



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY



UNIVERSIDAD DE LA REPÚBLICA
FACULTAD DE CIENCIAS SOCIALES

Tesis de Maestría en Políticas Públicas

Estudio panel sobre los factores asociados a las trayectorias educativas en Educación Media

¿Es posible construir un sistema de alertas tempranas en Uruguay?

Autor: Marcos Nahuel Alvez Legelén

Tutores: Dr. Santiago Cardozo Politi

Dra. Alina Inés Machado Fretes

Montevideo, Uruguay
Mayo 2024

Resumen

Esta tesis tiene por objetivo estudiar la configuración de las trayectorias educativas en Educación Media, haciendo énfasis en los diversos factores que inciden en ellas. Al mismo tiempo, para identificar de forma temprana a los estudiantes en riesgo de desvinculación, se elabora un modelo predictivo y se evalúa su potencial contribución a un Sistema de Alertas Tempranas.

Se utiliza un enfoque de trayectorias mediante el seguimiento de una muestra de estudiantes de 6° grado de Educación Primaria. El panel de datos incluye información sobre matriculación, progresión y abandono para cada estudiante, durante un período de observación de ocho años (2013 a 2021). Además de información sobre variables sociodemográficas, institucionales y vinculadas a la trayectoria escolar hasta el egreso de primaria.

Los resultados sugieren que la desvinculación de la Educación Media se asocia de forma significativa con: el sexo masculino, el bajo nivel socioeconómico, los centros educativos públicos, el rezago escolar, los desempeños académicos descendidos en matemática y lectura, y las bajas expectativas familiares sobre el nivel educativo que alcanzará el estudiante. Con base en información sobre estos factores al momento del egreso de educación primaria, es posible identificar, con altos niveles de sensibilidad y precisión, a los estudiantes que inician la Educación Media con altos riesgos de desvinculación. Estos hallazgos generan evidencia empírica que aportan al desarrollo del componente predictivo de un sistema de alertas tempranas en Uruguay.

Palabras clave: Trayectorias educativas, Desvinculación educativa, Políticas públicas, Modelo predictivo, Sistema de alertas tempranas

Clasificación JEL: C53, I21, I24, I28, J24

Abstract

This thesis aims to study the configuration of educational pathways in secondary school, focusing on the factors that impact them. Simultaneously, a predictive model is developed to identify students at risk of dropout at an early stage, and its potential contribution to an Early Warning System is evaluated.

A trajectory approach is used by tracking a sample of sixth-grade primary school students. The data panel includes information on enrollment, progression, and dropout for each student over an eight-year observation period (2013 to 2021). Plus, a set of sociodemographic, institutional and early school trajectories variables up to primary school graduation.

The results suggest that disengagement from secondary education is significantly associated with several factors. These include male gender, low socioeconomic level, public educational centers, school delay, declining academic performance in math and reading, and low family expectations regarding the student's educational attainment level. Based on information about these factors at the time of primary school graduation, it is possible to identify, with high levels of sensitivity and precision, the students who begin secondary school with high risks of dropping out. These findings generate empirical evidence that contribute to the development of the predictive component of an early warning system in Uruguay.

Keywords:

Educational pathways, Dropout, Public policies, Predictive model, Early Warning Systems

Clasificación JEL: C53, I21, I24, I28, J24

Índice de Contenido

Introducción	7
1. Marco Teórico	9
1.1. Capital humano y formación de habilidades.....	9
1.2. Elecciones educativas y enfoque de trayectorias	10
1.3. Sistemas de alertas tempranas	12
2. Antecedentes	13
2.1. Antecedentes Internacionales.....	13
2.2. Antecedentes Nacionales.....	16
3. Diseño y metodología.....	20
3.1 Objetivos del estudio.....	20
3.2 Preguntas que guían el estudio	22
3.3 Hipótesis	23
3.4 Estrategia de investigación y metodología.....	24
3.5 Metodología para la predicción de trayectorias de riesgo	28
Logit	28
Valoración del rendimiento del modelo.....	29
4. Caracterización de las trayectorias educativas en Educación Media.....	32
4.1 Situación del panel	32
4.2 Permanencia.....	33
4.3 Progresión en tiempo	36
4.4 Egreso	40
5. Desvinculación y tipología de trayectoria al final del período de observación.....	43
5.1 Desvinculación.....	43
5.2 Configuración de las trayectorias al final del período de observación	50
6. Determinantes de las trayectorias educativas	52
7. Modelo predictivo sobre la desvinculación como un instrumento que contribuye a un sistema de alertas tempranas.....	57
8. Síntesis y conclusiones	65
Referencias bibliográficas.....	68
Anexo.....	74

Índice de Gráficos.

Gráfico 1. Porcentaje de estudiantes según la situación registrada en marzo de cada uno de los 8 años que configuran el panel de datos.	33
Gráfico 2. Porcentaje de estudiantes matriculados sobre el total de la cohorte en de cada uno de los 8 años del panel de datos, según variables de interés.....	34
<i>Gráfico 3. Ratios entre el porcentaje de permanencia de las categorías extremas de las variables de interés.</i>	<i>36</i>
Gráfico 4. Porcentaje de inscriptos en el grado normativo sobre matriculados en cada t y sobre el total de la cohorte de 6° de Educación Primaria.	37
Gráfico 5. Porcentaje de estudiantes inscriptos en el grado normativo sobre el total de la cohorte de 6° de EP, según variables de interés.....	38
Gráfico 6. <i>Ratios entre el porcentaje de progresión en tiempo de las categorías extremas de las variables de interés.....</i>	<i>39</i>
Gráfico 7. Porcentaje de egresados en tiempo y con un año de rezago para el total de la cohorte de 6° de educación primaria, según variables de estudio.....	41
Gráfico 8. Porcentaje de estudiantes no inscriptos sobre el total de la cohorte de 6° de EP	44
Gráfico 9. Porcentaje de estudiantes no inscriptos sobre el total de matriculados del año anterior.	44
Gráfico 10. Diferencias en puntos porcentuales entre un año y el año anterior.	44
Gráfico 11. Porcentaje de desvinculados sobre el total de la cohorte.....	46
Gráfico 12. Porcentaje de estudiantes desvinculados (no inscriptos tres años consecutivos) sobre el total de la cohorte, análisis de supervivencia	47
Gráfico 13. Porcentaje de estudiantes de acuerdo a las diez combinaciones más frecuentes de matriculación/no matriculación para cada año del panel.....	49
Gráfico 14. Porcentaje de estudiantes según tipología de trayectoria al final del período de observación, para cada variable de estudio.....	50

Índice de Tablas.

Tabla 1. Matriz de confusión.....	30
Tabla 2. Efecto de cada variable sobre la probabilidad de realizar una trayectoria de rezago o de desvinculación, en comparación con una trayectoria de egreso, a los 8 años de haber cursado 6° de EP.	54
<i>Tabla 3. Regresión Logística sobre la probabilidad de desvinculación al cabo de 8 años en la EM para la Cohorte de 6° de EP en 2013. Logit 2.</i>	<i>59</i>
Tabla 4. Métricas del rendimiento del modelo Logit2 (L2), según el umbral de probabilidad elegido para clasificar a los desvinculados de los no desvinculados.....	61
Tabla 5. Rendimiento del modelo (L2) para un umbral de 0,180 y de 0,302 en la cohorte 2013 de estudiantes de 6° de Educación Primaria.....	62
 Tabla A. 1. Porcentaje de estudiantes para los que no se cuenta con información sobre su matriculación, fallecieron o migraron del país. Según año del Panel.....	74
Tabla A. 2. Medidas de resumen de las variables de interés.	75
Tabla A. 3. Correlación entre las variables de interés	76
Tabla A. 4. Porcentaje de inscriptos en EM respecto al total de la cohorte de 6° de primaria según variables de interés para cada año del panel.....	78

Tabla A. 5. Diferencias entre el porcentaje de inscriptos en EM según variables de interés para cada año del panel.	79
Tabla A. 6. Porcentaje de inscriptos en el grado normativo en EM respecto al total de la cohorte de 6° de primaria, según variables de interés para cada año del panel.	80
Tabla A. 7. Diferencias entre el porcentaje de inscriptos en el grado normativo según variables de interés para cada año del panel entre $t+1$ $t+6$	81
Tabla A. 8. Porcentaje de estudiantes según tipología de trayectoria al final del período de observación, para cada variable de estudio.	82
Tabla A. 9. Modelo logístico multinomial sobre la trayectoria al cabo de 8 años en la EM para la Cohorte de 6° de EP en 2013.	83
Tabla A. 10. Modelo logístico multinomial sobre la trayectoria al cabo de 8 años en la EM para la Cohorte de 6° de EP en 2013.	84
Tabla A. 11. Regresión Logística sobre la probabilidad de desvinculación al cabo de 8 años en la EM para la Cohorte de 6° de primaria en 2013. Logit 1.	85
Tabla A. 12. Métricas del rendimiento del modelo Logit 1 (L1), según el umbral de probabilidad elegido para clasificar a los desvinculados de los no desvinculados.	87

Índice de ilustraciones

Ilustración A. 1. Representación de la Función Logística	74
Ilustración A. 2. Curvas de ROC del Modelo Logit 2.	86
Ilustración A. 3. Curvas de ROC del Modelo Logit 1.	86

Estudio panel sobre los factores asociados a las trayectorias educativas en Educación Media

¿Es posible construir un sistema de alertas tempranas en Uruguay?

Introducción

En Uruguay, la Educación Media enfrenta grandes desafíos, sobre todo teniendo en cuenta que es obligatoria y considerada un derecho humano fundamental. Desde hace décadas, las tasas de egreso son bajas, tanto las de Educación Media Básica (EMB) como de Educación Media Superior (EMS). A pesar de los esfuerzos por mejorar las tasas de culminación de cada uno de los ciclos, aún persisten brechas significativas en términos de equidad y acceso a la educación.

Este fenómeno no solo afecta las trayectorias educativas de los estudiantes, sino también su inserción laboral y su participación en la sociedad en general. Según datos del observatorio educativo de la ANEP¹, en 2023 solo 4 de cada 10 personas de entre 18 y 20 años egresaron de la EM, mientras que para el grupo de edad de 21 a 23, 5 de cada 10 lograron egresar. Otro desafío para Uruguay es la desvinculación del sistema educativo formal, 2 de cada 10 estudiantes se desafilian de la EM (Fernández, 2009; Cardozo, 2016). Este problema es aún más grave cuando la desvinculación acontece de forma temprana. Según el Observatorio de ANEP, en el tramo etario de 15 a 17 años, hay un 10% de adolescentes que no asisten a ningún centro educativo.

Las trayectorias educativas en la enseñanza media no se limitan únicamente al progreso académico de los estudiantes, sino que también reflejan las complejas interacciones entre factores socioeconómicos, culturales e institucionales que influyen en el proceso de formación de los jóvenes. La educación formal es uno de los canales principales para que las personas desarrollen habilidades cognitivas y no cognitivas y acumulen capital humano. La importancia de dichas habilidades para la vida social y económica ha quedado demostrada en una amplia bibliografía. La inteligencia básica, las habilidades adquiridas, el autocontrol y la perseverancia, son importantes para el éxito en la vida y afectan los resultados del mercado laboral, la salud y la participación en la vida social, al mismo tiempo que disminuyen la tasa de embarazos adolescentes y la delincuencia (Heckman, Stixrud y Urzua, 2006; Heckman y La Fontaine, 2007; Heckman et al., 2010).

Por tanto, es crucial comprender por qué algunos jóvenes completan la Educación Media mientras que otros se desvinculan. Estudiar las trayectorias educativas no solo permite identificar las barreras y desafíos que enfrentan los estudiantes, sino también diseñar estrategias y políticas educativas más inclusivas y efectivas. En esta línea es fundamental establecer un Sistema de Alertas Tempranas (SAT) que permita identificar las trayectorias que presentan riesgos potenciales y brindar el apoyo necesario de manera oportuna.

¹ Los datos de egreso publicados por el Observatorio educativo de la ANEP se calculan a partir de la Encuesta Continua de Hogares del Instituto Nacional de Estadística.

Para que el SAT sea efectivo, los indicadores deben cumplir con dos características principales. Deben ser capaces de identificar adecuadamente las situaciones de riesgo, asegurando que ningún niño que necesita apoyo pase por debajo del radar, y deben ser precisos, para minimizar las "falsas alarmas". Para lograr esto, los *inputs* que informan a un SAT deben captar de manera efectiva las señales que sugieren la necesidad de intervención y evitar el uso ineficiente de recursos limitados (Leah P. Macfadyen, 2009; Bruce et al, 2011; Cardozo et al, 2023).

En este contexto, el presente estudio tiene como objetivo principal analizar y comprender las trayectorias educativas en la enseñanza media en Uruguay. Específicamente, se propone explorar los factores asociados a la permanencia en el sistema educativo, a la progresión por los grados, al egreso y, en especial, a la desvinculación. A través de un enfoque de trayectorias, se buscará identificar patrones, tendencias y desafíos que puedan orientar la formulación de políticas públicas más equitativas y oportunas. Por último, se propone evaluar la capacidad predictiva de un Sistema de Alertas Tempranas en la Educación Media a partir de datos de la Educación Primaria (EP).

Los resultados sugieren que la desvinculación de la Educación Media se asocia de forma significativa con: el sexo masculino, el bajo nivel socioeconómico, los centros educativos públicos, el rezago escolar, los desempeños académicos descendidos en matemática y lectura, y las bajas expectativas familiares sobre el nivel educativo que alcanzará el estudiante. Con base en información sobre estos factores al momento del egreso de educación primaria, es posible identificar, con altos niveles de sensibilidad y precisión, a los estudiantes que inician su transición a la EM con altos riesgos de desvinculación.

Esta investigación genera evidencia empírica para el desarrollo de sistemas de alerta temprana en Uruguay y aporta a la definición de políticas que aborden las necesidades de los estudiantes con trayectorias incompletas.

La tesis se organiza en ocho capítulos. En el [primero](#), se realizará una revisión de la literatura sobre las trayectorias educativas en la Educación Media, enmarcando el estudio dentro del contexto teórico existente. El [segundo capítulo](#) aborda los antecedentes nacionales e internacionales. El [tercero](#) detalla la fuente de datos, metodología y técnicas utilizadas. El [cuarto capítulo](#) presenta los resultados obtenidos en la caracterización de las trayectorias. El [quinto](#) analiza la desvinculación, mientras que el [sexto](#) examina los determinantes de las trayectorias mediante un modelo explicativo multinomial. El [séptimo capítulo](#) evalúa la posibilidad de desarrollar el componente predictivo de un SAT para las trayectorias educativas de la Educación Media en Uruguay. Finalmente, en el [octavo capítulo](#) se presenta la síntesis y conclusión del estudio.

1. Marco Teórico

1.1. Capital humano y formación de habilidades

Uno de los paradigmas dominantes en economía es el modelo del capital humano desarrollado por Becker (1964). Este enfoque concibe a la educación como una inversión que realiza el individuo para mejorar su inserción en el mercado de trabajo y obtener mejores ingresos que al no educarse. En tanto, el capital humano (KH) es visto como el stock de conocimiento o características de un individuo (innatas o adquiridas) que contribuyen a su productividad. Más KH conlleva a que el trabajador sea más productivo en todas las tareas, aunque la magnitud variará de acuerdo al tipo de tarea, de organización y de situaciones laborales. A pesar de desempeñar un papel complejo en el proceso productivo, el KH el stock de conocimientos o habilidades de una persona, comprende atributos como la calidad y años de educación recibida, la formación o entrenamiento del ámbito laboral, las actitudes en el trabajo entre otras competencias. Estas características se reconocen como una inversión en competencias valiosas que influye tanto en los incentivos a la inversión como en la estructura de los salarios e ingresos (Acemoglu y Autor, Notes Labour Economics, LSE 2009).

¿Hasta qué punto una persona se ve motivada a educarse? El modelo simple de dos periodos de inversión en educación es un marco teórico que explica cómo deciden los individuos a la hora de invertir en educación. Las personas dividen su vida en dos periodos: la juventud y la edad adulta. Durante el primero, los individuos pueden dedicar una parte de su tiempo a la escolarización, contribuyendo a su KH. La recompensa se manifestará en el mercado laboral en el periodo de la edad adulta. Las decisiones de escolarización se toman en función de los beneficios futuros, si se prevé un aumento de la demanda de mano de obra calificada, los individuos perciben un mayor rendimiento de la inversión en educación y por tanto se educarán. También se tiene en cuenta la influencia de la habilidad, las personas más calificadas demandarán más educación y alcanzarán un mayor KH debido al mayor rendimiento por el tiempo que dedican a estudiar. Aunque en principio todos los individuos deberían exigir la misma cantidad de educación, hay dos factores que contribuyen a diferenciar el nivel educativo de la población: el talento ("inteligencia pura" o un "mejor entorno familiar"), que impulsa la acumulación de KH de un individuo: y la riqueza familiar, que puede limitar a las familias pobres cuando los mercados financieros son imperfectos o inexistentes (Checchi, 2005).

Para este autor, el rol de la educación es crucial a la hora de cultivar habilidades esenciales para el funcionamiento eficaz de la sociedad, es un bien público compartido que no puede excluirse, contribuye al bienestar social en términos de salud, calidad de vida y satisfacción personal. Siguiendo a Checchi (2005) de acuerdo a una perspectiva desde el lado de la economía y la psicología, habilidad es aquello que contribuye al ingreso potencial de un individuo, es diversa, cambia a lo largo de la vida y varía en función de la formación y el entrenamiento.

En el trabajo de Heckman, Jagelka y Kautz (2019) se distingue entre habilidades cognitivas: inteligencia y conocimiento, medidos a través de tests de IQ y de pruebas de aprendizaje, y habilidades no cognitivas o habilidades blandas: hacen referencia a características de la personalidad como la motivación, la perseverancia, el autocontrol, la sociabilidad y la autonomía. Las habilidades cognitivas y no cognitivas cambian con la edad y la educación. La eficacia de las intervenciones para mejorar las habilidades depende de la etapa de vida en la que se apliquen, en la niñez es más eficiente invertir en las habilidades cognitivas, en cambio las habilidades no cognitivas son más maleables en la adolescencia. Los niños que demuestran una mayor motivación académica y una mayor apertura a la experiencia tienden a lograr un mayor aprendizaje, inclusive las capacidades no cognitivas compiten con los test IQ por el potencial que tienen en predecir el nivel educativo, el éxito en el mercado laboral, la salud y la delincuencia

La formación de habilidades está en el centro de estudio de la economía del desarrollo humano y la movilidad social. Según estas ramas de la economía, hay que tener en cuenta tres aspectos en la formación de habilidades: i) múltiples períodos en el ciclo de vida de los individuos (la niñez, la adolescencia y la adultez) y la presencia de periodos críticos, ii) múltiples conjuntos de capacidades para padres e hijos, y iii) múltiples formas de inversión. Un entorno enriquecedor en la primera infancia debe ir seguido de una buena escolarización y un buen entorno de aprendizaje en el trabajo, el aprendizaje y la motivación se presentan como procesos dinámicos y acumulativos. Invertir en los niños desfavorecidos es una política pública en la que no existe un trade-off entre equidad y eficiencia, reduce la desigualdad de origen y al mismo tiempo logra aumentar la productividad de la sociedad. Es destacable el rasgo de complementariedad de las inversiones en las distintas etapas de vida como una característica esencial del proceso de desarrollo de las capacidades humanas (Heckman y Mosso, 2014; Heckman y Masterov, 2007).

En cada generación, la culminación de los niveles educativos obligatorios² es una medida aproximada de la acumulación del KH que la sociedad acordó como esencial. La desvinculación de estos niveles esenciales, supone un desafío y un riesgo importante, tanto para los individuos que no completan sus estudios como para la sociedad en su conjunto.

1.2. Elecciones educativas y enfoque de trayectorias

Al considerar una trayectoria educativa como el camino que recorre una persona a través del sistema educativo, se tiene que para cada año se toman decisiones. Estas elecciones simplificadamente constan de resolver si se continuará estudiando, dónde estudiar y qué oferta educativa elegir. Para tomar cada uno de los caminos posibles que se habilitan cada año, los estudiantes cuentan con cierta información, cada uno sabe cómo le ha ido en los años anteriores y además se tiene cierta idea de lo que implica seguir o dejar de estudiar. El cuerpo teórico que se comenta a continuación refiere a los distintos tipos de costos y beneficios que se enfrentan los estudiantes y sus familias.

² En Uruguay, la Ley General de Educación, N° 18.437, establece que la educación inicial a partir de los cuatro años de edad, la Educación Primaria y la Educación Media son obligatorias.

La información que poseen los individuos incide al momento de tomar decisiones respecto a los eventos que van configurando las trayectorias académicas. Cada año que se completa, aporta nueva información y posibilita la oportunidad de continuar estudiando, lo cual estará vinculado a los costos operativos de estudiar, a los costos de oportunidad y a los ingresos esperados de cada año adicional (Heckman, Lochner y Todd, 2008). Además de considerar la incidencia de la incertidumbre en la toma de decisiones, también influyen, los costos psicológicos, el bajo interés por el aprendizaje, la incertidumbre, y las diferencias de habilidades (Heckman, Stixrud y Urzua 2006).

Esta tesis se basa en la idea general del marco conceptual anteriormente descrito, considerando que las condiciones sociales y económicas de los estudiantes, así como su desempeño académico en las etapas tempranas de la escolarización, actúan como factores de riesgo o incentivo para la continuidad educativa. Aunque no se abordan directamente las elecciones educativas, se analizan los resultados inmediatos de estas decisiones. Estos resultados se reflejan en la decisión de los estudiantes y sus familias sobre la inscripción al sistema educativo durante el período de observación que abarca el estudio, permitiendo así estudiar la permanencia o desvinculación de los estudiantes.

El avance por los grados educativos es un proceso secuencial que requiere la culminación de etapas educativas previas. Así, los logros educativos, funcionan como uno de los condicionantes de las opciones educativas que se habilitan y en gran medida determinará el avance por las etapas siguientes. A su vez, el logro de los estudiantes está condicionado por varios factores, algunos de ellos tienen mayor influencia en las primeras etapas de la escolarización, mientras que otros inciden preponderantemente en etapas posteriores (Méndez y Ramos, 2019).

Particularmente, las capacidades cognitivas se ven más afectadas a edades tempranas y las no cognitivas a edades más avanzadas. Esto significa que no es trivial la etapa educativa en la que se invierte sobre una u otra habilidad. Las diferentes etapas del ciclo vital del niño son periodos sensibles para invertir en ellos y posteriormente conllevan a resultados diferentes en la edad adulta. Dado el carácter acumulativo de las trayectorias escolares, la inversión y los apoyos en las primeras etapas de la escolarización son más eficientes y sobre todo más eficaces que las políticas remediales (Cunha y Heckman, 2008).

Un análisis en clave de trayectorias tiene la potencialidad de observar para una misma cohorte de estudiantes los eventos que van pautando los recorridos por el sistema educativo a lo largo del tiempo. Esto es de gran utilidad para entender los momentos y situaciones claves que configuran las trayectorias. En la literatura sobre las trayectorias educativas en Uruguay hay un consenso sobre dos fenómenos, el primero es que los resultados educativos son acumulativos (Cardozo et al, 2010; De Melo y Machado, 2018; ANEP, 2017). El segundo, es que existen etapas clave al inicio de la trayectoria educativa que moldearán al estudiante y determinarán posibles riesgos en las siguientes etapas de la trayectoria (Cardozo, 2016; Vásquez y Moreira, 2016).

Las trayectorias ocurren en un marco institucional y, por tanto, el tablero dónde se configuran las trayectorias es determinado por las reglas de juego y los objetivos que determina el sistema educativo. Además de transmitir conocimientos, las instituciones educativas ayudan al estudiante a identificar sus ventajas comparativas y absolutas, los educadores tienen el rol de identificar las distintas velocidades y formas en las que aprenden los estudiantes. La diferenciación de habilidades, puede generar mayores desigualdades de ingresos, es por ello, que la discusión de la organización de un sistema educativo, de los métodos de promoción y evaluación, tiene implicancias sobre la distribución del ingreso. Una parte de la discusión sobre los sistemas educativos radica en el diseño y en las respectivas transformaciones. El diseño implica resolver si las instituciones deben impartir la misma intensidad de conocimiento o no, cuál será la financiación, el grado de descentralización y el sector institucional de los centros, públicos, privados o mixtos (Stiglitz, 1975).

Las trayectorias escolares son las huellas que van dejando los estudiantes por su paso dentro del sistema educativo. Hay una concepción de que las trayectorias teóricas definidas en la currícula son de carácter lineal, tienen ciertos ritmos de avance y son anualizadas. El éxito de una trayectoria es aquella que no presenta eventos de repetición ni abandono, una trayectoria continua y completa que demuestre haber tenido los aprendizajes según lo establecido en la currícula. Las trayectorias teóricas se relacionan estrechamente con marcos conceptuales pedagógicos y didácticos, según las edades. Hay cierto orden temporal en el que deben acontecer los aprendizajes, hay un ritmo deseable para ello y también formatos escolares en el que se lleva adelante el proceso de aprendizaje. Los problemas surgen cuando existe un desfase entre las trayectorias reales y las diseñadas normativamente. No se puede pasar por alto los efectos de esas trayectorias desfasadas, son expectativas que no se cumplen e implica que hay formas de enseñar que no están previstas para la heterogeneidad en términos de ritmos de aprendizajes dentro del aula (Terigi, 2014).

La planificación e implementación de políticas educativas se desarrolla sobre una idea acerca de la configuración de las trayectorias. Existe la necesidad de incorporar a las políticas lo que demuestran algunos estudios sobre el bajo porcentaje de estudiantes que siguen una trayectoria normativa. Las trayectorias que se distancian de las normativas, se han mostrado como la pauta y no la excepción en la EM. Tienen diferentes recorridos, velocidades y grados de culminación (Cardozo, 2015; De Melo y Machado, 2018; Noboa 2020; ANEP, 2020).

1.3. Sistemas de alertas tempranas

Mediante el análisis de datos sobre diversos indicadores, como el rendimiento en los cursos, la asistencia y el compromiso, las instituciones pueden identificar a los estudiantes con dificultades o en riesgo, y realizar intervenciones oportunas y adecuadas. Este enfoque proactivo integra la recopilación de datos con métodos estadísticos y modelos predictivos, permitiendo tomar decisiones informadas tanto para la práctica educativa como para la política. Los resultados han

demostrado mejorar las tasas de permanencia y graduación de los estudiantes (Leah P. Macfadyen, 2009; U.S. Department of Education 2016).

Siguiendo a Bruce et al. (2011) los Sistemas de Alertas Tempranas (SAT) son iniciativas coordinadas a nivel escolar y/o regional, destinadas a ofrecer sistemas de datos capaces de identificar a los estudiantes con trayectorias educativas en riesgo, y que incluyen una intervención sistemática y personalizada. También implica el uso de indicadores con ciertos parámetros para identificar a los estudiantes que pueden estar en riesgo de no alcanzar resultados educativos esenciales. El objetivo es ofrecer intervenciones y apoyos oportunos para solventar las carencias educativas inmediatas y de largo plazo. Inclusive, algunos sistemas monitorean las intervenciones que se llevan adelante. Los SAT han evolucionado incluyendo indicadores de preparación escolar al comienzo de la trayectoria de un estudiante, e indicadores de preparación universitaria y profesional durante la trayectoria. El uso de los SAT ha identificado la asistencia, el comportamiento y el rendimiento en el curso como fuertes predictores de la finalización de la EM.

Con el auge de la analítica y los macro datos, la analítica del aprendizaje ha ganado capacidad para aportar información sustancial sobre el comportamiento del alumno. Los datos de alta calidad son la base de este tipo de análisis, y los entornos virtuales o los sistemas de gestión del aprendizaje pueden producir gran cantidad y calidad de datos necesarios (Azcona y Casey, 2017).

Según la literatura internacional se considera que hay tres grupos de factores que determinan los resultados educativos: los atributos del centro educativo, el contexto socioeconómico y cultural en el que está inmerso el niño o adolescente y, por último, las características propias de los alumnos (Carrasco, 2007). Estas tres dimensiones son las que se desarrollarán en la tesis.

2. Antecedentes

2.1. Antecedentes Internacionales

Un primer antecedente es el Estudio de Igualdad de las Oportunidades Educativas realizado por Coleman et al. (1966). La investigación fue pionera en los estudios sobre determinantes de los resultados educativos. Se evaluó la igualdad de oportunidades en materia educativa teniendo en cuenta características de los alumnos, como la raza, la religión y el lugar de nacimiento, así como diversos factores escolares. Los resultados mostraron una baja influencia del gasto educativo por estudiante en el rendimiento. El informe sugirió que los factores familiares, como el entorno hogareño y el nivel educativo de los padres, tenían un impacto significativo en el rendimiento estudiantil, en algunos casos superior al factor de recursos escolares. Los trabajos que continuaron esa línea de investigación fueron los de Weber (1971), Klitgaard y Hall (1974) y, más recientemente, la tradición de estudios sobre eficacia escolar, con base en la modelización multinivel. En ellos se concluye que hay dos factores relevantes para explicar las diferencias en los aprendizajes y la desvinculación, el primero es la organización de los centros educativos y el segundo, la calidad de

recursos materiales y humanos destinados (Fernández y Blanco 2004; BRYK y Lee 1992; Lee y Burkam 2003).

