gran importancia sobre el rendimiento de los alumnos. Ambas observaciones permiten concluir que las notas de educación media —sin algún tipo de corrección— parecen contener un elemento distorsionador, ya que las notas de distintos colegios no miden lo mismo.

Nuestros resultados muestran que los efectos predictivos de las variables de selección son estables a través de los años para distintas cohortes y son relevantes incluso en los años finales de la carrera. En resumen, el

³⁰ La corrección requiere de información sobre la correlación entre los distintos regresores en la población. Como se desconoce la correlación entre el SIMCE y los componentes de la prueba de selección universitaria, supusimos que esta correlación es nula. Creemos que correlación es positiva, de modo que nuestra corrección subestimaría el R^2 .

sistema de selección a la universidad vigente hasta el año 2002 entregaba buenos resultados y, aparentemente, mide algún factor de aptitud para el estudio de las carreras en la Escuela de Ingeniería. El cambio desde el sistema tradicional al nuevo PSU, al tener una sola prueba de conocimiento en ciencias, puede reducir la capacidad de discriminación del sistema de selección de la Escuela de Ingeniería de la Universidad de Chile.

REFERENCIAS

- Aravena. R., G. del Pino y E. San Martín. "Capacidad Predictiva de la PAA". *Perspectivas*, (6) 2, 2003.
- Arenas, Loreto. "Estudio de los Factores Explicativos del Rendimiento Académico de los Alumnos de la Escuela de Ingeniería y Ciencias de la Universidad de Chile". Memoria de Título, Depto. Ingeniería Industrial, junio 2003.
- Tukey, John W. *Exploratory Data Analysis*. Addison-Wesley, 1977.
- Vial, B., y R. Soto. "¿Predice la PAA el Rendimiento o Éxito en la Universidad?". *Administración y Economía*. Universidad Católica de Chile, 48, 2002. □