Las condiciones materiales a las que están expuestos los alumnos a la hora de estudiar y de desarrollar su vida en el hogar, así como los recursos educativos con los que cuentan, el puesto de trabajo de los padres y el clima educativo del hogar son factores que afectan en un sentido u otro los rendimientos (Brunner y Elacqua, 2003; Fernández y Blanco, 2004 y Blanco, 2008). Otros estudios apuntan al efecto organizacional de la escuela (Sheerens y Bosker, 1997, citado en Murillo, 2008) concuerdan en que se puede explicar entre un 12% y 18% la varianza de los rendimientos escolares mediante las características de los centros educativos.

Investigaciones que trabajan sobre el impacto de la desigualdad socioeconómica de un país en los resultados educativos, muestran que, a mayor desigualdad en la distribución del ingreso, menor es el rendimiento de los estudiantes (Hanushek y Wößmann, 2007; Fernández y Blanco, 2004). Asimismo, muestran que se genera un círculo vicioso en el que se ensanchan las brechas de desigualdad. Por otro lado, hay investigaciones que arriban a la idea de que la escuela estaría ayudando a disminuir la desigualdad en los aprendizajes que se vinculan a las diferencias socioeconómicas de los estudiantes (Mortimore, 1993 y Lee, Bryk y Smith 1993; citados en Carrasco, 2007; y UNESCO, 2008 citado en Da Rocha, 2011). De esta manera, los centros educativos lograrían revertir parte de las desigualdades de origen.

En Hanushek y Rivkin (2012) se analizan las implicancias de la distribución de la calidad del cuerpo docente en los centros. Destacan las considerables variaciones en la eficacia docente dentro de los centros educativos y las implicancias políticas que eso conlleva. Los principales resultados muestran que es crucial considerar cómo las familias (en su elección de centro) y los centros (en sus decisiones de asignación de aulas) afectan el rendimiento de los estudiantes.

En línea con investigaciones sobre recursos educativos, Das et al. (2013) exploran la conexión entre los recursos escolares y los resultados de los exámenes, centrándose en cómo reaccionan los hogares a los cambios en la financiación de las escuelas. Se emplea un modelo dinámico de optimización de los hogares para hacer predicciones y luego ponerlas a prueba en Zambia y en India. Las conclusiones indican que los hogares tienden a ajustar más su comportamiento en respuesta a las subvenciones previstas que a las inesperadas. Al mismo tiempo, se observa que las subvenciones inesperadas dan lugar a mejoras significativas en los resultados de los exámenes. Por último, las estimaciones del efecto del gasto público en educación sobre logros educativos, que no tienen en cuenta la re-optimización de las familias, podrían arrojar resultados distorsionados al estimar los parámetros de una función de producción educativa.

En Balfanz, Curran, Herzog (2007) se aborda la crisis de la graduación en la enseñanza secundaria en Estados Unidos, centrándose en las bajas tasas de graduación de estudiantes de bajos ingresos y pertenecientes a minorías. El estudio supervisó a toda la cohorte de alumnos de Filadelfia que comenzaron 6° en 1996, se recopilaron datos sobre las calificaciones, los resultados de los exámenes, las marcas de comportamiento, la asistencia, el estado de educación especial, el estado de aprendizaje de inglés y las categorías demográficas. La investigación reveló que durante los años de enseñanza media surgían varias señales de advertencia que predecían fuertemente la

probabilidad de abandono de secundaria. Estas señales incluían una calificación final de suspenso en matemáticas o inglés, una asistencia inferior al 80% durante el curso y una nota final de comportamiento "insatisfactorio" en al menos una clase. Los alumnos que mostraban más de una de estas señales tenían una probabilidad aún mayor de abandonar prematuramente. Sumado a esto, la mayoría de los alumnos que abandonaban los estudios mostraron signos tempranos de angustia.

El trabajo de Bruce et al. (2011) se centra en los SAT y su potencial para abordar la crisis de abandono escolar en la enseñanza secundaria y universitaria. Explora la necesidad de los SAT, el diseño, el potencial y los componentes básicos para que sean eficientes. Incluye estudios de casos en U.S.A como los de Chicago, Knoxville, Philadelphia, Diplomas Now y Louisiana. Luego del análisis, concluye que existen tres señales clave de alerta temprana, el "ABC" por sus siglas en inglés; *Attendance, Behavior, and Course performance*, cada uno con un umbral de referencia; i) Faltar 20 días o ausentarse el 10% de los días lectivos, ii) Dos o más casos de transgresiones de conducta moderadas o graves; iii) Incapacidad para leer al nivel del grado al final del tercer año, reprobado en inglés o matemáticas y no obtener la promoción a tiempo, entre otras.

A nivel terciario, algunas universidades han establecido programas para detectar a los estudiantes en situación de riesgo. *Course Signals* es un sistema de la Universidad de Purdue U.S.A que informa los profesores sobre la probabilidad de rendimiento de los estudiantes. El sistema utiliza datos sobre: el rendimiento, el esfuerzo, el historial académico y atributos del estudiante. El programa ha demostrado eficacia para mejorar la permanencia de los estudiantes. (Azcona y Casey, 2017)

En Bruch et al. (2020) se explora el uso de los datos de las escuelas y los organismos de bienestar infantil para predecir los riesgos académicos a corto plazo. Se examinan las relaciones entre factores predictivos, como la asistencia, el comportamiento y la participación en el bienestar infantil, y los resultados académicos. Se propusieron, desarrollar modelos predictivos para identificar a los estudiantes que corren riesgo académico. Se utilizaron tres algoritmos para el modelo predictivo: bosque aleatorio, regresión logística de red elástica y red neuronal recurrente. El estudio concluyó que las combinaciones de variables escolares y extraescolares puede predecir los problemas académicos con una precisión de moderada a alta. Las variables intraescolares, como los datos escolares, resultaron ser los predictores más potentes en todos los resultados. El rendimiento predictivo de los modelos no se redujo significativamente cuando se excluyeron las variables extraescolares, como la implicación en el bienestar infantil, los servicios de emergencia para personas sin hogar y la implicación en el sistema de justicia juvenil. Sin embargo, se observó que algunos acontecimientos y servicios extraescolares estaban relacionados individualmente con problemas académicos a corto plazo.

En síntesis, se presentaron investigaciones sobre dos líneas de trabajo, las que buscan explicar los factores que inciden en las trayectorias educativas y las que centran su análisis en el desarrollo de los SAT.

La primera línea de trabajo, agrupa investigaciones que identifican efectos de al menos uno de los tres factores siguientes: los centros educativos, la familia y los atributos del propio estudiante. Los referidos al papel que juegan los centros educativos sobre la trayectoria de los estudiantes, en general, coinciden en que es importante la calidad del cuerpo docente, la organización y los

recursos materiales con los que cuenta el centro. Los que tratan sobre la incidencia de la familia, identifican el clima educativo del hogar, el nivel educativo de los padres, y las condiciones económicas del hogar. para explicar el rendimiento académico y la desvinculación de los estudiantes. Por último, se resumieron los estudios cuyos resultados concuerdan en que las características demográficas (sexo, raza, lugar de nacimiento) y el historial académico del estudiante, impactan en las trayectorias educativas.

La segunda línea de trabajo, agrupa investigaciones que ponen el foco en desarrollar sistemas de alertas tempranas. Las variables relevantes según estos estudios comprenden: asistencia a clases, comportamiento, calificaciones sobre el rendimiento académico, características demográficas, aprendizajes en lectura, matemática e inglés, promoción a tiempo e historial académico del estudiante.

2.2. Antecedentes Nacionales

Algunos trabajos se centran en determinar los impactos del centro educativo por el cual transita el alumno, la gran mayoría utiliza como fuente de datos la base de PISA; Méndez y Zerpa (2011), Da Rocha et al. (2011), Bogliaccini y Rodríguez (2015) y Lambí et al. (2009). En este último, se aborda el problema de la desigualdad de oportunidades relacionadas al sistema educativo. Sus resultados indican que el mayor nivel de desigualdad de oportunidades se encuentra en la Educación Media Básica (EMB).

Por otro lado, Bogliaccini y Rodríguez (2015) encuentran tres aspectos sobre el diseño del sistema educativo uruguayo de EM, que impactan negativamente en el desempeño académico de los estudiantes: la asignación de docentes, la distribución de alumnos de acuerdo a la cercanía con el centro educativo y la provisión de materiales educativos y tecnológicos.

Otro estudio relacionado a características propias del sistema educativo, Borba et al. (2017), aborda los efectos de la extensión del tiempo pedagógico en educación primaria sobre la trayectoria posterior. Se evalúa el impacto de las escuelas de tiempo completo (ETC) utilizando datos de TERCE³, una sub muestra de ETC y un seguimiento realizado por la DICE de la ANEP. No se identificaron diferencias significativas en las trayectorias de los estudiantes en la EMB, entre los que asistieron a ETC y los que no.

En Lado (2019) también se utiliza como material empírico la base de datos de TERCE y un seguimiento realizado por la DICE para estudiar los efectos de factores asociados a la escuela primaria sobre los desempeños académicos de los estudiantes en la EMB. Para ello, estima un modelo probit. Los resultados muestran que las características de la labor docente, como el monitoreo de la dirección y el ambiente laboral del centro de educación primaria, están vinculados con la probabilidad de aprobar el primer año de EMB.

Desde el lado de los efectos de la familia, se puede mencionar el trabajo de Mendez-Errico y Ramos (2019) que investigan si el ingreso parental de largo plazo (aproximado a través de la educación los

³ Tercer Estudio Regional Comparativo y Explicativo, consistió en un estudio de logro de aprendizaje a gran escala aplicado en 2013 que abarcó 15 países de Latinoamérica

padres, raza, habilidades cognitivas y no cognitivas) y el ingreso familiar de corto plazo (entendido como el costo de oportunidad de la educación), inciden en la progresión escolar de los hijos. Usan como fuente de datos la ENAJ⁴ y la Encuesta Continua de Hogares (ECH) del INE⁵. Mediante un modelo probit secuencial, concluyen que hay efectos variados según la etapa de progresión escolar y que la repetición tiene efectos negativos y duraderos en los siguientes años, truncando los logros educativos.

Adentrando en las investigaciones que se enfocan en explorar las trayectorias, Cardozo (2010) analiza los procesos que llevan a la interrupción o culminación definitiva con el sistema educativo formal durante la EM. Se presta atención a la forma en que se configuran las trayectorias, identifica que las trayectorias no tienen un camino lineal, sino que presentan períodos que se van alternando entre matriculación, inactividad, abandono y reinserción. Es así, que las trayectorias van decantando según una trama compleja que agrupa eventos educativos y no educativos (situación laboral, familiar, entre otras) y cuyos impactos se van dando a lo largo de los años. Como fuente de datos usa la Encuesta Retrospectiva sobre trayectorias académicas y capitalización de conocimientos en el panel de alumnos de Uruguay evaluados por PISA en el año 2003.

Un abordaje similar es presentado en Failache et al. (2015). A través de un panel de jóvenes en edad liceal, iniciado en el 2004 -cuando los estudiantes cursaban 1° en las escuelas públicas-, se estudia los determinantes de la repetición y del abandono. Se encuentra que: para los varones, los problemas de conducta están asociados con el abandono, mientras que, para las niñas, la hiperactividad y las emociones son las variables asociadas a dicho fenómeno. Asimismo, confirman que la repetición predice fuertemente el abandono.

En ANEP (2022) se encuentra un antecedente que contribuye a la evidencia empírica sobre las trayectorias de los estudiantes en la educación media pública. Los principales hallazgos muestran que luego de seis años del egreso de primaria, continúan inscriptos en el sistema educativo poco más de 7 de cada 10 estudiantes, pero solamente 3 de cada 10 logran progresar en el tiempo normativamente previsto. Respecto a la permanencia, parecería que entre el 4to y 6to año del egreso de primaria es donde hay un mayor porcentaje de estudiantes que abandonan los estudios. En tanto, la progresión en tiempo muestra el mayor descenso entre 1er y 2do grado de EM. Finalmente, en términos generales cuando se compara entre las sucesivas cohortes, el desempeño de los dos indicadores mejora.

El Panel SERCE fue el primer estudio longitudinal realizado en Uruguay para reconstruir aspectos clave de las trayectorias desde la educación primaria hasta alrededor de los 15 años de edad, tanto del sector público como del privado. En 2006, una cohorte de alumnos de 3° participó en el (SERCE/LLECE-UNESCO) y se les hizo un seguimiento. En 2009, se les realizó una segunda evaluación del aprendizaje, cuando en su mayoría cursaban 6°. Al mismo tiempo, se examinó la desvinculación durante la transición a la EMB y en los primeros años de esa etapa. Finalmente, en 2012, para indagar sobre la desvinculación, se encuestó a un subconjunto de jóvenes, de alrededor

⁴ Encuesta Nacional de Adolescencia y Juventud

⁵ Instituto Nacional de Estadística

de 14-15 años, que no se encontraban matriculados en el sistema educativo formal. (De Melo, G., Failache, E. y Machado, A. 2015; De Melo, Gioia y Machado, Alina 2016)

En un sentido similar, estudios de seguimiento de cohortes basados en (PISA-L)⁶ han permitido examinar la estrecha conexión entre los factores tradicionales de desigualdad relacionados con el origen socioeconómico y sociocultural, el avance de las competencias académicas medido por evaluaciones estandarizadas y las trayectorias educativas subsiguientes. La acumulación de retrasos académicos a los 15 años, junto con bajos aprendizajes en lectura y matemática, disminuyen significativamente la probabilidad de completar la EM. Conjuntamente, el progreso académico y el cumplimiento de los estándares de competencia, por sí solos no garantizan un resultado educativo positivo. (Boado y Fernández, 2010; Fernández, 2009; Fernández et al, 2013; Cardozo, 2016).

La siguiente investigación es uno de los antecedentes más relevantes de la presente tesis. Presenta factores que inciden en las trayectorias asociados al sistema educativo, al hogar y también al propio estudiante, en De Melo y Machado (2018) se generó un panel a partir de la base de datos de SERCE y de la Evaluación Nacional de Aprendizajes de ANEP, conformando una base de datos con información de tres momentos del tiempo: 2006, 2009 y 2012. Mediante un modelo logit anidado analizan los predictores de las decisiones de asistir o de abandonar la EM, diferenciando los que asisten a educación técnica o a liceos de secundaria. Sobre variables institucionales concluyen que, a mayor experiencia laboral del director y si realiza una jornada horaria completa en el centro, las probabilidades de abandono disminuyen. Por su parte, encuentran un impacto de las variables blandas en las trayectorias: expectativas que tienen los estudiantes sobre sus estudios, la motivación y el disfrute por ciencias y matemática.

El segundo antecedente estrechamente relacionado con esta tesis es el análisis del Panel TERCE que realizó la DICE- ANEP⁷. También muestra resultados asociados a características del sistema educativo, de la familia y del estudiante. La primera ola del panel consta del seguimiento durante tres años (2014 a 2016) de una muestra de alumnos de 6° de educación primaria (EP) evaluados por TERCE en 2013. Encuentran que: I) A mayores expectativas familiares sobre el nivel educativo que alcanzará el hijo/a, mayor probabilidad de egresar de EMB. II) Transitar ofertas de educación técnica podría ser visto como trayectos socialmente segmentados. III) Desigualdades en los resultados educativos según el sector institucional (público/privado). IV) La permanencia y la progresión en tiempo tienen una importante asociación con el origen socioeconómico de los estudiantes. V) Desigualdades moderadas, entre los estudiantes de Montevideo y del interior y en función del sexo. VI) El rezago en primaria es un fuerte predictor de la trayectoria en EMB. VII) Niveles de lectura satisfactorios en 6° de EP son una condición necesaria pero no suficiente para progresar en tiempo hasta 3° de EMB.

⁶ Estudios longitudinales de los jóvenes evaluados por PISA

⁷ ANEP (2020). Transición y trayectorias en la Educación Media Básica en Uruguay. Análisis Longitudinal a partir del Panel TERCE. Montevideo: ANEP.

Por último, se presentan tres antecedentes recientes y relevantes por la metodología y resultados a los que arriban. Se constata el potencial que tiene usar información sobre las habilidades cognitivas y no cognitivas de un nivel educativo para predecir lo que sucederá en el siguiente nivel. De igual modo, estos estudios remarcan la importancia de los sistemas de alertas temprana.

En Cardozo, Silveira y Fonseca (2022) se deja evidente el poder predictivo que tienen las habilidades captadas por la Evaluación Infantil Temprana (EIT) sobre la repetición en primaria. Mediante técnicas de machine learning -regresión logística, redes bayesianas y modelos de árbol de clasificación- logran clasificar de forma correcta al 80% de los estudiantes según hayan repetido o no, en los primeros tres años de educación primaria. Además, determinaron que entre el 62% y el 64% de los niños que el modelo identificó con riesgo de rezago, efectivamente se rezagaron. Los tres métodos empleados lograron casi el mismo nivel de precisión, sensibilidad y especificidad.

En Cardozo (2023) se indaga el potencial que tienen las dimensiones cognitivas y no cognitivas del desarrollo infantil (DI) para predecir resultados educativos en educación primaria. Los resultados mostraron que los niños que egresaron de educación inicial con habilidades descendidas, tuvieron bajos resultados en el trayecto por la educación primaria, tanto en pruebas de lectura o matemática, como en las calificaciones y la aprobación del grado. Además, la asociación entre rezago en las habilidades del DI y el desempeño en primaria se mantiene alta, inclusive al 5to año del egreso de educación inicial. Desde el lado de las habilidades no cognitivas encuentran altas correlaciones entre el DI y las notas en conducta.

El estudio llevado adelante por la FCS-Udelar y la DICE-ANEP (Cardozo et al, 2023) buscó desarrollar y validar modelos explicativos y predictivos de riesgo escolar en educación primaria, contribuir a la definición e implementación de un SAT, y principalmente analizar la repetición y los aprendizajes de una cohorte de estudiantes desde educación inicial a 6° de EP. Asimismo, se exploran y sistematizan los recursos, criterios y prácticas que los equipos docentes utilizan para identificar situaciones de riesgo educativo y desarrollar estrategias de apoyo. Se mencionan limitaciones que han impedido que los SAT sean efectivos y se integren de manera adecuada en las escuelas, a saber: tendencia a realizar actividades rutinarias, superposición y omisión de información, y ausencia de protocolos de monitoreo para movilizar recursos de manera oportuna y organizada. A pesar de ello, se remarca la capacidad de innovación en las escuelas y un fuerte compromiso con la atención de la diversidad académica.

En resumen, estos antecedentes muestran los factores asociados a la configuración de las trayectorias según tres aspectos. En primer lugar, el sistema educativo en cuanto a las características del centro, ya sea público o privado, el rol del director, la organización territorial y el conjunto de reglas normativas que determinan la modalidad de cursado, aprobación y otras características sobre la oferta educativa. En segundo lugar, factores asociados a la familia como las expectativas sobre la trayectoria de los niños, el contexto socioeconómico, las decisiones que se toman principalmente sobre qué estudiar y dónde hacerlo, la educación de los padres, entre otras. En tercer lugar, los factores asociados al estudiante como el sexo, edad, habilidades cognitivas y no cognitivas, ascendencia racial, motivación por el estudio, entre otras. El marco en el que se gestan las trayectorias se ve inmerso en esta triada: sistema educativo, familia y estudiante. Estos factores son interdependientes entre ellos y dinámicos en el tiempo.

Algo importante a destacar de los antecedentes, es la fuerte relación entre las etapas de las trayectorias: las primeras marcan fuertemente a las siguientes. Asimismo, el carácter acumulativo y jerárquico del aprendizaje, pauta el avance y la movilidad por los niveles educativos. En este sentido, en los antecedentes se mencionan algunos trabajos sobre la importancia de contar con SAT y la información relevante de la que se nutren, para poder identificar a los alumnos en riesgo educativo.

3. Diseño y metodología.

La tesis aborda las trayectorias educativas en toda la Educación Media, identificando y explicando los determinantes de tres desenlaces posibles: trayectorias de egreso, de rezago o de desvinculación. Se le prestará especial atención al desenlace de desvinculación

3.1 Objetivos del estudio

Objetivo general

Mediante el seguimiento de una cohorte de estudiantes que se encontraba cursando el último grado de EP en 2013, el estudio pretende observar los factores asociados a tres tipologías de trayectorias en EM. Las tres trayectorias quedan definidas según la situación del estudiante al octavo año de haber cursado 6° de EP (2021). Este período corresponde a los años de cursado normativo en la EM y dos años más, para poder estudiar el fenómeno de rezago. Las trayectorias mencionadas son: i) Trayectoria normativa, ii) Trayectoria de rezago, iii) Trayectoria de desvinculación.

Las dos primeras corresponden a trayectorias que finalizan con el egreso del estudiante o con el estudiante rezagado, pero aún con cierto grado de vinculación con el sistema educativo. Por el contrario, la tercera, implica el alejamiento del sistema educativo. Como objetivo general, no solo interesa ver el desenlace final sino, cómo se van configurando año a año las trayectorias. Al mismo tiempo, interesa explorar el potencial que tiene utilizar información proveniente de Educación Primaria para detectar las trayectorias de riesgo en la Educación Media. Para ello, en el último capítulo se elabora y evalúa la potencialidad de un modelo predictivo como insumo para los sistemas de alerta temprana.

Objetivos específicos

1. Caracterizar las trayectorias educativas en Educación Media (Capítulo 4).
2. Estudiar cómo se comporta la no inscripción y el perfil de los estudiantes que se desvinculan de EM (Capítulo 5).
3. Analizar los determinantes de los tres tipos de trayectorias educativas: egreso, rezago y desvinculación (Capítulo 6).
4. Realizar una predicción sobre la desvinculación en EM para aportar al diseño de un instrumento que contribuya a los sistemas de alerta temprana (Capítulo 7).

1. Caracterizar las trayectorias educativas en Educación Media

El objetivo consiste en caracterizar las trayectorias educativas de la cohorte de alumnos evaluados por TERCE en 2013 en 6° de EP a lo largo de toda la EM. La ventana de observación corresponde a ocho años para la mayoría de los indicadores presentados y para otros corresponde solamente a los seis años normativos de la EM. Los procesos y las características de los estudiantes que se buscan caracterizar a lo largo del período de observación son los siguientes:

- **Matriculación/ no matriculación.** Se observará la trayectoria educativa que se configura año a año entre 2013 y 2021, de acuerdo a si el alumno se encontraba matriculado o no, dentro del sistema educativo formal. El indicador asociado a este análisis es el de permanencia en cada año de estudio, se define como el porcentaje de la cohorte de estudiantes que estaba matriculado en el sistema educativo formal
- **Progresión en tiempo.** Se continuará describiendo para cada año, el porcentaje y las características de los alumnos que logran avanzar por los grados educativos según el calendario establecido.
- **Egreso.** Se describirán las trayectorias de los estudiantes que logran finalizar la EM en el tiempo previsto según los planes de estudio o un año más tarde. El análisis descriptivo del perfil de estudiantes que presentan trayectorias normativas, permitirá tener un panorama sobre las oportunidades de egreso en la EM.

2. Estudiar cómo se comporta la no inscripción y el perfil de los estudiantes que se desvinculan de Educación Media

Este estudio tiene particular interés en ahondar sobre la desvinculación, es decir, estudiar específicamente cómo se comportan los estudiantes que no se inscriben. Se pretende aportar insumos para responder ¿cuáles son las señales que permiten anticipar la desvinculación? y ¿cuáles son las características de los alumnos que se desvinculan? Concretamente, este segundo objetivo pretende: i) calcular la no inscripción del total de la cohorte para todos los años del panel, ii) calcular la no inscripción en cada año respecto a los matriculados del año anterior, iii) definir y calcular la desvinculación, iv) describir el perfil de los estudiantes cuya trayectoria tiene un desenlace de desvinculación al final del periodo de estudio, v) presentar las 10 trayectorias más frecuentes en EM, según la matriculación de los estudiantes a lo largo de los ocho años del panel.

3. Analizar los determinantes de los tres tipos de trayectorias educativas: egreso, rezago y desvinculación

Se realiza un análisis multivariado y longitudinal para la identificación y ponderación de los factores que inciden en la probabilidad de egresar, rezagarse o desvincularse de la EM. Una vez que se hayan descrito las trayectorias educativas, es válido preguntarse ¿por qué quedan definidas de esa forma? y ¿cuáles son los factores que inciden en su determinación? El análisis intentará responder estas preguntas en función de las características del sistema educativo, de la familia y del propio estudiante. Las variables que componen cada dimensión son las siguientes:

- **Características del sistema educativo.** Sector institucional (público o privado) y tipo de oferta elegida por el estudiante y su familia para comenzar la EM (educación secundaria o educación técnica). Región del centro en donde el estudiante cursó el último año de educación primaria, Montevideo o Interior del País.
- **Características de la familia.** Expectativas respecto al nivel educativo que creen que alcanzarán sus hijos y el índice de nivel socioeconómico (NSE) del hogar. Este indicador es elaborado por TERCE y se compone del nivel educativo y laboral de la madre, el ingreso del hogar, los bienes y servicios de la vivienda, y la cantidad de libros disponibles en el hogar.
- **Características del estudiante.** El sexo, las habilidades cognitivas tales como el rezago en educación primaria y los aprendizajes captados por el nivel de desempeño en las pruebas de matemática y lectura de TERCE.

4. Realizar una predicción sobre la desvinculación en EM para aportar al diseño de un instrumento que contribuya a los sistemas de alerta temprana

Se desarrolla, ajusta y valora un modelo predictivo sobre la desvinculación en EM, haciendo uso de la información previa a la transición de los alumnos al segundo nivel de enseñanza. Las variables utilizadas en el modelo son las descritas en el del objetivo específico tres, mencionado anteriormente.

3.2 Preguntas que guían el estudio

Pregunta N° 1: Trayectorias.

- 1.1 ¿Cómo se configuran las trayectorias educativas en la Educación Media?

Pregunta N° 2: Desvinculación.

- 2.1 ¿Cuándo comienzan a manifestarse los eventos que llevan a la desvinculación?
- 2.2 ¿Los que no se inscriben en EM, ¿vuelven al sistema educativo formal?
- 2.3 ¿Cuál es la magnitud y características de los estudiantes que se desvinculan?

Pregunta N° 3: Factores explicativos de las trayectorias.

- 3.1 ¿Cuáles son los factores más relevantes en la determinación de las trayectorias en Educación Media?
- 3.2 ¿Cuál es el grado de incidencia de los factores encontrados y en qué sentido operan?

Pregunta N° 4: Elaboración y valoración de un modelo para predecir la desvinculación.

- 4.1 ¿Qué tan bueno es el rendimiento de un modelo para predecir la desvinculación en EM, usando información del nivel educativo anterior?
- 4.2 ¿Uruguay cuenta con las condiciones para desarrollar el componente predictivo de un SAT?

3.3 Hipótesis

Hipótesis 1: Caminos heterogéneos en la configuración de las trayectorias (Objetivo 1 y 2)

Formulación: El mapa de las trayectorias educativas queda definido de forma muy heterogénea, configurando varios caminos posibles, plasmando un entramado muy diverso con distintos grados de culminación.

Especificación: Existe un conjunto de trayectorias que no tienen grandes cambios, entre ellas están las trayectorias de los estudiantes que egresan. Por otro lado, hay un conjunto de estudiantes que se van y no vuelven, su trayectoria queda trunca por una cantidad de años considerable. Además, se van delineando trayectorias variadas que se forman por la elección de los estudiantes a la hora de elegir matricularse o no matricularse. Asimismo, transitan a diferentes velocidades. De esta forma, el entramado de la trayectoria se vuelve diverso por las opciones que tienen los estudiantes y sus familias, año a año, a la hora de transitar por la EM. El sistema educativo tiene diseñado ciertos caminos normativos que implican recorridos de los estudiantes en una cantidad de años pre-establecida. Sin embargo, las trayectorias reales de los jóvenes se ajustan poco a lo esperado

Hipótesis 2: Problemas al inicio de la trayectoria educativa determinan la desvinculación en Educación Media (Objetivo 3)

Formulación: El problema de la desvinculación del sistema educativo comienza a gestarse en etapas tempranas de la trayectoria educativa.

Especificación: Una de las principales razones por la que ciertos estudiantes se desvinculan del último nivel educativo definido como obligatorio, se debe a los problemas que enfrentaron al inicio de su trayectoria en educación primaria. Uno de ellos es el rezago que se genera por la repetición de uno o más grados y ocurre principalmente en los primeros grados de EP. Sumado a esto, una vez que cursan la EM, la acumulación de eventos de riesgo tales como: la repetición y la decisión de no inscribirse algún año, incrementan la probabilidad de que el estudiante se desvincule.

Hipótesis 3: Los aprendizajes en educación primaria, el sector institucional y la oferta elegida para comenzar la EM impactan en la configuración de las trayectorias (Objetivo 3)

Formulación: Los resultados académicos obtenidos en la prueba de TERCE por los alumnos de 6° de EP, el sector institucional (público o privado) y la oferta educativa (secundaria o educación técnica) en la que se inscribe un estudiante al inicio de la EM, son dos componentes que marcan las trayectorias.

Especificación: Bajos aprendizajes en la educación primaria ocasionan una débil trayectoria e incrementan la probabilidad de desvinculación en EM. Asimismo, aunque en menor medida, la combinación del sector institucional y la oferta en la que se inicia la EM, también afecta la probabilidad de desvinculación. Estudiantes que se inscriben en el sector privado tienen menos chances de desvincularse.

Hipótesis 4: Expectativas de la familia (Objetivo 3)

Formulación: Las expectativas que tienen las familias sobre el nivel educativo que lograrán finalizar sus hijos, incide en la configuración de las trayectorias en EM.

Especificación: El seguimiento del Panel TERCE tiene la ventaja de que acumula una cierta cantidad de años que permite contraponer alguna de las expectativas que reportaron tener las familias en 6° de EP, con el grado educativo que efectivamente logró el estudiante, ocho años más tarde. Esta hipótesis supone que las expectativas son en su gran medida coincidentes con lo que se verifica en la realidad y que a mayores expectativas del nivel al que se espera que finalice el estudiante, mayor es la probabilidad de que el alumno, al menos, llegue a matricularse en dicho nivel.

Hipótesis 5: El sistema educativo no logra revertir la desigualdad de origen (Objetivo 3)

Formulación: Factores sociodemográficos inciden en la configuración de las trayectorias determinando oportunidades o vulnerabilidades a lo largo del ciclo educativo.

Especificación: Existe un conjunto de factores sociodemográficos que parten de las ideas clásicas sobre las desigualdades y más concretamente sobre las desigualdades educativas que inciden en la probabilidad de que un estudiante finalice la EM. Inclusive, se espera encontrar que estas características determinen situaciones de vulnerabilidad o, por el contrario, oportunidades favorables para realizar trayectorias continuas y completas. La hipótesis supone que, a mayor nivel socioeconómico y cultural de la familia, el estudiante tiene más chances de no desvincularse de la EM, lo mismo ocurre si el estudiante es mujer o si realiza sus estudios en el interior del País. (ANEP, 2021 análisis longitudinal Panel TERCE, De Melo y Machado, 2018; Cardozo et al, 2018).

Hipótesis 6: Identificación temprana de la desvinculación en EM (Objetivo 4)

Formulación: Con base en información de la trayectoria escolar hasta 6° EP es posible identificar con altos niveles de sensibilidad y precisión a los estudiantes que se desvincularán en la EM.

Especificación: A partir información del Panel TERCE y el seguimiento de la cohorte por parte de la DICE-ANEP, es posible construir un modelo predictivo sobre la desvinculación en EM que tenga buenos parámetros de rendimiento y sea eficaz para crear las alertas que permitan hacer intervenciones oportunas. Este ejercicio es la base para dar cuenta de que es necesario contar con información para todas las cohortes y por tanto se deberían universalizar ciertos instrumentos con los que ya cuenta el País. Uruguay cuenta con las condiciones para desarrollar el componente predictivo de un SAT.

3.4 Estrategia de investigación y metodología

Fundamentación del diseño

La mayoría de los análisis que abordan el egreso, el rezago o la desvinculación en la EM se sustentan en información agregada para un cierto momento del tiempo, es decir, usando datos de

tipo transversal. Esto presenta dificultades a la hora de estudiar los fenómenos de las trayectorias porque entre otros aspectos, las trayectorias son dinámicas y se ven afectadas por situaciones cambiantes en el tiempo. En Uruguay pocos estudios han utilizado datos estadísticos de tipo longitudinal⁸. Este tipo de análisis permite tener lo que le sucedió a un mismo estudiante a lo largo del tiempo. Es de especial interés la vinculación entre las características educativas de la cohorte 2013 de 6° de EP en el marco del Panel TERCE y la trayectoria de esos alumnos en todo el ciclo de la EM. Contar con la trazabilidad de los alumnos permite estudiar los eventos y fenómenos que suceden en el inicio del ciclo escolar y que son determinantes para entender los procesos de culminación de las trayectorias.

La potencialidad de este tipo de análisis se basa en la posibilidad de realizar inferencia sobre lo que le sucede a una cohorte respecto a la subsiguiente, especialmente a las cohortes que aún están transitando la Educación Primaria y todavía no comenzar EM.

Base empírica: universo y naturaleza de los datos

El universo de estudio son los estudiantes que estaban cursando 6° de EP en 2013 a nivel nacional tanto en escuelas urbanas como rurales, ya sea en el sector público o en el privado.

Fuente de datos

La base empírica comprende información recabada a partir una muestra nacional de niños de 6° de educación primaria en 2013, en el marco de TERCE⁹ realizado por LLECE¹⁰, y el seguimiento de esa cohorte de niños llevado a cabo por la DIEE entre 2014 y 2021.

TERCE cuenta con información sobre logros académicos medidos por pruebas estandarizadas en cuatro áreas del conocimiento: matemática, ciencias, escritura y lectura. Asimismo, cuenta con datos socioeconómicos y socioculturales de los niños y su familia recabados mediante formularios que completaron los padres de los niños seleccionados en la muestra. Por último, se tiene el rezago acumulado en 6° de EP, según la cantidad de años que el niño haya repetido.

La ANEP llevó a cabo el seguimiento de las trayectorias en dos etapas¹¹. La primera, abarcó de 2014 a 2016, correspondió al seguimiento de los estudiantes en los años normativos de la EMB. La segunda ola de relevamiento de información abarcó del 2017 a 2021 y correspondió a la EMS. Para ambas olas, se relevó información para cada año y registrando para el mes de marzo, si el adolescente estaba o no matriculado en EM, la oferta que eligió (secundaria o técnica), el grado en el que se encontraba o si el alumno estaba egresado.

El procedimiento que llevó a cabo la DIEE para buscar información sobre los ocho años mencionados constó de tres pasos: primero, buscar año a año a todos los estudiantes de la muestra de TERCE en las bases de datos de matrícula de centros públicos de ANEP. En segundo lugar, buscar en los registros de matrícula de centros privados a quienes no estaban en centros públicos. Por

⁸ Cardozo, S 2015; De Melo, G. y Machado, A. 2018, Lado, A 2019

⁹ Tercer Estudio Regional Comparativo y Explicativo, consistió en un estudio de logro de aprendizaje a gran escala aplicado en 2013 que abarcó 15 países de Latinoamérica

¹⁰ Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación

¹¹ La DIEE-ANEP decidió hacer el seguimiento en dos etapas para poder analizar en la primera ola el tránsito entre 6° de primaria y la EMB.

último, a los que aún no logró encontrar, se chequeó que no hayan fallecido o migrado del país, para poder registrarlos como fuera del panel. Una vez finalizada la búsqueda de todos los adolescentes en el sistema educativo, se tomaron tres situaciones para cada año: estar matriculado, no estar matriculado y estar fuera del panel (fallecido migrado). Para los matriculados se registró: el grado entre 1° a 6°, y el tipo de centro: liceo público, liceo privado o centro de educación técnica. Otro dato que se utiliza a partir de la matriculación, es la ubicación geográfica del centro, esto permite ubicar la residencia departamental de cada niño debido a la proximidad de los centros educativos con la residencia de los estudiantes.

Los datos mencionados permiten realizar un estudio de las trayectorias analizando un conjunto de información sobre tres situaciones principales: i) la continuidad o abandono de los vínculos con el sistema educativo formal, ii) el avance o progresión por los diferentes grados de educación media y iii) la oferta e institución de EM elegida por el estudiante, secundaria (pública/privada) o educación técnica¹²

La información del panel quedó conformada por nueve años, 2013 a 2021, conteniendo datos sobre la trayectoria educativa captada por eventos simples medidos en marzo de cada año. Asimismo, contiene los datos recolectados por TERCE sobre la familia y las cuatro pruebas de desempeño en 6° de EP. El Panel permite vincular información familiar, institucional y de desempeño académico de la trayectoria del niño en el primero nivel de educación, con los eventos educativos acontecidos en la trayectoria del segundo nivel educativo. Los datos que se cuenta en la EM son: matriculación/no matriculación, promoción/repetición, sector institucional del centro (público/privado), tipo de oferta educativa (liceo o educación técnica), egreso de EM.

Imputaciones: para aquellos estudiantes que no se tenía certeza de que hayan egresado de acuerdo a las bases de datos de educación media, se los buscó en las bases de datos de Udelar, en las bases del nivel terciario de ANEP, como los estudiantes del Consejo de Formación en Educación (CFE) y en otras instituciones de estudios terciarios.¹³ En el caso de encontrarlos, se los imputó como egresos de EM.

Consideraciones sobre el impacto de la pandemia en la matriculación y la progresión en tiempo: Las trayectorias educativas no se vieron afectadas por la pandemia hasta el 2020 (t+7). Los datos del panel toman la matriculación de los estudiantes y las mismas se efectuaron previo a la declaración de la pandemia por SARS-COV-2. El año en el que pudo haber incidencia sobre la matriculación provocado por la pandemia, es en el 2021 (t+8). De todas formas, no hubo una constatación en el indicador de progresión en tiempo, en cambio, para el indicador de permanencia, que refiere a la matriculación de los estudiantes, el efecto de la pandemia fue muy leve (Anep,2022).

Diseño muestral

El Panel TERCE - cohorte 6°- estuvo conformado por los 4.336 niños que participaron de TERCE en 2013. Constituyen una muestra aleatoria de alcance nacional de los alumnos uruguayos de 6°.

¹² Casi la totalidad de la oferta de educación técnica en el País, es pública

¹³ Existen otras instituciones de Educación superior a los que no se accedió a la base de datos, pero representen una proporción muy menor de la matrícula de Educación Superior

Incluye al sector público y privado y a las modalidades de escuelas urbanas y rurales. El diseño de la muestra de LLECE/UNESCO es bietápico, primero se seleccionan centros y luego estudiantes¹⁴.

Estrategia de análisis

Se realizará un análisis retrospectivo que procurará caracterizar la trazabilidad de las trayectorias escolares de los jóvenes que integran la cohorte 6° del Panel TERCE. El diseño del análisis intenta ser funcional a la estructura de objetivos planteada en sus dos componentes principales: caracterizar las trayectorias, dar cuenta de la incidencia de ciertos fenómenos que las moldean y finalmente predecir la desvinculación en EM.

En primer lugar, el análisis buscará describir el quantum de los procesos estudiados, presentando los porcentajes de estudiantes que año a año se encuentran matriculados según el tipo de oferta, el sector institucional, los grados. Asimismo, se describirán las características de las trayectorias, los momentos clave, las trayectorias más frecuentes, los procesos de no inscripción y desvinculación. También se presentarán indicadores de permanencia y de progresión en tiempo.

En segundo lugar, se intentará dar luz a los fenómenos que inciden en las distintas trayectorias, mediante técnicas descriptivas de análisis de tablas de contingencias y observando año a año la situación del estudiante. Se realizará una primera exploración para encontrar asociaciones. El análisis buscará indagar de qué forma ciertas características de la educación primaria afectan la configuración de las trayectorias.

En tercer lugar, para determinar el impacto que tienen los factores que inciden en las trayectorias se utilizará el modelo de regresión logística multinomial. Este modelo, en función de un conjunto de variables explicativas, estima el cambio en los riesgos relativos de estar en alguna de las categorías definidas para la variable dependiente. En esta investigación esta variable de resultado toma tres posibilidades respecto a la situación del estudiante al cabo de ocho años de haber cursado 6° de EP: haber egresado de la EM, estar matriculado en la EM o haberse desvinculado del sistema educativo formal.

En cuarto y último lugar, se emplearán técnicas de machine learning para predecir si el estudiante se desvinculará o de lo contrario no se habrá desvinculado -habrá egresado o seguirá inscripto-. A su vez el análisis se centrará en determinar las potencialidades que hay en Uruguay para construir predicciones sobre la desvinculación que tanguen buenos parámetros de rendimiento y sean posible calcular para todas las cohortes de estudiantes.

¹⁴ Por más información ver el informe técnico UNESCO-OREALC (2016)

3.5 Metodología para la predicción de trayectorias de riesgo

Logit ¹⁵

La regresión logística es un modelo estadístico que tiene por detrás la función logística o *logit*. Esta función mapea la variable Y que toma valores entre 0 y 1 como una función sigmoidea de x . La función logística es representada gráficamente en la *Ilustración A. 1* del anexo. La forma funcional se representa en la ecuación (1):

$$(1) f(x) = \frac{1}{1+e^{-x}}$$

La regresión logística binaria con múltiples variables independiente es empleada en problemas de clasificación con dos resultados posibles. La variable dependiente toma dos valores, 1 si el estudiante se desvinculó y 0 en caso contrario. A diferencia de la ecuación (1) en donde solo hay una variable independiente, este modelo permite considerar un conjunto de variables. Para modelar los datos de entrada, se asume una relación lineal entre las diferentes variables independientes. Se puede escribir la probabilidad de desvinculación para cada individuo (P_i), condicionada a las variables X_j como se muestra en la ecuación (2)

$$(2) (P_i) = E(Y_i / X_{1i} + \dots + X_{ji}).$$

Luego se puede (2) como (3):

$$(3) P_i = \frac{1}{1+e^{-(B_0+B_1x_1+\dots+B_jx_{ji})}}$$

Seguidamente escribir (3) como (4) para representar la función de distribución logística acumulada:

$$(4) P_i = \frac{1}{1+e^{-Z_i}} = \frac{e^Z}{1+e^Z}$$

Donde $Z_i = B_0 + B_1x_1 + \dots + B_jx_i$

Si la probabilidad de desvincularse para cada individuo P_i está dada por (4), entonces $(1 - P_i)$ es la probabilidad de ocurrencia del otro fenómeno posible: no desvincularse. Esta probabilidad se representa en la ecuación (5):

$$(5) 1 - P_i = \frac{1}{1+e^Z}$$

Por último, en la ecuación (6), se presenta el cociente entre la probabilidad de ocurrencia del fenómeno 1 “desvincularse” sobre la probabilidad de ocurrencia del fenómeno 2 “no desvincularse”.

¹⁵ La bibliografía consultada para escribir las ecuaciones se basó en (Wooldridge, 2010; Arellano, 2003; Rabe-Hesketh y Skrondal, 2022)

$$(6) \frac{P_i}{1-P_i} = \frac{1+e^z}{1+e^{-z}} = e^z$$

Tomando logaritmos en la ecuación (6) se llega a la ecuación (7):

$$(7) L_i = \ln \left(\frac{P_i}{1-P_i} \right) = z_i = B_0 + B_1x_1 + \dots + B_jx_i$$

De esta forma se obtiene el modelo Logit con un resultado para cada individuo: L_i es el logaritmo natural del cociente entre las probabilidades. Aunque la función logística calcula un rango de valores entre 0 y 1, para cada individuo, se determinará de acuerdo a un cierto umbral (u), si el individuo es un desvinculado o un no desvinculado. Para determinar el umbral (u) que separa a los individuos en esas dos categorías, se toma en cuenta la maximización de ciertos parámetros que se presentan a continuación.

Valoración del rendimiento del modelo¹⁶

Luego de la predicción, a cada estudiante se le asigna la categoría de desvinculado (1) o no desvinculado¹⁷ (0). La evaluación del resultado global del modelo surgirá de comparar las predicciones de ambas categorías contra los valores reales u observados. Esta contraposición de resultados estimados y reales, tiene asociadas ciertas métricas: exactitud, precisión, especificidad y sensibilidad, que surgen de la matriz de confusión presentada en la *Tabla 1*. Si el modelo funcionara perfectamente se lograría clasificar a los que efectivamente se desvincularon como desvinculados y a la totalidad de los que no se desvincularon como no desvinculados.

La matriz de confusión queda conformada por las combinaciones entre la predicción y la situación real del estudiante. Puesto que el fenómeno a identificar es el riesgo educativo, la categoría de desvincularse es la que tiene más “importancia” y será tomada como etiqueta positiva, en tanto, a la no desvinculación, se le asigna la etiqueta negativa. A continuación, se presentan las diferentes posibilidades entre la categoría asignada a cada estudiante al momento del egreso de EP y lo que efectivamente le ocurrió, ocho años más tarde en la EM.

- 1) Se desvinculó y es clasificado como desvinculado: verdadero positivo (VP)
- 2) No se desvinculó y es clasificado como no desvinculado: verdadero negativo (VN)
- 3) Se desvinculó y es clasificado como no desvinculado: falso negativo (FN), es un error de tipo I, considerado como error grave
- 4) No se desvinculó y es clasificado como desvinculado: falso positivo (FP), es un error tipo II, considerado como un error no tan grave.

¹⁶ Esta sección se basó en (Irizarry,2020; Bishop,2006)

¹⁷ Los no desvinculados engloban a los egresados y a los rezagados

Tabla 1. Matriz de confusión.

		Valores observados		Total
		No desvinculado (rezagado o egresado)	Desvinculado	
Valores predichos	No desvinculado (rezagado o egresado)	VN	FN	VN + FN
	Desvinculado	FP	VP	FP + VP
Total		VN + FP	FN + VP	VP + FP + FN + VN

Fuente: Elaboración propia con base en [\(Irizarry,2020\)](#)

Las principales métricas para evaluar el modelo son la exactitud, la precisión, la sensibilidad y la especificidad. Todas varían entre 0 y 1, siendo cero el peor resultado y uno, la magnitud que representa el mejor desempeño del indicador.

- **Exactitud** = $(VP + VN) / (VP + FP + FN + VN)$. Es una medida que da cuenta de los aciertos globales del modelo, es la proporción de resultados verdaderos dividido entre el número total de casos. Refiere a la tasa de aciertos.
- **Precisión** = $VP / (VP + FP)$. Proporción de verdaderos positivos sobre el total predichos positivos.
- **Sensibilidad** = $VP / (VP + FN)$. Proporción de verdaderos positivos sobre los casos que experimentaron el evento.
- **Especificidad** = $VN / (VN + FP)$. Proporción de verdaderos negativos sobre los casos que no experimentaron el evento.

Otros indicadores surgen de operar con las anteriores, uno de ellos es la tasa de falsos positivos, refiere a la cantidad de falsos positivos sobre la cantidad total de negativos. Es la probabilidad de que haya una alerta sobre riesgo educativo que en realidad es falsa. Tasa de falsos positivos = $FP / (FP + VN)$. También se calcula como $1 - \text{especificidad}$

Otro indicador es el **F1 score**, resume el comportamiento de la precisión y la sensibilidad, es de utilidad cuando la distribución de las etiquetas tiene una diferencia considerable. **Puntaje F1** = $(2 \times \text{Sensibilidad} \times \text{Precisión}) / (\text{Sensibilidad} + \text{Precisión})$.

Un indicador similar que también funciona como una media armónica es el puntaje F2, la diferencia es que le otorga más peso a la sensibilidad que a la precisión. **Puntaje F2** = $(1 + 2^2) \times (\text{Sensibilidad} \times \text{Precisión}) / (2 \times \text{Sensibilidad} + \text{Precisión})$.

El último indicador a tener en cuenta es Kappa, muestra el porcentaje de aciertos totales una vez descontado el azar, considera las distribuciones marginales de la matriz de confusión, es decir, las probabilidades de cada una de las etiquetas si fueran asignadas al azar. Es la ganancia que se obtiene por encima de una asignación aleatoria, si Kappa es mayor que 0,8 se considera una concordancia excelente, entre 0,6 y 0,8 buena, entre 0,4 y 0,6 moderada, y por debajo de 0,4 una concordancia pobre. $Kappa = (Exactitud - Pr(e)) / (1 - Pr(e))$.¹⁸

El desafío es elegir el punto de corte o umbral (u) que transforma la probabilidad continua de las estimaciones, en valores binarios (0 o 1). El hecho de contar con la variable continua aporta más información porque mirando el valor predicho para cada estudiante, se puede indicar que tiene más, o menos probabilidad de desvincularse que otro, y se puede hablar en términos de riesgo de desvinculación. Igualmente, es necesario decidir qué valor determina el umbral (u) para la asignación de la respuesta binaria 1 o 0.

El último criterio que se utilizará es el valor AUC¹⁹, muy utilizado en la Teoría de detección de señales para discernir qué es información y qué es ruido, corresponde al área por debajo de la Curva de ROC²⁰. Esta curva muestra la sensibilidad (la proporción de verdaderos positivos) frente a la proporción de falsas alarmas (1-especificidad) para un sistema de clasificación binario. Cada punto en la curva representa un par de sensibilidad y 1-especificidad según varía el umbral de discriminación. Es una medida que muestra la eficacia del modelo, es independiente de los criterios que cada investigador utilice de acuerdo a su problema y es independiente de la tasa de verdaderos positivos. Valores de AUC cercanos a 1 indican un buen rendimiento y valores cercanos a 0,5 un rendimiento aleatorio.

Para finalizar la sección de metodología, es pertinente mencionar que existen varios modelos para predecir un evento, a modo de ejemplo se listan algunos que no se utilizaron, pero serían válidos. Inclusive, sería interesante hacer la comparación de las predicciones resultantes de los siguientes métodos: I) KNN (k- Nearest Neighbors): es un algoritmo no paramétrico de aprendizaje supervisado. Predice la clasificación de un punto de datos basándose en la mayoría de las clases de sus k vecinos más cercanos en el espacio de características. II) Árbol de Decisión: es un modelo en forma de árbol donde cada nodo interno representa una decisión basada en una característica, cada rama representa un resultado de esa decisión y cada hoja una etiqueta de clase. III) Redes Bayesianas: son modelos gráficos probabilísticos que representan un conjunto de variables aleatorias y sus dependencias condicionales a través de un grafo acíclico dirigido. (Bishop, 2006).

¹⁸ $Pr(e)$ = Coeficiente de exactitud aleatorio, es el porcentaje de aciertos que cabría esperar del azar

¹⁹ Área por debajo de la curva, por sus siglas en inglés: "Area Under the Curve"

²⁰ acrónimo de Receiver Operating Characteristic, por sus siglas en inglés

4. Caracterización de las trayectorias educativas en Educación Media

4.1 Situación del panel

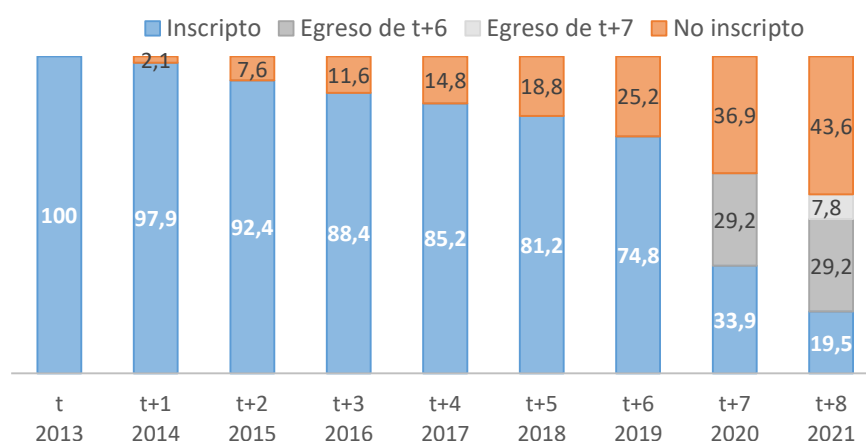
Como se mencionó anteriormente, se sigue durante ocho años a una cohorte de estudiantes representativa del sistema educativo uruguayo que cursaron 6° de EP en 2013. La configuración final del panel en cada año, presenta una baja mortandad (perdida de casos) ver Tabla A. 1 del anexo. Para el análisis de las trayectorias se dejan por fuera estos casos, para cada año oscilan entre 0,4% y 1,04%.

Como muestra el Gráfico 1, la ventana de observación comienza en $t =$ marzo del 2013 y termina en $t+8=$ marzo de 2021. Para el mes de marzo de cada uno de los ocho años que componen el panel, se configuraron las siguientes situaciones para cada estudiante: i) estar inscripto, ii) no estar inscripto, iii) ser un egresado correspondiente al año lectivo $t+6$ y iv) ser un egresado del año lectivo $t+7$.

En $t+1$ (2014) estaban inscriptos el 97,9% de los estudiantes de la cohorte de 6° de EP. Al cabo de seis años se encontraban inscriptos el 74,8%. Por su parte, en $t+8$ (2021): el 19,5% continuaba matriculado, el 29,2% logró egresar en tiempo, el 7,8% egresó con un año de rezago y el 43,6% no se encontraba inscripto. A lo largo del capítulo se comentarán los resultados en mayor profundidad con base en los indicadores de permanencia, progresión en tiempo y egreso. Además, se presentarán los resultados según diferentes variables de interés. Las variables de interés se utilizarán a lo largo de toda la tesis como variables que moldean las trayectorias de los alumnos. Refieren al momento del egreso de Educación Primaria y se agrupan en tres dimensiones. i) características del sistema educativo: sector institucional, oferta educativa y región del centro, ii) características de la familia: NSE y expectativas familiares sobre la trayectoria del niño y iii) características del propio estudiante: sexo, rezago y desempeños en la prueba de matemática y de lectura de TERCE. En la *Tabla A. 2* del anexo se presentan las variables y las medidas de resumen, y en la *Tabla A. 3* la correlación entre ellas.

El egreso de cierto año lectivo, se registró en marzo del siguiente año. Así, la posibilidad de egresar de $t+6$, correspondiente al año 2019 y se registra al otro año, en marzo de $t+7$. En tanto, los egresos correspondientes al año lectivo $t+7$, se registran en marzo de $t+8$, tal como muestra el *Gráfico 1*.

Gráfico 1. Porcentaje de estudiantes según la situación registrada en marzo de cada uno de los 8 años que configuran el panel de datos.



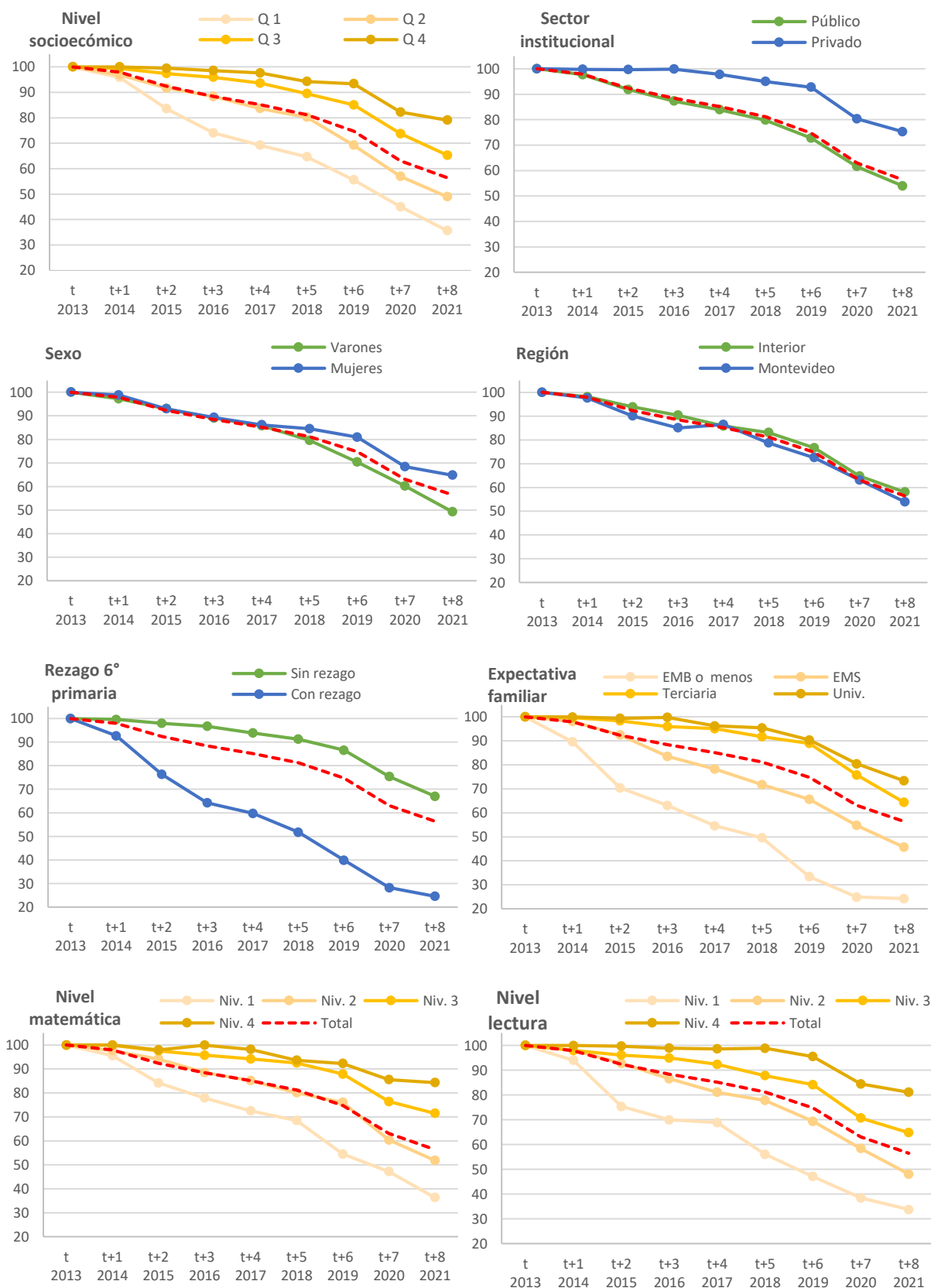
Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

4.2 Permanencia

Este apartado describe las trayectorias en función del indicador de permanencia presentado en *Objetivos específicos*. En el *Gráfico 2* se observa para el total de la cohorte²¹, el porcentaje de inscripción en EM y la inscripción abierta según las variables de interés que caracterizan a los estudiantes de 6° de EP. Para cada año se tiene la “fotografía de matriculación”, un alumno no matriculado en cierto período puede estar matriculado en los años siguientes. En la *Tabla A. 4* del anexo se presentan los porcentajes de matriculación y en la *Tabla A. 5*, las diferencias entre los extremos de las categorías de las variables de interés. Estrictamente hasta t+6 se muestran los porcentajes de inscripción, mientras que en t+7 y t+8, se suma a la inscripción, el porcentaje de estudiantes que hayan egresado. De esta forma los últimos dos años del panel están reflejando una situación de permanencia o egreso.

²¹ El 100% de los casos representa a todos los estudiantes de la cohorte de 6° de primaria, menos, a los casos missing que varían año a año, como se indicó no superan el 1,04% de la cohorte de estudiantes.

Gráfico 2. Porcentaje de estudiantes matriculados sobre el total de la cohorte en de cada uno de los 8 años del panel de datos, según variables de interés.



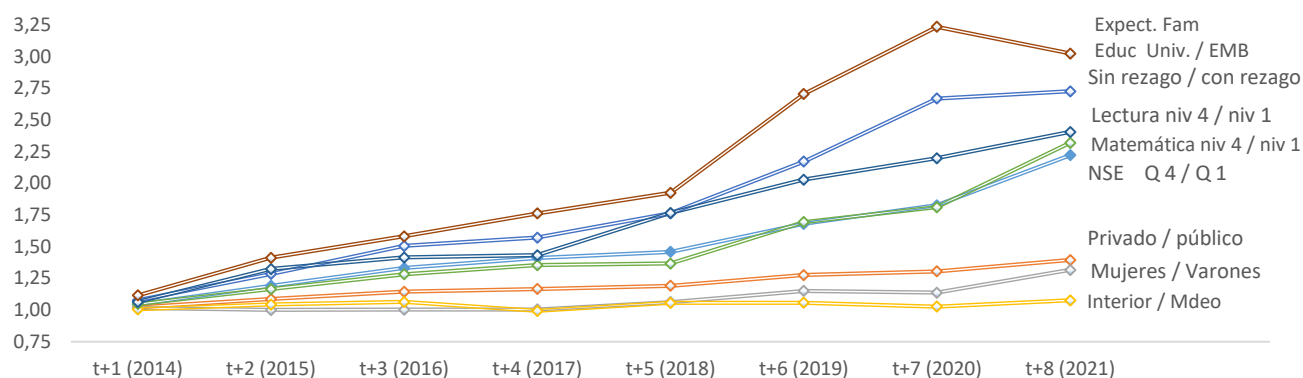
Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021

Para el total de la cohorte hay una caída que mantiene su ritmo hasta alcanzar 3 de cada 4 estudiantes que permanecen inscriptos en t+6 (2019), último año normativo del calendario académico de EM. Luego, se aprecia un descenso más pronunciado hasta alcanzar en t+8 (2021), el 56,4% de adolescentes. Ese porcentaje se compone por la suma de los que estaban dentro del sistema educativo con rezago (19,4%) y los que ya habían egresado (37%). La caída de la permanencia deja evidenciado los eventos de no inscripción que van aumentando año a año y muestran como en su conjunto, la cohorte de estudio se va desvinculando. En el capítulo 5.1 se estudiará este fenómeno con mayor profundidad.

En la lectura bivariada de los gráficos presentados, se aprecian las desigualdades de origen: a mayor NSE, mayor es la permanencia en el sistema educativo. A su vez, si el estudiante cursó 6° en el sector privado, luego, en la Educación Media, permanece más, que aquel que cursó 6° en el sector público. Por su parte, cabe remarcar que hay una correspondencia casi uno a uno entre quienes pertenecen al cuartil superior del índice de nivel socioeconómico (Q4) y quienes asistieron al sector privado. Para toda la serie no hay grandes diferencias según sexo y región, aunque en los últimos años, las niñas permanecen cerca de 10 puntos porcentuales más que los varones. La situación de rezago en 6° de EP muestra una rápida separación de caminos: aquellos que comienzan la EM sin rezago, permanecen en mayor proporción que los que sí presentan esa condición. Estos dos caminos se abren tempranamente y se bifurcan aún más, a medida que pasa el tiempo.

Las expectativas de las familias respecto a la culminación del estudiante tienen un comportamiento similar al comportamiento del rezago, en el sentido de que también presentan grandes brechas que inician en el primer año de la transición entre la EP y EM, y se van ensanchando hasta alcanzar, por ejemplo, en (t+8), diferencias del orden de los 50 puntos porcentuales entre las categorías de menor y mayor expectativa. Por último, se presenta la permanencia según los niveles de desempeño en la prueba de aprendizaje de matemática y lectura de TERCE. Más del 90 % de los estudiantes de nivel 4 se encontraban inscriptos en EM al cabo de seis años de haber cursado el último grado de EP, en cambio, solamente la mitad de los estudiantes de nivel 1 se encontraban inscriptos en t+6.

Gráfico 3. Ratios entre el porcentaje de permanencia de las categorías extremas de las variables de interés.



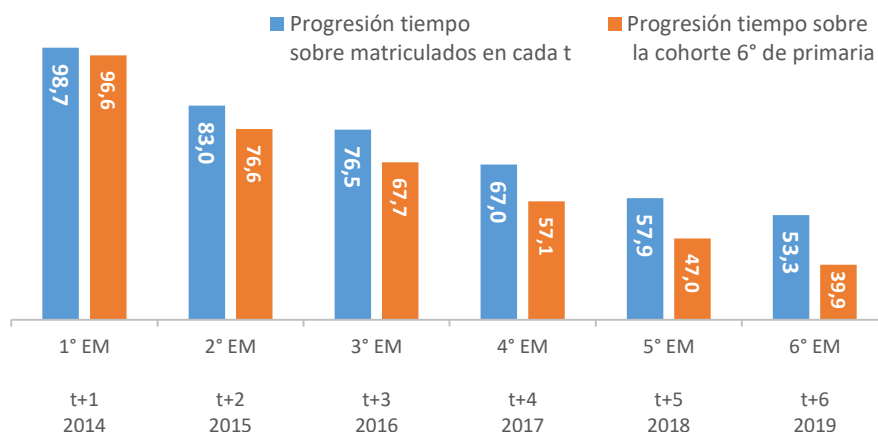
Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021

El Gráfico 3 presenta el ratio entre las categorías extremas de cada variable. A partir de t+5 se puede decir que por cada estudiante inscripto para el cual su familia cree que el máximo nivel alcanzado será EMB, se tiene al menos, el doble de estudiantes inscriptos cuyas familias creen que el niño culminará la universidad. Para los últimos tres años del panel, el ratio aumenta hasta alcanzar valores cercanos a tres. La siguiente variable que muestra grandes diferencias es la situación de rezago, desde t+6, por cada estudiante con rezago se encuentran inscriptos, más del doble sin haber repetido durante la EP. Le sigue, el ratio entre el nivel 4 y el nivel 1 de las pruebas de aprendizaje (matemática y lectura), y el ratio entre el cuartil 4 y el 1 del NSE. Finalmente, las ratios con valores más bajos (inferiores a 1,5) corresponden al sector institucional, al sexo y a la región. Para estos subgrupos no hay grandes diferencias en los ocho años del panel.

4.3 Progresión en tiempo

Una vez presentado el análisis de matriculación, que es la condición necesaria para que exista un proceso educativo, interesa revisar la velocidad en que se transita por la EM. Normativamente, se espera que un estudiante se encuentre inscripto en 6° de EM, al cabo de seis años de que inicia sus estudios en el segundo nivel de enseñanza. Para ello, deberá avanzar por los grados sin repetir ningún año. Este indicador se construye para cada año entre t+1 (2014) y t+6 (2019). Los porcentajes registrados se encuentran en la Tabla A. 6 del anexo y las diferencias según las variables de interés se reportan en la Tabla A. 7.

Gráfico 4. Porcentaje de inscriptos en el grado normativo sobre matriculados en cada t y sobre el total de la cohorte de 6° de Educación Primaria.



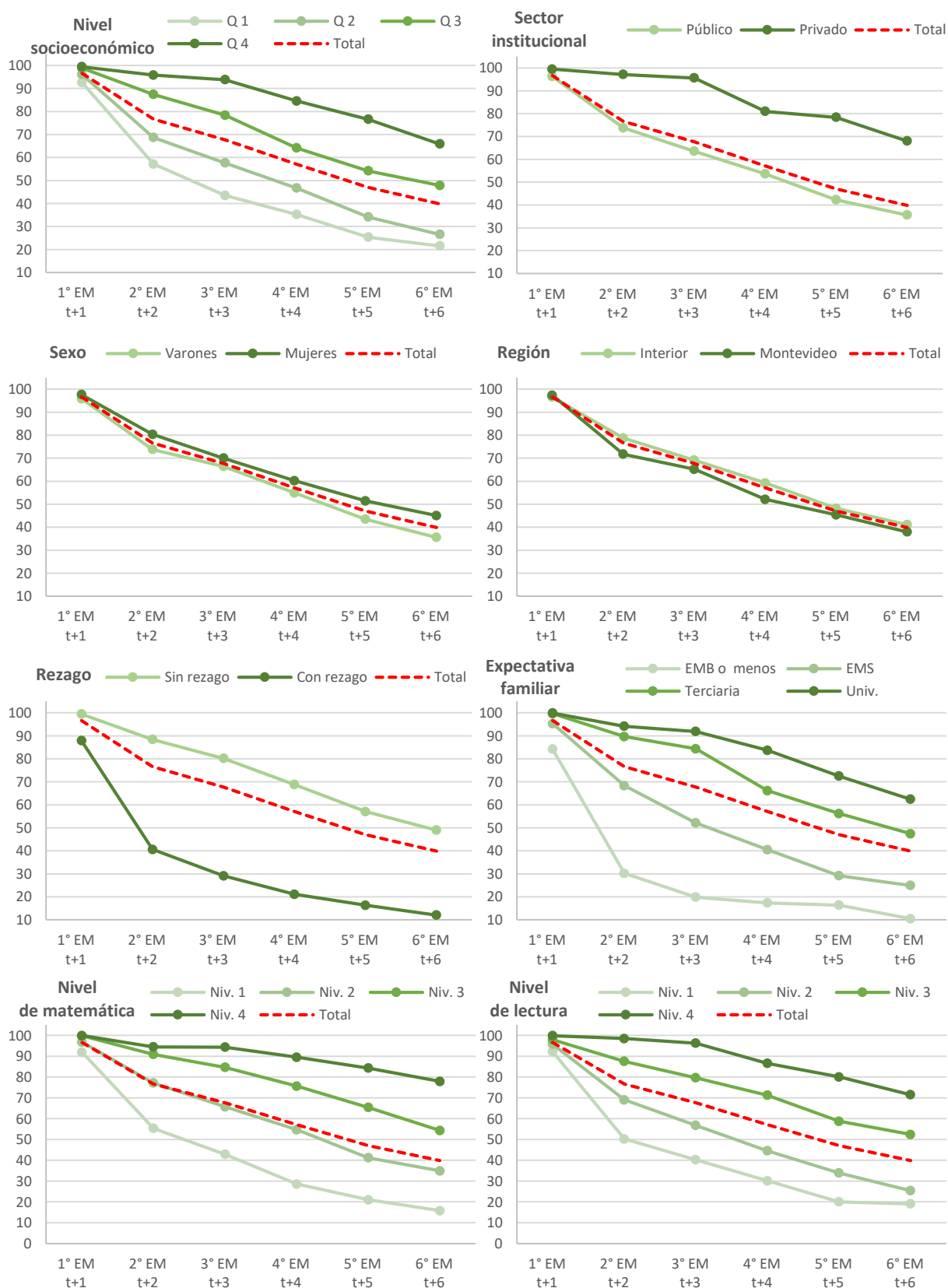
Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021

El *Gráfico 4* muestra el porcentaje de estudiantes que progresan en tiempo para dos poblaciones: los matriculados en cada año del panel y el total de la cohorte de estudio. Para cualquiera de las dos poblaciones, se aprecia la disminución en el porcentaje de estudiantes. Del total de inscriptos en t+6 (2019), la mitad se encontraba matriculados en 6° de EM. Por otro lado, respecto al total de la cohorte de 6° de EP, solamente 4 de cada 10 progresaron en tiempo.

Para el total de la cohorte, hay un descenso pronunciado de 20 puntos porcentuales entre t+1 y t+2 (1° y 2° respectivamente). Esto sugiere que los primeros años de la transición entre los ciclos de EP y EM resultan especialmente críticos para la progresión en tiempo, incluso cuando no se considera la desvinculación, un aspecto documentado tempranamente para Uruguay por Castro (1949). En los siguientes períodos, de t+3 a t+6, la progresión en tiempo cae, aproximadamente 10 puntos porcentuales por año. En t+6 el indicador muestra el porcentaje máximo de estudiantes que está matriculado en el último grado de EM y, por tanto, implica que como máximo, solamente el 40% está en condiciones de egresar en tiempo. El análisis sobre el egreso se retomará más adelante.

Para las implicancias de tener el resultado del indicador de permanencia también es interesante, analizar lo que sucede en t+3 y t+6, ya que marcan el último año de EMB y EMS respectivamente. Llegan a tercer grado en el tiempo normativamente previsto, menos del 70% de los estudiantes, y únicamente el 40% llega al último grado de EMS según el ritmo pautado por el calendario normativo. En el *Gráfico 5* se presenta la progresión en tiempo según las variables de interés.

Gráfico 5. Porcentaje de estudiantes inscriptos en el grado normativo sobre el total de la cohorte de 6° de EP, según variables de interés.



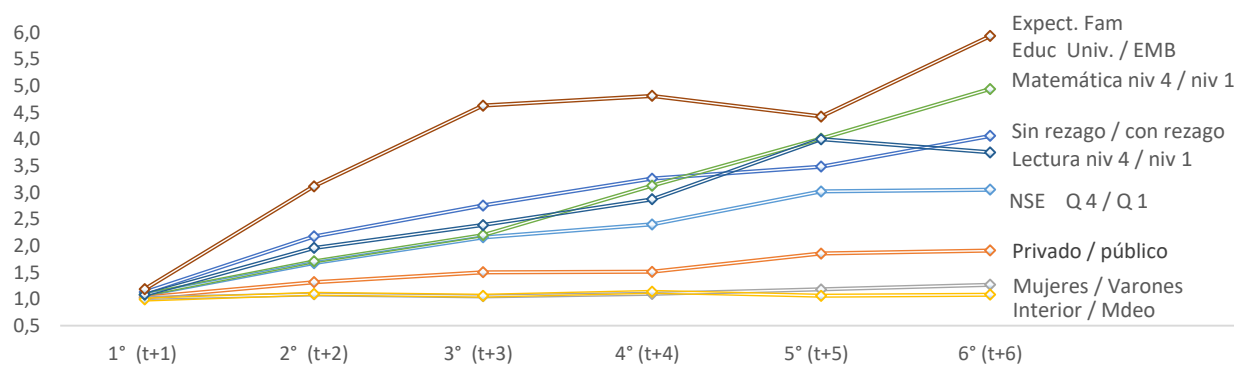
Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Nuevamente en la apertura por región se aprecian muy pocas diferencias. Lo mismo sucede respecto a la apertura según sexo. Sin embargo, aunque las diferencias a favor de las mujeres en el indicador de progresión en tiempo sean leves, se aprecian tempranamente desde el primer año de la EM, en cambio en el indicador de permanencia lo hacían desde t+4. En líneas generales, el ordenamiento de las categorías presentadas sobre la progresión en tiempo, es el mismo que para el indicador de permanencia, pero las brechas se agrandan y acontecen antes en el tiempo.

Respecto al NSE y al sector institucional, prácticamente la totalidad de los estudiantes de cuartil 4, y la totalidad de los estudiantes que asistieron a escuela privada en 6° EP, logran progresar en tiempo por la EMB (de t+1 a t+3). En cambio, en la EMS (de t+4 a t+6), 7 de cada 10 estudiantes de cada una de las poblaciones mencionadas logra avanzar hasta 6° de EM en el tiempo esperado. Es decir, la población que se encuentra en el cuartil más favorable, no tiene asegurado su egreso en tiempo de la EMS. Más severa es la situación que registran los estudiantes del cuartil más bajo, en donde apenas 2 de cada 10 logran llegar a 6° de EM en t+6. Las últimas dos aperturas muestran que, de cada 10 niños con rezago en EP, solamente 1, logra progresar en tiempo por la EM. Lo mismo sucede para los niños cuyos padres piensan que como máximo, culminarán la EMB. Es altamente probable que estos resultados estén estrechamente asociados, debido a la correlación de las expectativas de las familias con los desempeños académicos previos del niño. Este aspecto se retomará en el análisis multivariado del capítulo 6.

Una vez analizadas las diferencias entre las categorías extremas representadas por las ratios en el Gráfico 6. Se encuentra que el ordenamiento de las variables de interés en el indicador de progresión en tiempo, es similar, al ordenamiento del indicador de permanencia.

Gráfico 6. *Ratios entre el porcentaje de progresión en tiempo de las categorías extremas de las variables de interés*



Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

La mayor diferencia, está en la variable expectativa familiar y las menores diferencias, se presentan en el sector institucional, el sexo y la región. La variable que se ubica en una posición distinta en ambos indicadores es el nivel de matemática, aparece como la segunda variable con mayores diferencias. Por cada estudiante inscripto en 6° de EM que pertenece al nivel 1 de matemática, hay 5 estudiantes del nivel 4 inscriptos en ese grado. En el escalón siguiente se encuentra el ratio entre

los que tuvieron rezago y los que no, y el ratio entre los de nivel 4 y nivel 1 en lectura, en ambos casos la diferencia es 4 veces superior en 6° EM.

4.4 Egreso

La culminación de la Educación Media marca el fin de la trayectoria establecida como obligatoria en Uruguay. Es el resultado de haber cumplido con los requisitos impuestos por los reglamentos de las diferentes ofertas educativas que culminan con la aprobación 6° de EM.²²

A continuación, se reporta el indicador de egreso en tiempo, corresponde al porcentaje de estudiantes que culminaron la EMS luego de cursar 6 años y aprobaron todos los cursos de la propuesta curricular sobre el total de la cohorte de estudio. En t+6, habrán empezado a cursar 6° de EM y en marzo de t+7 se los registra como egresados de t+6. También se analiza el egreso con un año de rezago en EM, corresponde a los estudiantes que a marzo de t+8 se encontraban como egresados del año lectivo t+7. El *Gráfico 7* presenta estos dos indicadores para el total de la cohorte y para las variables de estudio. El 29,2% de los alumnos de la cohorte de 6° de EP egresa en tiempo de la EM y 7,8% lo hace con un año de rezago generado en EM.

²² Hay algunas ofertas de Educación Media que finalizan en grados anteriores a sexto, pero para este estudio se tomó 6° como el último requisito para finalizar los dos niveles de la EM: la EMB y la EMS.

Gráfico 7. Porcentaje de egresados en tiempo y con un año de rezago para el total de la cohorte de 6° de educación primaria, según variables de estudio



Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021

El egreso en tiempo es un fenómeno poco ocurente en las trayectorias de los estudiantes y a su vez está muy estratificado. Se presenta en mayor proporción en el cuartil 4, en aquellos estudiantes que cursaron 6° en el sector privado, en los que obtuvieron el nivel 4 en la prueba de TERCE, y en los estudiantes cuyas familias tienen altas expectativas respecto al nivel al que lograrán finalizar sus hijos. Para la población de estudiantes con esas características, el indicador se encuentra entre 50% y 65%, un valor de todos modos bajo, si se considera que el nivel es obligatorio y que se trata de estudiantes que, en principio, no presentan vulnerabilidades sociales ni académicas. Luego, le siguen los estudiantes que no presentan rezago en EP, en donde cerca del 36% logra egresar en tiempo. Por último, tanto aquellos de Montevideo y del interior, como las mujeres y los varones, presentan un egreso en tiempo cercano al promedio nacional. Igualmente, hay diferencias por sexo de 13 puntos porcentuales a favor de las mujeres.

Otro punto a mencionar es que la ganancia en el porcentaje de egresados que se tiene por dejar pasar un año más, es muy baja. Tres de cada diez estudiantes egresaron en tiempo, y luego de un año, de los siete estudiantes restantes, solamente uno, logró egresar. Asimismo, la estratificación de quienes egresan con rezago es similar.

Si se observa el conjunto de los que egresan, la población de los alumnos que alcanzan desempeños académicos más altos en 6° EP tienen más chances de egresar en tiempo. El 64,8% de los que obtuvieron nivel 4 en matemática y el 57,9% de los obtuvieron nivel 4 en lectura, egresaron en tiempo.

El egreso es una de las tres situaciones posibles en la trayectoria educativa al finalizar la ventana de observación del panel, esto deja dos situaciones posibles para el resto de los estudiantes, seguir matriculados o haberse desvinculado del sistema. Precisamente, en el siguiente capítulo, se analizará la trayectoria de desvinculación.

5. Desvinculación y tipología de trayectoria al final del período de observación

5.1 Desvinculación

Se presentan tres abordajes respecto a la no inscripción. El primero son los estudiantes no inscriptos en cada año, sobre el total de la cohorte. El segundo, es el porcentaje de estudiantes en cada año, sobre el total de estudiantes matriculados el año anterior, este abordaje refiere a la desvinculación inter-anual o también vista como tasa instantánea de desvinculación. Por último, se presenta el porcentaje de estudiantes sobre el total de la cohorte que no estuvieron matriculados durante tres años consecutivos. Para cada año entre $t+3$ y $t+8$, este indicador sigue el mismo criterio para definir a un estudiante desvinculado al final del período de observación. En este estudio se define a un individuo desvinculado como aquel que no registra inscripción en los tres últimos años del panel ($t+6$, $t+7$, $t+8$). Además, se calcula el mismo indicador, pero para dos años consecutivos sin inscripción y así, poder observar las diferencias entre estar dos o tres años consecutivos sin registrar matriculación.

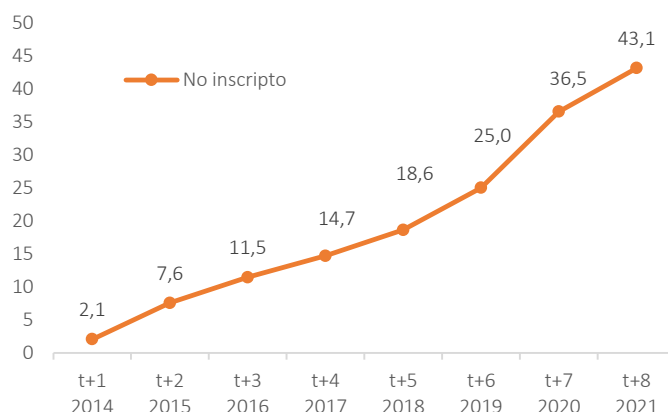
No inscriptos en cada año respecto al total de la cohorte

Este indicador nos permite tener una idea de lo que le va pasando a toda la cohorte en su conjunto a lo largo de la EM, en cuanto a la no inscripción. La no inscripción a lo largo de la EM es un fenómeno acumulativo, año a año crece el porcentaje de estudiantes que no se inscriben y como muestra el *Gráfico 8*, parece ser a un ritmo constante en toda la serie, cada año, el indicador crece aproximadamente 5 puntos porcentuales.

El primer año ($t+1$) un 2,1% no se matriculó en la educación formal luego de haber cursado el último grado de EP. Al tercer año ($t+3$), 1 de cada 10 estudiantes no se encontraba matriculado. Esto marca una pauta de riesgo temprana, en $t+3$, hay una cantidad considerable de estudiantes que ese año estarán fuera del sistema educativo. El propio evento de no inscripción significa un riesgo importante para la trayectoria y muy posiblemente sea la antesala de la desvinculación, este desenlace es un proceso que se va dando poco a poco (Cardozo, 2010). Esta tesis no permite estudiar el abandono intra-anual, es decir, un estudiante que se matricula en un cierto año, pero deja de ir a clases durante ese mismo año. El abandono intra-anual es una fuerte señal de que el estudiante no concretará su inscripción en alguno de los años siguientes. (Cardozo, 2016; Cardozo, Schenck y Severino, 2018)

Siguiendo con el análisis del *Gráfico 8* para cada año. En $t+4$ sigue creciendo la no inscripción hasta alcanzar en $t+5$, a casi 2 de cada 10 estudiantes, para el último año normativo de EM ($t+6$) se tiene que, 1 de cada 4 no estaba matriculado. Entre $t+6$ y $t+7$, el ritmo de crecimiento es mayor que en los anteriores períodos de la serie, crece en 11,5 puntos porcentuales. Finalmente, en el último año del panel se tiene que 4 de cada 10 alumnos no estaban inscriptos y tampoco habían egresado.

Gráfico 8. Porcentaje de estudiantes no inscriptos sobre el total de la cohorte de 6° de EP



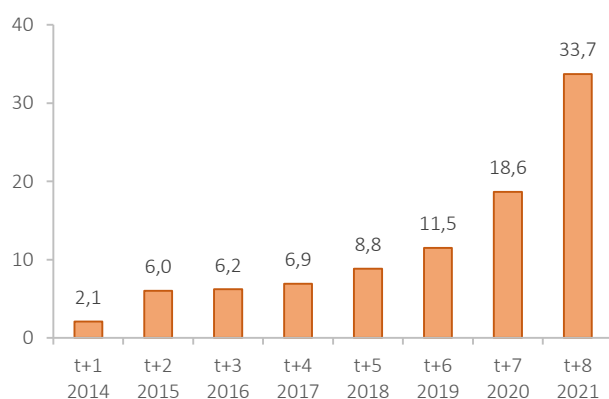
Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Hasta esta sección, quedan analizadas las cuatro situaciones que muestra el *Gráfico 1* al inicio de este capítulo: Inscriptos, egresados en tiempo, egresados con rezago y no inscriptos. En el último año del panel (t+8 o 2021) de cada 10 alumnos representativos de la cohorte de estudiantes uruguayos que estaban cursando 6° en 2013: 3 habían egresado en tiempo, 1 egresado con rezago, 2 continuaban inscriptos y 4 no estaban matriculados.²³

Desvinculación respecto al año anterior.

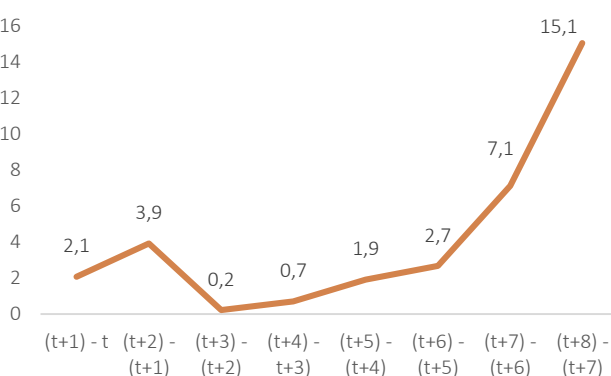
La tasa de riesgo instantáneo nos permite analizar la matriculación en cada año respecto al anterior, también se le suele llamar abandono interanual. El *Gráfico 9* muestra el porcentaje de estudiantes que, habiendo estado matriculados en un año, al siguiente, ya no se encontraban matriculados, y el *Gráfico 10* muestra la diferencia en puntos porcentuales entre un año y otro.

Gráfico 9. Porcentaje de estudiantes no inscriptos sobre el total de matriculados del año anterior.



Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021

Gráfico 10. Diferencias en puntos porcentuales entre un año y el año anterior.



Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021

²³ De esos 4 estudiantes no matriculados en t+8, puede suceder que no se hayan matriculado pero que en el correr de t+8 estén dando exámenes, lo siguientes apartados dan a entender que esta situación es muy poco frecuente

En la transición entre EP y EM se pierde al 2% de los estudiantes. Luego, del total de estudiantes inscriptos el primer año de EM ya no se encontraban matriculados al siguiente año, en t+2, el 6%. Misma situación ocurre al tercer y cuarto año. En cambio, en t+5, casi 9% no se matriculó con respecto al año anterior. El indicador sigue creciendo cada vez a un ritmo mayor. En el *Gráfico 10* se aprecian las diferencias entre un año y otro, registrándose un gran cambio en la pendiente de la curva entre t+6 y t+7. Finalmente, de t+7 a t+8 se registra la mayor diferencia alcanzando los 15 puntos porcentuales. Estos dos últimos años corresponden al período posterior a los seis años normativos de la EM, lo que sugiere que quienes no egresan en tiempo o un año más tarde de lo normativo, tienen alto riesgo de no inscribirse en los años siguientes.

En resumen, se puede concluir que entre el primer y segundo año de EM se pierden más alumnos que en la transición de EP a EM, del segundo al cuarto año la magnitud de pérdida de alumnos se estabiliza en torno a 6 puntos porcentuales por año. Entre el cuarto y el sexto año, ronda un 10% pero finalizado el período normativo de la EM, se registra una pérdida inter-anual mucho mayor, en donde 2 de cada 10 estudiantes inscriptos en t+6 ya no estaban en t+7. Finalmente 1 de cada 3 inscriptos en t+7 no estaba en t+8. Esto implica que, si el estudiante no logró finalizar sus cursos en el período normativo o un año después, no persiste en terminarlos²⁴, aspecto también señalado en (Fernández, 2009).

Desvinculación sobre el total la cohorte de 6° de Educación Primaria

El indicador de desvinculación en cada año considera dos situaciones. La primera es la más estricta, se compone por el porcentaje de estudiantes no inscriptos durante tres años consecutivos sobre el total de la cohorte. El primer momento donde puede ocurrir es en t+3 (2016) y el último, es en t+8 donde la definición coincide con una de las tres posibilidades en la tipología de trayectoria al final del panel: egresado, rezagado o desvinculado. La segunda es más laxa, se compone por el porcentaje de estudiantes que no estuvieron matriculados dos años consecutivos, el primer momento se registra en t+2 (2015). El *Gráfico 11* ilustra los dos criterios.

²⁴ Al menos desde el cursado tradicional en donde se requiere estar inscripto.

Gráfico 11. Porcentaje de desvinculados sobre el total de la cohorte.



Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

En el primer criterio, al cabo de tres años de haber cursado 6° en 2013, el 1,5% de los estudiantes se encontraban desvinculados, en t+4 crece a 5%, en t+5 a 8,4%. Al cabo de seis años se tiene a 1 de cada 10 estudiantes desvinculados y dos años más tarde, en t+8, 2 de cada 10 estaban fuera del sistema educativo. Estos estudiantes no se llegaron a inscribir en el último año de EM (2019), ni en los dos siguientes.

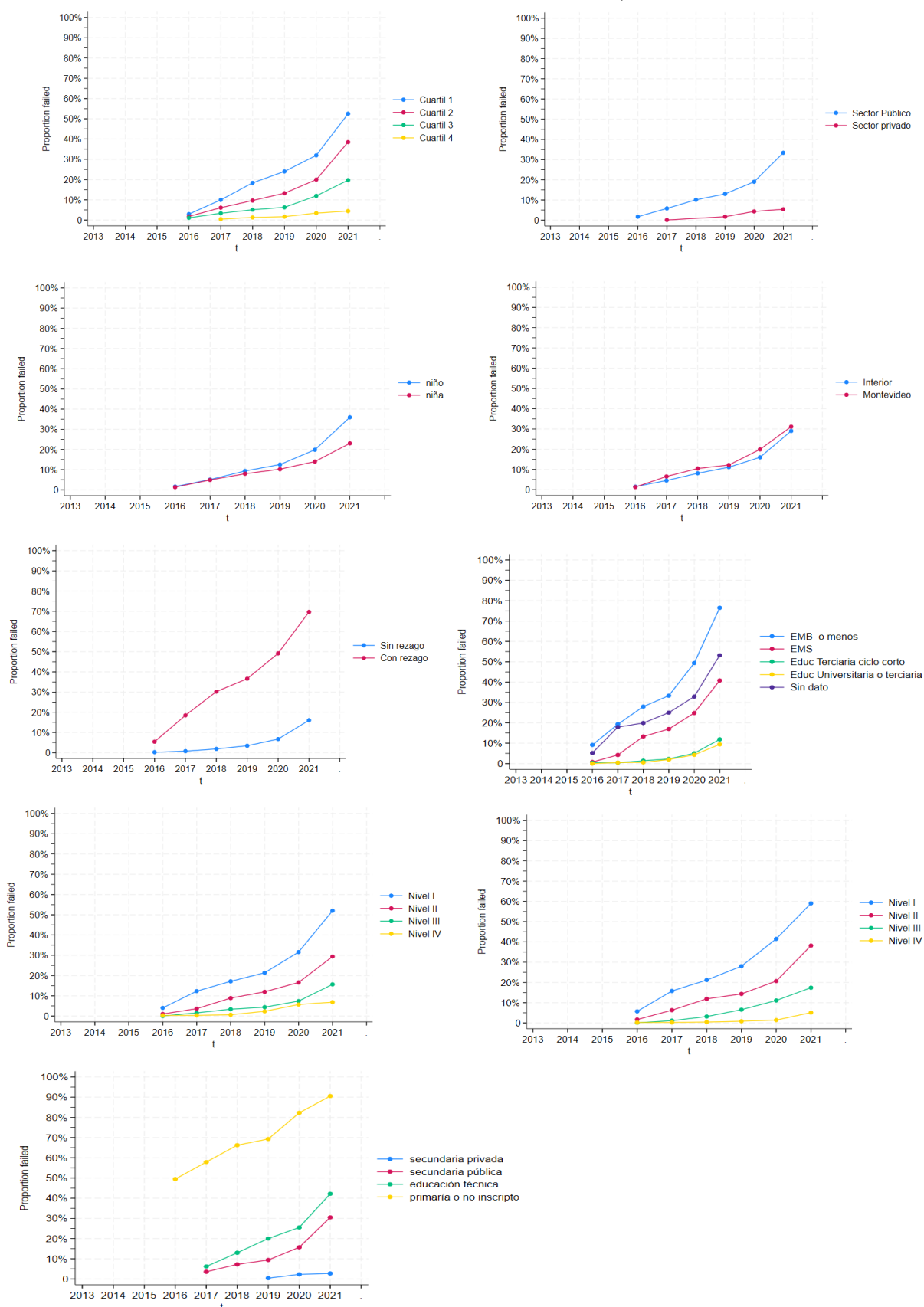
Para el segundo criterio, en donde un desvinculado es aquel que no se inscribe durante dos años seguidos, se encuentra que, tempranamente en t+3 (2016), en vez de haber 1,5% de desvinculados como en el criterio de 3 años consecutivos, el 5,8% de los estudiantes se había desvinculado. En t+6, el 15,9% no se matriculó ni al 5to ni al 6to año luego de haber cursado EP. Para los dos últimos años del panel, la desvinculación alcanza al 23% y 32% respectivamente. Para cualquiera de los dos criterios, por cada año que transcurre, la magnitud del problema se incrementa en una proporción similar, es decir, el desgranamiento de la matrícula es muy parecida en ambos casos. Los porcentajes de desvinculación del segundo criterio acontecen un año antes. Habría un indicio de que estar 2 años sin matriculación implica que el estudiante es muy probable que no se re-vincule con el sistema educativo.

Análisis de Supervivencia

Lo que sigue, muestra como la desvinculación también es un fenómeno estratificado, el *Gráfico 12* presenta el primer criterio de desvinculación (3 años consecutivos sin matriculación) según las variables de estudio. La particularidad es que, al ser un análisis de supervivencia, lo que se muestra es el porcentaje de estudiantes que no estuvieron matriculados durante tres años consecutivos por primera vez, independientemente de si volvieron a matricularse más tarde en su trayectoria.²⁵ Específicamente, se reporta la proporción acumulada de estudiantes que se desvincularon, es decir, el inverso del porcentaje de supervivientes.

²⁵ Técnicamente, esto implica que los sujetos que experimentaron el evento salen del conjunto en riesgo y no son considerados para el análisis de las duraciones o años posteriores.

Gráfico 12. Porcentaje de estudiantes desvinculados (no inscriptos tres años consecutivos) sobre el total de la cohorte, análisis de supervivencia



Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

La estratificación mantiene relación con los indicadores anteriormente presentados. Las chances de desvincularse crecen, cuanto menor es el NSE, menor el nivel de lectura y matemática y menor sean las expectativas familiares. A su vez, en los varones, la desvinculación se presenta con mayor intensidad, así como en los que asistieron al sector público en 6° de EP, en comparación con los que asistieron al sector privado. La mayor brecha se evidencia según la situación de rezago. La mitad de los que acumularon rezago en 6° de EP, al cabo de siete años –en el 2020- confirmaron su situación como desvinculados, es decir, no se inscribieron en t+5, t+6 ni t+7, y muy probablemente tampoco se encuentren inscriptos en t+8. Estos adolescentes que se terminaron desvinculando y que presentaron rezago en EP, representan al 12,5% de la cohorte. En cambio, para el mismo período, de los que no presentan rezago al iniciar la EM, solamente 1 de cada 10 se encontraba desvinculado. De igual modo, las expectativas familiares también presentan grandes brechas. En t+7 (2020), el 50% de los niños cuyas familias esperan que cursen como máximo hasta EMB, se habían desvinculado, frente al 5% de los niños cuyas familias esperan que finalicen educación terciaria o universitaria.

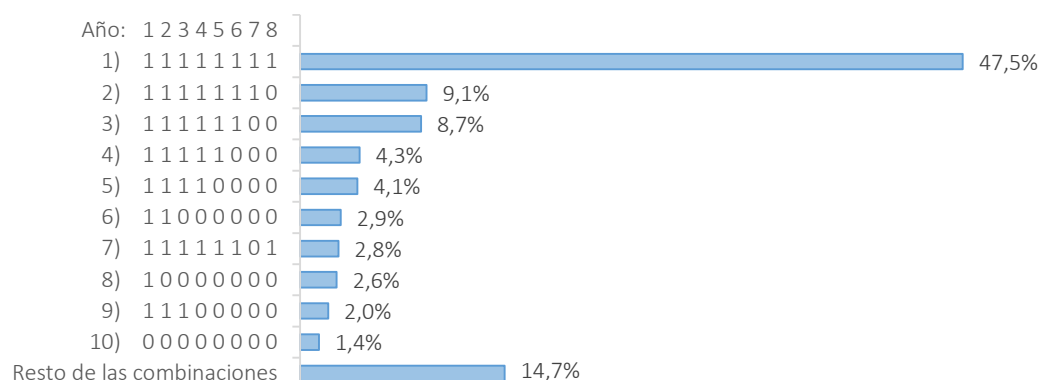
Finalizando la descripción de cada variable, no hay diferencias relevantes según región, pero sí, de acuerdo a la combinación de oferta y sector que eligieron los niños y sus familias para comenzaron a cursar EM. Los que eligieron educación técnica se desvinculan más que los que eligieron secundaria pública, y los que comenzaron a cursar secundaria en el sector privado casi que no presentan desvinculación. Este análisis se retomará en el capítulo 6 para mostrar el análisis multivariado en donde se tiene en cuenta el efecto de cada variable sobre la desvinculación, manteniendo las otras constantes.

Tipos de trayectorias según combinaciones posibles de matriculación (estar o no estar inscripto). Análisis de las 10 trayectorias más frecuentes

A continuación, se listan las 10 trayectorias más frecuentes respecto a la matriculación. En cada uno de los 8 años del panel un estudiante es un 1 si se matriculó o un 0 si no lo hizo²⁶. El *Gráfico 13*, muestra las trayectorias que se configuraron en la cohorte de estudiantes para cada año del panel.

²⁶ Existen $2^8=256$ combinaciones teóricamente posibles de matriculación y no matriculación a lo largo de 8 años.

Gráfico 13. Porcentaje de estudiantes de acuerdo a las diez combinaciones más frecuentes de matriculación/no matriculación para cada año del panel



Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Nota: Cada columna refiere a un año del panel, la primera al 2014 (1) y la última al 2021 (8).

Las filas están numeradas de acuerdo a la posición del ranking de las trayectorias más frecuentes.

En los primeros seis años, el 1 significa estar inscripto y el 0 no estar inscripto, pero en el año siete y ocho, el 1, además de estar inscripto, puede ser haber egresado. Así, el 1, representa una situación de estar en el sistema o haber culminado, y el 0, de no encontrarse matriculado. Lo primero a destacar es que el 85,3% de los estudiantes se encuentran en alguno de estas 10 trayectorias más frecuentes, el restante 14,7% corresponde a las restantes combinaciones posibles entre 1 y 0 (estar inscripto o haber egresado vs no estar inscripto)

La trayectoria más frecuente con un 47,5% de estudiantes que hicieron este camino, implica que estuvieron inscriptos durante todo el período o egresaron.

En segundo y tercer lugar, se encuentren dos situaciones que representan cada una aproximadamente el 9% de las trayectorias: haber estado inscripto los primeros siete años y haberse ido sin egresar al octavo año, o bien, haber estado inscripto los primeros seis años e irse y no volver los siguientes dos años del panel. En 4to y 5to lugar están las trayectorias de aquellos estudiantes que se inscribieron los primeros cinco años o los primeros cuatro años y luego no regresaron. Ambos tipos de trayectorias tienen un peso de 4% cada una. En 6to, 8vo y 9no lugar con una frecuencia de ocurrencia de entre 2% y 3% cada una, están las trayectorias de quienes, se inscribieron dos años y no se inscribieron más, de los que solo se matricularon el primer año y el porcentaje de estudiantes que se matricularon únicamente los tres primeros años.

Al menos hasta el 7mo año del panel, el 56%²⁷ de los estudiantes o bien seguían inscriptos o habían egresado. Al restante 44% - que son los estudiantes que no registraron inscripción en algún año-, le pasan dos situaciones: 6 de cada 10, ya no vuelven al sistema educativo formal, en cambio, los restantes 4, vuelven a registrar inscripción. Esto quiere decir, que cuando no se inscriben, es probable que se terminen desvinculando

El peso que tiene el tipo de trayectorias en donde “se van y no vuelven”, no es menor, al menos el 26%²⁸ de la cohorte, una vez que se van, no se matriculan por 2 años seguidos o más. La situación de una posible re-vinculación es menos probable por cada año que el adolescente no registra

²⁷ Es la suma de 47,5% que corresponde a la trayectoria más frecuente y 9,1% de la 2da

²⁸ El 26% es la sumatoria de la trayectoria 3,4,5,6, 8 y 10.

inscripción (Cardozo, 2016). Inclusive, al menos cerca del 13% de los estudiantes de la cohorte no registra matriculación desde (t+4) o en forma previa, esto implica que se fueron de forma prematura del sistema educativo.

En el ranking de las 10 trayectorias más frecuentes quedan por describir dos lugares más, la que se encuentra en el 10mo lugar, representa al 1,4% de los estudiantes y refiere a una situación muy problemática, son aquellos estudiantes que no registran ningún año de inscripción. Por otro lado, La 7ma es la que luego de un evento de no inscripción, el estudiante vuelve al año inmediato, esto le ocurre al 2,8% de los estudiantes y es la única situación de intermitencia dentro de las 10 trayectorias más frecuentes. Las trayectorias que refieren al resto de combinaciones de matriculación y no matriculación representan al 14,7% de la cohorte, en su gran mayoría, son situaciones de intermitencia entre periodos de matriculación y no matriculación.

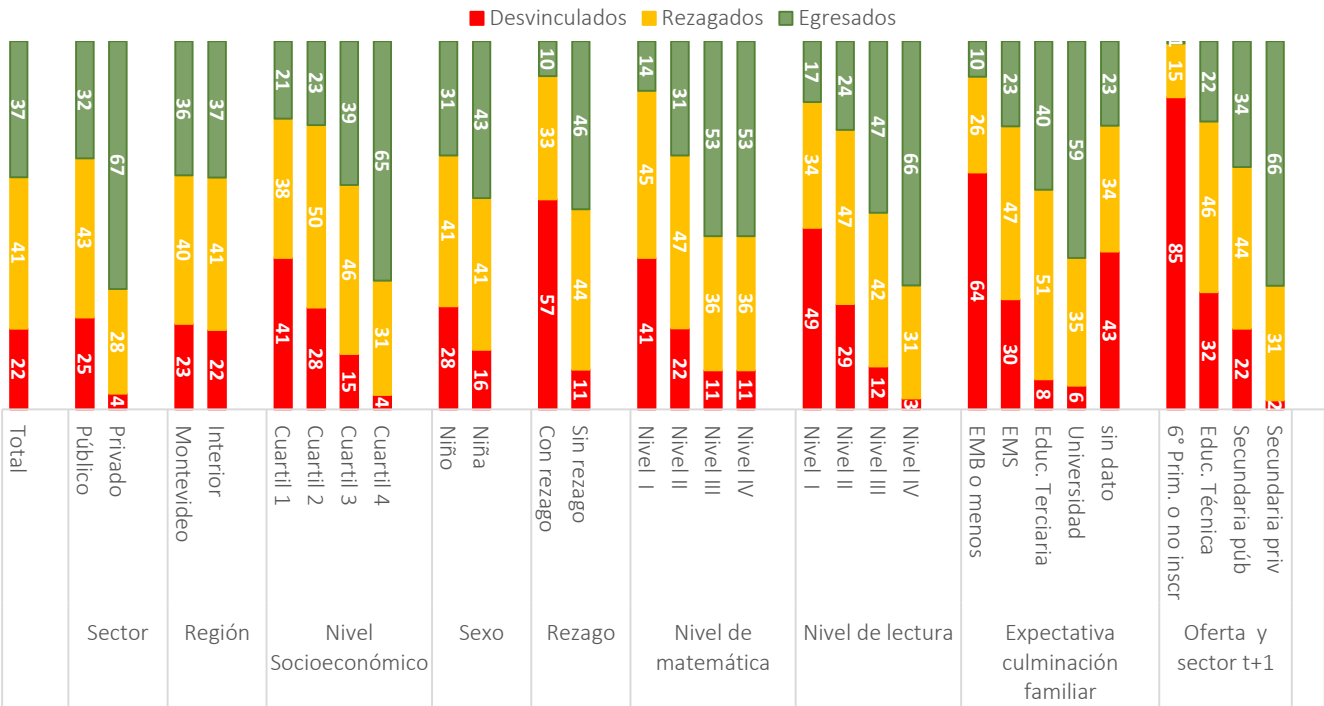
5.2 Configuración de las trayectorias al final del período de observación

Al finalizar la ventana de observación que inicia en marzo 2013 (t) y finaliza en marzo2021 (t+8), hay tres posibilidades para categorizar el desenlace de la trayectoria de los estudiantes:

- **Egresados:** estudiantes egresados en tiempo, marzo de 2020 (t+7) o un año más tarde.
- **Rezagados:** estudiantes matriculados en al menos uno de los tres años finales del panel, en marzo de 2019 (t+6), marzo del 2020 (t+7) y/o marzo del 2021 (t+8).
- **Desvinculados:** estudiantes que no registran inscripción en los tres últimos años del panel, t+6, t+7, t+8.

En el *Gráfico 14* se resume la situación de la cohorte que se viene estudiando a la luz de la tipología presentada.

Gráfico 14. Porcentaje de estudiantes según tipología de trayectoria al final del período de observación, para cada variable de estudio



Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

En la *Tabla A. 8* del anexo se presenta para cada variable de interés, el porcentaje de egresados, rezagados y desvinculados. Además, el ratio entre egresados y desvinculados.

Para el total de la cohorte en marzo de t+8, último momento en que se registran datos de cada individuo, se tiene que el 37% egresó, el 41% se encontraba rezagado²⁹ y el 22% se desvinculó³⁰.

En general se mantienen las desigualdades planteadas anteriormente, a modo de ejemplo: por cada egresado del cuartil 1 se tiene a dos desvinculados. En cambio, por cada 16 egresados del cuartil superior hay 1 solo desvinculado. La relación entre un varón egresado y otro desvinculado es casi de 1 a 1, en cambio, por cada niña desvinculada hay 2,6 egresadas.

En el capítulo siguiente se retoman las variables de interés en el marco de un análisis multivariado. Se busca explicar los factores asociados a los tres resultados mencionados que se configura en t+8 (2021). Se usarán las variables de interés que caracterizan a los estudiantes de 6° de EP en el momento t, (2013), es decir, 8 años antes.

²⁹ Refiere a los matriculados en al menos un año de los tres últimos del panel. (2019 -2021), es decir, a marzo de 2021 no habían egresado ni tampoco se habían desvinculado.

³⁰ Refiere a los que no registran matriculación en los tres últimos años del panel.

6. Determinantes de las trayectorias educativas

En este capítulo se analiza el efecto de cada variable de interés analizada anteriormente, sobre las trayectorias educativas. Concretamente, se analiza el efecto de cada variable sobre la probabilidad de que un estudiante siga una trayectoria de egreso (categoría de referencia) vs. a) una trayectoria de rezago y b) una trayectoria de desvinculación. En la *Tabla A. 9* del anexo se muestran en detalle las estimaciones.

La regresión logística multinomial se utiliza para predecir las probabilidades de múltiples resultados categóricos, es una extensión del modelo logístico binario con $M > 2$ categorías. En este estudio $M=3$ y representa a cada uno de los resultados de la trayectoria: Egreso, Rezago, Desvinculación. La regresión consiste el ajuste de $M-1$ (2) ecuaciones, una ecuación para por cada categoría de la variable dependiente a lo que se resta la categoría de referencia, que es la trayectoria de egreso. De esa forma, se estima el cambio en los riesgos relativos de pertenecer cada una de las categorías de la variable de resultado, en función de un conjunto de variables explicativas.

Variables explicativas:

- X_1 = Región del centro en 6° EP. (Montevideo/ Interior)
- X_2 = Nivel socioeconómico de la familia del estudiante en 6° EP. (Cuartil 1 a 4)
- X_3 = Sexo. (Varón/ Niña)
- X_4 = Rezago en 6° EP (Sin rezago/ Con rezago)
- X_5 = Nivel de matemática en 6° EP. (Nivel 1 a 4 de la prueba TERCE)
- X_6 = Expectativa familiar en 6° EP. (EMB o menos, EMS, educación terciaria, universidad)
- X_7 = Oferta/sector en 1° EM. (Secundaria pública, secundaria privada, educación terciaria)

El ajuste multinomial viene dado por³¹:

Para $Y=2$ (rezago) en comparación con $Y=1$ (egreso):

$$\log \frac{P(Y=1)}{P(Y=2)} = \beta_{2,0} + \beta_{2,1}X_1 + \beta_{2,2}X_2 + \beta_{2,3}X_3 + \beta_{2,4}X_4 + \beta_{2,5}X_5 + \beta_{2,6}X_6 + \beta_{2,7}X_7^{32}$$

Para $Y=3$ (desvinculación) en comparación con $Y=1$ (egreso):

$$\log \frac{P(Y=1)}{P(Y=3)} = \beta_{3,0} + \beta_{3,1}X_1 + \beta_{3,2}X_2 + \beta_{3,3}X_3 + \beta_{3,4}X_4 + \beta_{3,5}X_5 + \beta_{3,6}X_6 + \beta_{3,7}X_7$$

$\beta_{2,0}$ y $\beta_{3,0}$ son los interceptos de las ecuaciones para $Y=2$ y $Y=3$ respectivamente.

$\beta_{2,j}$ y $\beta_{3,j}$ (para $j=1$ a 7) son los coeficientes correspondientes a las variables explicativas X_1 a X_7 para las categorías $Y=2$ y $Y=3$.

³¹ La bibliografía consultada para escribir las ecuaciones se basó en (Rabe-Hesketh y Skrondal, 2022).

³² A diferencia de los modelos de regresión ordinaria, los modelos de verosimilitud no necesitan incluir un término de error explícito en la ecuación. Las desviaciones entre los valores observados y los predichos se incorporan en la función de verosimilitud.

La probabilidad estimada de cada resultado condicionado a X , viene dada por:

$$P(Y=1|X) = \frac{1}{1+e^{\eta_2}+e^{\eta_3}}$$

$$P(Y=2|X) = \frac{e^{\eta_2}}{1+e^{\eta_2}+e^{\eta_3}}$$

$$P(Y=3|X) = \frac{e^{\eta_3}}{1+e^{\eta_2}+e^{\eta_3}}$$

Donde:

$$\eta_2 = \beta_{2,0} + \beta_{2,1}X_1 + \beta_{2,2}X_2 + \dots + \beta_{2,7}X_7$$

$$\eta_3 = \beta_{3,0} + \beta_{3,1}X_1 + \beta_{3,2}X_2 + \dots + \beta_{3,7}X_7$$

A continuación, en la *Tabla 2*, se resumen los resultados del análisis multivariado, mediante un modelo logístico multinomial.

Tabla 2. Efecto de cada variable sobre la probabilidad de realizar una trayectoria de rezago o de desvinculación, en comparación con una trayectoria de egreso, a los 8 años de haber cursado 6° de EP.

		Trayectorias en t+8. Egreso vs ...			
		Rezago		Desvinculación	
		Coef.	Sign.	Coef.	Sign.
Región en t	Montevideo (vs. Interior)	0,619	***	0,805	**
Nivel socioeconómico en t	Cuartil 2 (vs.cuartil1)	0,381		- 0,167	
	Cuartil 3 (vs.cuartil1)	- 0,064		- 0,869	**
	Cuartil 4 (vs.cuartil1)	- 0,633	**	- 1,966	***
Sexo en t	Niña (vs. Varón)	- 0,529	***	- 1,403	***
Rezago en t	Con rezago (vs. Sin rezago)	0,828	***	2,292	***
Nivel de matemática en t	Nivel II (vs. Nivel 1)	- 0,650	***	- 1,032	***
	Nivel III (vs. Nivel 1)	- 1,323	***	- 1,777	***
	Nivel IV (vs. Nivel 1)	- 1,758	***	- 2,285	***
Expectativas de culminación familiar en t	EMS (vs. EMB o menos)	0,006		- 1,350	**
	Educ. Terciaria (vs. EMB o menos)	- 0,146		- 2,222	***
	Educ. Universitaria (vs. EMB o menos)	- 0,470		- 1,898	***
Oferta y sector en t + 1	Secundaria priv (vs. Secundaria púb)	- 0,445	*	- 1,484	***
	Educ. Técnica (vs. Secundaria púb)	0,107		0,004	
	Repitió primaria o no se inscribió en EM (vs. Secundaria púb)	1,852	***	3,037	***

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

*** = Significación al 99%; ** = Significación al 95%; * = Significación al 90%

Para la mejor lectura de los resultados, se agrupan según el siguiente orden que surge de la significación estadística de las variables.

- 1) Características sociodemográficas: nivel socioeconómico, región y sexo.
- 2) Desempeño en educación primaria: nivel de competencia en matemática y situación de rezago.
- 3) Oferta educativa: oferta y sector institucional al inicio de la EM.
- 4) Expectativas de la familia sobre la trayectoria del niño.

Nivel socioeconómico: Al igual que lo sugerían los análisis descriptivos, las trayectorias en la EM están estrechamente vinculadas al NSE, una vez que se controla por el resto de las variables. Según los resultados del modelo, los estudiantes del cuartil 4 tienen riesgos significativamente menores de rezagarse en la trayectoria por la EM y aún riesgos más bajos de desvincularse del sistema educativo. Es decir, tienen una probabilidad significativamente mayor de egresar. No hay

diferencias significativas entre los tres cuartiles más bajos respecto a las chances relativas de rezagarse vs. la de seguir una trayectoria de egreso. Respecto a la desvinculación, los estudiantes del cuartil 1 (Q1) muestran riesgos mayores que los del cuartil 3, pero no hay diferencias significativas entre el Q1 y el Q2.

Región y sexo. A igualdad en el resto de los factores explicativos, los riesgos de rezago como los de desvinculación en la EM son mayores para los alumnos que cursaron 6° de EP en centros de Montevideo en comparación con los del interior³³. Lo mismo ocurre según el sexo, las niñas tienen menos chances de rezagarse y desvincularse que los varones.

Nivel de competencia en matemática. Asimismo, el análisis muestra, una fuerte correlación entre las competencias académicas al final de la educación primaria y la trayectoria posterior. Rezagarse y desvincularse en la EM es significativamente menos probable entre los alumnos de nivel 2 en la prueba de matemática de TERCE en comparación con los de nivel 1, y es aún menos probable para los de nivel 3 y 4. Esto re-afirma los resultados del análisis descriptivo del capítulo anterior. Sugiere, además, la existencia de cierto nivel de competencias mínimas a la hora de iniciar la EM que son condición necesaria, para luego, egresar en tiempo o con un año de rezago. Como se muestra en la *Tabla A. 10* del anexo, se llega a resultados similares empleando la regresión multinomial utilizando como variable explicativa el nivel de lectura en vez del nivel de matemática. Para el caso del rezago, al usar el nivel de lectura se tiene que solamente para el nivel 4 las diferencias son significativas respecto al nivel 1. Para la desvinculación, las diferencias son significativas para el nivel 3 y 4.

Rezago. Otra variable que da cuenta de cómo le fue al niño en la EP en cuanto a sus aprendizajes, es si presenta o no rezago en 6°. El análisis descriptivo no permitía aislar el efecto de las condiciones socioeconómicas del hogar, el desempeño académico ni controlar por el resto de las variables del modelo, pero una vez hecha la regresión, es decir, considerando el efecto neto del rezago en primaria, los resultados indican que haber repetido en primaria, predice fuertemente, el posterior rezago en la EM y predice aún mejor, la desvinculación.

Oferta/sector en el primer año de la EM. Existen riesgos significativamente menores de rezago³⁴ y de desvinculación³⁵ para quienes cursaron el primer año de EM en centros educativos privados en comparación con estudiantes de secundaria pública. Los que estaban cursando primaria en t+1 o no se habían inscripto, tienen, tal como era previsible, más riesgo de rezagarse y de desvincularse. En tanto, no se registran diferencias significativas entre los estudiantes que iniciaron su trayectoria en EM en liceos públicos, vs los que iniciaron la EM en escuelas técnicas.

Expectativas familiares: Finalmente, el análisis multivariado confirma lo que se describió en el análisis descriptivo y sugiere que las expectativas de las familias sobre el nivel educativo que alcanzarán sus hijos tienen un peso relevante en las trayectorias de EM. Según las estimaciones, la probabilidad de desvinculación aumenta a medida que las expectativas de los padres se vuelven más bajas, es decir, son inversamente proporcionales. Si la familia piensa que el hijo terminará la EMS, entonces ese niño correrá menos chances de desvincularse de la EM en comparación a los niños cuyos padres piensan que solamente terminarán la EMB. Nada se puede decir en este estudio

³³ Las diferencias según región no se apreciaban en el análisis bivariado presentado en el capítulo 4

³⁴ Al 90% de significación

³⁵ Al 95% de significación

sobre los mecanismos que están operando detrás de esta asociación, sin embargo, se podría ensayar alguna hipótesis. Por ejemplo, una crítica podría plantear que las expectativas son adaptativas, las familias en 6° de EP ya se hicieron una idea acerca del nivel académico y de los gustos por el estudio que tienen sus hijos. Por ello, las expectativas podrían estar correlacionadas con lo que efectivamente sucede años más tarde en la EM. Otra hipótesis sería que, a mayores expectativas, también serán mayores los apoyos e incentivos para que el niño avance por los niveles educativos. No obstante, este trabajo confirma que existe una correlación elevada entre las expectativas familiares sobre el desempeño de los niños y los logros que efectivamente alcanzan.

7. Modelo predictivo sobre la desvinculación como un instrumento que contribuye a un sistema de alertas tempranas

Este capítulo busca responder en qué medida es útil la información que se tiene del estudiante hasta 6° de EP, para predecir una trayectoria cuyo desenlace es la desvinculación de la EM.

Si se muestra que, empleando información del último grado de EP (de relativamente fácil disponibilidad para el sistema educativo), se logra predecir la desvinculación posterior con altos niveles de sensibilidad, precisión y especificidad, entonces existen buenas condiciones para desarrollar un sistema de alertas tempranas. Dicho SAT permitiría diseñar acciones específicas para proteger las trayectorias de los estudiantes que inician la Educación Media con altos riesgos de abandonar los estudios.

El modelo logístico explicado en el apartado 3.5 se utiliza para estimar los riesgos de desvinculación de cada estudiante durante la EM y, en función de estos riesgos, clasificarlos como desvinculados o no desvinculados potenciales. Esta predicción se hace únicamente a partir de información disponible hasta el término de la educación primaria (t) y utilizando la oferta/sector en la que se inscribió el estudiante en t+1. Luego de etiquetar a cada estudiante se contrastará la clasificación hecha con lo que efectivamente le ocurrió a cada uno. Esto permitirá valorar qué tan útiles son las predicciones realizadas para contribuir a un SAT.

Primero se estima el modelo Logit 1 (de aquí en adelante L1) presentado en la *Tabla A. 11*, muy similar al modelo explicativo Mlogit del *capítulo 6*, difiere únicamente en que la variable de resultado se vuelve dicotómica: desvinculados y no desvinculados. Estos últimos comprenden a los rezagados y egresados de EM. Las variables explicativas son las mismas que el Mlogit.

En segundo lugar, se estima el modelo Logit 2 (de aquí en adelante L2) presentado en la

Tabla 3, El modelo L2 es una especificación alternativa al L1, en donde se cambia la forma de incluir algunas variables independientes³⁶. El NSE y el puntaje de matemática en la prueba TERCE, son incluidas como variables continuas para obtener mayor precisión sobre la posición del estudiante y su familia. Además, se incorpora el puntaje de lectura en la prueba TERCE³⁷, también como variable continua. Las restantes variables permanecen incambiadas: región, sexo, expectativas familiares y oferta/sector.

³⁶ Esto permite un mejor rendimiento del modelo, tanto globalmente como localmente, es decir, en el entorno de los umbrales que maximizan la sensibilidad y precisión.

³⁷ Como muestra la *Tabla A. 3* del anexo, la correlación entre el puntaje en matemática y lectura es de 0,6 por tanto no se incluyó en el modelo Mlogit en el cual, el objetivo principal fue la interpretación del modelo. En cambio, en el modelo Logit 2, se incluyeron ambos puntajes porque el modelo tiene fines predictivos, en estos casos no es tan necesario el cuidado de la correlación entre las variables. El modelo logit es capaz de manejar ciertas multicolinealidades sin que estas afecten significativamente la capacidad predictiva del modelo

Los modelos L1 y L2 corresponden a una regresión logística con la siguiente especificación:

$$(1) \ln \left(\frac{y=1}{y=0} \right) = G(\beta_0 + X\beta) = G(Z)$$

$$(2) \text{riesgo} = r = \frac{e^Z}{1 + e^Z}$$

y es una variable binaria que vale 1 si el alumno experimentó el evento (desvinculación) y 0 en caso contrario (rezagado o egresado). X es el conjunto de variables predictivas que varía entre el modelo L1 y L2.

En ambas especificaciones, se entrenó al modelo empleando todas las observaciones menos una que se excluyó para utilizarla como test. En cada iteración se saca la observación que se quiere predecir, ajustando el modelo con las restantes observaciones. A esta técnica se le llama *leave-one-out cross validation*. La probabilidad estimada en toda la muestra se representa de la siguiente manera:

$$r = \hat{p}(x_1, x_2, \dots, x_k), \{0 \leq r \leq 1\}$$

Y es la variable de resultado, $(x_1, x_2, \dots, x_k)$ representa la realización de las variables predictivas, y r es la probabilidad estimada de desvinculación. La *Tabla 3* presenta las estimaciones del L2, que es el modelo elegido para hacer las predicciones. Por su parte, en la *Tabla A. 11* del anexo se muestran las estimaciones del L1.

Tabla 3. Regresión Logística sobre la probabilidad de desvinculación al cabo de 8 años en la EM para la Cohorte de 6° de EP en 2013. Logit 2.

Logistic regression		Number of obs	=	4,286		
		Wald chi2(32)	=	232,64		
		Prob > chi2	=	0,000		
Log pseudolikelihood = -18768,534		Pseudo R2	=	0,3475		
Trayectoria en t + 8 dicotómica: desvinculación / no desvinculación	Coefficient	Robust std. err.	P> z		[95% conf. interval]	
Región t						
Montevideo (vs. Interior)	0,333	0,315	0,290		-0,284	0,949
NSE t (continua)	-0,597	0,151	0,000	***	-0,894	-0,301
Sexo t						
Niña (vs. Varón)	-0,956	0,225	0,000	***	-1,396	-0,515
Rezago t						
Con rezago (vs. Sin rezago)	1,503	0,215	0,000	***	1,082	1,924
Puntaje matemática t (continua)	-0,002	0,001	0,159		-0,005	0,001
Puntaje lectura t (continua)	-0,005	0,001	0,000	***	-0,008	-0,002
Expectativas culminación familiar t						
EMS (vs. EMB)	-1,219	0,413	0,003	***	-2,029	-0,409
Educ. Terciaria (vs. EMB)	-1,964	0,501	0,000	***	-2,945	-0,982
Educ. Universitaria (vs. EMB)	-1,360	0,490	0,005	***	-2,320	-0,400
Sin dato (vs. EMB)	-0,567	0,456	0,214		-1,460	0,327
Oferta/sector t+1						
Secundaria privada (vs. Secundaria pub.)	-0,986	0,400	0,014	**	-1,769	-0,203
Educación técnica (vs. Secundaria pub.)	-0,087	0,232	0,706		-0,541	0,367
6° Primaria o no inscripto (vs. Secundaria pub.)	1,365	0,587	0,020	**	0,214	2,516
_Constante	4,483	1,035	0,000	***	2,454	6,513

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

*** = Significación al 99%; ** = Significación al 95%; * = Significación al 90%

Primero, se estima la probabilidad (r) de que cada niño experimente en el futuro un evento educativo considerado problemático en función de las variables x , en este caso, la probabilidad de que un alumno de 6° de EP se desvincule de la EM. La desvinculación implica que el estudiante no se haya inscripto en t+6 ni en ninguno de los dos años siguientes del panel (t+7 y t+8), en el mejor de los casos, la trayectoria de un estudiante que se desvincula, se compuso solamente de 5 años dentro de la EM.

Seguidamente, y para poder hacer la clasificación, es necesario determinar un umbral (u), a partir del cual, la probabilidad a priori de que ocurra un desenlace de desvinculación, sea considerada lo suficientemente alta como para realmente encender una alerta. Este umbral convierte los valores estimados en 0 o 1. Hay varios criterios para definir u , en la aplicación práctica en el marco de una

política pública, cabe la posibilidad de elegir el umbral de acuerdo a los recursos con los que se cuenta y los objetivos que se persiguen.

Aquí se presentará para varios valores de u , el rendimiento del modelo en cuanto a sus parámetros de calidad, y qué es lo que se está priorizando a la hora de elegir uno u otro umbral.

Lo interesante del método descrito es que puede ser aplicado según el problema que se quiera abordar desde la política pública. Por ejemplo, se podría hacer una estimación para identificar a aquellos alumnos que, al cabo de cierta cantidad de años, no lograron progresar hasta cierto grado educativo. En esta investigación se utiliza el método para abordar la desvinculación.

Un SAT para el caso de la educación debería priorizar predicciones sensibles (S), además de específicas (E) y precisas (P). Lo ideal sería poder identificar tempranamente a los estudiantes que requieren algún tipo de intervención para poder cambiar una posible trayectoria de desvinculación y lo importante es que la menor cantidad de estudiantes pasen por debajo del radar (sensibilidad). Al mismo tiempo, es importante minimizar las "falsas alarmas" sobre alumnos que, a pesar de las predicciones, progresan adecuadamente durante su escolarización. Esto se logra mediante predicciones precisas y específicas.

No hay consenso sobre los parámetros exactos para definir un modelo de clasificación como deficiente, aceptable, bueno o excelente. Sin embargo, en la bibliografía existen directrices generales que ayudan a contextualizar los resultados obtenidos en una aplicación específica.

En primer lugar, para pruebas de identificación de factores de riesgo en el ámbito educativo, algunos autores clasifican un buen modelo cuando presenta una sensibilidad y especificidad de al menos 0,7. (Catts et al., 2009; Jenkins et al., 2007; Kilgus et al., 2014; Jojo et al., 2022).

En segundo lugar, como se mencionó en el apartado 3.5, el valor de AUC que muestra el área bajo la curva de ROC, informa sobre el rendimiento global del modelo en términos de sensibilidad y especificidad. AUC, es una medida resumen de la sensibilidad y especificidad de todo el modelo independientemente de la combinación entre (S) y (E) que surja del umbral elegido. En la literatura sobre el tema suele aceptarse que un valor de AUC por debajo de 0,7 indica un rendimiento malo, entre 0,7 y 0,8 aceptable o bueno, entre 0,8 y 0,9 muy bueno y superior a 0,9 excelente. (Compton et al., 2006; Cerda et al., 2012; Martínez et al., 2023)

En tercer lugar, están los indicadores F1 y F2 también mencionados en el apartado 3.5, ambas son medidas que toman en cuenta la sensibilidad y la precisión. El F1 es una media armónica entre la precisión y la sensibilidad, ofrece una visión equilibrada del rendimiento del modelo especialmente cuando hay desbalance en las clases. El F2 también es una media armónica, pero le otorga el doble de peso a la sensibilidad. Siguiendo estos dos indicadores, el umbral que maximiza el valor de F1 (o, alternativamente, a F2) es el que logra maximizar de forma conjunta la (S) y la (P) de la predicción.

La *Tabla 4*, presenta las métricas para evaluar el rendimiento del modelo según el umbral elegido. Las filas marcadas en negrita muestran las dos combinaciones que maximizan el criterio de puntaje F1 y F2 y determinan un umbral de riesgo de 0,302 y 0,18 respectivamente. De aquí en más, se trabaja con el modelo Logit 2 porque para cualquier umbral, el rendimiento del modelo es mejor que el Logit 1.

Tabla 4. Métricas del rendimiento del modelo Logit2 (L2), según el umbral de probabilidad elegido para clasificar a los desvinculados de los no desvinculados.

Umbral	Exactitud	Precisión	Sensibilidad	Especificidad	Puntaje F1	Puntaje F2	Kappa
0,050	0,540	0,320	0,969	0,420	0,481	0,689	0,411
0,100	0,681	0,399	0,898	0,619	0,552	0,718	0,591
0,150	0,737	0,447	0,825	0,713	0,580	0,706	0,664
0,170	0,760	0,474	0,821	0,744	0,601	0,716	0,693
0,180	0,774	0,491	0,818	0,761	0,613	0,722	0,710
0,190	0,777	0,495	0,792	0,773	0,609	0,707	0,715
0,200	0,781	0,501	0,782	0,781	0,610	0,703	0,719
0,250	0,810	0,549	0,761	0,824	0,638	0,706	0,757
0,300	0,838	0,608	0,734	0,867	0,665	0,705	0,792
0,302	0,839	0,611	0,734	0,869	0,667	0,706	0,794
0,305	0,838	0,609	0,726	0,869	0,663	0,699	0,792
0,310	0,838	0,611	0,726	0,870	0,663	0,700	0,793
0,400	0,843	0,654	0,605	0,910	0,628	0,614	0,799
0,450	0,852	0,702	0,568	0,932	0,628	0,591	0,811
0,500	0,852	0,718	0,540	0,940	0,616	0,568	0,811
0,550	0,853	0,751	0,491	0,954	0,594	0,528	0,811

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

En la

Tabla A. 2 del anexo se presenta la Curva de ROC, cuyo valor de AUC es 0,827, indicando un muy buen rendimiento del modelo.

El primer criterio para decidir el umbral (0,302) fue el puntaje F1 que alcanza su máximo valor en 0,667. El rango entre 0,6 y 0,8 indica un buen rendimiento, muestra que el modelo está alcanzando un equilibrio razonable entre precisión (0,611) y sensibilidad (0,734) y a su vez, alcanza una especificidad (0,869). De todas formas, existen posibilidades de mejoras. Por otro lado se presenta Kappa³⁸, con un valor cercano a 1, indicando que la relación entre las clasificaciones observadas y las predicciones del modelo son mejores que la resultante del azar (Landis,J y Koch.G ,1977)

El segundo criterio fue el puntaje F2 que, en la maximización, le otorga el doble de peso a la sensibilidad respecto a la precisión, en este escenario se tiene un umbral de 0,18. A medida que se baja el umbral se estará captando a más desvinculados, por ello la sensibilidad es de 0,818, a costo de perder precisión (0,491). Esta precisión, en términos de la aplicación práctica a un SAT, podría suponer un número excesivamente alto de falsos positivos, y, en ese sentido, implicaría destinar recursos, siempre escasos, a estudiantes que posiblemente no lo requirieran. Por último, el valor Kappa sigue siendo aceptable para el rendimiento del modelo en este nuevo umbral más bajo.

A modo de realizar un ejercicio, La *Tabla 5* ejemplifica con la cantidad de niños que se estaría trabajando, si se hacen las predicciones para los aproximadamente 52.500³⁹ estudiantes que integraron la cohorte 2013 de 6° de EP.

³⁸ Para la definición de este parámetro ver [apartado 3.5](#)

³⁹ Según datos del anuario estadístico del MEC 2013, la cohorte de 6° estuvo conformada por 52.553 estudiantes, 44.273 de educación pública y 8.280 de educación privada.

Tabla 5. Rendimiento del modelo (L2) para un umbral de 0,180 y de 0,302 en la cohorte 2013 de estudiantes de 6° de Educación Primaria

Umbral		Cohorte de 6° de E.P. 2013	Desvinculados en EM	Predicción de desvinculados	Exactitud	Precisión	Sensibilidad	Especificidad	Falsos positivos (falsas alarmas)	Falsos negativos (debajo del radar)
0,180	%	100%	21,9%	37,8%	77,4%	49,1%	81,8%	76,1%	23,9%	18,2%
	N	52.500	11.522	19.219	40.618	9.429	9.429	31.188	9.790	2.093
0,302	%	100%	21,9%	26,4%	83,9%	61,1%	73,4%	86,9%	13,1%	26,6%
	N	52.500	11.522	13.834	44.055	8.456	8.456	35.600	5.378	3.067

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

De los 52.500 alumnos de 6° de EP en 2013, el 21,9% (11.522) se desvincularon. Para un umbral de 0,302, la sensibilidad obtenida indica que, de esos estudiantes, el modelo logra predecir correctamente al 73,4% (8.456/11.522) por lo tanto, pasan por debajo del radar un 26,6% de los estudiantes que luego se terminan desvinculando (3.067). La especificidad expresa que del total de estudiantes que no se desvincularon, se pueden identificar al 86,9% (35.600) y, por tanto, se generan un 13,1% (5.378) falsas alarmas. Tanto las falsas alarmas como los que pasan por debajo del radar, probablemente sean estudiantes que en su gran mayoría estén cerca del umbral que los clasifica como 1 o 0 (desvinculado, no desvinculado). Es decir, se va a estar apoyando a niños que de todas formas tienen una posible trayectoria vulnerable, no se desvinculan, pero se rezagan. En cuanto a los que no se logra identificar como desvinculados, probablemente, serían aquellos que están cerca de ser alumnos que no se desvincularían, su comportamiento es similar a los que se rezagan. En resumen, la sensibilidad y la especificidad presentan valores que informan un buen rendimiento del modelo, ambas superan al 0,7.

Por su parte, la precisión es 61,1% (8.456/13.834). La precisión es la razón entre los que el modelo acertó que se iban a desvincular (8.456) y el total de los individuos que el modelo predijo que se desvincularían (13.834). Por último, la exactitud es el porcentaje de aciertos totales, la misma es de 83,9% (44.055/52.500).

Para un umbral de 0,18, el modelo logra identificar a una cantidad mayor de estudiantes que se desvincularon, obteniendo una sensibilidad de 81,8% (9.429/11.522). Quedan sin detectar, el 18,2% de estudiantes (2.093/11.522). Como es de esperar, la especificidad se reduce, de los que no se desvincularon, se acertó al 76,1% (31.188/40.978⁴⁰). Por tanto, se generan casi el doble de falsas alarmas, 13,1% (9.790) con un umbral de 0,18 en comparación con un 23,9% (5.378) de falsas alarmas con un umbral de 0,302.

A su vez, la precisión, como es de esperar, también disminuye. De los 19.219 estudiantes que el modelo predijo que se iban a desvincular, solamente a 9.429 efectivamente se desvincularon, 49,1% (9.429/19.219). Por último, la exactitud es de 77,4% (40.618/52.500). En resumen, la elección de este escenario que maximiza el valor de F2, valora mucho más, ser sensible que ser preciso. Es decir, prioriza identificar al mayor porcentaje de estudiantes que se desvincularán y no tanto, el

⁴⁰ Total de la cohorte - cantidad de estudiantes que se desvincularon. 52.500 - 11.522 = 40.978

hecho de que se está siendo ineficiente por darle apoyos a quienes de todas formas no se desvincularán.

Como ya se mencionó, es probable que por más que no se desvinculen igualmente tengan trayectorias que se apartan de las normativas, estudiantes que presenten rezago en EM o que inclusive si la ventana de observación continuara podríamos observar que se desvinculan en los siguientes años. Hay una tensión entre detectar a la mayor cantidad de estudiantes vs no darle apoyos a quien no lo precisaría, de hecho, en este escenario la cantidad de falsas alarmas (9.790) es mayor a la cantidad de alarmas correctamente encendidas (9.429), motivo por el que sería conveniente usar F1 para definir el umbral

Una vez presentados los valores que toman las métricas para evaluar el rendimiento del modelo, más allá de que en el ámbito de la política pública, cada decisor podrá elegir el umbral, este estudio sugiere una cierta combinación de P, S y E que varían entre el umbral determinado por F2 (0,18) y F1 (0,302). Si se tuviera que elegir un solo umbral se entiende más pertinente elegir el hallado según F1, en donde la cantidad de niños a intervenir es menor (13.834 vs 19.219). Uno de los motivos es porque las intervenciones a realizar deben ser potentes, tener un impacto significativo, es decir, intervenciones que requieran gran cantidad de horas, de recursos, que se implementen correctamente, etc. Se trata de estudiantes que se desvincularán, su situación al iniciar la EM ya se encuentra vulnerada. Cubrir menos cantidad de estudiantes permitirá un shock de política mayor vs tomar el doble de niños contando con los mismos recursos. De todos modos, en un escenario potencial de aplicación de un instrumento como el sugerido, no existirían, en principio, restricciones para actualizar las predicciones del modelo en virtud de evidencia nueva, por ejemplo, sobre los resultados de los alumnos en $t+1$. Esta actualización requeriría de políticas de monitoreo continuas y de carácter universal.

El mismo análisis se realizó para evaluar el rendimiento el modelo (L1), las métricas se presentan en la *Tabla A. 12* del anexo. Con el criterio del puntaje F1 para elegir el umbral, se tiene una sensibilidad de 0,656, una especificidad de 0,887, una precisión de 0,621, una exactitud de 0,837 y un valor de kappa de 0,791. El valor de AUC es de 0,823, en la *Ilustración A. 3* del anexo se presenta la Curva de ROC correspondiente. A nivel global no es un mal modelo, pero localmente, cuando se requiere determinar un umbral siguiendo la maximización de sensibilidad y precisión de F1, el modelo no alcanza 0,7 de (S) ni de (P). La idea de presentar las métricas para el (L1) es que inclusive, partiendo de un modelo simple de interpretar, para el cual previamente se realizó un análisis del efecto de cada variable sobre las trayectorias, aun así, el modelo está muy cerca de cumplir con buenos estándares de rendimiento. Se optó por realizar las predicciones con el modelo L2 que es una versión mejorada del L1. Al incluir el puntaje de lectura y principalmente al usar el NSE y el puntaje de matemática como variables continuas, el modelo pasó a cumplir con los estándares de calidad.

En un sistema de alertas tempranas, luego de tener identificados a los estudiantes que posiblemente se desvincularán, llega el momento de diseñar intervenciones para intentar que los alumnos no se desvinculen. Se volcarán recursos, apoyos o lo que se entienda pertinente para esos niños y sus familias. A la hora de pensar las políticas para apoyar las trayectorias de estudiantes en riesgo y ver sobre qué factores intervenir, este modelo arroja algunas señales. Existen técnicas de predicción que se comportan como una caja negra en donde no se sabe cuáles son los canales que

están operando tras la desvinculación. En cambio, una ventaja del modelo presentado es que nos podemos guiar en el mismo modelo para diseñar los apoyos.

Dicho esto, sería pertinente hacer foco en los aprendizajes de lectura y matemática, principalmente de ésta última competencia, porque en este estudio, el desempeño en matemática resultó ser mejor predictor de la desvinculación, que los desempeños en lectura. Asimismo, focalizarse en la población socioeconómicamente más vulnerable y trabajar sobre las expectativas que se formulan las familias sobre el nivel educativo que alcanzarán sus hijos. Por último, otro factor relevante que se desprende de las predicciones sobre desvinculación, es trabajar sobre el rezago acumulado en educación primaria, es decir, sobre la repetición^{41 42}.

Oportunidad usar registros administrativos y pruebas de aprendizaje nacionales.

Tal como se reporta en Cardozo et al. (2023), Uruguay cuenta con la información, las capacidades técnicas e institucionales y los sistemas de información para poder construir un SAT. Los sistemas de información, han ido evolucionando, (GURÍ en EP y Portafolio en EM). A su vez, se han desarrollado instrumentos para el seguimiento y acompañamiento de las trayectorias educativas: el Registro Único del Alumno, GURÍ-familia para la comunicación entre docentes y la familia de los estudiantes, acompañamientos en el marco del Sistema de Protección de Trayectorias Educativas, entre otros instrumentos. Otro aspecto a destacar es que, en los últimos años, el sistema educativo privado se ha integrado a los sistemas de información de la ANEP, aportando datos individualizados por alumno.

La información necesaria para construir modelos predictivos sobre los fenómenos de interés en la EM, existe de manera universal sin la necesidad de depender de una muestra de estudiantes. La excepción, son las evaluaciones de aprendizaje, porque ni las pruebas de UNESCO (SERCE, TERCE o ERCE), las de PISA, ni las pruebas de Aristas (INEEd) son censales ni anuales. Sin embargo, la ANEP aplica las evaluaciones adaptativas (SEA, SEA+ y LEO) para distintas áreas del conocimiento, que podrían universalizarse. Asimismo, brindan información no solo sobre el nivel de competencias en el último grado de la EP, sino, sobre la trayectoria de aprendizajes desde 3°. El uso de estas pruebas podría mejorar las predicciones. En el caso de que no se logre la universalización de las pruebas mencionadas, se puede recurrir a las calificaciones que otorgan los docentes a sus alumnos. Se ha demostrado que hay una correlación importante entre las calificaciones y los desempeños de los estudiantes en pruebas de aprendizaje (Rimfeld et al., 2019; Zakharov, A. y Carnoy, M.,2015; Westphal, A., Vock, M., y Kretschmann, J.,2021).

En definitiva, el ejercicio predictivo presentado en este capítulo, es factible de universalizar para todas las cohortes de estudiantes. Cada variable del modelo sería reemplazada por otra que surge de los sistemas de información con los que ya cuenta la ANEP y otros organismos del Estado sobre cada niño y sus familias. Inclusive se podrían agregar otras como el nivel de asistencias o la conducta del niño en EP.

⁴¹ Hay estudios que discuten sobre la repetición como herramienta pedagógica (Habiaga, V. 2019; Hattie, J. 2008; Filgueira, Rodríguez, y Fuentes, 2006)

⁴² En el 2023 se comenzó aplicar un nuevo reglamento que solo permite la repetición en 2°, 4° y 6°.

8. Síntesis y conclusiones

La preocupación central de esta tesis fue aportar insumos para abordar el problema de la desvinculación y de los bajos egresos en Educación Media. El enfoque analítico y metodológico desarrollado permitió observar estos problemas desde un enfoque de trayectorias académicas que se configuran a lo largo del tiempo. Este enfoque ha sido muy útil para explorar las pautas y ritmos de los vínculos que establecen los jóvenes con el sistema educativo formal a lo largo de la Educación Media. El valor del seguimiento de trayectorias reside en la posibilidad de inferir conclusiones que surgen del seguimiento de una cohorte de alumnos y es especialmente relevante para aquellas nuevas cohortes que ingresarán a la Educación Media.

Tal como fue formulado en la hipótesis 1: la estilización de las trayectorias en el sistema educativo uruguayo dista de tener un carácter normativo. Aproximadamente, de cada diez estudiantes que se encontraban en 2013 cursando 6° de Educación Primaria, al cabo de ocho años: cuatro lograron egresar de la Educación Media, dos se habían desvinculado y cuatro se encontraban rezagados.

Esta investigación, junto a otras, (De Melo y Machado, 2018; ANEP, 2021; Cardozo, 2016), confirma la hipótesis 2: los problemas que se materializan en la Educación Media, como el rezago, las bajas tasa de egreso y la interrupción de los estudios, son el complejo resultado de procesos que, en buena medida, se originan en etapas educativas anteriores.

Una trayectoria cuyo desenlace es la desvinculación de un estudiante en Educación Media, está fuertemente asociada con el rezago escolar, con los desempeños académicos descendidos en Educación Primaria, con el sector institucional (público), con bajas expectativas familiares, con bajo NSE y, con el sexo (masculino) del estudiante. Estos resultados confirman las hipótesis 3 y 4, y a su vez, están en línea con lo reportado en De Melo y Machado (2018), ANEP (2021) y en los estudios de PISA-L (Boado y Fernández, 2010; Fernández et al, 2013; Cardozo, 2016). Los factores mencionados son relevantes a la hora de diseñar instrumentos de política pública que permitan reconocer tempranamente trayectorias de riesgo y actuar sobre ellas de forma eficaz y oportuna.

Particularmente en este estudio se encontró que los desempeños en matemática evaluados por TERCE en 6° de Educación Primaria, funcionan como un mejor predictor que los desempeños en lectura, al intentar predecir una trayectoria de egreso vs de rezago y una trayectoria de egreso vs una de desvinculación. De igual modo, Fernández (2009) encontró que el resultado de matemática en las pruebas PISA es un potente predictor de la desvinculación. Otra particularidad a señalar, es que en esta tesis no se encontraron diferencias significativas entre secundaria pública y educación técnica pública al comparar trayectorias de desvinculación vs las de no desvinculación para poblaciones “comparables”, es decir, una vez que se aísla el efecto de otras variables.

Al igual que se reporta en (ANEP, 2022), se encuentra que la progresión en tiempo es especialmente crítica durante el primer año de Educación Media. Dos de cada diez estudiantes no logran avanzar a 2° grado, para los años siguientes, el ritmo de pérdida de estudiantes que no logran avanzar el grado, es de uno cada diez por año. Por su parte, al cabo de seis años del egreso de Educación Primaria, solamente cuatro de cada diez se inscriben en 6° grado de Educación Media. Estos alumnos son los que potencialmente podrían egresar en tiempo, pero efectivamente, lo logran tres, de esos cuatro estudiantes. El perfil de los alumnos que logra finalizar la Educación Media, muestra señales de desigualdad de oportunidades en el sistema educativo y además indica ciertas condiciones mínimas que deben tener los estudiantes al egresar de Educación Primaria,

para poder finalizar la Educación Media en tiempo: alumnos que hayan obtenido buenos niveles de aprendizaje, que pertenezcan al cuartil más alto del NSE, que no hayan acumulado rezago en la Educación Primaria y alumnos cuyas familias tengan altas expectativas sobre el nivel de estudios que alcanzarán. De esta forma se comprueba la hipótesis 5: la Educación Media no logra revertir la desigualdad de origen.

A pesar de la inequidad con la que se presenta el riesgo de rezagarse o desvincularse en los adolescentes que cursan Educación Media, ambos problemas, pero principalmente el rezago, también acontece para aquellos estudiantes que no presentan grandes vulnerabilidades socioeconómicas ni académicas. Esto plantea la necesidad de revisar el diseño del sistema educativo, particularmente en la transición entre ciclos.

El análisis de las trayectorias de los estudiantes según las combinaciones que tienen de estar o no estar inscriptos a lo largo de los ocho años del panel, permite concluir que el peso que tiene el tipo de trayectorias en donde “se van y no vuelven” no es menor. Aproximadamente el 26% de estudiantes que integran la cohorte, una vez que se van, no se matriculan por al menos dos años seguidos. La situación de una posible re-vinculación es menos probable por cada año que el adolescente no se encuentra en el sistema educativo (Cardozo, 2016). Inclusive cerca del 13% de los estudiantes de la cohorte no registra matriculación, al menos, desde (t+4), esto implica que se fueron de forma prematura. Ante las situaciones de no inscripción, se vuelve necesario aumentar las capacidades de acción del sistema educativo para que las intervenciones sean eficaces y oportunas.

Otro resultado a destacar es que Independientemente de si se consideran dos o tres años sin inscripción para categorizar a un estudiante como desvinculado, aquellos alumnos que no lograron finalizar sus cursos, no permanecen dentro del sistema educativo para intentar terminarlos, tal como se reporta en Fernández (2010) y Cardozo (2016).

Sobre la sexta y última hipótesis, la tesis evidencia que el sistema educativo uruguayo, cuenta con gran parte de la información poder construir, al menos, uno de los principales componentes de un sistema de alerta temprana: la capacidad predictiva. Empleando información de los sistemas de registros administrativos sobre los estudiantes y sus familias durante el transcurso de la Educación Primaria, sería viable predecir la desvinculación en la Educación Media con buenos niveles de sensibilidad, precisión y especificidad, tal como surge de los modelos presentados para la muestra de estudiantes de 6° de EP utilizada en este trabajo. Las variables necesarias, que caracterizan la situación del estudiante al momento de egresar de Educación Primaria, serían: región, acumulación de rezago, nivel socioeconómico, expectativas familiares, sexo, sector/oferta elegida para cursar 1° de Educación Media y el resultado de pruebas de aprendizaje (de matemática y/o de lectura) y en el caso de no contar con el resultado de dichas pruebas, sustituirlas por calificaciones docentes.

A modo de cierre, se destacan tres futuras líneas de investigación:

- Estudiar las migraciones de estudiantes entre el sector público y el privado, y entre educación secundaria y técnica, a lo largo de la trayectoria en Educación Media.
- Utilizar algoritmos de aprendizaje automático no paramétricos como el KNN, redes bayesianas y árbol de decisión, para evaluar la sensibilidad y precisión de las predicciones.

- Realizar predicciones para la totalidad de una cohorte de estudiantes de 6° de EP, empleando registros administrativos. Por ejemplo, la cohorte de 6° de EP del año 2017, que teóricamente comenzó la EM en el 2018. A marzo del año 2025, para esa cohorte, habrán pasado ocho años y, por tanto, se podrá contrastar las predicciones de los modelos empleados con lo que efectivamente le ocurrió a cada estudiante.

Referencias bibliográficas

- Acemoglu, D. and Autor, D. (2009) Lectures in Labor Economics, Chapter 1.
<https://economics.mit.edu/sites/files/LecturesinLaborEconomics.pdf>
- Andrey Zakharov, Martin Carnoy, Are teachers accurate in predicting their students' performance on high stakes' exams? The case of Russia, International Journal of Educational Development, Volume 43,2015, Pages 1-11,ISSN 0738-0593. <https://doi.org/10.1016/j.ijedudev.2015.04.007>.
- ANEP (2014). El Plan de Tránsito entre Ciclos Educativos. Informe de Seguimiento y Evaluación de la Segunda Cohorte (2012-2013), DIEE, Montevideo.
- ANEP (2017). Evaluación de impacto de las escuelas de Tiempo Completo en Uruguay 2013-2016, ANEP-DIEE-CEIP-PAEPU, Montevideo.
- ANEP (2019). Análisis de las trayectorias educativas de los egresados de la Educación Primaria entre 2013 y 2017, ANEP-DIEE, Montevideo.
- ANEP (2020). Rendición de Cuentas 2019 de ANEP. Tomo I, Capítulo 6, Montevideo.
- Arellano, Manuel, 2003. "Panel Data Econometrics," OUP Catalogue, Oxford University Press, number 9780199245291. <https://ideas.repec.org/b/oxp/obooks/9780199245291.html>
- Becker, G.S. (1964). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. University of Chicago Press.
- Bishop, Christopher M. Pattern Recognition and Machine Learning. Springer, 2006.
<https://doi.org/10.1007/978-0-387-45528-0>
- Borba, E, Cardozo, S., Castela, G. y Forteza, D. (2017), Evaluación de impacto de las escuelas de Tiempo Completo en Uruguay, Consejo Directivo Central de Administración Nacional de Educación Pública (CODICEN), Montevideo, Uruguay.
- Boado, M. y Fernández, T. (2010). Trayectorias académicas y laborales de los jóvenes en Uruguay. el Panel PISA 2003-2007, FCS-UDELAR, Montevideo.
- Bogliaccini, J. A. y Rodríguez, F. (2015), 'Regulación del sistema educativo y desigualdades de aprendizaje en el Uruguay.', Revista de la CEPAL (116), 87–101
- Borba, E., Cardozo, S., Castela, G., Peri, A., Salsamendi, G. y Silveira, A. (2018). Desarrollo infantil y riesgo de repetición en 1ero. ANEP-DIEE, documento de trabajo, Montevideo
- Blanco, E. (2008). "Factores escolares asociados a los aprendizajes en la educación primaria mexicana: un análisis multinivel". Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación. Vol. 4, n.º 1, España.
- BRYK, Anthony y LEE, Valerie (1992) Als Politics the Problem and Markets the Answer?. An Essay Review of Politics, Markets and America=s Schools@. In Economics of Education Review vol. 11 number 4. Pag. 439-451.
- Bruce, M., Bridgeland, J. M., & Fox, J. (2011). On track for success: The use of early warning indicator and intervention systems to build a grad nation. Civic Enterprises.

Brunner, J.J. y Elacqua, G. (2003). Informe de Capital Humano. Universidad Adolfo Ibáñez, Escuela de Gobierno, Santiago.

Bruch, J., Gellar, J., Cattell, L., Hotchkiss, J., & Killewald, P. (2020). Using data from schools and child welfare agencies

Cardozo Politi, S. (2015.). El comienzo del fin: desandando los caminos de la desafiliación escolar en la Educación Media. UR. FCS.

Cardozo Politi, S.; Biramontes, T.; Fernández, T.; Fonseca, B.; Peri, A.; Salsamendi, G. y Silveira, A. (2023). De la cuna a las aulas. Desarrollo infantil, alerta temprana y trayectoria escolar. Montevideo: DS-FCS-Udelar, DIEE-DSPE-ANEP; ANII-FMV, Doble clic · Editoras.

Cardozo, S. (2023). Desarrollo infantil y preparación para la escuela en Uruguay: efectos del rezago cognitivo y no cognitivo al término de la educación inicial sobre el desempeño escolar en la enseñanza primaria. *Revista de Sociología de la Educación* 16(1), 57-75.
<https://doi.org/10.7203/RASE.16.1.24777>

Cardozo, S. (2010). “El comienzo del fin: las decisiones de abandono durante la Educación media y su incidencia en las trayectorias” en Fernández, T. (ed) *La desafiliación en la Educación Media y Superior de Uruguay: conceptos, estudios y políticas*. CSIC-Udelar

Cardozo, S. (2016). Trayectorias educativas en la educación media PISA-L 2009-2014 [Informe]. INEE - Grupo de estudios sobre Transiciones Educación-Trabajo (TET). Montevideo. Recuperado de <https://www.ineed.edu.uy/images/pdf/trayectorias-educativas.pdf>

Cardozo, S.; Silveira, A. y Fonseca, B. (2022). Detección temprana del riesgo escolar. Predicción de trayectorias de rezago en la educación primaria en Uruguay mediante técnicas de machine learning. *Revista Latinoamericana de Estudios Educativos*, 52(2), 297-326.
<https://doi.org/10.48102/rlee.2022.52.2.391>

Cardozo, S., Schenck, M. y Severino, R. (2018). Las trayectorias de transición a la educación media básica de las dos primeras cohortes del Plan de Tránsito entre Ciclos Educativos en Fernández, T. y Mancebo, M. E. (editores.), “Política de Inclusión en la Educación Básica de Uruguay (2005-2015)”, FCS-UDELAR, Montevideo.

Cardozo, S., Fernández, T., Míguez, M. N. y Patrón, R. (2014), Transición entre ciclos: marco analítico, en Fernández, T. y Ríos, A. (ed.) “El tránsito entre ciclos en la educación media y superior de Uruguay”.

Carrasco, G. (2007). Calidad y Equidad en las escuelas peruanas: un estudio del efecto escuela en la prueba de matemática-PISA 2000. Convenio de Investigación 2005. Consorcio de Investigación Económica y Social (CIES).

Castro, J. (1996) Coordinación entre Primaria y Secundaria. Montevideo: ANEP- CODICEN.

Casey, Kevin & Azcona, David. (2017). Utilizing student activity patterns to predict performance. *International Journal of Educational Technology in Higher Education*. 14. 10.1186/s41239-017-0044-3.

Catts, H. W.; Petscher, Y.; Schatschneider, C.; Sittner Bridges, M. y Mendoza, K. (2009). Floor effects associated with universal screening and their impact on the early identification of reading disabilities. *Journal of Learning Disabilities*, 42: 163-176. <https://doi.org/10.1177/0022219408326219>

Cerda, J., y Cifuentes, L. (2012). Uso de curvas ROC en investigación clínica: Aspectos teórico-prácticos. *Revista chilena de infectología*, 29(2), 138-141. <https://doi.org/10.4067/S0716-10182012000200003>

Checchi (2005). Chapter 2. The economics of education. Human capital, family background and inequality. Cambridge University Press 2005.

Cunha, F. y Heckman, J. J. (2008), 'Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation', *Journal of Human Resources* 43(4), 738–782. URL: <http://jhr.uwpress.org/content/43/4/738.short>

Coleman, J. et al. (1966). Equality of educational opportunity. Washington US Government Printing Office. Citado por Murillo, F. (2000) en "La red iberoamericana de investigación sobre eficacia escolar y mejora de la escuela: un ejemplo de colaboración internacional en investigación educativa", *Revista española de educación comparada*, n.º 6, pp. 67-86.

Compton, D.; Fuchs, D.; Fuchs, L. y Bryant, J. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology*, 98: 394-409. <https://doi.org/10.1037/0022-0663.98.2.394>

Da Rocha, P., Martinez, J. P. y Rímoli, P. (2011), La escuela importa: Incidencia de los factores asociados a la escuela pública sobre las competencias de los estudiantes de PISA 2006.

Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K., &Sundararaman, V. (2013). School inputs, household substitution, and test scores. *American Economic Journal: Applied Economics*, 5(2), 29-57.

De Armas, G. y Retamoso, A. (2010) La universalización de la educación media en Uruguay: tendencias, asignaturas pendientes y retos a futuro, UNICEF: Montevideo.

De Melo, G.; Failache, E., y Machado, A. (2015) Adolescentes que no asisten a Ciclo Básico: caracterización de su trayectoria académica, condiciones de vida y decisiones de abandono, Serie Documentos de Trabajo, DT 04/2015, Instituto de Economía, FCEA-UDELAR: Montevideo.

De Melo, G. y Machado, A. (2016) Trayectorias educativas. Evidencias para Uruguay, Documento de Trabajo. INEEd: Montevideo.

De Melo, G. y Machado, A. (2018), 'Educational trajectories. Evidence from Uruguay', *International Journal of Educational Research* 92(March), 110–134. URL: <https://doi.org/10.1016/j.ijer.2018.09.018>

Fernández, T., (2009). La Desafiliación en la Educación Media en Uruguay. Una aproximación con base en el panel de estudiantes evaluados por PISA 2003. Reice. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 7(4), 164-179.

Fernández, T., Alonso, C., Boado, M., Cardozo, S., Ríos, A. y Menese, P. (2013). Reporte técnico PISA-L 2003-2012. Metodología de la segunda encuesta de seguimiento a los jóvenes evaluados por PISA en 2003, UDELAR, Montevideo.

Fernández, S., Woitschach, P., Fernández, M. y Álvarez -Díaz, R. (2018), 'Análisis de la Oportunidad de Aprendizaje en el estudio TERCE de la UNESCO', *Revista Investigación Educativa* 36(2), 509–528.

Fernández, Tabaré; Blanco, Emilio (2004) "¿Cuánto importa la escuela?. El caso de México en el contexto de América Latina". En la Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en la Educación, vol. 2 núm 1, Julio de 2004. Universidad de Deusto (País Vasco, España) y Universidad Autónoma de Madrid.

Filgueira, C., Rodríguez, F., y Fuentes, Á. (2006). Viejos instrumentos de la inequidad educativa: repetición en primaria y su impacto sobre la equidad en Uruguay. Serie Documentos de Trabajo del IPES / Colección Monitor Social N°9. Universidad Católica del Uruguay. Recuperado de <https://carreras.ucu.edu.uy/aucdocumento.aspx?1020,3224>

Flavia Terigi. "Trayectorias escolares e inclusión educativa: del enfoque individual al desafío para las políticas educativas". En Marchesi, Álvaro; Blanco, Rosa; y Hernández, Laura (eds.). Avances y desafíos de la educación inclusiva en Iberoamérica. Madrid: Organización de Estados Iberoamericanos, 2014.

Habiaga, V. (2019). La repetición escolar y sus atolladeros: Aportes para la reflexión pedagógica. Repositorio Institucional del Consejo de Formación en Educación. <http://repositorio.cfe.edu.uy/handle/123456789/377>

Hanushek, E. y Wößmann, L. (2007). The Role of Education Quality in Economic Growth. World Bank Policy Research Working Paper 4122.

Hanushek, E.A., S.G. Rivkin (2012). The Distribution of Teacher Quality and Implications for Policy. Annual Review of Economics 4: 131-157.

Hattie, J. (2008). Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement. Routledge.

Heckman, James, Stixrud, Jora and Urzua, Sergio, (2006), The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior, Journal of Labor Economics, 24, issue 3, p. 411-482

Heckman, James J. & LaFontaine, Paul A., 2007. "The American High School Graduation Rate: Trends and Levels," IZA Discussion Papers 3216, Institute of Labor Economics (IZA).

Heckman, J.J., Humphries, J.E., Urzúa, S.S., & Veramendi, G.F. (2010). The effects of educational choices on labor market, health, and social outcomes.

Heckman, Jagelka, T., Kautz, T. (2019), "Some contributions of economics on the study of personality", NBER WP, 26459.

Heckman, J. y Masterov, D. (2007). "The productivity argument for investing in young children". Applied Economic Perspectives and Policy, 29(3), 446-493.

Heckman, J. J., & Mosso, S. (2014). The Economics of Human Development and Social Mobility. Annual Review of Economics, 6(1), 689–733.

Heckman, J.; Lochner, L. y Todd, P. (2008), "Earnings Functions and Rates of Return", en Journal of Human Capital, Vol. 2, No. 1 (Spring 2008), pp. 1-31

INEEd (2015), Evaluación y tránsito educativo. Estudio de propuestas de evaluación en las aulas de educación primaria y media en Uruguay. Resumen ejecutivo, INEE, Montevideo

INEEd (2018). Aristas 2017. Informe de resultados de tercer y sexto de educación primaria, INEE, Montevideo.

Irizarry, R. A. (2020). Introduction to Data Science: Data Analysis and Prediction Algorithms with R. Chapman y Hall/CRC.

Jojo, S.; Jittawiriyankoon, C. y John, V. K. (2022). Efficacy of an online reading intervention program “Literacy Pro” in classifying students as “at risk” and “not at risk” among ESL school students – A logistic regression analysis. *Creative Education*, 13: 952-970, <https://doi.org/10.4236/ce.2022.133063>

Kilgus, S. P.; Methe, S. A.; Maggin, D. M. y Tomasula, J. L. (2014). Curriculum-based measurement of oral reading (R-CBM): A diagnostic test accuracy meta-analysis of evidence supporting use in universal screening. *Journal of School Psychology*, 52: 377-405. <https://doi.org/10.1016/j.jsp.2014.06.002>

Lado, A. (2019). Efectos de la Escuela Primaria en los Desempeños de los Estudiantes en la Educación Media Básica, Tesis de Maestría, FCEA-UDELAR, Montevideo.

Landis, J. R., y Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159-174.

Lambí, C., Perera, M. y Messina, P. (2009), ‘Desigualdad de oportunidades y el rol del sistema educativo en los logros de los jóvenes uruguayos’, Estudios de la Edición 2008 del Fondo Concursable Carlos Filgueira. *Infancia, Adolescencia y Políticas Sociales* (598 2), 39–118.

Lee, Valerie & Burkam, David. (2003). Dropping Out of High School: The Role of School Organization and Structure. *American Educational Research Journal - AMER EDUC RES J.* 40. 353-393. 10.3102/00028312040002353.

Ley Nº 18.437.LEY GENERAL DE EDUCACIÓN. Publicada D.O. 16 ene/009 - Nº 27654. <https://www.impo.com.uy/educacion/>

Macfadyen, Leah & Sorenson, Peter. (2010). Using LiMS (the Learner Interaction Monitoring System) to Track Online Learner Engagement and Evaluate Course Design. *Educational Data Mining 2010 - 3rd International Conference on Educational Data Mining*. 301-302.

Mara, S., Alesina, L., Cabrio, S., Erramouspe, R., Ibañez, W. y Pazos, L. (2000). Estudio de evaluación de impacto de la educación inicial en Uruguay, MECAEP-ANEP-BIRF, Montevideo.

Martínez Pérez, J. A., y Pérez Martín, P. S. (2023). La curva ROC. *Medicina de Familia. SEMERGEN*, 49(1), 101821. <https://doi.org/10.1016/j.semerg.2022.101821>

Ministerio de Educación y Cultura. Anuario estadístico de educación 2013. Montevideo, Uruguay, Ministerio de Educación y Cultura, 2013. Web. <https://www.gub.uy/ministerio-educacion-cultura/datos-y-estadisticas/datos/anuario-estadistico-educacion-2013>.

Mendez-Errico, L. y Ramos, X. (2014), Schooling progression in Uruguay: Why some children are left behind?, PhD thesis, Universitat Autònoma de Barcelona.

Méndez, L., Ramos, X. (2019) “Selection and educational attainment: Why some children are left behind? Evidence from a middle-income country”. Serie Documentos de Trabajo, DT 01/2019. Instituto de Economía, Facultad de Ciencias Económicas y Administración, Universidad de la República, Uruguay.

Murillo, J. (2008). Hacia un modelo de Eficacia Escolar. Estudio Multinivel sobre los factores de eficacia en las Escuelas Españolas. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*. Vol. 6, n.º 1, pp. 4-28. España.

Neild, R. C., & Balfanz, R. (2006). Unfulfilled Promise: The Dimensions and Characteristics of Philadelphia's Dropout Crisis, 2000–2005. [1]

Noboa, L. (2020.). Cursos de vida y destinos tempranos de inclusión, vulnerabilidad y exclusión educativo-laboral de los jóvenes evaluados en PISA 2009. Tesis de maestría. Universidad de la República (Uruguay). Facultad de Ciencias Sociales.

Observatorio de la Educación - ANEP. (2024). Observatorio de la educación. Recuperado de <https://observatorio.anep.edu.uy>

Rabe-Hesketh, S. y Skrondal, A. (2022). Multilevel and longitudinal modeling using STATA. Stata Press.

Rimfeld K, Malanchini M, Hannigan LJ, Dale PS, Allen R, Hart SA, Plomin R. Teacher assessments during compulsory education are as reliable, stable and heritable as standardized test scores. *J Child Psychol Psychiatry*. 2019 Dec;60(12):1278-1288. doi: 10.1111/jcpp.13070. Epub 2019 May 12. PMID: 31079420; PMCID: PMC6848749.

Stiglitz, J. E. (1975). The Theory of Screening, Education, and the Distribution of Income. *American Economic Review*, 65, 283-300.

UNESCO (2015) Informe de resultados TERCE. Logros de aprendizaje, Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación: Santiago de Chile

UNESCO-OREALC (2016), Reporte técnico. Tercer estudio regional comparativo y explicativo, TERCE, Santiago, Chile.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. (2016, septiembre). Issue Brief: Early Warning Systems.

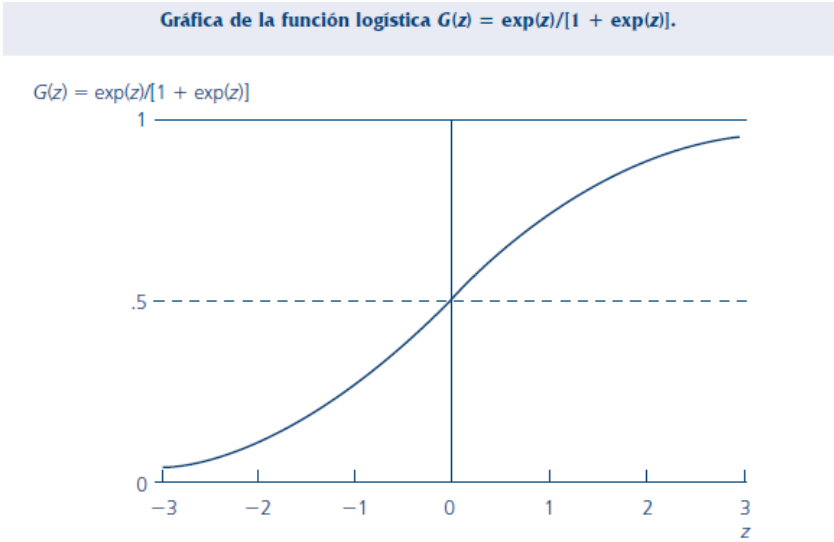
Vásquez Echeverría, A. y Moreira, K. (2016). Preparación para la escolarización. Dimensiones y Medición. Facultad de Psicología, Universidad de la República, Uruguay.

Westphal A, Vock M and Kretschmann J (2021) Unraveling the Relationship Between Teacher-Assigned Grades, Student Personality, and Standardized Test Scores. *Front. Psychol*. 12:627440. doi: 10.3389/fpsyg.2021.627440

Wooldridge, Jeffrey M. Introducción a la econometría. Un enfoque moderno, 4a. edición. ISBN-13: 978-607-481-312-8

Zakharov, Andrey & Carnoy, Martin, 2015. "Are teachers accurate in predicting their students' performance on high stakes' exams? The case of Russia," *International Journal of Educational Development*, Elsevier, vol. 43(C), pages 1-11. <https://ideas.repec.org/a/eee/injoed/v43y2015icp1-11.html>

Ilustración A. 1.Representación de la Función Logística



Fuente: tomado de Wooldridge, Jeffrey M. Introducción a la econometría. Un enfoque moderno, 4a. edición

Tabla A. 1. Porcentaje de estudiantes para los que no se cuenta con información sobre su matriculación, fallecieron o migraron del país. Según año del Panel.

	Inscripto	No inscripto	Egreso de t+6	Egreso de t+7	Fuera del panel (falleció, emigró, o missing)	Total
t (2013)	100	0,0	-	-	0,00	100
t+1 (2014)	97,5	2,1	-	-	0,40	100
t+2 (2015)	91,7	7,6	-	-	0,68	100
t+3 (2016)	87,5	11,5	-	-	1,02	100
t+4 (2017)	84,3	14,7	-	-	0,96	100
t+5 (2018)	80,5	18,6	-	-	0,92	100
t+6 (2019)	74,0	25,0	-	-	0,98	100
t+7 (2020)	33,6	36,5	28,9	-	1,04	100
t+8 (2021)	19,3	43,1	28,9	7,7	1,04	100

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Tabla A. 2. Medidas de resumen de las variables de interés.

Valores de las categorías	Nombre y categoría	Porcentaje	Error estándar	[intervalo de confianza 95%]	Mean	Std. dev.	Min	Max
Trayectoria con 3 resultados					1,85	0,75	1	3
1	Egresados	36,9	1,6	33,84 - 40,14				
2	Rezagados	41,1	1,7	37,91 - 44,40				
3	Desvinculados	22,0	1,5	19,22 - 24,95				
Trayectoria de desvinculación					0,22	0,41	0	1
0	Rezagado o egresado	78,05	1,46	75,05 - 80,78				
1	Desvinculado	21,95	1,46	19,22 - 24,95				
Región en 6° EP					0,23	0,42	0	1
0	Interior	77,4	1,4	74,61 - 79,98				
1	Montevideo	22,6	1,4	20,02 - 25,39				
Nivel Socioeconómico en 6° EP					2,50	1,12	1	4
1	Cuartil 1	25,1	1,6	22,12 - 28,24				
2	Cuartil 2	25,1	1,5	22,23 - 28,16				
3	Cuartil 3	24,9	1,5	22,13 - 27,89				
4	Cuartil 4	25,0	1,3	22,52 - 27,59				
Sexo en 6° EP					0,51	0,50	0	1
0	Niño	49,0	1,7	45,72 - 52,33				
1	Niña	51,0	1,7	47,67 - 54,28				
Rezago en 6° EP					0,24	0,43	0	1
0	Sin rezago	75,8	1,5	72,70 - 78,66				
1	Con rezago	24,2	1,5	21,34 - 27,30				
Nivel de matemática en 6° EP					2,24	0,98	1	4
1	Nivel I	26,2	1,5	23,30 - 29,31				
2	Nivel II	36,2	1,6	33,13 - 39,47				
3	Nivel III	24,9	1,5	22,13 - 27,90				
4	Nivel IV	12,7	1,0	10,79 - 14,80				
Nivel de lectura en 6° EP					2,52	0,95	1	4
1	Nivel I	11,3	1,1	9,27 - 13,63				
2	Nivel II	46,4	1,7	43,13 - 49,76				
3	Nivel III	21,2	1,4	18,63 - 23,99				
4	Nivel IV	21,1	1,3	18,70 - 23,74				
Expectativa culminación familiar en 6° EP					3,20	1,16	1	5
1	EMB o menos	4,7	0,7	3,48 - 6,26				
2	EMS	30,6	1,6	27,57 - 33,87				
3	Educ.Terciaria ciclo corto	18,9	1,3	16,47 - 21,67				
4	Educ. Universitaria o terciaria	31,3	1,5	28,46 - 34,32				
5	Sin dato	14,5	1,2	12,18 - 17,05				
Sector en en 6° EP					0,14	0,35	0	1
0	Público	85,6	0,7	84,08 - 86,98				
1	Privado	14,4	0,7	13,02 - 15,92				
Oferta y sector en 1° EM					2,05	0,69	1	5
1	Secundaria privada	18,1	1,0	16,24 - 20,20				
2	Secundaria pública	61,5	1,6	58,38 - 64,61				
3	Educación técnica	17,4	1,4	14,85 - 20,16				
4	Educ. primaria o no inscripto	2,9	0,6	1,97 - 4,28				
5	Sin dato	0,1	0,0	0,02 - 0,18				

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Tabla A. 3. Correlación entre las variables de interés

variables de resultado				Variables categóricas de a 6° EP (explicativas)										de 1° EM
variables de resultado	Trayectoria de 3 resultados		Logit ¹ y 2	Región	Cuartiles NSE	Sexo	Rezago	Nivel de Mat.	Nivel de Lec. ¹	Expectativa culminación familiar	Sector institucional	Oferta y sector		
	Trayectoria de desvinculación	1												
				1										
Variables explicativas categóricas	Región	0,0124	0,0164											
	Cuartiles Nivel Socioeconómico	-0,4051	-0,3361	0,1489	1									
	Sexo	-0,1557	-0,1431	-0,012	-0,016	1								
	Rezago	0,4695	0,479	0,0043	-0,294	-0,041	1							
	Nivel de matemática	-0,4205	-0,3035	0,0648	0,3372	-0,041	-0,323	1						
de 6° EP	Nivel de lectura	-0,4219	-0,3389	0,094	0,3579	0,044	-0,334	0,5558	1					
	Expectativa culminación familiar	-0,17	-0,1088	0,1748	0,2845	-0,059	-0,164	0,1999	0,1972	1				
	Sector institucional	-0,262	-0,1751	0,4009	0,4613	-0,014	-0,16	0,2682	0,299	0,211	1			
de 1° EM	Oferta y sector	0,3657	0,3121	-0,297	-0,486	-0,068	0,2913	-0,3209	-0,3302	-0,2388	-0,5548	1		
Variables explicativas continuas de 6° EP														
Variables explicativas continuas	Nivel Socio Económico	-0,4222	-0,3335	1										
	Puntaje en matemática	-0,4328	-0,3241	0,396	1									
de 6° EP	Puntaje en lectura	-0,4548	-0,3680	0,4035	0,6276	1								
				730,2446	102,3846	255,79	1088,935							

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Tabla A. 4. Porcentaje de inscriptos en EM respecto al total de la cohorte de 6° de primaria según variables de interés para cada año del panel.

Estudiantes inscriptos en 6° de educación primaria		Porcentaje de estudiantes inscriptos en EM. sobre el total de la cohorte de 6° de primaria en t							
	t 2013	t+1 2014	t+2 2015	t+3 2016	t+4 2017	t+5 2018	t+6 2019	t+7 2020	t+8 2021
Total	100	97,9	92,4	88,4	85,2	81,2	74,8	63,1	56,4
Nivel Socioeconómico									
Cuartil 1	100	95,9	84,3	74,6	69,4	65,0	55,6	44,9	36,0
Cuartil 2	100	96,8	90,8	86,6	81,9	78,6	69,0	54,3	49,3
Cuartil 3	100	98,9	94,9	93,9	91,9	86,9	80,8	70,6	62,3
Cuartil 4	100	100,0	99,5	98,6	97,6	94,4	93,7	82,6	78,3
Sector institucional									
Público	100	97,6	91,1	86,5	83,1	78,9	71,7	60,2	53,3
Privado	100	99,8	99,7	99,9	97,4	94,9	93,0	80,1	74,8
Sexo									
Varones	100	97,2	92,1	87,8	84,3	78,3	69,1	59,0	48,4
Mujeres	100	98,6	92,6	89,0	86,0	84,0	80,2	67,0	64,1
Región									
Interior	100	98,0	93,2	89,5	85,0	82,1	75,4	63,4	57,0
Montevideo	100	97,7	89,4	84,7	85,7	78,3	72,4	62,0	54,3
Rezago en 6° primaria									
Sin rezago	100	99,6	97,9	96,7	93,8	91,1	86,3	74,6	67,0
Con rezago	100	92,7	75,0	62,3	58,1	50,1	38,4	27,0	23,4
Desempeño en matemática (TERCE)									
Nivel 1	100	95,3	83,5	77,0	72,1	67,8	54,1	45,2	36,1
Nivel 2	100	97,7	93,6	88,1	84,7	79,5	75,1	59,3	51,1
Nivel 3	100	100,0	97,0	95,3	93,3	91,6	87,1	76,1	71,4
Nivel 4	100	100,0	98,0	99,6	97,5	93,2	92,1	85,4	84,3
Desempeño en lectura (TERCE)									
Nivel 1	100	93,8	73,9	68,6	67,6	53,9	45,2	37,4	33,0
Nivel 2	100	97,9	91,8	85,5	80,3	77,0	68,3	56,9	48,4
Nivel 3	100	98,1	96,2	95,1	92,0	87,6	84,1	70,6	64,3
Nivel 4	100	100,0	99,7	98,8	98,3	98,5	95,4	84,9	80,9
Expectativas familiares del nivel educativo que alcanzará el niño/a									
EMB o menos	100	89,9	72,5	65,6	57,8	49,3	31,0	23,4	22,7
Educ. Media Superior	100	98,1	92,5	82,8	78,4	72,2	66,3	55,4	46,7
Educ. Terciaria	100	99,7	98,3	96,0	95,0	91,8	88,9	75,1	63,9
Educ. Universitaria	100	100,0	99,4	99,8	96,2	95,4	90,5	80,3	73,5

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Tabla A. 5. Diferencias entre el porcentaje de inscriptos en EM según variables de interés para cada año del panel.

	t+1 2014	t+2 2015	t+3 2016	t+4 2017	t+5 2018	t+6 2019	t+7 2020	t+8 2021
Nivel Socioeconómico								
Cuartil 4 / cuartil 1	1,0	1,2	1,3	1,4	1,5	1,7	1,8	2,2
Cuartil 4 - cuartil 1	4,1	15,2	24,0	28,3	29,5	38,1	37,7	42,4
Sector institucional								
Privado / público	1,0	1,1	1,2	1,2	1,2	1,3	1,3	1,4
Privado - público	2,2	8,6	13,4	14,3	16,0	21,3	19,8	21,5
Sexo								
Mujeres / Varones	1,0	1,0	1,0	1,0	1,1	1,2	1,1	1,3
Mujeres - Varones	1,4	0,6	1,2	1,8	5,7	11,2	8,0	15,7
Región								
Interior / Mvd	1,0	1,0	1,1	1,0	1,0	1,0	1,0	1,1
interior - Mvd	0,3	3,8	4,9	-0,7	3,8	3,0	1,4	2,7
Rezago en 6° primaria								
Sin rezago / con rezago	1,1	1,3	1,6	1,6	1,8	2,2	2,8	2,9
Sin rezago - con rezago	6,9	22,8	34,4	35,7	41,1	47,9	47,6	43,5
Desempeño en matemática en TERCE								
Nivel 4 / nivel 1	1,0	1,2	1,3	1,4	1,4	1,7	1,9	2,3
Nivel 4 - nivel 1	4,7	14,6	22,6	25,4	25,4	37,9	40,3	48,1
Desempeño en lectura en TERCE								
Nivel 4 / nivel 1	1,1	1,3	1,4	1,5	1,8	2,1	2,3	2,4
Nivel 4 - nivel 1	6,2	25,8	30,1	30,7	44,7	50,2	47,4	47,9
Expectativas familiares del nivel educativo que alcanzará el niño/a								
Educ Univ. / EMB	1,1	1,4	1,5	1,7	1,9	2,9	3,4	3,2
Educ Univ. - EMB	10,0	26,9	34,2	38,4	46,1	59,4	56,9	50,8

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Tabla A. 6. Porcentaje de inscriptos en el grado normativo en EM respecto al total de la cohorte de 6° de primaria, según variables de interés para cada año del panel.

	6° primaria t	1° EM t+1	2° EM t+2	3° EM t+3	4° EM t+4	5° EM t+5	6° EM t+6
Total	100	96,6	76,6	67,7	57,1	47,0	39,9
Nivel Socioeconómico							
Cuartil 1	100	92,6	57,2	43,5	35,2	25,4	21,6
Cuartil 2	100	96,1	68,7	57,6	46,7	34,1	26,6
Cuartil 3	100	98,9	87,4	78,4	64,2	54,2	47,9
Cuartil 4	100	99,5	95,8	93,8	84,6	76,6	65,9
Sector institucional							
Público	100	96,3	73,7	63,6	53,6	42,3	35,6
Privado	100	99,5	97,1	95,6	81,0	78,4	68,0
Sexo							
Varones	100	95,8	73,8	66,4	54,9	43,5	35,5
Mujeres	100	97,6	80,4	70,0	60,2	51,4	45,0
Región							
Interior	100	96,5	78,7	69,1	59,2	48,2	41,1
Montevideo	100	97,3	71,7	65,2	52,1	45,3	37,9
Rezago en 6° primaria							
Sin rezago	100	99,5	88,4	80,2	68,8	57,1	49,0
Con rezago	100	87,9	40,6	29,1	21,1	16,4	12,1
Desempeño en matemática (TERCE)							
Nivel 1	100	91,9	55,4	42,9	28,6	21,0	15,8
Nivel 2	100	96,9	77,1	65,7	54,7	41,2	34,9
Nivel 3	100	99,8	90,9	84,7	75,7	65,4	54,3
Nivel 4	100	99,9	94,5	94,4	89,5	84,3	77,9
Desempeño en lectura (TERCE)							
Nivel 1	100	92,2	50,3	40,3	30,1	20,0	19,1
Nivel 2	100	95,8	69,1	56,8	44,6	33,9	25,4
Nivel 3	100	98,0	87,5	79,7	71,3	58,8	52,4
Nivel 4	100	99,9	98,5	96,3	86,6	80,1	71,5
Expectativas familiares del nivel educativo que alcanzará el niño/a							
EMB o menos	100	84,2	30,2	19,9	17,4	16,4	10,5
Educ. Media Superior	100	95,3	68,3	52,1	40,5	29,2	24,9
Educ. Terciaria	100	99,6	89,7	84,4	66,2	56,2	47,5
Educ. Universitaria	100	99,9	94,2	91,9	83,7	72,5	62,4

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Tabla A. 7. Diferencias entre el porcentaje de inscriptos en el grado normativo según variables de interés para cada año del panel entre t+1 t t+6.

	1° EM t+1	2° EM t+2	3° EM t+3	4° EM t+4	5° EM t+5	6° EM t+6
Nivel Socioeconómico						
Cuartil 4 / cuartil 1	1,1	1,7	2,2	2,4	3,0	3,1
Cuartil 4 - cuartil 1	6,9	38,6	50,3	49,3	51,2	44,3
Sector institucional						
Privado / público	1,0	1,3	1,5	1,5	1,9	1,9
Privado - público	3,2	23,4	32,0	27,4	36,1	32,4
Sexo						
Mujeres / Varones	1,0	1,1	1,1	1,1	1,2	1,3
Mujeres - Varones	1,8	6,6	3,6	5,3	7,9	9,5
Región						
Interior / Mvd	1,0	1,1	1,1	1,1	1,1	1,1
interior - Mvd	-0,8	7,0	4,0	7,1	2,8	3,2
Rezago en 6° primaria						
Sin rezago / con rezago	1,1	2,2	2,8	3,3	3,5	4,1
Sin rezago - con rezago	11,6	47,8	51,1	47,7	40,7	36,9
Desempeño en matemática en TERCE						
Nivel 4 / nivel 1	1,1	1,7	2,2	3,1	4,0	4,9
Nivel 4 - nivel 1	8,0	39,1	51,5	60,9	63,3	62,1
Desempeño en lectura en TERCE						
Nivel 4 / nivel 1	1,1	2,0	2,4	2,9	4,0	3,8
Nivel 4 - nivel 1	7,7	48,2	56,0	56,4	60,0	52,5
Expectativas familiares del nivel educativo que alcanzará el niño/a						
Educ Univ / EMB	1,2	3,1	4,6	4,8	4,4	5,9
Educ Univ - EMB	15,7	64,0	72,0	66,3	56,2	51,9

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Tabla A. 8. Porcentaje de estudiantes según tipología de trayectoria al final del período de observación, para cada variable de estudio.

	Egresados	Rezagados	Desvinculados	Total	Egresados / Desvinculados
Total	36,9	41,1	22,0	100	1,7
Sector					
Público	31,8	43,3	24,9	100	1,3
Privado	67,3	28,4	4,3	100	15,7
Región					
Interior	37,1	41,4	21,6	100	1,7
Montevideo	36,5	40,3	23,2	100	1,6
Nivel Socioeconómico					
Cuartil 1	21,0	37,9	41,1	100	0,5
Cuartil 2	22,8	49,5	27,7	100	0,8
Cuartil 3	39,0	46,0	15,0	100	2,6
Cuartil 4	65,0	31,1	3,9	100	16,8
Sexo					
Niño	31,0	41,0	28,0	100	1,1
Niña	42,6	41,3	16,1	100	2,6
Rezago					
Sin rezago	45,7	43,6	10,8	100	4,3
Con rezago	9,5	33,5	57,0	100	0,2
Nivel de matemática					
Nivel I	13,5	45,4	41,1	100	0,3
Nivel II	31,0	47,0	22,0	100	1,4
Nivel III	52,9	36,4	10,6	100	5,0
Nivel IV	52,9	36,4	10,6	100	5,0
Nivel de lectura					
Nivel I	16,6	34,1	49,3	100	0,3
Nivel II	24,1	47,3	28,6	100	0,8
Nivel III	46,6	41,7	11,7	100	4,0
Nivel IV	66,4	30,7	2,9	100	22,7
Expectativa culminación familiar					
EMB o menos	9,64	25,97	64,39	100	0,1
Educ. Media Superior	23,15	46,9	29,95	100	0,8
Educ. terciaria de ciclo corto	40,36	51,47	8,17	100	4,9
Educ. universitaria o terciaria	58,85	34,68	6,47	100	9,1
Sin dato	23	34,17	42,84	100	0,5
Oferta y sector t+1					
Secundaria privada	66,46	31,12	2,41	100	27,6
Secundaria pública	34,19	43,86	21,95	100	1,6
Educación técnica	21,84	46,3	31,86	100	0,7
6°Educ. Prim. o no inscripto en EM	0,64	14,56	84,81	100	0,01

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Tabla A. 9. Modelo logístico multinomial sobre la trayectoria al cabo de 8 años en la EM para la Cohorte de 6° de EP en 2013

Multinomial logistic regression		Number of obs		=	4.283								
		Wald chi2(32)		=	387,72								
		Prob > chi2		=	0,0000								
Log pseudolikelihood = -43557,316		Pseudo R2		=	0,2521								
Trayectoria final en t+8													
Egresados (base outcome)													
Rezagados t+8							Desvinculados t+8						
	Coeff	Robust std. err.	P> z	Sign.	[95% conf. interval]		Coeff	Robust std. err.	P> z	Sign	[95% conf. interval]		
Región t													
Montevideo (vs. Interior)	0,619	0,214	0,004	***	0,199	1,038	0,805	0,348	0,021	**	0,122	1,487	
Cuartiles NSE t													
Cuartil 2 (vs.cuarti1)	0,381	0,290	0,189		-0,187	0,949	-0,167	0,342	0,625		-0,838	0,504	
Cuartil 3 (vs.cuarti1)	-0,064	0,299	0,831		-0,649	0,521	-0,869	0,369	0,019	**	-1,592	-0,145	
Cuartil 4 (vs.cuarti1)	-0,633	0,309	0,040	**	-1,238	-0,028	-1,966	0,463	0,000	***	-2,873	-1,059	
Sexo t													
niña (vs. Varón)	-0,529	0,176	0,003	***	-0,873	-0,185	-1,403	0,259	0,000	***	-1,910	-0,895	
Rezago t													
Con rezago (vs. Sin rezago)	0,828	0,310	0,008	***	0,219	1,436	2,292	0,323	0,000	***	1,660	2,925	
Nivel matematica t													
Nivel II (vs. Nivel I)	-0,650	0,252	0,010	***	-1,144	-0,156	-1,032	0,305	0,001	***	-1,631	-0,433	
Nivel III (vs. Nivel 1)	-1,323	0,272	0,000	***	-1,857	-0,790	-1,777	0,377	0,000	***	-2,516	-1,037	
Nivel IV (vs. Nivel 1)	-1,758	0,313	0,000	***	-2,371	-1,145	-2,285	0,536	0,000	***	-3,337	-1,234	
Expectativas culminación familiar t													
EMS (vs. EMB o menos)	0,006	0,650	0,993		-1,267	1,279	-1,350	0,660	0,041	**	-2,644	-0,055	
Educ. Terciaria (vs. EMB o menos)	-0,146	0,660	0,825		-1,439	1,147	-2,222	0,715	0,002	***	-3,623	-0,820	
Educ. Universitaria (vs. EMB o menos)	-0,470	0,653	0,472		-1,751	0,810	-1,898	0,691	0,006	***	-3,252	-0,543	
Sin dato (vs. EMB o menos)	-0,319	0,691	0,645		-1,674	1,036	-0,915	0,708	0,196		-2,302	0,473	
Oferta/sector t+1													
Secundaria priv. (vs. Secundaria pub.)	-0,445	0,250	0,075	*	-0,935	0,044	-1,484	0,434	0,001	***	-2,335	-0,634	
Educ. Técnica (vs. Secundaria pub.)	0,107	0,306	0,728		-0,493	0,706	0,004	0,337	0,990		-0,656	0,664	
6°Educ. Prim. o no inscripto en EM (vs. Secundaria pub)	1,852	0,701	0,008	***	0,477	3,227	3,037	0,680	0,000	***	1,704	4,370	
cons.	1,510	0,732	0,039		0,075	2,945	2,654	0,776	0,001		1,133	4,175	

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Nota: El modelo incluye los niveles de desempeño en **Matemática**

*** = Significación al 99%; ** = Significación al 95%; * = Significación al 90%

Tabla A. 10. Modelo logístico multinomial sobre la trayectoria al cabo de 8 años en la EM para la Cohorte de 6° de EP en 2013

Multinomial logistic regression	Number of obs		=	4.283								
	Wald chi2(32)		=	362,52								
	Prob > chi2		=	0,0000								
	Pseudo R2		=	0,2448								
Log pseudolikelihood = -43557,316												
Trayectoria final en t+8												
Egresados (base outcome)												
Rezagados t+8												
Desvinculados t+8												
	Coeff	Robust std. err.	P> z	Sign.	[95% conf. interval]		Coeff	Robust std. err.	P> z	Sign.	[95% conf. interval]	
Región t												
Montevideo (vs. Interior)	0,643	0,207	0,002	***	0,238	1,048	0,820	0,329	0,013	**	0,176	1,464
Cuartiles NSE t												
Cuartil 2 (vs.cuarti1)	0,318	0,295	0,280		-0,259	0,896	-0,207	0,344	0,548		-0,881	0,468
Cuartil 3 (vs.cuarti1)	-0,090	0,310	0,771		-0,698	0,518	-0,867	0,386	0,025	**	-1,623	-0,111
Cuartil 4 (vs.cuarti1)	-0,655	0,314	0,037	**	-1,271	-0,039	-1,999	0,458	0,000	***	-2,897	-1,102
Sexo t												
niña (vs. Varón)	-0,425	0,174	0,015	***	-0,767	-0,083	-1,244	0,260	0,000	***	-1,754	-0,734
Rezago t												
Con rezago (vs. Sin rezago)	0,887	0,330	0,007	***	0,240	1,535	2,283	0,344	0,000	***	1,609	2,956
Nivel Lectura t												
Nivel II (vs. Nivel I)	-0,014	0,360	0,970		-0,719	0,692	-0,331	0,376	0,378		-1,068	0,406
Nivel III (vs. Nivel 1)	-0,588	0,372	0,114		-1,317	0,141	-1,482	0,446	0,001	***	-2,356	-0,609
Nivel IV (vs. Nivel 1)	-0,956	0,376	0,011	***	-1,692	-0,219	-2,255	0,526	0,000	***	-3,286	-1,224
Expectativas culminación familiar t												
EMS (vs. EMB o menos)	-0,198	0,615	0,747		-1,403	1,007	-1,509	0,638	0,018	**	-2,759	-0,258
Educ. Terciaria (vs. EMB o menos)	-0,291	0,616	0,637		-1,498	0,916	-2,339	0,683	0,001	***	-3,677	-1,000
Educ. Universitaria (vs. EMB o menos)	-0,666	0,615	0,279		-1,872	0,540	-1,968	0,668	0,003	***	-3,277	-0,659
Sin dato (vs. EMB o menos)	-0,446	0,646	0,490		-1,712	0,820	-1,025	0,668	0,125		-2,335	0,285
Oferta/sector t+1												
Secundaria priv. (vs. Secundaria pub.)	-0,486	0,244	0,046	*	-0,963	-0,008	-1,384	0,410	0,001	***	-2,188	-0,581
Educ. Técnica (vs. Secundaria pub.)	0,150	0,320	0,638		-0,476	0,777	0,023	0,358	0,948		-0,679	0,725
6°Educ. Prim. o no inscripto en EM (vs. Secundaria pub)	1,908	0,685	0,005	***	0,565	3,251	3,153	0,632	0,000	***	1,915	4,392
cons.	1,149	0,732	0,117		-0,286	2,583	-1,965	1,128	0,081		0,894	3,941

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Nota: El modelo incluye los niveles de desempeño en **Lectura**

*** = Significación al 99%; ** = Significación al 95%; * = Significación al 90%

Tabla A. 11. Regresión Logística sobre la probabilidad de desvinculación al cabo de 8 años en la EM para la Cohorte de 6° de primaria en 2013. Logit 1.

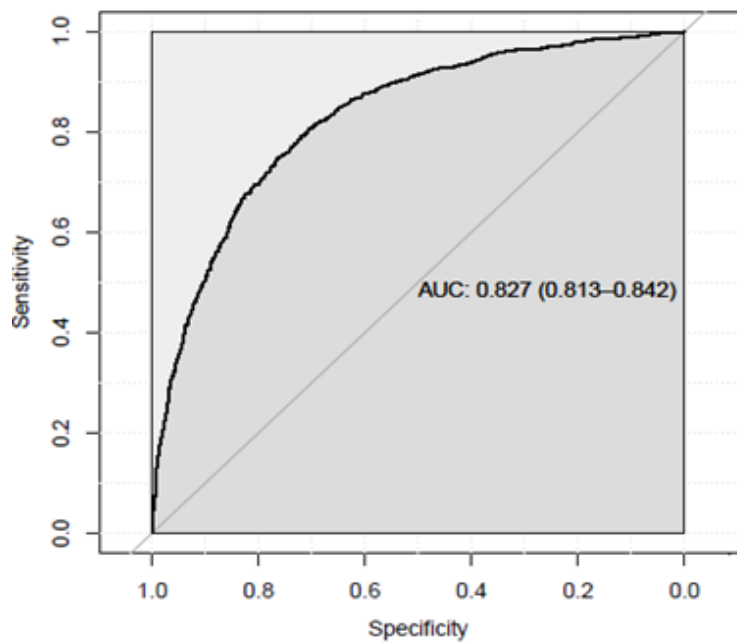
Logistic regression		Number of obs		= 4.283	
		Wald chi2(32)		= 232,64	
Log pseudolikelihood = -18768,534		Prob > chi2		= 0,000	
		Pseudo R2		= 0,3475	
trayectoria_fin_t_8_desvin	Coefficient	Robust std. err.	P> z	Sign.	[95% confinterval]
Region t					
Montevideo (vs. Interior)	0,331	0,317	0,296		-0,290 0,952
Cuartiles NSE t					
Cuartil 2 (vs.cuarti1)	-0,452	0,264	0,086	*	-0,969 0,065
Cuartil 3 (vs.cuarti1)	-0,818	0,281	0,004	***	-1,369 -0,267
Cuartil 4 (vs.cuarti1)	-1,594	0,422	0,000	***	-2,420 -0,767
Sexo t					
niña (vs.varón)	-1,015	0,220	0,000	***	-1,446 -0,583
Rezago t					
Con rezago (vs. Sin rezago)	1,662	0,210	0,000	***	1,250 2,073
Nivel matemática t					
Nivel II (vs.Nivel 1)	-0,514	0,236	0,029	**	-0,977 -0,051
Nivel III (vs.Nivel 1)	-0,825	0,316	0,009	***	-1,444 -0,205
Nivel IV (vs.Nivel 1)	-1,183	0,504	0,019	**	-2,172 -0,195
Expectativas culminación familiar t					
EMS (vs.EMB o menos)	-1,367	0,377	0,000	***	-2,105 -0,628
Educ Terciaria (vs.EMB o menos)	-2,113	0,468	0,000	***	-3,030 -1,197
Educ Universitaria (vs.EMB o menos)	-1,621	0,454	0,000	***	-2,510 -0,731
Sin dato (vs.EMB o menos)	-0,684	0,416	0,100	*	-1,499 0,130
Oferta/sector t+1					
Secundaria priv (vs. Secundaria púb)	-1,217	0,406	0,003	***	-2,014 -0,421
Educ. Técnica (vs. Secundaria púb)	-0,070	0,226	0,756		-0,512 0,372
6° Primaria o no insc (vs. Secundaria púb)	1,336	0,606	0,027	**	0,149 2,523
Sin dato (vs. Secundaria púb)	-1,368	0,903	0,130		-3,138 0,402
_cons	0,925	0,436	0,034		0,070 1,780

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Nota: El modelo incluye los niveles de desempeño en **Matemática**

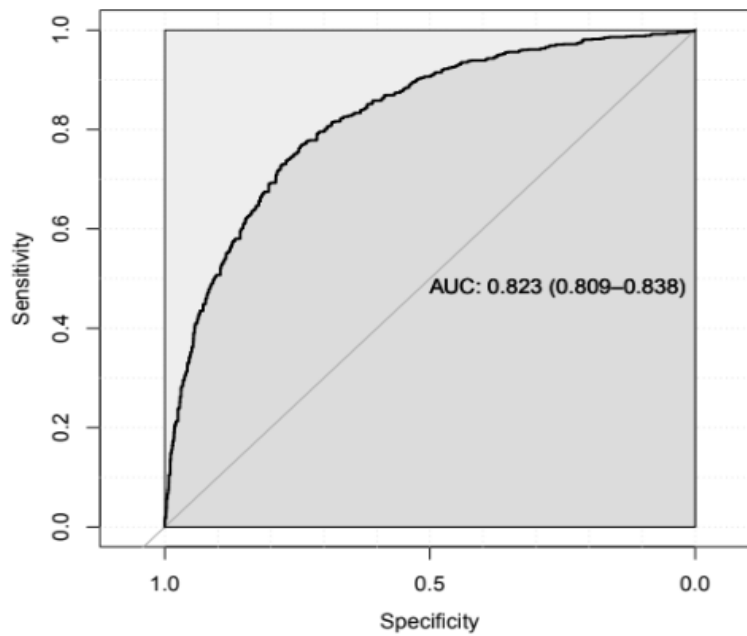
*** = Significación al 99%; ** = Significación al 95%; * = Significación al 90%

Ilustración A. 2. Curvas de ROC del Modelo Logit 2



Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Ilustración A. 3. Curvas de ROC del Modelo Logit 1.



Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Tabla A. 12. Métricas del rendimiento del modelo Logit 1 (L1), según el umbral de probabilidad elegido para clasificar a los desvinculados de los no desvinculados.

Umbral	Exactitud	Precisión	Sensibilidad	Especificidad	Puntaje F1	Puntaje F2	Kappa
0,100	0,672	0,391	0,879	0,614	0,541	0,70	0,58
0,150	0,733	0,442	0,830	0,706	0,577	0,71	0,66
0,170	0,760	0,474	0,827	0,741	0,602	0,72	0,69
0,180	0,763	0,477	0,819	0,748	0,603	0,72	0,70
0,190	0,766	0,480	0,808	0,754	0,603	0,71	0,70
0,200	0,778	0,496	0,796	0,773	0,611	0,71	0,72
0,250	0,806	0,542	0,751	0,822	0,630	0,70	0,75
0,300	0,820	0,573	0,701	0,853	0,630	0,67	0,77
0,310	0,825	0,584	0,695	0,861	0,635	0,67	0,78
0,350	0,832	0,610	0,657	0,882	0,632	0,65	0,79
0,36	0,837	0,621	0,656	0,888	0,638	0,649	0,791
0,37	0,837	0,626	0,637	0,893	0,632	0,635	0,791
0,38	0,837	0,626	0,634	0,894	0,630	0,632	0,791
0,39	0,835	0,628	0,614	0,898	0,621	0,616	0,789
0,40	0,841	0,646	0,611	0,906	0,628	0,618	0,796
0,45	0,846	0,690	0,539	0,932	0,605	0,564	0,802

Fuente: elaboración propia con base en el Seguimiento de la cohorte de 6° 2013. Panel TERCE 2013-2021.

Índice de abreviaturas

ANEP Administración Nacional de Educación Pública

AUC Area Under the Curve

DIEE División de investigación, Evaluación y Estadística

EM Educación Media

EMB Educación Media Básica

EMS Educación Media Superior

ERCE Estudio Regional Comparativo y Explicativo

ETC Escuelas de Tiempo Completo

GURI Gestión Unificada de Registros e Información

IECON Instituto de Economía

KH Capital Humano

INEED Instituto Nacional de Evaluación Educativa

LLECE Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación

PISA Programme for International Student Assessment

PISA-L Estudio longitudinal de los jóvenes evaluados por PISA

ROC Receiver Operating Characteristic

SAT Sistema de Alertas Tempranas

SERCE Segundo Estudio Regional Comparativo y Explicativo

TERCE Tercer Estudio Regional Comparativo y Explicativo

UDELAR Universidad de la República

UNESCO United Nations Educational, Scientific and Cultural Organization